This is to certify that the

dissertation entitled

Relations Among Conditional Entropy,
Topological Entropy and
Pointwise Preimage Entropy

presented by

Wen-Chiao Cheng

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Mathematics

Major professor

Date April 29, 2003

PLACE IN RETURN BOX to remove this checkout from your record.
TO AVOID FINES return on or before date due.
MAY BE RECALLED with earlier due date if requested.

DATE DUE	DATE DUE	DATE DUE

RELATIONS AMONG CONDITIONAL ENTROPY, TOPOLOGICAL ENTROPY AND POINTWISE PREIMAGE ENTROPY

By

WEN-CHIAO CHENG

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Mathematics

2004

ABSTRACT

RELATIONS AMONG CONDITIONAL ENTROPY, TOPOLOGICAL ENTROPY AND POINTWISE PREIMAGE ENTROPY

By

Wen-Chiao,Cheng

Entropy was introduced as a conjugacy invariant for measure-preserving transformations and continuous transformations. However, in 1995 Hurley, Nitecki, and Przytycki introduced several other entropy-like invariants for non-invertible maps. The purpose of this dissertation is to define and study two new invariants for non-invertible maps. Our new invariants are motivated by the some of those presented by Hurley, Nitecki, and Przytycki.

In Chapter 2 and Chapter 3 we introduce the standard notions of measure-theoretic entropy and topological entropy and we recall their basic properties. See[14]. We also describe two of the preimage entropy invariants studied by Hurley, Nitecki, and Przytycki. After that we introduce new invariants of non-invertible maps which we call the *upper pre-image entropy* and the *metric pre-image entropy* and study their properties. Among other things we obtain analogs of the well-known Variational Principle for Topological Entropy and Shannon-McMillan-Breiman theorem for these new invariants. The proofs require adaptation and modification of a number techniques in the literature of ergodic theory and topological dynamics.

ACKNOWLEDGMENTS

First of all I would like to thank my academic advisor, Dr. S. Newhouse, for his guidance, help and encouragement throught my studies at Michigan State University. I am especially grateful for his excellent lectures and the freedom he granted me in preparing my thesis. His knowledge and mathematical insight have been invaluable. I was constantly inspired by his enthusiasm and professional approach to research.

I would like to thank Dr. Z. Nitecki for providing useful materials regarding the preimage entropy and my dissertation committee members Professor Michael Frazier, Professor William Sledd, Professor Cliff Weil and Professor Zhengfang Zhou for their valuable suggestions and their time.

Contents

Chapter 1

Introduction

In 1958 Kolmogorov introduced the concept of entropy into ergodic theory, and this has been the most successful invariant so far. For example, in 1942 it was known that the two-sided $(\frac{1}{2}, \frac{1}{2})$-shift and the two-sided $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$-shift both have countable Lebesgue spectrum and hence are spectrally isomorphic, but it was not known whether they are conjugate. This was resolved in 1958 when Kolmogorov showed that they had entropies of $\log 2$ and $\log 3$, respectively, and hence are not conjugate. The notion of entropy now used is slightly differently from that used by Kolmogorov-the improvement was made by Sinai in 1959.

The topological entropy $h(T)$ of a continuous map T of a compact metric space to itself is a measure of its dynamical complexity. It was first defined by Adler, Konheim and McAndrew, and later given several equivalent definitions by Bowen and others (see [2] for an exposition) and these definitions led to results connecting topological and measure-theoretic entropies. Roughly speaking, the topological entropy of T measures the exponential growth rate with n of the number of different forward orbit segments of length n that can be distinguished to at least some finite tolerance.

When the mapping T under consideration is a homeomorphism, then extending this procedure into the past instead of the future results in the entropy $h(T^{-1})$ of the inverse mapping, which equals $h(T)$. However, when the map is not invertible, different ways of "extending the procedure into the past" lead to several new entropy-like invariants for non-invertible maps.

More recently, the preimage relation entropy $h_r(T)$ of a compact metric space was introduced by Langevin and Walczak (see [8].) and shown to be a new tool for studying the topology and dynamics of compact metric spaces. Later, Hurley and Nitecki, Przytycki (see [5] and [9].) introduced several other entropy-like invariants for non-invertible maps. One of these, which we call preimage branch entropy $h_i(T)$, is closely related to h_r(T). The other pair of entropy invariants is based on how many branches of the inverse of the iterated map T^{-n} at a point x can be distinguished by measurements of finite accuracy. We call them pointwise preimage entropies and denote $h_p(T)$ and $h_m(T)$. In [5] and [9], Hurley established the following relationships among these five invariants: $h_p(T) \leq h_m(T) \leq h(T) \leq h_i(T) + h_m(T) \leq h_r(T) + h_m(T)$, where T is a continuous mapping on a compact metric space.

In this dissertation we concentrate on $h_p(T)$ and $h_m(T)$, since in the context of pointwise preimage entropy, the definitions of $h_p(T)$ and $h_m(T)$ are in some sense analogous to (and were motivated by) Bowen's notion of "local entropy" (see [2].). In Chapter 2 and Chapter 3 we introduce measure-theoretic and topological entropies and show the variational principle as conclusion. That is to say when X is a compact metric space and $T : X \to X$ is continuous, the theorem shows that the supremum of measure-theoretic entropy $h_\mu(T)$ is equal to the topological entropy $h(T)$ of T. After that, we investigate basic properties of pointwise preimage entropies $h_p(T)$ and $h_m(T)$, such as forward generator and metric ρ compatible with the topology. Finally, we modify the definition of those two invariants in order to show the preimage S-M-B theorem. We also show the relationship between metric preimage entropy and upper preimage entropy. The main technique used is the construction of the Variational Principle by M. Misiurewicz. Finally, we follow the method introduced by Shannon, McMillan and Brieman and use it to show the asymptotic behavior of metric preimage entropy and ergodic decomposition.

Chapter 2

Measure-Theoretic Entropy

In this chapter we discuss measure-preserving transformations, measure-theoretic entropy and some of their basic properties. Finally we will show the existence of invariant measures. We refer to Peter Walters' book for this chapter. See[14].

2.1 Measure-Preserving Transformation

Definition 2.1.1 Suppose $(X_1, \mathcal{B}_1, m_1), (X_2, \mathcal{B}_2, m_2)$ are probability spaces.

(a) A transformation $T : X_1 \to X_2$ is measurable if $T^{-1}(\mathcal{B}_2) \subset \mathcal{B}_1$.

(b) A transformation $T : X_1 \to X_2$ is measure-preserving if T is measurable and $m_1(T^{-1}(B_2)) = m_2(B_2), \forall B_2 \in \mathcal{B}_2$.

(c) We say that $T : X_1 \to X_2$ is an invertible measure-preserving transformation if T is measure-preserving, bijective, and T^{-1} is also measure-preserving.

Remarks:

(1) We should write $T : (X_1, \mathcal{B}_1, m_1) \to (X_2, \mathcal{B}_2, m_2)$ since the measure-preserving property depends on the $\mathcal{B}'s$ and m's.

(2) If $T : X_1 \to X_2$ and $S : X_2 \to X_3$ are measure-preserving so is $S \circ T : X_1 \to X_3$.

(3) Measure-preserving transformations are the structure preserving maps (morphisms) between measure spaces.

(4) Let $(X_i, \bar{\mathcal{B}}_i, \bar{m}_i)$ denote the completion of $(X_i, \mathcal{B}_i, m_1), i = 1, 2$. If $T : (X_1, \mathcal{B}_1, m_1) \to$

$(X_2,\mathcal{B}_2,m_2)$ is measure-preserving, then so is $T:(X_1,\bar{\mathcal{B}}_1,\bar{m}_1)\to(X_2,\bar{\mathcal{B}}_2,\bar{m}_2)$.

(5) We shall be mainly interested in the case $(X_1,\mathcal{B}_1,m_1)=(X_2,\mathcal{B}_2,m_2)$ since we wish to study the iterates T^n. When $T: X\to X$ is a measure-preserving transformation of $(X,\mathcal{B},m)$ we say that T preserves m or that m is T-invariant.

In practice it would be difficult to check, using Definition 1.1, whether a given transformation is measure-preserving or not since one usually does not have explicit knowledge of all the members of $\mathcal{B}$. However we often do have explicit knowledge of a semi-algebra $\mathcal{T}$ generating $\mathcal{B}$. (For example, when X is the unit interval $\mathcal{T}$ may be the semi-algebra of all subintervals of X, and when X is a direct product space $\mathcal{T}$ may be the collection of all measurable rectangles.) The following result is therefore desirable in checking whether transformations are measure-preserving or not.

Theorem 2.1.1 Suppose $(X_1,\mathcal{B}_1,m_1),(X_2,\mathcal{B}_2,m_2)$ are probability spaces and $T: X_1\to X_2$ is a transformation. Let $\mathcal{T}_2$ be a semi-algebra which generates $\mathcal{B}_2$. If for each $A_2\in\mathcal{T}_2$ we have $T^{-1}(A_2)\in\mathcal{B}_1$ and $m_1(T^{-1}(A_2))=m_2(A_2)$, then T is measure-preserving.

Examples of Measure-Preserving Transformations

(1) The identity map I on $(X,\mathcal{B},m)$ is obviously measure-preserving.

(2) Let K be the unit circle and $\mathcal{B}$ be the σ-algebra of Borel subsets of K and let m be Haar measure. Let a be any fixed point in K and define $T:K\to K$ by $T(z)=a\cdot z$. Then T is measure-preserving since m is Haar measure. The transformation T is called a rotation of K.

(3) The transformation $T(x)=a\cdot x$ defined on any compact group G (where a is a fixed element of G) preserves Haar measure. Such transformations are called rotations of G.

(4) Any continuous endomorphism of a compact group onto itself preserves Haar measure. For example $T(z)=z^n$ preserves Haar measure on the unit circle if n is any non-zero integer.

(5) Any affine transformation of a compact group G preserves Haar measure. An

affine transformation is a map of the form $T(x) = a \cdot A(x)$ where a is any fixed element in G and $A : G \to G$ is a surjective endomorphism. It follows that T is measure-preserving because it is the composition of a rotation and an endomorphism. When dealing with affine transformations as measure-preserving transformations we always assume the measure involved is normalised Haar measure.

(6) Let $k \geq 2$ be a fixed integer and let $(p_0, p_1, ..., p_{k-1})$ be a probability vector with non-zero entries (i.e., $p_i > 0$ each i and $\sum_{i=0}^{k-1} p_i = 1$). Let $(Y, 2^Y, \mu)$ denote the measure space where $Y = \{0, 1, \cdots, k-1\}$ and the point i has measure p_i. Let $(X, \mathcal{B}, m) = \prod_{-\infty}^{\infty}(Y, 2^Y, \mu)$. Define $T : X \to X$ by $T(\{x_n\}) = \{y_n\}$ where $y_n = x_{n+1}$. If $\mathcal{F}$ denotes the semi-algebra of all measurable rectangles, then $m(T^{-1}A) = m(A), \forall A \in \mathcal{F}$. By Theorem 2.1.1, T is measure$-$preserving. We call T the two-sided $(p_0, p_1, \cdots, p_{k-1})$-shift. This is an example of an invertible measure-preserving transformation. We sometimes use the notation $(\cdots, x_{-1}\hat{x}_0 x_1, \cdots)$ for a point of X (the ˆ indicates the 0-th position in the product) and then T can be written $T((\cdots, x_{-1}\hat{x}_0 x_1, \cdots)) = (\cdots, x_{-1} x_0 \hat{x}_1 x_2, \cdots)$. The set Y is called the state space of the shift.

2.2 Partition and Entropy

Throughout this chapter $(X, \mathcal{B}, m)$ will denote a probability space.

Definition 2.2.1 A partition of $(X, \mathcal{B}, m)$ is a disjoint collection of elements of $\mathcal{B}$ whose union is X.

Here we shall be interested in finite partitions. They will be denoted by Greek letters, e.g., $\zeta = \{A_1, A_2, \cdots, A_k\}$. If ζ is a finite partition of $(X, \mathcal{B}, m)$, then the collection of all elements of $\mathcal{B}$ which are unions of elements of ζ is a finite sub-σ-algebra of $\mathcal{B}$. We denote it by $A(\zeta)$.

Definition 2.2.2 Suppose ζ and η are two partitions of $(X, \mathcal{B}, m)$. We write $\zeta \leq \eta$ to mean that each element of ζ is a union of elements of η. We have $\zeta \leq \eta \Leftrightarrow A(\zeta) \subseteq A(\eta)$.

Definition 2.2.3 Let $\zeta = \{A_1, A_2, ..., A_n\}, \eta = \{C_1, C_2, ..., C_k\}$ be two finite partitions of $(X, \mathcal{B}, m)$. Their join is the partition

$$\zeta \vee \eta = \{A_i \cap C_j : 1 \leq i \leq n, 1 \leq j \leq k\}.$$

Definition 2.2.4 Suppose $T : X \to X$ is a measure-preserving transformation. If $\zeta = \{A_1, A_2, ..., A_k\}$, then $T^{-n}\zeta$ denotes the partition $\{T^{-n}A_1, ..., T^{-n}A_k\}$ and if $\mathcal{A}$ is a sub-σ-algebra of $\mathcal{B}$, then $T^{-n}\mathcal{A}$ denotes the sub-σ-algebra $\{T^{-n}a : a \in \mathcal{A}\}$.

Definition 2.2.5 Let $\zeta = \{A_1, A_2, ..., A_k\}$ be a partition of $(X, \mathcal{B}, m)$. The entropy of ζ is the value $H(\zeta) = -\sum_{i=1}^{k} m(A_i) \log m(A_i)$.

Remarks:

(1) If $\zeta = \{A_1, ..., A_k\}$ where $m(A_i) = \frac{1}{k}, \forall i$ then

$$H(\zeta) = -\sum_{i=1}^{k} \frac{1}{k} \log \frac{1}{k} = \log k.$$

We will find that $\log k$ is the maximum value for the entropy of a partition with k sets.

(2) $H(\zeta) \geq 0$.

(3) If $T : X \to X$ is measure-preserving then $H(T^{-1}\zeta) = H(\zeta)$.

Conditional entropy is useful in deriving properties of entropy, and we discuss it now before we consider the entropy of a transformation.

Let A and C be two partitions on (X, B, m) with

$A = \{A_1, ..., A_k\}$ and $C = \{C_1, ..., C_p\}$

Definition 2.2.6 The entropy of A given C is the number

$$H(A \mid C) = -\sum_{j=1}^{p} m(C_j) \sum_{i=1}^{k} \frac{m(A_i \cap C_j)}{m(C_j)} \log \frac{m(A_i \cap C_j)}{m(C_j)}$$

omitting the j-terms when $m(C_j) = 0$.

So to get $H(A \mid C)$ one considers C_j as a measure space with normalized measure $m(\cdot) \mid m(C_j)$ and calculates the entropy of the partition of the set C_j induced by A and then averages the answer taking into account the size of C_j.

Theorem 2.2.1 Let $(X, \mathcal{B}, m)$ be a probability space. If A, C, D are partitions on X, then:

(1) $H(A \vee C \mid D) = H(A \mid D) + H(C \mid A \vee D)$.

(2) $H(A \vee C) = H(A) + H(C \mid A)$.

(3) $H(A) \geq H(A \mid D)$.

(4) $H(A \vee C \mid D) \leq H(A \mid D) + H(C \mid D)$.

(5) $H(A \vee C) \leq H(A) + H(C)$.

(6) If T is measure-preserving, then

$$H(T^{-1}A \mid T^{-1}C) = H(A \mid C), \text{ and } H(T^{-1}A) = H(A).$$

Now we extend this conditional entropy to more general situations. We let $\zeta = \{A_1, A_2, ...,\}$ be a countable partition of X into measurable sets. For each $x \in X$, denoted by $\zeta(x)$ the element of ζ to which x belongs. Then the information function associated to ζ is defined to be

$$I_\zeta(x) = -\log m(\zeta(x)) = -\sum_{A \in \zeta} \log m(A) \chi_A(x),$$

so that $I_\zeta(x)$ takes the constant value $-\log m(A)$ on the cell A of ζ. Clearly

$$H(\zeta) = \int_X I_\zeta(x)\, dm(x)$$

.

It is useful to consider conditional information and entropy, which take into account information that may already be in hand. Let $\Im$ be a sub-σ-algebra of $\mathcal{B}$. We can recall that for $\phi \in L^1(X)$, the conditional expectation $E(\phi \mid \Im)$ of ϕ given $\Im$ is an $\Im$-measurable function on X which satisfies

$$\int_F E(\phi \mid \Im)\, dm = \int_F \phi\, dm$$

for all $F \in \Im$; $E(\phi \mid \Im)(x)$ represents our expected value for ϕ if we are given the foreknowledge $\Im$. Thus we let $m(A \mid \Im) = E(\chi_A \mid \Im)$ and define the conditional information function of a countable partition ζ given a σ-algebra $\Im \subset \mathcal{B}$ to be

$$I_{\zeta|\Im}(x) = -\sum_{A \in \zeta} \log m(A \mid \Im)\chi_A(x) = -\sum_{A \in \zeta} \log m(A \mid \Im) m(A \mid \Im)$$

The conditional entropy of ζ given $\Im$ is defined by

$$H(\zeta \mid \Im) = \int_X I_{\zeta|\Im}(x)\, dm$$

Lemma 2.2.1 If α and β are countable measurable partitions of X and $\Im$ is a sub-σ-algebra of $\mathcal{B}$, then

$$I_{\alpha\vee\beta|\Im} = I_{\alpha|\Im} + I_{\beta|A(\alpha)\vee\Im}$$

where $A(\alpha)$ is the σ-algebra generated by α.

2.3 Entropy of a Measure-Preserving Transformation

Definition 2.3.1 Suppose $T : X \to X$ is a measure-preserving transformation of the probability space $(X, \mathcal{B}, m)$. If ζ is a partition of X, then

$$h(T, \zeta) = \lim_{n\to\infty} \frac{1}{n} H(\bigvee_{i=0}^{n-1} T^{-1}\zeta)$$

is called the entropy of T with respect to ζ.

This means that if we think of an application of T as a passage of one day of time, then $\bigvee_{i=0}^{n-1} T^{-i}\zeta$ represents the combined experiment of performing the original experiment, represented by ζ, on n consecutive days. Then $h(T, \zeta)$ is the average information per day that one gets from performing the original experiment daily forever. Now we can give the final stage of the defintion of the entropy of a measure-preserving transformation.

Definition 2.3.2 If $T : X \to X$ is a measure-preserving transformation of the probability space $(X, \mathcal{B}, m)$ then $h(T) = \sup h(T, \zeta)$, where the supremum is taken over all finite partitions ζ of X, is called the entropy of T.

If, as above, we think of an application of T as a passage of one day of time, then $h(T)$ is the maximum average per day obtainable by performing the same experiment daily.

Theorem 2.3.1 If $\{a_n\}_{n\geq 1}$ is a sequence of real numbers such that $a_{n+p} \leq a_n + a_p$ for all n, p then

$$\lim_{n\to\infty} \frac{a_n}{n}$$

exists and equals

$$\inf_n \frac{a_n}{n}.$$

Corollary 2.3.1 If $T : X \to X$ is measure-preserving and α is a finite partition of X, then $\lim_{n\to\infty} \frac{1}{n} H(\bigvee_{i=0}^{n-1} T^{-i}\alpha)$ exists.

Definition 2.3.3 Let T_i be a measure-preserving transformation of the probability space $(X_i, C_i, m_i), i = 1, 2$. We say that T_1 is *conjugate* to T_2 if there is a measure-algebra isomorphism $\phi : (\hat{C}_2, \hat{m}_2) \to (\hat{C}_1, \hat{m}_1)$ such that $\phi T_2 = T_1 \phi$.

Theorem 2.3.2 Entropy is a conjugacy invariant and hence an isomorphism invariant.

Theorem 2.3.3 Suppose A, C are finite partitions of $(X, \mathcal{B}, m)$ and T is a measure-preserving transformation of the probability space $(X, \mathcal{B}, m)$. Then

(1) $h(T, A) \leq H(A)$.

(2) $h(T, A \vee C) \leq h(T, A) + h(T, C)$.

(3) $h(T, A) \leq h(T, C) + H(A)$.

(4) $h(T, T^{-1}A) = h(T, A)$.

(5) If $k \geq 1$, $h(T, A) = h(T, \bigvee_{i=0}^{k-1} T^{-i}A)$.

(6) If T is invertible and $k \geq 1$, then $h(T, A) = h(T, \bigvee_{i=-k}^{k} T^{i}A)$.

Theorem 2.3.4 Let T be a measure-preserving transformation of the probability space (X,$\mathcal{B}$,m). Then:

(1) For $k > 0, h(T^k) = kh(T)$.

(2) If T is invertible then $h(T^k) = |k|h(T), \forall k \in Z$.

2.4 Some Methods for Calculating $h(T)$

It is difficult to calculate $h(T)$ from its definition because one would need to calculate $h(T, A)$ for every finite partition A. We consider what conditions on A are needed to ensure $h(T) = h(T, A)$. The result leads to methods of calculating $h(T)$ for specific examples of measure-preserving transformations and they also lead to proofs of further properties of $h(T)$.

Lemma 2.4.1 Let $r \geq 1$ be a fixed integer. For each $\epsilon > 0$ there exists $\delta > 0$ such that if $\xi = \{A_1, ..., A_r\}, \eta = \{C_1, ..., C_r\}$ are any two partitions of $(X, \mathcal{B}, m)$ into r sets with $\sum_{i=1}^{r} m(A_i \triangle C_i) < \delta$, then $H(\xi \mid \eta) + H(\eta \mid \xi) < \epsilon$.

Let C be a finite sub-σ-algebra of $\mathcal{B}$, say $C = \{C_i : i = 1, 2, ..., n\}$, then the non-empty sets of the form $B_1 \cap B_2 ... \cap B_n$, where $B_i = C_i$ or $X \backslash C_i$, form a finite partition of X. We denote it by $\alpha(C)$ and we define $h(T, C) = h(T, \alpha(C))$. If D is another finite sub-σ-algebra, then $H(C \mid D) = H(\alpha(C) \mid \alpha(D))$.

Lemma 2.4.2 Let $(X, \mathcal{B}, m)$ be a probability space and $\mathcal{B}_0$ be an algebra such that the σ-algebra generated by $\mathcal{B}_0$(denoted by $\mathcal{B}(\mathcal{B}_0)$) satisfies $\mathcal{B}(\mathcal{B}_0) = \mathcal{B}$. Let $\mathcal{C}$ be a finite sub-algebra of $\mathcal{B}$. Then for every $\epsilon > 0$, there exists a finite algebra $\mathcal{D}$, $\mathcal{D} \subset \mathcal{B}_0$ such that $H(\mathcal{D}|\mathcal{C}) + H(\mathcal{C}|\mathcal{D}) < \epsilon$.

Lemma 2.4.3 If $\{\mathcal{A}_n\}$ is an increasing sequence of finite sub-algebras of $\mathcal{B}$ and $\mathcal{C}$ is a finite sub-algebra with $\mathcal{C} \subset \bigvee_n \mathcal{A}_n$, then $H(\mathcal{C} | \mathcal{A}_n) \to 0$ as $n \to \infty$.

Thoerem 2.4.1 (Kolmogorov-Sinai Theorem)

Let T be an invertible measure-preserving transformation of the probability space $(X, \mathcal{B}, m)$ and let $\Re$ be a finite sub-algebra of $\mathcal{B}$ such that $\bigvee_{n=-\infty}^{\infty} T^n \Re = \mathcal{B}$. Then $h(T) = h(T, \Re)$.

Lemma 2.4.4 If T is a measure-preserving transformation of the probability space $(X, \mathcal{B}, m)$ and if $\mathcal{A}$ is a finite sub-algebra of $\mathcal{B}$ with $\bigvee_{i=0}^{\infty} T^{-i} \mathcal{A} = \mathcal{B}$ then $h(T) = h(T, \mathcal{A})$.

We shall now calculate the entropy of our examples.

(1) If $I : (X, \beta, m) \to (X, \beta, m)$ is the identity, then $h(I) = 0$. This is because $h(I, A) = \lim \frac{1}{n} H(A) = 0$. Also, if $T^p = I$ for some $p \neq 0$, then $h(T) = 0$. In particular any measure-preserving transformation of a finite space has zero entropy.

(2) **Theorem 2.4.2** Any rotation,$T(z) = az$, of the unit circle K has zero entropy.

(3) **Theorem 2.4.3** Any rotation of a compact metric abelian group has entropy zero.

Definition 2.4.1 Let $(X, \mathcal{B}, m)$ be a probability space. A measure-preserving transformation T of $(X, \mathcal{B}, m)$ is called ergodic if the only members B of $\mathcal{B}$ with $T^{-1}B = B$ satisfy $m(B) = 0$ or $m(B) = 1$.

Corollary 2.4.1 *Any* ergodic transformation with discrete spectrum has zero entropy.

(4) If A is an endomorphism of the n-torus K^n, then $h(A) = \sum \log |\lambda|$ where the summation is over all eigenvalues of the matrix $[A]$ with absolute value greater than one.

(5) **Theorem 2.4.4** The two-sided $\{p_0, \ldots, p_{k-1}\}$-shift has entropy $-\sum_{i=0}^{k-1} p_i \log p_i$.

Remark: The 2-sided $(\frac{1}{2}, \frac{1}{2})$-shift has entropy $\log 2$; the 2-sided $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$-shift has entropy $\log 3$. Thus these transformations can not be conjugate.

2.5 Bogolioubov Theorem

We call the members of $M(X)$ Borel probability measures on X. Each $x \in X$ determines a member δ_x of $M(X)$ defined by $\delta_x(A)$=1 if $x \in A$ and $\delta_x(A) = 0$, otherwise.

Lemma 2.5.1 Let m, μ be two Borel probability measures on the metric space X. If $\int_X f\, dm = \int_X f\, d\mu, \forall f \in C(X)$, then $m = \mu$.

We define a map $\hat{T} : M(X) \to M(X)$ give by $(\hat{T}\mu)(B) = \mu(T^{-1}B)$. We sometimes write $\mu \circ T^{-1}$ instead of $\hat{T}\mu$. We shall have the following.

Lemma 2.5.2

$$\int f\, d(\hat{T}\mu) = \int f \circ T\, d\mu, \forall f \in C(X).$$

We are interested in those members of $M(X)$ that are invariant measures for T.

Let $M(X,T) = \{\mu \in M(X) | \hat{T}\mu = \mu\}$. This set consists of all $\mu \in M(X)$ making T a measure-preserving transformation of $(X, \mathcal{B}, \mu)$. The following gives us a method of constructing members of $M(X,T)$.

Theorem 2.5.1 (Bogolioubov Theorem) Let $T : X \to X$ be continuous. If $\{\sigma_n\}_{n=1}^{\infty}$ is a sequence in $M(X)$ and we define a new sequence $\{\mu_n\}_{n=1}^{\infty}$ by $\mu_n = \frac{1}{n}\sum_{i=0}^{n-1} \hat{T}^i \sigma_n$, then any limit point μ of $\{\mu_n\}$ is a member of $M(X,T)$.(Such limit points exist by the compactness of $M(X)$.)

Corollary 2.5.1 If $T : X \longrightarrow X$ is a continuous map of a compact metric space X, then $M(X,T)$ is non-empty.

Chapter 3

Topological Entropy

Adler, Konheim, and McAndrew introduced topological entropy as an invariant of topological conjugacy and also as an analogue of measure theoretical entropy. To each continuous transformation $T : X \to X$ of a compact topological space a non-negative real number or ∞, denoted by $h(T)$, is assigned. Later Dinaburg and Bowen gave a new, but equivalent, definition and this definition led to proofs of the result connecting topological and measure-theoretic entropies. For these materials we recommend Peter Walters' book. See [14].

3.1 Definition Using Open Covers

Let X be a compact topological space. We shall be interested in open covers of X which we denote by $\alpha, \beta, ...$

Definition 3.1.1 If α, β are open covers of X their join $\alpha \vee \beta$ is the open cover by all sets of the form $A \cap B$ where $A \in \alpha, B \in \beta$. Similarly we can define the join $\bigvee_{i=1}^{n} \alpha_i$ of any finite collection of open covers of X.

Definition 3.1.2 An open cover β is a refinement of an open cover α, written $\alpha < \beta$, if every member of β is a subset of a member of α.

Hence $\alpha < \alpha \vee \beta$ for any open covers α, β. Also if β is a subset of α, then $\alpha < \beta$.

Definition 3.1.3 If α is an open cover of X and $T : X \to X$ is continuous, then $T^{-1}\alpha$ is the open cover consisting of all sets $T^{-1}A$ where $A \in \alpha$.

Definition 3.1.4 If α is an open cover of X, let $\aleph(\alpha)$ denote the number of sets in a finite subcover of α with smallest cardinality. We define the entropy of α by $H(\alpha) = \log \aleph(\alpha)$.

Remarks:

(1) $H(\alpha) \geq 0$

(2) $H(\alpha) = 0$ iff $\aleph(\alpha) = 1$ iff $X \in \alpha$.

(3) If $\alpha < \beta$, then $H(\alpha) \leq H(\beta)$.

(4) $H(\alpha \vee \beta) \leq H(\alpha) + H(\beta)$.

(5) If $T : X \to X$ is a continuous map, then $H(T^{-1}\alpha) \leq H(\alpha)$. If T is also surjective, then $H(T^{-1}\alpha) = H(\alpha)$.

Theorem 3.1.1 If α is an open cover of X and $T : X \longrightarrow X$ is continuous, then $\lim_{n\to\infty} \frac{1}{n} H(\bigvee_{i=0}^{n-1} T^{-i}\alpha)$ exists.

Definition 3.1.5 If α is an open cover of X and $T : X \to X$ is a continuous map, then the entropy of T relative to α is given by

$$h(T, \alpha) = \lim_{n\to\infty} \frac{1}{n} H(\bigvee_{i=0}^{n-1} T^{-i}\alpha)$$

Remarks:

(6) $h(T, \alpha) \geq 0$.

(7) If $\alpha < \beta$, then $h(T, \alpha) \leq h(T, \beta)$.

(8) $h(T, \alpha) \leq H(\alpha)$.

Definition 3.1.6 If $T : X \to X$ is continuous, the topological entropy of T is given by:

$$h(T) = \sup_{\alpha} h(T, \alpha)$$

where α ranges over all open covers of X.

Remarks:

(9) $h(T) \geq 0$.

(10) In the defintion of $h(T)$ one can take the supremum over finite open covers of X.

(11) $h(I) = 0$ where I is the identity map of X.

(12) If Y is a closed subset of X and $TY = Y$ then $h(T|Y) \leq h(T)$.

The next result shows that topological entropy is an invariant of topological conjugacy.

Theorem 3.1.2 If X_1, X_2 are compact spaces and $T_i : X_i \to X_i$ are continuous for $i = 1, 2$, and if $\phi : X_1 \to X_2$ is a continuous map with $\phi X_1 = X_2$ and $\phi T_1 = T_2 \phi$, then $h(T_1) \geq h(T_2)$. If ϕ is a homeomorphism, then $h(T_1) = h(T_2)$.

In the next section we shall give a definition of $h(T)$ that does not require X to be compact and we give a definition of $h(T)$ in this more general setting. However, one result that is false when X is not compact is the following.

Theorem 3.1.3 If $T : X \to X$ is a homeomorphism of a compact space X, then $h(T) = h(T^{-1})$.

3.2 Bowen's Definition

In this section we give the definition of topological entropy using separating and spanning sets. This was done by Dinaburg and by Bowen, but Bowen also gave the definition when the space X is not compact and this will prove useful later. We shall give the definition when X is a metric space but the definition can easily be formulated when X is a uniform space. See [3].

In this section (X, d) is a metric space, not necessarily compact. The open ball centre x radius r will be denoted by $B(x; r)$ and the closed ball by $\bar{B}(x, r)$. Our definitions will depend on the metric d on X; we shall see later what the dependence on d is.

Throughout this section T will denote a fixed continuous function. If n is a natural number, we can define a new metric d_n on X by $d_n(x, y) = \max_{0 \leq i \leq n-1} d(T^i(x), T^i(y))$. (The notation does not show the dependence on T.) The open ball centre x and radius r in the metric d_n is $\bigcap_{i=0}^{n-1} T^{-i} B(T^i x; r)$.

Definition 3.2.1 Let n be a natural number, $\epsilon > 0$ and let K be a compact subset of X. A subset F of X is said to a (n,ϵ) span K with respect to T if $\forall x \in K \exists y \in F$ with $d_n(x,y) \leq \epsilon$, i.e.,

$$K \subset \bigcup_{y\in F} \bigcap_{i=0}^{n-1} T^{-i}\bar{B}(T^i y;\epsilon).$$

If n is a natural number, $\epsilon > 0$ and K is a compact subset of X let $r_n(\epsilon, K)$ denote the smallest cardinality of any $(n,\epsilon)-$spanning set for K with respect to T.

Remark: Clearly $r_n(\epsilon, K) < \infty$ because the compactness of K implies the covering of K by the open sets $\bigcap_{i=0}^{n-1} T^{-i}B(T^i x;\epsilon), x \in X$, has a finite subcover.

Definition 3.2.2 If $\epsilon > 0$ and K is a compact subset of X, let

$$r(\epsilon, K, T) = \limsup_{n\to\infty} \frac{1}{n} \log r_n(\epsilon, K).$$

We write $r(\epsilon, K, T, d)$ if we wish to emphasis the metric d.

Definition 3.2.3 If K is a compact subset of X, let $h(T,K) = \lim_{\epsilon\to 0} r(\epsilon, K, T)$. The topological entropy of T is $h(T) = \sup_K h(T,K)$, where the supremum is taken over the collection of all compact subsets of X. We sometimes write $h_d(T)$ to emphasis the dependence on d.

Before giving any interpretations or explanations of this definition we shall give an equivalent but "dual" definition. This definition will use the idea of separated sets which is dual to the notation of spanning sets.

Definition 3.2.4 Let n be a natural number, $\epsilon > 0$ and K be a compact subset of X. A subset E of K is said to be (n,ϵ) separated with respect to T if $x, y \in E, x \neq y$, implies $d_n(x,y) > \epsilon$, i.e., for $x \in E$ the set $\bigcap_{i=0}^{n-1} T^{-i}\bar{B}(T^i x;\epsilon)$ contains no other point of E.

Definition 3.2.5 If n is a natural number, $\epsilon > 0$ and K is a compact subset of X, let $s_n(\epsilon, K)$ be the largest cardinality of any (n,ϵ) separated subset of K with respect to T. We write $s_n(\epsilon, K, T)$ to emphasis T if we need to.

Remark: We have $r_n(\epsilon, K) \leq s_n(\epsilon, K) \leq r_n(\frac{\epsilon}{2}, K)$ and hence $s_n(\epsilon, K) < \infty$.

Definition 3.2.6 If $\epsilon > 0$ and K is a compact subset of X put

$$s(\epsilon, K, T) = \limsup_{n\to\infty} \frac{1}{n} \log s_n(\epsilon, K)$$

We sometimes write $s(\epsilon, K, T, d)$ when we need to emphasis *the metric d.*

Remark: The ideas for the definition come from the work of Kolmogorov on the size of a metric space. If (X, ρ) is a metric space, then a subset F is said to an ϵ-span of X if $\forall x \in X \exists y \in F$ with $\rho(x, y) \leq \epsilon$, and a subset E is said to be ϵ-separated if whenever $y, z \in E, y \neq z$, then $\rho(y, z) > \epsilon$. The ϵ-entropy of (X, ρ) is then the logarithm of the minimum number of elements of an ϵ-spanning set and the ϵ-capacity is the logarithm of the maximum number of elements in an ϵ-separated set. So in the definition 3.2.6, we are considering the metric spaces (K, d_n) and $r_n(\epsilon, K)$ is the ϵ-entropy of (K, d_n) and $s_n(\epsilon, K)$ is the ϵ-capacity of (K, d_n) where the ϵ-entropy is the logarithm of the minimum number of elements of an ϵ-spanning set and the ϵ-capacity is the logarithm of maximum number of elements in an ϵ-separated set. Therefore,

$$\begin{aligned} h(T, K) &= \lim_{\epsilon\to 0} \limsup_{n\to\infty} \frac{1}{n}(\epsilon\text{-entropy of } (K, d_n)) \\ &= \lim_{\epsilon\to 0} \limsup_{n\to\infty} \frac{1}{n}(\epsilon\text{-capacity of } (K, d_n)) \end{aligned}$$

We shall now observe that the definition of $h(T)$ in this section coincides with that given in Section 3.1 when T is a continuous map of a compact metrisable space. For the moment let us denote by $h^*(T)$ and $h^*(T, \alpha)$ the numbers occuring in the definition of topological entropy using open covers. In a metric space (X, d) we define the diameter of a cover to be $\operatorname{diam}(\alpha) = \sup_{A\in\alpha} \operatorname{diam}(A)$, where $\operatorname{diam}(A)$ denotes the diameter of the set A. If α, γ are open covers of X and $\operatorname{diam}(\alpha)$ is less than a Lebesgue number for γ then $\gamma < \alpha$. The following result is often useful for calculating $h^*(T)$.

Theorem 3.2.1 Let (X, d) be a compact metric space. If $\{\alpha_n\}_1^\infty$ is a sequence of open covers of X with $\operatorname{diam}(\alpha_n) \to 0$, then if $h^*(T) < \infty, \lim_{n\to\infty} h^*(T, \alpha_n)$ exists and equals $h^*(T)$, and if $h^*(T) = \infty$, then $\lim_{n\to\infty} h^*(T, \alpha_n) = \infty$.

The next result gives the basic relationship between the two ways of defining topological entropy.

Theorem 3.2.2 Let $T : X \to X$ be a continuous map of a compact metric space (X, d).

(1) If α is an open cover of X with Lebesgue number δ, then

$$\aleph(\bigvee_{i=0}^{n-1} T^{-i}\alpha) \le r_n(\delta/2, X) \le s_n(\delta/2, X).$$

(2) If $\epsilon > 0$ and γ is an open cover with $\mathrm{diam}(\gamma) \le \epsilon$, then

$$r_n(\epsilon, X) \le s_n(\epsilon, X) \le \aleph(\bigvee_{i=0}^{n-1} T^{-i}\gamma).$$

Theorem 3.2.3 If $T : X \to X$ is a continuous map of the compact metric space (X,d), then $h(T) = h^*(T)$; i.e., the two definitions of topological entropy coincide.

Theorem 3.2.4

(1) If (X, d) is a metric space, T is a continuous map and $m > 0$, then $h_d(T^m) = m \cdot h_d(T)$.

(2) Let (X_i, d_i), $i = 1, 2$ be a compact metric space and T_i is continuous on X_i. Define a metric d on $X_1 \times X_2$ by $d((x_1, x_2), (y_1, y_2)) = \max\{d_1(x_1, y_1), d_2(x_2, y_2)\}$. Then $h_d(T_1 \times T_2) = h_{d_1}(T_1) + h_{d_2}(T_2)$.

3.3 Calculation of Topological Entropy

Theorem 3.2.1 provided the only method we have given so far for calculating the topological entropy of examples. The following is an analogue of the Kolmogorov-Sinai theorem and provides a method of calculating topological entropy for some examples.

Theorem 3.3.1 Let $T : X \to X$ be an expansive homeomorphism of the compact metric space (X, d).

(1) If α is a generator for T then $h(T) = h(T, \alpha)$.

(2) If δ is an expansive constant for T, then $h(T) = r(\delta_0, T) = s(\delta_0, T)$ for all $\delta_0 < \delta/4$.

Corollary 3.3.1 An expansive homeomorphism has finite topological entropy.

Theorem 3.3.2 The two-sided shift on $X = \Pi_{-\infty}^{\infty} Y$, where $Y = \{0, 1, 2, \cdots, k-1\}$, has topological entropy $\log k$.

Theorem 3.3.3 Let $T : X \to X$ be the two-sided shift on $X = \Pi_{-\infty}^{\infty} Y$ where $Y = \{0, 1, \cdots, k-1\}$. Then

(1) If X_1 is a closed subset of X with $TX_1 = X_1$, then $h(T|_{X_1}) = \lim_{n\to\infty} \frac{1}{n} \log \theta_n(X_1)$, where $\theta_n(X_1)$ is the number of n-tuples $[i_0, i_1, \ldots, i_{n-1}]$ such that the set $\{\{x_n\}_{-\infty}^{\infty} \in X_1 | x_0 = i_0, \ldots, x_{n+1} = i_{n+1}\}$ is non-empty.

(2) Let $T_A : X_A \to X_A$ denote the topological Markov chain given by an irreducible $k \times k$ matrix A whose entries belong to $\{0,1\}$. Then $h(T_A) = \log \lambda$ where λ is the largest positive eigenvalue of A.

The corresponding one-sided results are true. Part (2) holds also when A is reducible by arranging the matrix A in lower diagonal block form as in the theory of Markov chains.

Remark: There is a transformation with topological entropy equal to any given positive real number.

We already know that a rotation T of a compact metric group G has zero topological entropy because there is a metric on G making T an isometry. In fact we now show any homeomorphism of K has zero entropy where K is the unit circle.

Theorem 3.3.4 If $T : K \to K$ is a homeomorphism of the unit circle, then $h(T) = 0$.

Corollary 3.3.1 Any homeomorphism of $[0,1]$ has zero topological entropy.

3.4 The Variational Principle

In this section we describe the basic relationship between topological entropy and measure-theoretic entropy. If T is a continuous map of a compact metric space, then $h(T) = \sup\{h_\mu(T) | \mu \in M(X,T)\}$. The inequality $\sup\{h_\mu(T) | \mu \in M(X,T)\} \leq h(T)$ was proved by L.W. Goodwyn in 1968. In 1970 E. I. Dinaburg proved equality when X has finite covering dimension and later in 1970, T. N. T. Goodman proved equality in the general case. See [3].

We shall need the following simple lemma, where we use ∂B to denote the boundary of a set B.

Lemma 3.4.1 Let X be a compact metric space and $\mu \in M(X)$.

(1) If $x \in X$ and $\delta > 0$ there exists $\acute{\delta} < \delta$ such that $\mu(\partial B(x;\delta)) = 0$.

(2) If $\delta > 0$, there is a finite partition $\xi = \{A_1, \cdots, A_k\}$ such that $\mathrm{diam}(A_j) < \delta$ and $\mu(\partial A_j) = 0$ for each j.

We now collect together some results we will use in the proof of the variational principle. In this section X will always denote a compact metric space and $\mathcal{B}(X)$ the σ-algebra of Borel subsets.

Remarks:

(1) If $\mu_i \in M(X)$, $1 \le i \le n$, and $p_i \ge 0, \sum_{i=1}^n p_i = 1$, then

$$H_{\sum_{i=1}^n p_i\mu_i}(\xi) \ge \sum_{i=1}^n p_i H_{\mu_i}(\xi)$$

for any finite partition ξ of $(X, \mathcal{B}(X))$.

(2) Suppose q, n are natural numbers and $1 < q < n$. For $0 \le j \le q-1$ put $a(j) = [\frac{(n-j)}{q}]$. Here $[b]$ denotes the integer part of $b > 0$. We have the following facts

- $a(0) \ge a(1) \ge \cdots \ge a(q-1)$.
- Fix $0 \le j \le q-1$. Then $\{0,1,2,...,n-1\} = \{j + rq + i / 0 \le r \le a(j)-1, 0 \le i \le q-1\} \cup S$ where $S = \{0, 1, ..., j-1, j + a(j)q, j + a(j)q + 1, ..., n-1\}$. Since $j + a(j)q \ge j + [(\frac{(n-j)}{q}) - 1]q = n - q$, we have the cardinality of S is at most $2q$.
- For each $0 \le j \le q-1$, $(a(j)-1)q + j \le [(\frac{(n-j)}{q}) - 1] + j = n - q$. The numbers $\{j + rq | 0 \le j \le q-1, 0 \le r \le a(j)-1\}$ are all distinct and are all no greater than $n - q$.

(3) If $\mu \in M(X,T)$ and if $\mu(\partial A_i) = 0, 0 \le i \le n-1$, then

$$\mu(\partial(\bigcap_{i=0}^{n-1} T^{-i}A_i)) = 0 \text{ since } \partial(\bigcap_{i=0}^{n-1} T^{-i}A_i) \subset \bigcup_{i=0}^{n-1} T^{-i}\partial A_i.$$

Theorem 3.4.1 (Variational Principle) Let $T : X \to X$ be a continuous map of a compact metric space X. Then

$$h(T) = \sup\{h_\mu(T) | \mu \in M(X,T)\}.$$

3.5 Measures with Maximal Entropy

The variational principle gives a natural way to pick out some members of $M(X,T)$.

Definition 3.5.1 Let $T : X \to X$ be a continuous transformation on a compact metric space X. A member μ of $M(X,T)$ is called a measure of maximal entropy for T if $h_\mu(T) = h(T)$. And we let $M_{\max}(X,T)$ denote the collection of all measures with maximal entropy for T.

Theorem 3.5.1 Let $T : X \to X$ be a continuous transformation of a compact metric space. Then

(1) $M_{\max}(X,T)$ *is convex.*

(2) If $h(T) < \infty$ *the extreme* points of $M_{\max}(X,T)$ are precisely the ergodic members of $M_{\max}(X,T)$.

(3) If $h(T) < \infty$ and $M_{\max}(X,T) \neq \emptyset$ then $M_{\max}(X,T)$ contains an ergodic measure.

(4) If $h(T) = \infty$ then $M_{\max}(X,T) \neq \emptyset$.

(5) If the entropy map is upper semi-continuous, then $M_{\max}(X,T)$ is compact and non-empty.

Definition 3.5.2 A continuous transformation $T : X \to X$ of a compact metric space X is said to have a unique measure with maximal entropy if $M_{max}(X,T)$ consists of exactly one member. Such transformations are also called intrinsicially ergodic.

Remarks:

(1) If T is uniquely ergodic and $M(X,T) = \{\mu\}$ then T has a unique measure with maximal entropy, because the variational principle gives $h_\mu(T) = h(T)$ in this case.

(2) If $h(T) = \infty$ and T has a unique measure with maximal entropy, then T is uniquely ergodic, because if $M_{\max}(X,T) = \{\mu\}$ and $m \in M(X,T)$, then $h_{\frac{\mu}{2}+\frac{m}{2}}(T) = \infty$ so $m = \mu$.

(3) If $M_{\max}(X,T) = \{\mu\}$ then μ is ergodic. If $h(T) = \infty$ this follows from (2) and if $h(T) < \infty$ it follows from Theorem 3.5.1.

Chapter 4

Pointwise Preimage Entropy

In this chapter we first introduce pointwise preimage entropies $h_p(T)$ and $h_m(T)$ which are defined in [5] and [9]. After that, we investigate basic properties of those two invariants, show the existence of forward generator and metric ρ compatible with the topology.

4.1 Definitions of Pointwise Preimage Entropy

Definition 4.1.1 Suppose $T : X \to X$ and $x \in X$. For k=1,2,3,..., the k^{th} preimage set of x under T is the subset $T^{-k}(x)$ of X where $T^{-k}(x) = \{z \in X | T^{k}(z) = x\}$. For N=1,2,..., the N^{th} branch at x is denoted by $B_N(x,T) \subset X^N$ and is defined in the following:

$$B_N(x,T) = \{(z_N, z_{N-1}, ..., z_0) | T(z_{i+1}) = z_i, 0 \leq i \leq N-1 \text{ and } z_0 = x\}$$

To formulate a topological definition, we let $O(X)$ be the collection of all open covers of this compact metric space X (finite or infinite). Given $U \in O(X)$, let U^N be the open cover of X^N by product sets $U_1 \times U_2 \times ... \times U_N, U_i \in U$. For a subset $S_N \subset X^N$, define $\aleph(U, N, S_N)$ to be the least cardinality among subcollections of U^N which can cover S_N.

Definition 4.1.2 (Pointwise preimage entropies).

Let $T : X \to X$ be a continuous mapping from a compact space X to itself, define

$$h_p(T) = \sup_{x \in X}\{\sup_{U \in O(X)}[\limsup_{N\to\infty}\frac{1}{N}\log \aleph(U, N, B_N(x,T))]\}$$

and

$$h_m(T) = \sup_{U \in O(X)}\{\limsup_{N\to\infty}\frac{1}{N}\log[\sup_{x\in X}\aleph(U, N, B_N(x,T))]\}$$

Remark 4.1.1 Continuity of T and compactness of X insure that $B_N(x,T)$ is compact, and hence that the numbers $\aleph(U,N,B_N(x,T))$ are all finite and bounded for fixed N over $x \in X$.

Like the topological entropy, we can show the metric definitions of our invariants by reinterpreting the numbers $\aleph(U,N,S_N)$ in terms of ϵ-spanning and ϵ-separated sets. Given any metric space (X,d), we say a subset $S \subset X$ is ϵ-separated for some $\epsilon > 0$ if distinct points of S are at least ϵ-apart:$s \neq t \in S, \Rightarrow d(s,t) \geq \epsilon$, and say that $R \subset A \subset X$ ϵ-spans A if for every $a \in A$, there exists $r \in R$ with $d(a,r) < \epsilon$. Let

$$r(\epsilon, d, A) = \min\{\mathrm{card}(R) \mid R \text{ is } \epsilon\text{-spans } A\},$$
$$s(\epsilon, d, A) = \max\{\mathrm{card}(S) \mid S \text{ is } \epsilon\text{-separated } A\}.$$

Theorem 4.1.1[9] If (X,d) is a compact metric space, for any positive integer N let d^N be the metric on X^N given by

$$d^N((x_1,...,x_N),(y_1,...,y_N)) = \max_{1\leq i\leq N} d(x_i,y_i)$$

Then for $T : X \to X$ continuous, the invariants from definition 4.1.2 can be calculated via the following.

$$h_p(T) = \sup_{x\in X}\{\lim_{\epsilon\to 0}(\limsup_{N\to\infty}\frac{1}{N}\log(s(\epsilon, d^N, B_N(x,T))))\},$$

and

$$h_m(T) = \lim_{\epsilon\to 0}\{\limsup_{N\to\infty}\frac{1}{N}\log(\sup_{x\in X} s(\epsilon, d^N, B_N(x,T)))\}$$

In either formula,

$$s(\epsilon, d^N, B_N(x,T))$$

can be replaced by $r(\epsilon, d^N, B_N(x,T))$

In topological entropy we define a new metric d_n on X by

$$d_n(x, y) = \max_{0 \le i \le n-1} d(T^i(x), T^i(y))$$

A subset F of X is said to (n, ϵ) span K if for all $x \in K$, exist $y \in F$ with $d_n(x, y) \le \epsilon$ and let $r_n(\epsilon, K)$ denote the smallest cardinality of any (n, ϵ)-spanning set for K. Similar definition for (n, ϵ) separated set and $s_n(\epsilon, K)$. We denote $\aleph(U)|_Y$ to be the smallest cardinality of subsets in U which covers Y. See [14].

Remark 4.1.2 Let $U \in$O(X), some easy consequences are the following
$r(\epsilon, d^N, B_N(x, T)) = r_N(\epsilon, T^{-N}(x))$, $s(\epsilon, d^N, B_N(x, T)) = s_N(\epsilon, T^{-N}(x))$
and $\aleph(U, N, B_N(x, T)) = \aleph(\bigvee_{n=0}^{N} T^{-n}U)|_{T^{-N}(x)}$.

Remark 4.1.3 If T is a homeomorphism, then $h_p(T) = h_m(T) = 0$.

Remark 4.1.4[7] and [8] If X is the circle or any closed interval, then $h_p(T) = h_m(T) = h(T)$.

Remark 4.1.5[4] There exists $T : X \to$X continuous, X a zero-dimensional compact metric space, for which $h_p(T) = 0$ and $h_m(T) > 0$.

Theorem 4.1.2[9] If $T_1 : X \to X$ and $T_2 : Y \to Y$ are topologically conjugate, then $h_p(T_1) = h_p(T_2)$ and $h_m(T_1) = h_m(T_2)$.

Remark 4.1.6 Like topological entropy property, the next one is trivial. Also, in section 4.3, we will concern another metric compatible with the topology of X and represent its pointwise preimage entropy with respect to this metric.

Theorem 4.1.3 If $\acute{d}$ is another metric on compact set X which defines the same topology as d, then the pointwise preimage entropy with respect to d are equal to the pointwise preimage entropy with respect to $\acute{d}$.

In theorem 4.1.2, if T_2 is a factor of T_1 then $h(T_2) \le h(T_1)$. However this inequality can not hold for pointwise preimage entropy. The easiest example of increase under factors for the pointwise preimage entropies is obtained via inverse limits. Recall that the inverse limit of the map $f : X \to X$ is the shift σ_f defined on the sequence space

$$\sum_f = \{\{x_i\}_{i=0}^{\infty} \mid f(x_i) = x_{i-1}, i = 1, 2, ...\}$$

by

$$\sigma_f(x_0, x_1, ...) = (f(x_0), f(x_1), ...) = (f(x_0), x_0, x_1, ...)$$

The product topology on $\sum_f \subset X^N$ makes $\sum_f$ compact and σ_f a homeomorphism. Furthermore, if f is surjective, then it is a factor of its inverse limit via the projection

$$\varphi(\{x_i\}_{i=0}^{\infty}) = x_0.$$

By Remark 4.1.3, we have

$$h_p(\sigma_f) = h_m(\sigma_f) = 0.$$

Thus, any map f with $h_p(f) = h_m(f) > 0$ gives an example showing

Remark 4.1.7 There exist maps $f : X \to X, g : Y \to Y$ with f a factor of g and

$$h_m(f) = h_p(f) > h_m(g) = h_p(g).$$

An easy example is the standard expanding map of the circle:set $S^1 = R/Z$ and define f(x+Z)=2x+Z. It is easy to check that

$$h_p(f) = h_m(f) = \log 2.$$

We turn now to Cartesian products and additivity. Subadditivity of all two invariants is relatively easy to prove:

Lemma 4.1.1 For any continuous maps $T_i : X_i \to X_i, i = 1, 2,$ we have

$$h_\alpha(T_1 \times T_2) \leq h_\alpha(T_1) + h_\alpha(T_2)$$

where $\alpha = p$ or m.

Topological entropy is multiplicative under iterates and we can show that the same is true for pointwise preimage entropy.

Lemma 4.1.2 Suppose $T : X \to X$ is continuous, where X *is a compact metric space.* Then for every $k \in N$, we have

$$h_\alpha(T^k) = k \cdot h_\alpha(T)$$

where $\alpha = p$ or m.

4.2 Forward Generator

After finishing the first version of this section, we found that D.Fiebig, U.Fiebig and Z. Nitecki used the tool of graph theory to get similar but better results. See [4].

Definition 4.2.1 Let X be a compact metric space and $T : X \to X$ a continuous function. A finite open cover α of X is a forward generator for T if for every sequence $(A_n)_0^\infty$ of members of α the set $\bigcap_{n=0}^\infty T^{-n}\bar{A}_n$ contains at most one point of X.

Lemma 4.2.1 Let $T : X \to X$ be a continuous function with a compact metric space (X, d). Let α be a forward generator for T. Then for any $\epsilon > 0$, $\exists N > 0$ such that each set in $\bigvee_{n=0}^N T^{-n}\alpha$ has diameter less than ϵ.

PROOF

Suppose the theorem does not hold. Then $\exists\epsilon > o$ such that $\forall j > 0, \exists x_j, y_j, d(x_j, y_j) > \epsilon$ and $\exists A_{j,i} \in \alpha, 0 \leq i \leq j$ with $x_j, y_j \in \bigcap_{i=0}^j T^{-i}A_{j,i}$. There is a subsequence $\{j_k\}$ of natural numbers such that $x_{jk} \to x$ and $y_{jk} \to y$ since X is compact. We have $x \neq y$. Consider the sets $A_{jk,0}$. Infinitely many of them coincide since α is finite. Thus $x_{jk}, y_{jk} \in A_0$, say, for infinitely many k and hence $x, y \in \bar{A}_0$. Similarly, for each n, infinitely many $A_{jk,n}$ coincide and we obtain $A_n \in \alpha$ with $x, y \in T^{-n}\bar{A}_n$. Thus

$$x, y \in \bigcap_0^\infty T^{-n}\bar{A}_n,$$

contradicting the fact that α is a forward generator. $\diamond$

Definition 4.2.2 Let T be a continuous function of a compact metric space (X, d)

to itself is said to be forward expansive if $\exists\delta>0$ and x$\neq y \in$ X, then $\exists n \in N$ with d$(T^n x, T^n y)> \delta$. We call δ a forward expansive constant.

Lemma 4.2.2 Let T be a continuous function of a compact metric space (X, d) to itself. Then T is forward expansive iff T has a forward generator.

PROOF

Let δ be a forward expansive constant for T and let α be a finite cover by open balls of radius $\delta/2$. Suppose that $x, y \in \bigcap_0^\infty T^{-n}\bar{A}_n$ where $A_n \in \alpha$. Then $d(T^n(x), T^n(y)) \leq \delta$ for all $n \in$ N$\bigcup$[0] so, by assumption $x = y$. Then α is a forward generator.

Conversely, suppose α is a forward generator. Let δ be a Lebesgue number for α, If $d(T^n(x), T^n(y)) \leq \delta$ for all $n \in$N$\bigcup$[0], then for all $n \in$N exists $A_n \in \alpha$ with $T^n(x), T^n(y) \in A_n$ and so, $x, y \in \bigcap_0^\infty T^{-n}A_n$.

Since this intersection contains at most one point we have $x = y$. Hence T is forward expansive. $\diamondsuit$

Example: $\{\{1, 2, ...m\}^N, \sigma\}$ where σ is left-shift.

As section 4.1,we let $O(X)$ be the collection of all covers on X. Let $U \in O(X)$ and $x \in$X, we denote

$$h_p(T, U) = \sup_{x\in X}\{\limsup_{N\to\infty} \frac{1}{N} \log \aleph(U, N, B_N(x, T))\},$$

and

$$h_m(T, U) = \limsup_{N\to\infty} \frac{1}{N} \log\{\sup_{x\in X} \aleph(U, N, B_N(x, T))\}.$$

Let $Y \subseteq$X, $\aleph(U)|_Y$ be the smallest cardinality of subsets in U which covers Y. For any fixed x in X, we have $\aleph(U, N, B_N(x, T)) = \aleph(\bigvee_{n=0}^N T^{-n}U)|_{T^{-N}(x)}$.

Theorem 4.2.1 Let $T : X \to X$ be a forward expansive continuous function of the compact metric space (X,d). If α is a forward generator for T, then

$$h_p(T) = h_p(T, \alpha) \text{ and } h_m(T) = h_m(T, \alpha).$$

PROOF

Since α is a forward generator, for any $U \in O(X)$, we can choose N large enough such that $U < \bigvee_{n=0}^{N} T^{-n}\alpha$,

$$\log \aleph(\bigvee_{n=0}^{k} T^{-n}U)|_{T^{-k}(x)} \leq \log \aleph(\bigvee_{n=0}^{k} T^{-n} \bigvee_{n=0}^{N} T^{-n}\alpha)|_{T^{-k}(x)}, \text{ for any } k.$$

Then

$$\begin{aligned}
\limsup_{k\to\infty} \frac{1}{k} \log \aleph[\bigvee_{n=0}^{k} T^{-n}U]|_{T^{-k}(x)} &\leq \limsup_{k\to\infty} \frac{1}{k} \log \aleph(\bigvee_{n=0}^{k} T^{-n}(\bigvee_{n=0}^{N} T^{-n}\alpha)|_{T^{-k}(x)} \\
&= \limsup_{k\to\infty} \frac{1}{k} \log \aleph(\bigvee_{n=0}^{k+N} T^{-n}\alpha)|_{T^{-k}(x)} \\
&\leq \limsup_{k\to\infty} \frac{1}{k} \log \aleph(\bigvee_{n=0}^{k+N} T^{-n}\alpha)|_{T^{-(k+N)}(x)} \\
&= \limsup_{k\to\infty} \frac{k+n}{k} \frac{1}{k+N} \log \aleph(\bigvee_{n=0}^{k+N} T^{-n}\alpha)|_{T^{-(k+N)}(x)} \\
&\leq \limsup_{k\to\infty} \frac{k+N}{k} \limsup_{k\to\infty} \frac{1}{k+N} \log \aleph(\bigvee_{n=0}^{k+N} T^{-n}\alpha)|_{T^{-(k+N)}(x)} \\
&= \limsup_{k\to\infty} \frac{1}{k+N} \log \aleph(\bigvee_{n=0}^{k+N} T^{-n}\alpha)|_{T^{-(k+N)}(x)}
\end{aligned}$$

So we can get $h_p(T,U) \leq h_p(T,\alpha)$ for any fixed x in X, this implies

$$h_p(T) = \sup_{x\in X}\{\sup_{U\in O(X)} h_p(T,U)\} \leq \sup_{x\in X} h_p(T,\alpha)$$

Then $h_p(T) = h_p(T,\alpha)$.

Similarly, since $U < \bigvee_{n=0}^{N} T^{-N}\alpha$, $\aleph(\bigvee_{n=0}^{k} U)|_{T^{-k}(x)} \leq \aleph(\bigvee_{n=0}^{k} \bigvee_{n=0}^{N} T^{-n}\alpha)|_{T^{-k}(x)}$,

$$\begin{aligned}
h_m(T,U) &= \limsup_{k\to\infty} \frac{1}{k} \log[\sup_{x\in X} \aleph(\bigvee_{n=0}^{k} T^{-n}U]|_{T^{-k}(x)} \\
&\leq \limsup_{k\to\infty} \frac{1}{k} \log[\sup_{x\in X} \aleph(\bigvee_{n=0}^{k} T^{-n}(\bigvee_{n=0}^{N} T^{-n}\alpha)]|_{T^{-k}(x)},
\end{aligned}$$

By a similar calculation, we can get $h_m(T,U) \leq h_m(T,\alpha)$. Finally we get $h_m(T) = h_m(T,\alpha)$. $\diamond$

Theorem 4.2.2[4] Let $T: X \to X$ be forward expansive with $h(T) = \log\lambda$. Then there is $x \in X$ with card$(T^{-n}x) \geq \lambda^n$ for all n and in particular

$$h_p(T) = h_m(T) = h(T).$$

4.3 Metric ρ Compatible with the Topology

Lemma 4.3.1 Let $T: X \to X$ be a continuous map of a compact metric space (X, d).

(1) **If** α is an open cover of X with Lebesgue number δ, then for any $x \in X$,

$$\aleph(\vee_{i=0}^{n-1} T^{-i}\alpha|_{T^{-n}(x)}) \leq r_n(\delta/2, T^{-n}(x)) \leq s_n(\delta/2, T^{-n}(x))$$

(2) **If** $\epsilon > 0$ and γ is an open cover with diam$(\gamma) \leq \epsilon$, then for any $x \in X$,

$$r_n(\epsilon, T^{-n}(x)) \leq s_n(\epsilon, T^{-n}(x)) \leq \aleph(\vee_{i=0}^{n-1} T^{-i}\gamma|_{T^{-n}(x)}).$$

PROOF (1) It's abvious that $r_n(\epsilon, T^{-n}(x)) \leq s_n(\epsilon, T^{-n}(x))$. Consider any $T^{-n}(x)$ of x and let F be a $(n, \delta/2)$ spanning set for $T^{-n}(x)$ of cardinality $r_n(\delta/2, T^{-n}(x))$. Then

$$T^{-n}(x) \subseteq \bigcup_{x\in F} \bigcap_{i=0}^{n-1} T^{-i}\bar{B}(T^i x; \delta/2)$$

and since for each i, $\bar{B}(T^i x; \delta/2)$ is a subset of a member of α, so we have $\aleph(\vee_{i=0}^{n-1} T^{-i}\alpha|_{T^{-n}(x)}) \leq r_n(\delta/2, T^{-n}(x))$.

(2) Let E be a (n, ϵ) separated set of cardinality $s_n(\epsilon, T^{-n}(x))$ for $T^{-n}(x)$. No member of the cover $\vee_{i=0}^{n-1} T^{-i}\gamma|_{T^{-n}(x)}$ can contain two elements of E so $s_n(\epsilon, T^{-n}(x)) \leq \aleph(\vee_{i=0}^{n-1} T^{-i}\gamma|_{T^{-n}(x)})$. $\Diamond$

Lemma 4.3.2 Let $T: X \to X$ be a forward expansive continuous function for a compact metric space (X,d) with forward expansive constant e. Then

$$\begin{aligned} h_p(T) &= \sup_{x\in X}\{\limsup_{n\to\infty} \frac{1}{n} \log r_n(\delta_o, T^{-n}(x))\} \\ &= \sup_{x\in X}\{\limsup_{n\to\infty} \frac{1}{n} \log s_n(\delta_o, T^{-n}(x))\} \end{aligned}$$

and

$$\begin{aligned} h_m(T) &= \limsup_{n\to\infty} \frac{1}{n} \log \sup_{x\in X} r_n(\delta_\circ, T^{-n}(x)) \\ &= \limsup_{n\to\infty} \frac{1}{n} \log \sup_{x\in X} s_n(\delta_\circ, T^{-n}(x)) \end{aligned}$$

for all $\delta_\circ < e/4$.

PROOF

We can let $\delta_\circ < e/4$. For all $x \in X$, choose $x_1, x_2, ..., x_k$ such that $T^{-n}(x) \subseteq \bigcup_{i=1}^{k} B(x_i; \frac{e}{2} - 2\delta_\circ)$. This cover $\alpha = \{B(x_i; e/2) | 1 \le i \le k\}$ is a forward generator with the Lebesgue number $2\delta_\circ$. So by Lemma 4.3.1,

$$h_m(T, \alpha) \le \limsup_{n\to\infty} \frac{1}{n} \log \sup_{x\in X} r_n(\delta_0, T^{-n}(x)) \le h_m(T)$$

Similar calculation for $h_p(T)$. $\diamond$

Lemma 4.3.3 Let T be a forward expansive continuous function from compact metric space (X, d) to itself with forward expansive constant e. Then for all $\epsilon > 0, \exists N > 0$, such that $d(T^i x, T^i y) \le e$ for all i with $0 \le i \le N$, this implies $d(x, y) \le \epsilon$.

PROOF

We may assume there exists $\epsilon > 0$. For all $N = 1, 2, 3, \ldots$, we can find $x_n, y_n \in X$ s.t. $d(T^n(x_n), T^n(y_n)) \le e, 0 \le n \le N$ and $d(x_n, y_n) \ge \epsilon$. Choose $x_{ni} \to x$ and $y_{ni} \to y$. Then $d(x, y) > \epsilon$, but $d(T^i x, T^i y) \le e$ for all $i = 1, 2, 3, \ldots$. This contradicts the forward expansive property of T. $\diamond$

Lemma 4.3.4 Let X be a compact metric space and $T : X \to X$ a forward expansive with forward expansive constant e. Then for $0 < \epsilon < e/2$ and $\delta > 0$ there exists $C_{\delta,\epsilon}$ such that for all positive integer n and all x in X, we have

$$s_n(\delta, T^{-n}(x)) \le C_{\delta,\epsilon} s_n(\epsilon, T^{-n}(x)).$$

PROOF

For $0 < \epsilon < e/2$ and $0 < \delta$, by Lemma 4.3 and uniform continuity of T on X, there exists a positive integer N and $\alpha > 0$ such that if $d_N(x,y) \leq 2\epsilon$, then $d(x,y) \leq \delta$

and if $d(x,y) \leq \alpha$, then $d_N(x,y) \leq \delta$

Now fix x and let n be big enough, assume E is a maximal (n,δ)-separated set of $T^{-n}(x)$ and F is a maximal (n,ϵ)-separated set of $T^{-n}(x)$, then for $\acute{x} \in E$ there is a $z(\acute{x}) \in F$ such that $d_n(\acute{x}, z(\acute{x})) < \epsilon$. Let $E_z = \{\acute{x} \in E : z(\acute{x}) = z\}$, then $\mathrm{card}(E) \leq \sum_{z\in F} \mathrm{card}(E_z)$. But if $x, y \in E_z$, then $d_n(x,y) \leq 2\epsilon$ by definition of E_z, hence

$$d(T^i(x), T^i(y)) \leq \delta \text{ for } i = 0, 1, 2, ..., n - N.$$

Since $x, y \in E, d_n(x,y) > \delta$ and if $d(T^{n-N+1}(x), T^{n-N+1}(y)) \leq \alpha$, this implies that $d_N(T^{n-N+1}(x), T^{n-N+1}(y)) \leq \delta$, then $d_n(x,y) \leq \delta$. This is a contradiction. So $d(T^{n-N+1}(x), T^{n-N+1}(y)) > \alpha$ and $T^{n-N+1}(x), T^{nN+1}(y) \in X$, X is compact. This implies that $\mathrm{card}(E_z)$ is bounded by some constant $C_{\delta,\epsilon}$. Therefore, $\mathrm{card}(E_z) \leq C_{\delta,\epsilon}\mathrm{card}(F)$. $\diamondsuit$

Remark: We can show Lemma 4.3.2 from Lemma 4.3.4.

During the remainder of this section we will assume that T is a forward expansive continuous function of a compact metric space (X,d) onto itself with forward expansive constant e >0.

Now for any integer $n \geq 0$, we define:

$$W_n = \{(x,y) \in X \times X : d(T^i x, T^i y) \leq e \text{ for } 0 \leq i \leq n\}$$

It's obvious that $\bigcap_{n=0}^{\infty} W_n = \triangle$ where $\triangle = \{(x,x) : x \in X\}$.

Take ϵ small enough such that $3\epsilon \leq e$. Choose N from the above Lemma 4.3 with respect to ϵ. We define $V_n = W_{nN}$ for n=0,1,2,3,...

and $(x,y) \in V_{n+1} \circ V_{n+1} \circ V_{n+1}$ means there exists $u, v \in X$ s.t $(x,u), (u,v)$ and $(v,y) \in V_{n+1}$.

Lemma 4.3.5 The sequence V_n is a nested sequence of symmetric neighborhoods of $\triangle$ whose intersection is $\triangle$ and such that $V_{n+1} \circ V_{n+1} \circ V_{n+1} \subseteq V_n$ for all $n \geq 0$.

PROOF

Let $(x, y) \in V_{n+1} \circ V_{n+1} \circ V_{n+1}$, then exists $u, v \in X$ s.t

$(x, u), (u, v)$ and $(v, y) \in V_{n+1}$

i.e. $d(T^i x, T^i u) \leq e, 0 \leq i \leq (n+1)N$,

$d(T^i u, T^i v) \leq e, 0 \leq i \leq (n+1)N$

and $d(T^i v, T^i y) \leq e, 0 \leq i \leq (n+1)N$

By Lemma 4.3.3,

$$d(T^i x, T^i u) \leq \epsilon, d(T^i u, T^i v) \leq \epsilon \text{ and } d(T^i v, T^i y) \leq \epsilon \text{ for } 0 \leq i \leq nN.$$

The triangle inequality can imply

$$d(T^i x, T^i y) \leq 3\epsilon \leq e \text{ for } 0 \leq i \leq nN.$$

This implies that $(x, y) \in V_n$. $\Diamond$

Metrization Lemma 4.3.6[6] Let V_n be a sequence of symmetric neighborhoods of the diagonal, $\triangle$, of X× X with $V_0 = X \times X$ such that $V_{n+1} \circ V_{n+1} \circ V_{n+1} \subset V_n$ for each n and $\bigcap_0^\infty V_n = \triangle$. Then there is a metric D compatible with the topology of such that the following condition holds for n≥ 1,

$$V_n \subset \{(x, y) : D(x, y) < 1/2^n\} \subset V_{n-1}$$

We define $N_d(A; \epsilon) = \{x; d(x, A) < \epsilon\}$ where d is a metric on A.

There following consequence comes from Lemma 4.3.5 and Lemma 4.3.6

Lemma 4.3.7 There is a metric ρ compatible with the topology of X such that

$$N_\rho(\triangle; 1/2^{n+1}) \subseteq V_n \subseteq N_\rho(\triangle; 1/2^n)$$

for all positive integer n.

Lemma 4.3.8 There is a metric ρ compatible with the topology of X and there is λ, $0 < \lambda < 1$, such that

$$N_\rho(\triangle; \lambda^{m+2N}) \subseteq W_m \subseteq N_\rho(\triangle; \lambda^{m-N})$$

for all positive integers m.

PROOF

Consider any positive integer $m = nN + j, 0 \leq j < N$, it is easy to see that

$$V_{n+1} = W_{(n+1)N} = W_{nN+N} \subseteq W_{nN+j} = W_m \subseteq W_{nN} = V_n.$$

Therefore $V_{n+1} \subseteq W_m \subseteq V_n$. From Lemma 4.3 we can get $N(\triangle\ ; 1/2^{n+2}) \subseteq V_{n+1}$ and $V_n \subseteq N(\triangle; 1/2^n)$. Now we let $\lambda = (\frac{1}{2})^{\frac{1}{N}}$.

$$\begin{aligned}
&N_\rho(\triangle; \lambda^{m+2N}) \subseteq N_\rho(\triangle; \lambda^{m+2N-j}) = N_\rho(\triangle; \lambda^{(n+2)N}) \\
&= N_\rho(\triangle; (\tfrac{1}{2})^{n+2}) \subseteq V_{n+1} \subseteq W_m \subseteq V_n \\
&\subseteq N_\rho(\triangle; \tfrac{1}{2^n}) = N_\rho(\triangle; (\tfrac{1}{2})^{n\frac{N}{N}}) \\
&= \mathrm{N}_\rho(\triangle; \lambda^{nN}) = N_\rho(\triangle; \lambda^{m-j}) \subseteq N_\rho(\triangle; \lambda^{m-N})
\end{aligned}$$

This finished the proof of the Lemma. ◇

Theorem 4.3.1 Assume T is a forward expansive continuous function of a compact metric space (X, d) onto itself with forward expansive constant $e > 0$. Then there is a metric ρ compatible with the topology of X and there is λ, $0 < \lambda < 1$, such that

$$h_p(T) = \sup_{x \in X}\{\limsup_{k \to \infty} \frac{1}{k} \log r_1(\lambda^k, T^{-k}(x))\}$$

and

$$h_m(T) = \limsup_{k \to \infty} \frac{1}{k} \log\{\sup_{x \in X} r_1(\lambda^k, T^{-k}(x))\}$$

with respect to this metric ρ.

PROOF

For any $x \in$X we consider $T^{-k}(x)$. Let E be a (k, e)-spanning set of $T^{-k}(x)$ with minimum cardinality. For any $y \in T^{-k}(x)$, there exists $z \in E$ s.t. $d(T^i y, T^i z) \leq e$ for $0 \leq i < k$. So $(y, z) \in W_{k-1}$. From Lemma 4.8 we can find a metric ρ on X and $\lambda, 0 < \lambda < 1$, such that $(x, y) \in N_\rho(\triangle; \lambda^{k-1-N})$. This means that there exists an F which is $(1,\lambda^{k-1-N})$-spanning set with metric ρ and $\sharp(F) \leq \sharp(E)$. Therefore $r_1(\lambda^{k-1-N}, T^{-k}(x)) \leq r_k(e, T^{-k}(x))$.

On the other hand consider F to be a $(1,\lambda^{k-1+2N})$-spanning set of $T^{-k}(x)$ with respect to this metric ρ and with minimum cardinality. Thus for any $y \in T^{-k}(x)$, there exists z$\in F$ such that $(y, z) \in N_\rho(\Delta; \lambda^{k-1+2N})$. So $(x, z) \in W_{k-1}$ by Lemma 4.8. This means $d(T^i y, T^i z) \leq e, 0 \leq i < k$. and this implies that we can find E which is (k, e)-spanning set of $T^{-k}(x)$ with card$(E) \leq$ card(F). Therefore $r_k(e, T^{-k}(x)) \leq r_1(\lambda^{k-1+2n}, T^{-k}(x))$. Since ρ is fixed, let e be small enough and using Lemma 4.2, we have

$$h_p(T) = \sup_{x \in X}\{\limsup_{k \to \infty} \frac{1}{k} \log r_1(\lambda^k, T^{-k}(x))\}$$

and

$$h_m(T) = \limsup_{k \to \infty} \frac{1}{k} \log \sup_{x \in X} r_1(\lambda^k, T^{-k}(x)).$$

$\diamondsuit$

Chapter 5

Modified Pointwise Preimage entropy

In this chapter we modify the original pointwise preimage entropy and call it upper preimage entropy. Then we show the relationship between conditional entropy and upper preimage entropy. Finally, we follow the S-M-B method and use it to show the asymptotic behavior of metric preimage entropy and ergodic decompsition.

5.1 New Definitions

We continue to consider a continuous self-map T of the compact metric space (X, d). Given a subset $K \subset X$, a $\delta > 0$, and an positive integer n, we set

$$r(n, \delta, K) = r(n, \delta, K, T) = \max\{\text{card}(E) : E \subseteq K, E \text{ is } (n, \delta) - \text{separated}\}.$$

Definition 5.1.1(Upper Preimage Entropy)

$$h_{top}(T \mid \xi^-) = \lim_{\delta \to 0} \limsup_{n \to \infty} \frac{1}{n} \log \sup_{k \geq 0, x \in X} r(n, \delta, T^{-k}x)$$

$$= \sup_{\alpha \text{ open cover}} \limsup_{n \to \infty} \frac{1}{n} \log \sup_{k \geq 0, x \in X} \aleph\left(\bigvee_{i=0}^{n-1} T^{-i}\alpha\right)|_{T^{-k}x}$$

Remark 5.1.1 $h_p(T) \leq h_m(T) \leq h_{top}(T \mid \xi^-) \leq h(T)$

Example: Consider $S : \{1,2\}^N \to \{1,2\}^N$ and $T : \{1,2\}^Z \to \{1,2\}^Z$ where S and T

are left-shifts. Then we can get

$$h_p(S \times T) = h_m(S \times T) = h_{top}(S \times T \mid \xi^-) = \log 2 < h(S \times T) = 2\log 2$$

Next, let ξ denote the point partition of X which we also identify with the σ-algebra $\mathcal{B}$ of Borel measurable sets.

For $n > 0$, we set

$$\xi^{-n} = T^{-n}\xi$$

Given a finite partition α, let $\alpha^n = \vee_{i=0}^{n-1} T^{-i}\alpha$. For a T-invariant probability μ, let

$$H_\mu(\alpha^n \mid \xi^{-k})$$

denote the conditional entropy of α^n given the σ-algebra $T^{-k}\mathcal{B}$. We call this the entropy of α^n given the preimage partition ξ^{-k}.

Note that, since $H_\mu(\cdot \mid \cdot)$ is increasing in the first variable and decreasing in the second variable, we have $n \geq m, l \geq k$ implies

$$H_\mu(\alpha^n \mid \xi^{-l}) \geq H_\mu(\alpha^m \mid \xi^{-k}).$$

Set

$$H_\mu(\alpha^n \mid \xi^-) = H_\mu(\alpha^n \mid \xi^{-\infty}) = \sup_{k \geq 0} H_\mu(\alpha^n \mid \xi^{-k}) = \lim_{k\to\infty} H_\mu(\alpha^n \mid \xi^{-k}).$$

Lemma 5.1.1 The function $a_n = H_\mu(\alpha^n \mid \xi^-)$ is subadditive.

PROOF

We need to show

$$a_{n+m} \leq a_n + a_m.$$

We have

$$\begin{aligned}
a_{n+m} &= \lim_{k\to\infty} H_\mu(\alpha^{n+m} \mid \xi^{-k}) \\
&= \lim_{k\to\infty} H_\mu(\alpha^n \vee T^{-n}\alpha^m \mid \xi^{-k}) \\
&= \lim_{k\to\infty} (H_\mu(\alpha^n \mid \xi^{-k}) + H_\mu(T^{-n}\alpha^m \mid \alpha^n \vee \xi^{-k})) \\
&\leq \lim_{k\to\infty} H_\mu(\alpha^n \mid \xi^{-k}) + \lim_{k\to\infty} H_\mu(T^{-n}\alpha^m \mid \xi^{-k}) \\
&= \lim_{k\to\infty} H_\mu(\alpha^n \mid \xi^{-k}) + \lim_{k\to\infty} H_\mu(\alpha^m \mid \xi^{-k}) \\
&= a_n + a_m
\end{aligned}$$

$\diamond$

Definition 5.1.2(Metric Preimage Entropy)

$$h_\mu(T \mid \xi^-, \alpha) = h_\mu(\alpha \mid \xi^-) = \lim_{n\to\infty} \frac{1}{n} H_\mu(\alpha^n \mid \xi^-) = \inf_{n\to\infty} \frac{1}{n} H_\mu(\alpha^n \mid \xi^-)$$

and

$$h_\mu(T \mid \xi^-) = \sup_\alpha h_\mu(\alpha \mid \xi^-) = \sup_\alpha h_\mu(T \mid \xi^-, \alpha)$$

Lemma 5.1.2 Metric preimage entropy $h_\mu(T \mid \xi^-)$ is a measure-theoretic conjugacy invariant and upper preimage entropy $h_{top}(T \mid \xi^-)$ is a topological conjugacy invariant.

PROOF

It is easy to show the measure-theoretic conjugacy invariant. Here we prove the topological conjugacy invariant. First we let X_1, X_2 be compact spaces and $T_i : X_i \to X_i$ be continuous for $i = 1, 2$ and $\phi : X_1 \to X_2$ be a homeomorphism map with $\phi T_1 = T_2 \phi$. First we let α be an open cover of X_2. Then if $\phi(y) = x$, we have for $k \geq 0, \aleph(\phi^{-1}\alpha)|_{T_1^{-k}(y)} = \aleph(\alpha)|_{T_2^{-k}(x)}$.

Hence,

$$\begin{aligned}
h_{top}(T_2 \mid \xi^-, \alpha) &= \limsup_{n\to\infty} \frac{1}{n} \log \sup_{k\geq 0, x\in X_2} \aleph(\bigvee_{i=0}^{n-1} T_2^{-i}\alpha)|_{T_2^{-k}(x)} \\
&= \limsup_{n\to\infty} \frac{1}{n} \log \sup_{k\geq 0, y\in X_1} \aleph(\phi^{-1} \bigvee_{i=0}^{n-1} T_2^{-i}\alpha)|_{T_1^{-k}(y)} \\
&= \limsup_{n\to\infty} \frac{1}{n} \log \sup_{k\geq 0, y\in X_1} \aleph(\bigvee_{i=0}^{n-1} \phi^{-1} T_2^{-i}\alpha)|_{T_1^{-k}(y)} \\
&= \limsup_{n\to\infty} \frac{1}{n} \log \sup_{k\geq 0, y\in X_1} \aleph(\bigvee_{i=0}^{n-1} T_1^{-i}\phi^{-1}\alpha)|_{T_1^{-k}(y)} \\
&= h_{top}(T_1 \mid \xi^-, \phi^{-1}\alpha)
\end{aligned}$$

Hence $h_{top}(T_2 \mid \xi^-) \leq h_{top}(T_1 \mid \xi^-)$. If ϕ is a homeomorphism then $\phi^{-1}T_2 = T_1\phi^{-1}$ so by the above, $h_{top}(T_1 \mid \xi^-) \leq h_{top}(T_2 \mid \xi^-)$. $\diamond$

Lemma 5.1.3 Let ζ and η be two finite partitions of X, then

$$h_\mu(\zeta \mid \xi^-) \leq h_\mu(\eta \mid \xi^-) + H_\mu(\zeta \mid \eta)$$

PROOF

$$H_\mu(\bigvee_{i=0}^{n-1} T^{-i}\zeta \mid \xi^{-k}) \leq H_\mu((\bigvee_{i=0}^{n-1} T^{-i}\zeta \vee \bigvee_{i=0}^{n-1} T^{-i}\eta) \mid \xi^{-k})$$
$$= H_\mu(\bigvee_{i=0}^{n-1} T^{-i}\eta \mid \xi^{-k}) + H_\mu(\bigvee_{i=0}^{n-1} T^{-i}\zeta \mid \bigvee_{i=0}^{n-1} T^{-i}\eta \vee \xi^{-k})$$
$$\leq H_\mu(\bigvee_{i=0}^{n-1} T^{-i}\eta \mid \xi^{-k}) + H_\mu(\bigvee_{i=0}^{n-1} T^{-i}\zeta \mid \bigvee_{i=0}^{n-1} T^{-i}\eta)$$

Let $k \longrightarrow \infty$, and $H_\mu(\bigvee_{i=0}^{n-1} T^{-i}\zeta \mid \bigvee_{i=0}^{n-1} T^{-i}\eta) \leq n \cdot H_\mu(\zeta \mid \eta)$

This implies that $H_\mu(\zeta_0^n \mid \xi^-) \leq H_\mu(\eta_0^n \mid \xi^-) + n \cdot H_\mu(\zeta \mid \eta)$.

Divide by n and let n go to infinity, then $h_\mu(\zeta \mid \xi^-) \leq h_\mu(\eta \mid \xi^-) + H_\mu(\zeta \mid \eta)$. $\diamond$

Lemma 5.1.4 For any fixed k,

$$h_\mu(T \mid \xi^-, \alpha) = h_\mu(T \mid \xi^-, \bigvee_{i=0}^{k} T^{-i}\alpha).$$

PROOF

$$h_\mu(T \mid \xi^-, \bigvee_{i=0}^{k} T^{-i}\alpha) = \lim_{n\to\infty} \frac{1}{n} H_\mu((\alpha^k)^n) \mid \xi^-)$$
$$= \lim_{n\to\infty} \frac{1}{n} H_\mu(\bigvee_{i=0}^{n-1} T^{-i}(\bigvee_{i=0}^{k} T^{-i}\alpha) \mid \xi^-)$$
$$= \lim_{n\to\infty} \frac{1}{n} H_\mu(\bigvee_{i=0}^{k+n-1} T^{-i}\alpha) \mid \xi^-)$$
$$= \lim_{n\to\infty} \frac{k+n-1}{n} \frac{1}{k+n-1} H_\mu(\bigvee_{i=0}^{k+n-1} T^{-i}\alpha / \xi^-)$$
$$= h_\mu(T \mid \xi^-, \alpha)$$

$\diamond$

Lemma 5.1.5 If $\{A_n\}$ is an increasing sequence of finite partitions of X and C is a partition with $C \leq \bigvee_{i=0}^{\infty} A_i$, then $H_\mu(C \mid A_n) \to 0$ as n$\to \infty$.

Let C be a finite sub-σ-algebra of $\mathcal{B}$, say $C = \{C_i : i = 1, 2, ..., n\}$, then the non-empty sets of the form $B_1 \cap B_2 ... \cap B_n$, where $B_i = C_i$ or $X \backslash C_i$, form a finite partition of X. We denote it by $\alpha(C)$ and we define $h_\mu(T \mid \xi^-, C) = h_\mu(T \mid \xi^-, \alpha(C))$.

Theorem 5.1.1 (Kolmogorov-Sinai)

Let T be a m.p.t. of $(X, \mathcal{B},\mu)$ *and* $\Re$ *be a finite sub-σ-algebra s.t.* $\vee_{n=0}^{\infty} T^{-n}(\alpha(\Re))=\mathcal{B}$, *then*

$$h_\mu(T \mid \xi^-) = h_\mu(T \mid \xi^-, \Re)$$

PROOF Let C be any partition, we show that $h_\mu(T \mid \xi^-, C) \leq h_\mu(T \mid \xi^-, \alpha(\Re))$
For $n \geq 1$, by Lemma 5.1.3 and Lemma 5.1.4,

$$\begin{aligned} h_\mu(T \mid \xi^-, C) &\leq h_\mu(T \mid \xi^-, \vee_{i=0}^{n} T^{-i}\alpha(\Re)) + H_\mu(C \mid \vee_{i=0}^{n} T^{-i}\alpha(\Re)) \\ &= h_\mu(T \mid \xi^-, \alpha(\Re)) + H_\mu(C \mid \vee_{i=0}^{n} T^{-i}\alpha(\Re)) \end{aligned}$$

Let $A_n = \vee_{i=0}^{n} T^{-i}\alpha(A)$ be Lemma 5.1.5, $H_\mu(C \mid A_n) \to 0$ as $n \to \infty$. $\diamond$

Lemma 5.1.6 Let $(X,\mathcal{B},\mu)$ be a probability space. If $\mathcal{B}_0$ is a sub-algebra of $\mathcal{B}$ with $\mathcal{B}(\mathcal{B}_0)=\mathcal{B}$ then for m.p.t. $T : X \to X$ we have $h_\mu(T \mid \xi^-) = \sup h_\mu(T \mid \xi^-, A)$ where the supremum is taken over all finite sub-algebras A of $\mathcal{B}_0$.

PROOF

Let $\epsilon > 0$. Let $C \subseteq\mathcal{B}$ be finite. Then there exists a finite $D \subseteq\mathcal{B}_0$ such that $H_\mu(C \mid D) < \epsilon$.

Thus

$$\begin{aligned} h_\mu(T \mid \xi^-, C) &\leq h_\mu(T \mid \xi^-, D) + H_\mu(C \mid D) \\ &\leq h_\mu(T \mid \xi^-, D) + \epsilon \end{aligned}$$

Therefore $h_\mu(T \mid \xi^-, C) \leq \epsilon + \sup\{h_\mu(T \mid \xi^-, D) : D \subset\mathcal{B}_0, D \text{ finite}\}$ and thus $h_\mu(T \mid \xi^-) \leq \sup\{h_\mu(T \mid \xi^-, D) : D \subset\mathcal{B}_0, D \text{ is finite}\}$. The opposite inequality is obvious. $\diamond$

Lemma 5.1.7 Let $(X,\mathcal{B},\mu)$ be a probability space and let $\{A_n\}_1^\infty$ be an increasing sequence of finite sub-algebras of $\mathcal{B}$ such that $\bigvee_{n=1}^{\infty} A_n =\mathcal{B}$. If $T : X \to X$ is m.p.t., then

$$h_\mu(T \mid \xi^-) = \lim_{n\to\infty} h_\mu(T \mid \xi^-, A_n).$$

Lemma 5.1.8 Let α_i be a finite partition and $\Im_i$ be a sub-σ-algebra of $(X_i,\mathcal{B}_i, m_i)$, for $i = 1, 2$, then

$$E(\chi_{A\times B} \mid \Im) = E(\chi_A \mid \Im_1) \cdot E(\chi_B \mid \Im_2), a.e.$$

where $\Im = \Im_1 \times \Im_2, A \in \alpha_1,\ B \in \alpha_2$, $\Im_1$ is a sub-σ-algebra of $\mathcal{B}_1$ and $\Im_2$ is a sub-σ-algebra of $\mathcal{B}_2$.

PROOF

We have $I_{\alpha|\Im}(x, y) = -\sum_{A\times B\in\alpha} \log E(\chi_{A\times B} \mid \Im)\chi_{A\times B}(x, y)$ and let $\mu = m_1 \times m_2$, since $\int_{F_1} E(\chi_A \mid \Im_1)dm_1 = \int_{F_1} \chi_A dm_1, \int_{F_2} E(\chi_B \mid \Im_2)dm_2 = \int_{F_2} \chi_B dm_2$, for $F_1 \in \Im_1, F_2 \in \Im_2$.

Then

$$\begin{aligned}&\int\int_{F_1\times F_2} E(\chi_{A\times B} \mid \Im)d\mu = \int\int_{F_1\times F_2} \chi_{A\times B} d\mu\\ &= \mu(A \cap F_1 \times B \cap F_2) = m_1(A \cap F_1) \cdot m_2(B \cap F_2)\end{aligned}$$

And

$$\begin{aligned}&\int_{F_1} E(\chi_A \mid \Im_1)dm_1 \cdot \int_{F_2} E(\chi_B \mid \Im_2)dm_2 = m_1(A \cap F_1) \cdot m_2(B \cap F_2)\\ &= \int\int_{F_1\times F_2}(E(\chi_A \mid \Im_1) \cdot E(\chi_B \mid \Im_2))d\mu\end{aligned}$$

This implies that $E(\chi_{A\times B} \mid \Im) = E(\chi_A \mid \Im_1) \cdot E(\chi_B \mid \Im_2)$ a.e. $\diamond$

Theorem 5.1.2 Let $(X_1,\mathcal{B}_1, m_1)$ and $(X_2,\mathcal{B}_2, m_2)$ be probability spaces and let $T_1 : X_1 \to X_1, T_2 : X_2 \to X_2$ be m.p.t. Then

$$h_\mu(T_1 \times T_2 \mid \xi^-) = h_{m_1}(T_1 \mid \xi^-) + h_{m_2}(T_2 \mid \xi^-)$$

where $\mu = m_1 \times m_2$.

PROOF

Let $\mathcal{F}_0$ denote the algebra of finite unions of measurable rectangles. Then $\mathcal{B}(\mathcal{F}_0)=\mathcal{B}_1\times\mathcal{B}_2$. By Lemma 5.1.6,

$$h_\mu(T_1 \times T_2 \mid \xi^-) = \sup\{h_\mu(T_1 \times T_2 \mid \xi^-, C) : C \subset\mathcal{F}_0, \text{C finite}\}.$$

But if C is finite and $C \subset\mathcal{F}_0$, then $C \subset \alpha_1 \times \alpha_2$ for some finite $\alpha_1 \subset\mathcal{B}_1, \alpha_2 \subset\mathcal{B}_2$, Hence

$$h_\mu(T_1 \times T_2 \mid \xi^-) = \sup\{h_\mu(T_1 \times T_2 \mid \xi^-, \alpha_1 \times \alpha_2) : \alpha_1 \subset \mathcal{B}_1, \alpha_2 \subset \mathcal{B}_2, \alpha_1, \alpha_2 \text{ finite}\}$$

Let $\alpha = \alpha(\alpha_1 \times \alpha_2) = \alpha(\alpha_1) \times \alpha(\alpha_2)$,

$$\begin{aligned}
\int\int I_{\alpha|\xi^{-k}}(x,y)d\mu &= -\sum_{A\times B\in\alpha} \int\int \log E(\chi_{A\times B} \mid \xi^{-k})\chi_{A\times B}(x,y)\, d\mu \\
&= -\sum_{A\times B\in\alpha} \int\int (\log E(\chi_A \mid \xi^{-k}) + \log E(\chi_B \mid \xi^{-k}))\chi_{A\times B}(x,y)\, dm_1 dm_2 \\
&= -\sum_{A\times B\in\alpha} (\int\int \log E(\chi_A \mid \xi^{-k})\chi_{A\times B}(x,y)\, dm_1 dm_2 \\
&\quad + \int\int \log E(\chi_B \mid \xi^{-k})\chi_{A\times B}(x,y))\, dm_1 dm_2) \\
&= -\sum_{A\in\alpha(\alpha_1)} \int \log E(\chi_A \mid \xi^{-k})\chi_A\, dm_1 \\
&\quad - \sum_{B\in\alpha(\alpha_2)} \int \log E(\chi_B \mid \xi^{-k})\chi_B\, dm_2
\end{aligned}$$

This implies

$$H_\mu(\alpha \mid \xi^{-k}) = H_{m_1}(\alpha(\alpha_1) \mid \xi^{-k}) + H_{m_2}(\alpha(\alpha_2) \mid \xi^{-k}).$$

Then

$$H_\mu(T_1 \times T_2 \mid \xi^-, \alpha) = H_{m_1}(T \mid \xi^-, \alpha(\alpha_1)) + H_{m_2}(T \mid \xi^-, \alpha(\alpha_2)).$$

So that

$$h_\mu(T_1 \times T_2 \mid \xi^-) = h_{m1}(T_1 \mid \xi^-) + h_{m2}(T_2 \mid \xi^-).$$

$\diamond$

Theorem 5.1.3 Let $T_i : X_i \to X_i, i = 1,2$, be a continuous map on the compact metric space X_i. Then $h_{top}(T_1 \times T_2 \mid \xi^-) = h_{top}(T_1 \mid \xi^-) + h_{top}(T_2 \mid \xi^-)$.

PROOF

Let d_i be the metric on X_i. We use the metric $d((x_1,x_2),(y_1,y_2)) = \max(d_1(x_1,y_1), d_2(x_2,y_2))$ on $X_1 \times X_2$.

If F_i is an (n,ϵ)-spanning set for $T^{-k}x_i \in X_i$ then $F_1 \times F_2$ is an (n,ϵ)- spanning set for $T_1^{-k}(x_1) \times T_2^{-k}(x_2)$. Hence

$$r(n,\epsilon, (T_1 \times T_2)^{-k}(x_1,x_2)) \leq r(n,\epsilon, T^{-n}(x_1)) \cdot r(n,\epsilon, T^{-k}(x_2))$$

which implies

$$\log \sup_{k\geq 0, x_1\times x_2\in X_1\times X_2} r(n,\epsilon,(T_1\times T_2)^{-k}(x_1,x_2)) \leq \log \sup_{k\geq 0, x_1\in X_1} r(n,\epsilon,T_1^{-k}(x_1)) + \log \sup_{k\geq 0, x_2\in X_2} r(n,\epsilon,T_2^{-k}(x_2)).$$

Therefore

$$h_d(T_1\times T_2|\xi^-) \leq h_{d_1}(T_1|\xi^-,X_1) + h_{d_2}(T_2|\xi^-,X_2).$$

Now we show the other inequality.

For all T_1-invariant measure μ_1 and T_2-invariant measure μ_2, by variational principle (Theorem 5.2.1) we have

$$h_{d_1}(T_1 \mid \xi^-, X_1) \geq h_{\mu_1}(T_1 \mid \xi^-) \text{ and } h_{d_2}(T_2 \mid \xi^-, X_2) \geq h_{\mu_2}(T_2 \mid \xi^-)$$

Then

$$\begin{aligned} h_{d_1}(T_1 \mid \xi^-, X_1) + h_{d_2}(T_2 \mid \xi^-, X_2) &\geq h_{\mu_1}(T_1 \mid \xi^-) + h_{\mu_2}(T_2 \mid \xi^-) \\ &= h_{\mu_1\times\mu_2}(T_1 \times T_2 \mid \xi^-) \end{aligned}$$

This implies

$$\begin{aligned} h_{d_1}(T_1 \mid \xi^-, X_1) + h_{d_2}(T_2 \mid \xi^-, X_2) &\geq \sup_{\mu_1\times\mu_2} h_{\mu_1\times\mu_2}(T_1 \times T_2 \mid \xi^-) \\ &= h_d(T_1 \times T_2 \mid \xi^-) \end{aligned}$$

So we can get the equality. ◇

Theorem 5.1.4 Let T be a measure-preserving transformation of the probability space $(X,\mathcal{B},\mu)$. Then the map $\mu \to h_\mu(\alpha,T)$ is affine where α is any finite partition of X. Hence, so is the map $\mu \to h_\mu(T \mid \xi^-)$.

PROOF(See [3].)

For any integer n, constant $0<\lambda<1$ and measures μ,μ_1,μ_2, with

$$\mu = \lambda\mu_1 + (1-\lambda)\mu_2,$$

we have

$$0 \leq H_\mu(\alpha^n)) - \lambda H_{\mu_1}(\alpha^n) - (1-\lambda)H_{\mu_2}(\alpha^n) \leq \log 2$$

Hence,

$$h_\mu(\alpha) = \lambda h_{\mu_1}(\alpha) + (1-\lambda)h_{\mu_2}(\alpha).$$

Now, we proceed to $h_\mu(T, \xi^-)$.

Fix a finite partition α and an increasing sequence $\beta_1 < \beta_2 < \ldots$ converging to ξ as $i \to \infty$.

Then, for a positive integer n, we have

$$H_\mu(\alpha \mid \xi^-) = \lim_{i\to\infty} \lim_{m\to\infty} H_\mu(\alpha \mid T^{-m}\beta_i)$$

Next, consider the finite partition α, β. For any measure μ, we have

$$H_\mu(\alpha \mid \beta) = H_\mu(\alpha \vee \beta) - H_\mu(\beta).$$

Using the finite partitions α^n and $T^{-m}\beta_i$, we have

$$0 \le H_\mu(\alpha^n \vee T^{-m}\beta_i)) - \lambda H_{\mu_1}(\alpha^n \vee T^{-m}\beta_i) - (1-\lambda)H_{\mu_2}(\alpha^n \vee T^{-m}\beta_i) \le \log 2 \quad (5.1)$$

and

$$0 \ge -[H_\mu(T^{-m}\beta_i) - \lambda H_{\mu_1}(T^{-m}\beta_i) - (1-\lambda)H_{\mu_2}(T^{-m}\beta_i)] \ge -\log 2 \quad (5.2)$$

The second term of (5.2) is non-positive, so adding it to the second term of (5.1) does not increase the latter's value, so

$$H_\mu(\alpha^n \mid T^{-m}\beta_i) - \lambda H_{\mu_1}(\alpha^n \mid T^{-m}\beta_i) - (1-\lambda)H_{\mu_2}(\alpha^n \mid T^{-m}\beta_i) \le \log 2$$

Similarly, adding the second term of (5.1) to that of (5.2) does not decrease the latter's value, so

$$-\log 2 \le H_\mu(\alpha^n \mid T^{-m}\beta_i) - \lambda H_{\mu_1}(\alpha^n \mid T^{-m}\beta_i) - (1-\lambda)H_{\mu_2}(\alpha^n \mid T^{-m}\beta_i)$$

Putting these two inequalities together gives

$$-\log 2 \le H_\mu(\alpha^n \mid T^{-m}\beta_i) - \lambda H_{\mu_1}(\alpha^n \mid T^{-m}\beta_i) - (1-\lambda)H_{\mu_2}(\alpha^n \mid T^{-m}\beta_i) \le \log 2$$

Letting $i \to \infty$ and then $m \to \infty$ gives

$$-\log 2 \le H_\mu(\alpha^n \mid \xi^-) - \lambda H_{\mu_1}(\alpha^n \mid \xi^-) - (1-\lambda)H_{\mu_2}(\alpha^n \mid \xi^-) \le \log 2$$

Now, dividing by n and letting $n \to \infty$ gives that

$$h_\mu(\alpha \mid \xi^-) = \lambda h_{\mu_1}(\alpha \mid \xi^-) + (1-\lambda) h_{\mu_2}(\alpha \mid \xi^-)$$

as required. $\diamondsuit$

5.2 Variational Principle

Lemma 5.2.1 For each positive integer r, $h_\mu(T^r \mid \xi^-) = r \cdot h_\mu(T \mid \xi^-)$

PROOF

First show that $h_\mu(T^r \mid \xi^-, \bigvee_{i=0}^{r-1} T^{-i}A) = r \cdot h_\mu(T \mid \xi^-, A)$ for any finite partition A.

We have

$$\begin{aligned} r \cdot h_\mu(T \mid \xi^-, A) &= \lim_{n\to\infty} \frac{r}{nr} \sup_{k\geq 0} H_\mu\Big(\bigvee_{i=0}^{nr-1} T^{-i}A \mid T^{-rn-k}(\xi)\Big) \\ &= \lim_{n\to\infty} \frac{1}{n} \sup_{k\geq 0} H_\mu\Big(\bigvee_{j=0}^{n-1} T^{-rj}\Big(\bigvee_{i=0}^{r-1} T^{-i}A\Big) \mid T^{-rn-k}(\xi)\Big) \\ &= h_\mu\Big(T^r \mid \xi^-, \bigvee_{i=0}^{r-1} T^{-i}A\Big) \end{aligned}$$

Thus

$r \cdot h_\mu(T \mid \xi^-) = r \cdot \sup_A h_\mu(T \mid \xi^-, A) = \sup_A h_\mu(T^r \mid \xi^-, \bigvee_{i=0}^{r-1} T^{-i}A)$
$\leq \sup_c h_\mu(T^r \mid \xi^-, c) = h_\mu(T^r \mid \xi^-)$ where c is any finite partition.

On the other hand,
$h_\mu(T^r \mid \xi^-, A) \leq h_\mu(T^r \mid \xi^-, \bigvee_{i=0}^{r-1} T^{-i}A) = r \cdot h_\mu(T^r \mid \xi^-, A)$
This implies that $h_\mu(T^r \mid \xi^-) = r \cdot h_\mu(T \mid \xi^-)$. $\diamondsuit$

Definition 5.2.1

$$\begin{aligned} h_{top}(T \mid \xi^-) &= \lim_{\epsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} \log \sup_{k\geq 0, x\in X} r(n, \epsilon, T^{-k}(x)) \\ &= \sup_{\text{open cover } \beta} \limsup_{n\to\infty} \frac{1}{n} \log \sup_{k\geq 0, x\in X} \aleph\Big(\bigvee_{i=0}^{n-1} T^{-i}\beta|_{T^{-k}(x)}\Big) \end{aligned}$$

where $r(n,\epsilon,T^{-k}x)$ is the minimal cardinality of (n,ϵ) spanning set in $T^{-k}(x)$ or the max cardinality of (n,ϵ) separating set in $T^{-k}(x)$ and $\aleph(\beta)|_{T^{-k}(x)}$ is the minimal cardinality of subcover of β which can cover $T^{-k}(x)$.

Lemma 5.2.2 $h_{top}(T^m \mid \xi^-) = m \cdot h_{top}(T \mid \xi^-)$ for all positive integer m.

PROOF

Here we consider the spanning set.

Since for $x \in X, r(n,\epsilon,T^m,T^{-km}(x)) \leq r(mn,\epsilon,T,T^{-k}(x))$ for all $k \geq 1$

We have $\frac{1}{n}\log\sup_{k\geq 0, x\in X} r(n,\epsilon,T^m,T^{-km}(x)) \leq \frac{m}{mn}\log\sup_{k\geq 0,x\in X} r(mn,\epsilon,T,T^{-k}(x))$

Therefore $h_{top}(T^m \mid \xi^-) \leq m \cdot h_{top}(T \mid \xi^-)$

Since T is uniformly continuous, $\forall\epsilon > 0, \exists\delta > 0$ such that $d(x,y) < \delta$ implies $max_{0\leq i\leq m-1} d(T^i x, T^i y) < \epsilon$ for all x,y in X.

So an $(n,\delta,T^{-km}x)$-spanning set w.r.t. T^m is also an (nm,ϵ)-spanning set for $T^{-k}x$ w.r.t. T.

Hence $r(n,\delta,T^m,T^{-km}x) \geq r(mn,\epsilon,T,T^{-k}x)$, so

$$\frac{m}{mn}\log\sup_{k\geq 0,x\in X} r(mn,\epsilon,T,T^{-k}(x)) \leq \frac{1}{n}\log\sup_{k\geq 0,x\in X} r(n,\delta,T^m,T^{-km}(x))$$

Therefore, $m \cdot h_{top}(T \mid \xi^-) \leq h_{top}(T^m \mid \xi^-)$. $\diamond$

Lemma 5.2.3 For all ϵ there exists x and $k > 0$ such that E_n is an (n,ϵ) separated set in $T^{-k}(x)$ with $\mathrm{card}(E_n) = \sup_{k\geq 0,x\in X} r(n,\epsilon,T^{-k}(x))$.

PROOF

Because $r(n,\epsilon,T^{-k}(x)) \leq r(n,\epsilon,X)$, it's a finite bounded positive integer for all x in X and positive integer k. We can easily find such E_n. $\diamond$

Lemma 5.2.4 Let $E \subset T^{-n}(x)$, then $E \subset T^{-n-k}(T^k x)$ for all positive integer k.

Remark 5.2.1 [See [14], Remark 8.2.2]Assume $1 < q < n$, for $0 \leq j \leq q-1$, put $a(j) = [\frac{(n-j)}{q}]$,where [b] denotes the integer part of b. We have :

(1) Fix $0 \leq j \leq q-1$. Then

$$\{0,1,2,...,n-1\} = \{j + rq + i | 0 \leq r \leq a(j)-1, 0 \leq i \leq q-1\} \bigcup S$$

where $S = \{0, 1, ..., j-1, j + a(j)q, j + a(j)q + 1, ..., n-1\}$ and the cardinality of S is at most $2q$.

(2) The numbers $\{j + rq | 0 \leq j \leq q-1, 0 \leq r \leq a(j)-1\}$ are all distinct and are all no greater than $n-q$.

Theorem 5.2.1 Assume $\mu_j \to \mu$ with $\mu(\partial\alpha) = 0$, then

$$\lim_{j\to\infty} H_{\mu_j}(\bigvee_{i=0}^{q-1} T^{-i}\alpha \mid \xi^{-k}) = H_\mu(\bigvee_{i=0}^{q-1} T^{-i}\alpha \mid \xi^{-k})$$

for all finite partition α.

Lemma 5.2.5 Assume $\mu_j \to \mu$ with $\mu(\partial\alpha) = 0$, then

$$\lim_{n\to\infty}\lim_{k\to\infty} H_{\mu_n}(\bigvee_{i=0}^{j-1} T^{-i}\alpha \mid \xi^{-k}) = \lim_{k\to\infty} H_\mu(\bigvee_{i=0}^{q-1} T^{-i}\alpha \mid \xi^{-k}) = H_\mu(\bigvee_{i=0}^{q-1} T^{-i}\alpha \mid \xi^{-})$$

PROOF

By Lemma 5.3.6 we know that

$$\begin{aligned}\lim_{n\to\infty}\lim_{k\to\infty} H_{\mu_n}(\bigvee_{i=0}^{q-1} T^{-i}\alpha \mid \xi^{-k}) &= \lim_{n\to\infty} H_{\mu_n}(\bigvee_{i=0}^{q-1} T^{-i}\alpha \mid \lim_{k\to\infty} \xi^{-k}) \\ &= H_\mu(\bigvee_{i=0}^{q-1} T^{-i}\alpha \mid \lim_{k\to\infty} \xi^{-k}) \\ &= \lim_{k\to\infty} H_\mu(\bigvee_{i=0}^{q-1} T^{-i}\alpha \mid \xi^{-k}) \\ &= H_\mu(\bigvee_{i=0}^{q-1} T^{-i}\alpha \mid \xi^{-})\end{aligned}$$

$\diamond$

Measurable decompositions are necessary to show the variational principle. See [12]. Let ζ be an arbitrary decomposition of the Lebesgue space X and $X|_\zeta$ be the factor space. Let the factor map $\pi : X \to X|_\zeta$ be $\pi(x) = C$ where $x \in C \in \zeta$. Then

$$\mu(A) = \int_{X|\zeta} \mu_C(A \cap C)\, d\pi_*\mu$$

Where μ_C is the conditional measure on C.

Lemma 5.2.6 Let α be a partition of $(X, \mathcal{B}, \mu)$, consider the factor map $\pi_k : X \to$

$X|_{T^{-k}\xi}$ and $\mu_{x,k}$ is the conditional measure of μ on $T^{-k}(x)$ - Then

$$H_\mu(\bigvee_{i=0}^{n-1} T^{-i}\alpha \mid \xi^{-k}) = \int_{X|T^{-k}\xi} H_{\mu_{x,k}}(\bigvee_{i=0}^{n-1} T^{-i}\alpha|_{T^{-k}(x)})\, d\pi_{k*}\mu$$

Lemma 5.2.7: Let $\eta = \{B_0, B_1, ..., B_k\}$ be a partition of X such that $\beta = \{B_0 \cup B_1, ..., B_0 \cup B_k\}$ is an open cover of X. Then

$$\aleph(\bigvee_{i=0}^{n-1} T^{-i}\eta)|_Y \le \aleph(\bigvee_{i=0}^{n-1} T^{-i}\beta)|_Y \cdot 2^n$$

for any subset Y of X.

PROOF:

Consider the subcover $\hat{\beta}$ of β with cardinality $\aleph(\bigvee_{i=0}^{n-1} T^{-i}\beta)|_Y$,

let $A_i = (B_0 \cup B_{i_0}) \cap (T^{-1}B_0 \cup T^{-1}B_{i_1}) \cap ... \cap (T^{-(n-1)}B_0 \cup T^{-(n-1)}B_{i_{n-1}}) \in \hat{\beta}$

Now we decompose A_i into the partition

$$\hat{A}_i = \{B_{j_0} \cap T^{-1}B_{j_1} \cap ... \cap T^{-(n-1)}B_{j_{n-1}} : j_k = 0 \text{ or } i_k, 0 \le k \le n-1\}$$

Then $\aleph(\hat{A}_i) = 2^n$ and we have

$$\{\bigvee_{i=0}^{n-1} T^{-i}\eta \cap Y : \bigvee_{i=0}^{n-1} T^{-i}\eta \cap Y \ne \phi\} \subseteq \cup_{A_i \in \hat{\beta}} \hat{A}_i$$

This implies that

$$\aleph(\bigvee_{i=0}^{n-1} T^{-i}\eta)|_Y \le \aleph(\bigvee_{i=0}^{n-1} T^{-i}\beta)|_Y \cdot 2^n$$

◇

Now we are ready to show the relation between upper preimage entropy and metric preimage entropy. Here, the main technique used is the construction made by M. Misiurewicz.

Theorem 5.2.2 (Variational Principle)

Let $T : X \to X$ be a continuous map of a compact metric space X, then

$$h_{top}(T \mid \xi^-) = \sup_{\mu \in M(X,T)} h_\mu(T \mid \xi^-)$$

PROOF

Let $\mu \in M(X,T)$. We show that $h_\mu(T \mid \xi^-) \le h_{top}(T \mid \xi^-)$. Let $\zeta = \{A_1, ..., A_k\}$

be a finite partition of X. Choose $\epsilon > 0$ so that $\epsilon < \frac{1}{k\log k}$. Then we can choose compact sets $B_j \subset A_j$, $1 \leq j \leq k$, with $\mu(A_j \setminus B_j) < \epsilon$ and $B_i \cap B_j = \phi$ if $i \neq j$. Let $\eta = \{B_0, B_1, ..., B_k\}$ where $B_0 = X \setminus \bigcup_{j=1}^{k} B_j$. We have $\mu(B_0) < k\epsilon$, and

$$H_\mu(\zeta \mid \eta) = -\sum_{i=0}^{k} \mu(B_i) \sum_{j=1}^{k} \frac{\mu(B_i \cap A_j)}{\mu(B_i)} \log \frac{\mu(B_i \cap A_j)}{B_i}$$

$$= -\mu(B_0) \sum_{j=1}^{k} \frac{\mu(B_0 \cap A_j)}{\mu(B_0)} \log \frac{\mu(B_0 \cap A_j)}{B_0} \quad \text{since if } i \neq 0, \frac{\mu(B_i \cap A_j)}{\mu(B_i)} = 0 \text{ or } 1$$

$$\leq \mu(B_0) \log k$$

$$< k\epsilon \log k < 1$$

So we have $H_\mu(\zeta \mid \eta) < 1$.

Then $\beta = \{B_0 \cup B_1, ..., B_0 \cup B_k\}$ is an open cover of X. We have if $n, k \geq 1$, $H_{\mu_{x,k}}(\bigvee_{i=0}^{n-1} T^{-i}\eta|_{T^{-k}(x)}) \leq \log \aleph(\bigvee_{i=0}^{n-1} T^{-i}\eta|_{T^{-k}(x)})$ where $\mu_{x,k}$ is the conditional measure of μ on $T^{-k}(x)$ and $\aleph(\bigvee_{i=0}^{n-1} T^{-i}\eta|_{T^{-k}(x)})$ denotes the number of nonempty set in the partition $\bigvee_{i=0}^{n-1} T^{-i}\eta$ under $T^{-k}(x)$. Let $\pi_k : X \to X|_{T^{-k}\xi}$ be the facotr map, by Lemma 5.2.6 and Lemma 5.2.7

$$H_\mu(\bigvee_{i=0}^{n-1} T^{-i}\eta \mid \xi^{-k}) = \int_{X|T^{-k}\xi} H_{\mu_{x,k}}(\bigvee_{i=0}^{n-1} T^{-i}\eta|_{T^{-k}(x)})\, d\pi_{k*}\mu$$

$$\leq \log(\sup_{x\in X} \aleph(\bigvee_{i=0}^{n-1} T^{-i}\eta)|_{T^{-k}(x)})$$

$$\leq \log \sup_{x\in X, r\geq 0} \aleph(\bigvee_{i=0}^{n-1} T^{-i}\beta)|_{T^{-r}(x)} \cdot 2^n)$$

Let k go to infinity, divide by n and n approach to infinity, therefore

$$h_\mu(\eta \mid \xi^-) \leq h_{top}(T \mid \xi^-, \beta) + \log 2 \leq h_{top}(T \mid \xi^-) + \log 2$$

So by Lemma 5.1.3

$$h_\mu(\zeta \mid \xi^-) \leq h_\mu(\eta \mid \xi^-) + H_\mu(\zeta \mid \eta)$$

$$\leq h_{top}(T \mid \xi^-) + \log 2 + 1$$

This gives $h_\mu(T \mid \xi^-) \leq h_{top}(T \mid \xi^-) + \log 2 + 1$ for all $\mu \in M(X, T)$.

This inequality holds for T^n so $n \cdot h_\mu(T \mid \xi^-) \leq n \cdot h_{top}(T, \xi^-) + \log 2 + 1$. We divide by n and let n approach to infinity. Hence $h_\mu(T \mid \xi^-) \le h_{top}(T/\xi^-)$.

Now we show the other inequality.

Let $\epsilon > 0$ be given. We should find some invariant measure μ such that

$$h_\mu(T \mid \xi^-) \ge \limsup_{n\to\infty} \frac{1}{n} \log \sup_{x\in X, k\ge 0} r(n,\epsilon,T^{-k}(x))$$

Let $s(\epsilon, X, T)$ be the right side of this inequality.

As in Lemma 4.5.4 let $E_{n,k}(x)$ be such an (n,ϵ) separated set for $T^{-k}(x)$ of maximal cardinality $s_{n,k}(\epsilon, X)$. Let $\sigma_{n,k,x} \in M(X)$ be the atomic measure concentrated uniformly on the points of $E_{n,k}(x)$, i.e. $\sigma_{n,k,x} = \frac{1}{s_{n,k}(\epsilon,X)} \sum_{y\in E_{n,k}(x)} \delta_y$. Let $\mu_{n,k} \in M(X)$ be defined by $\mu_{n,k} = \frac{1}{n}\sum_{i=0}^{n-1} \sigma_{n,k,x} \circ T^{-i}$. Since $M(X)$ is compact we can choose a subsequences $\{n_j, k_j\}$ of natural numbers such that

$$\lim_{j\to\infty} \frac{1}{n_j} \log s_{n_j,k_j}(\epsilon, X) = \limsup_{n\to\infty} \frac{1}{n} \log \sup_{k\ge 0, x\in X} r(n,\epsilon,T^{-k}(x))$$

and $\{\mu_{n_j,k_j}\}$ converges in $M(X)$ to some $\mu \in M(X)$. We know that μ is an invariant measure.

Now we choose a partition $\alpha = \{A_1, A_2, ..., A_k\}$ of X so that $\operatorname{diam}(A_i) < \epsilon$ and $\mu(\partial A_i) = 0$ for $1 \le i \le k$. Since no member of $\bigvee_{i=0}^{n-1} T^{-i}\alpha$ can contain more than one member of $E_{n,k}(x)$, then as in Remark 5.2.1

$$\begin{aligned}
\log s_{n,k}(\epsilon, X) &= H_{\sigma_{n,k,x}}\Big(\bigvee_{i=0}^{n-1} T^{-i}\alpha\Big) \\
&= H_{\sigma_{n,k,x}}\Big(\bigvee_{i=0}^{n-1} T^{-i}\alpha \mid \xi^{-k}\Big) \\
&\le H_{\sigma_{n,k,x}}\Big(\bigvee_{i=0}^{n-1} T^{-i}\alpha \mid \xi^{-k-m}\Big) \text{ for all positive integer m} \\
&= H_{\sigma_{n,k,x}}\Big(\bigvee_{i=0}^{n-1} T^{-i}\alpha \mid \xi^{-}\Big) \\
&\le \sum_{r=0}^{a(j)-1} H_{\sigma_{n,k,x}}\Big(T^{-(rq+j)}\Big(\bigvee_{i=0}^{q-1} T^{-i}\alpha \mid \xi^-\Big) + \sum_{l\in S} H_{\sigma_{n,k,x}}(T^{-l}\alpha) \\
&\le \sum_{r=0}^{a(j)-1} H_{\sigma_{n,k,x}\circ T^{-(rq+j)}}\Big(\bigvee_{i=0}^{q-1} T^{-i}\alpha \mid \xi^-\Big) + 2q\log(l)
\end{aligned}$$

Sum this inequality over j from 0 to $q-1$ and by Remark 5.2.1, then

$$q \log s_{n,k}(\epsilon, X) \le \sum_{p=0}^{n-1} H_{\sigma_{n,k,x} \circ T^{-p}}(\bigvee_{i=0}^{q-1} T^{-i}\alpha \mid \xi^-) + 2q^2 \log(l).$$

If we divide by n and use Remark 5.2.1 and the concavity of $-x \log x$ we can get

$$\frac{q}{n} \log s_{n,k}(\epsilon, X) \le H_{\mu_{n,k}}(\bigvee_{i=0}^{q-1} T^{-i}\alpha \mid \xi^-) + \frac{2q^2}{n} \log(l) \tag{5.3}$$

Since the members of $\bigvee_{i=0}^{q-1} T^{-i}\alpha$ have boundaries of μ-measure zero, by Lemma 5.2.5 we can claim that

$$\lim_{j\to\infty} H_{\mu_{n_j}, k_j}(\bigvee_{i=0}^{q-1} T^{-i}\alpha \mid \xi^-) = H_\mu(\bigvee_{i=0}^{q-1} T^{-i}\alpha \mid \xi^-)$$

Therefore replacing (n, k) by (n_j, k_j) in (5.3) and letting j go to infinity we have

$$qs(\epsilon, X, T) \le H_\mu(\bigvee_{i=0}^{q-1} T^{-i}\alpha \mid \xi^-) = H_\mu(\alpha_0^{q-1} \mid \xi^-).$$

where $s(\epsilon, X, T) = \lim_{j\to\infty} \frac{1}{n_j} \log s_{n_j,k_j}(\epsilon, X)$. We can divide by q and let q go to infinity to get

$$s(\epsilon, X, T) \le h_\mu(T \mid \xi^-, \alpha) \le h_\mu(T \mid \xi^-)$$

$\diamondsuit$

5.3 Preimage S-M-B Theorem

For each finite partition α of X, let $\mathcal{B}(\alpha)$ be the σ-algebra generated by α.

Definition 5.3.1 We define

(1) $\alpha_i^j = \bigvee_{k=i}^{j} T^{-k}\alpha$

(2) $\lim_{k\to\infty}\mathcal{B}(\alpha_1^l \vee T^{-k}(\xi)) =_{def} \mathcal{B}(\lim_{k\to\infty}(\alpha_1^l \vee T^{-k}(\xi)) = \bigcap_{k=1}^{\infty}\mathcal{B}(\alpha_1^l \vee T^{-k}(\xi))$

Now let (X,$\mathcal{B}$,μ) be a probability space, $\{(B_n)\}$ be a sequence of sub-σ-algebra of $\mathcal{B}$ and $\{X_n\}$ be a sequence of random variables. Then $\{(X_n, B_n) : n = 1, 2, ...\}$ is a martingale if

(1) $B_n \subset B_{n+1}$

(2) X_n is measurable w.r.t. B_n

(3) $E[|X_n|] < \infty$

(4) $E[X_{n+1}/B_n] = X_n$ a.e.

Theorem 5.3.1[1] Every L^1 bounded martingale converges a.e.

$\{(X_{-n}, B_{-n}) : n = 1, 2, 3, \ldots\}$ is a reversed martingale if (1),(2),(3) and (4) hold for $n \geq 1$.

Remark 5.3.1: For a reversed martingale, $\lim_{n\to\infty} X_{-n} = X$ exists and is integrable.

Lemma 5.3.1[[1],Theorem 35.9] As above, we have for all $A \in \alpha$

$$\lim_{k\to\infty} E(\chi_A \mid \alpha_1^n \vee T^{-k}(\xi)) = E(\chi_A \mid \lim_{k\to\infty} (\alpha_1^n \vee T^{-k}(\xi))) \text{ a.e.}$$

Lemma 5.3.2 Let $g_n = \lim_{k\to\infty} I_{\alpha|\alpha_1^n \vee T^{-k}(\xi)}$ for all n=1,2,3,... and $g^* = \sup_{n\geq 1} g_n$, then for each $\lambda \geq 0$ and each $A \in \alpha$, we have

$$\mu\{x \in A : g^*(x) > \lambda\} \leq e^{-\lambda}$$

PROOF

For each $A \in \alpha$, and n=1,2,... Consider

$$g_n^A = \lim_{k\to\infty} I_{A|\alpha_1^n \vee T^{-k}(\xi)} = -\lim_{k\to\infty} \log E(\chi_A \mid \alpha_1^n \vee T^{-k}(\xi))$$
$$= -\log E(\chi_A \mid \lim_{k\to\infty} (\alpha_1^n \vee T^{-k}(\xi)))$$

This shows that g_n^A exists and consider

$$B_n^A = \{x : g_1^A(x), \ldots, g_{n-1}^A(x) \leq \lambda, g_n^A(x) > \lambda\}$$

Since $B_l^A \in \mathcal{B}(\lim_{k\to\infty}(\alpha_1^l \vee T^{-k}(\xi)))$

$$\mu(B_l^A \cap A) = \int_{B_l^A} \chi_A \, d\mu$$
$$= \int_{B_l^A} E(\chi_A \mid \lim_{k\to\infty} (\alpha_1^l \vee T^{-k}(\xi)))\, d\mu$$
$$= \int_{B_l^A} e^{-g_l^A}\, d\mu \leq e^{-\lambda}\mu(B_l^A)$$

Therefore

$$\mu\{x \in A : g^*(x) > \lambda\} = \sum_{k=1}^{\infty} \mu(B_k^A \cap A) \leq e^{-\lambda} \sum_{k=1}^{\infty} \mu(B_k^A) \leq e^{-\lambda}$$

$\diamond$

Lemma 5.3.3 [[10],Corollary 6.2.2] $g^* \in L^1$.

Lemma 5.3.4 $\lim_{k\to\infty} I_{\alpha_0^l | T^{-k}(\xi)}$ exists.

PROOF

Since $I_{\alpha_0^l|T^{-k}(\xi)} = -\sum_{A\in\alpha_0^l} \log E(\chi_A \mid T^{-k}(\xi))\chi_A$,

And $\mathcal{B}(T^{-k}(\xi)) \supset \mathcal{B}(T^{-(k+1)}(\xi))$ for all $k \geq 1$

We also have $E(E(\chi_A \mid T^{-k+1}(\xi)) \mid_{T^{-k}(\xi)}) = E(\chi_A \mid T^{-k}(\xi))$

And $E(E(\chi_A \mid T^{-k}(\xi))) < \infty$ for all positive integer k.

by reversed martingale theorem, the limit exists. $\diamond$

Lemma 5.3.5 Let $g_n = \lim_{k\to\infty} I_{\alpha|\alpha_1^n \vee T^{-k}(\xi)}$,

then

$$g = \lim_{n\to\infty} g_n = \lim_{n\to\infty}\lim_{k\to\infty} I_{\alpha|\alpha_1^n\vee T^{-k}(\xi)} \text{ exists a.e in } L^1.$$

PROOF

Since $\mathcal{B}(\alpha_1^n \vee T^{-k}(\xi)) \subset \mathcal{B}(\alpha_1^{n+1} \vee T^{-k}(\xi))$ for all $k \geq 1$,

then $\mathcal{B}(\lim_{k\to\infty}(\alpha_1^n \vee T^{-k}(\xi))) \subset \mathcal{B}(\lim_{k\to\infty}(\alpha_1^{n+1} \vee T^{-k}(\xi)))$

And

$$\begin{aligned} g_n &= \lim_{k\to\infty} I_{\alpha|\alpha_1^n\vee T^{-k}(\xi)} \\ &= I_{\alpha|\lim_{k\to\infty}(\alpha_1^n \vee T^{-k}(\xi))} \text{ by Lemma 4.6.1} \\ &= -\sum_{A\in\alpha} \log E(\chi_A \mid \lim_{k\to\infty}(\alpha_1^n \vee T^{-k}(\xi)))\chi_A \end{aligned}$$

with

$$E(E(\chi_A \mid \lim_{k\to\infty} \alpha_1^{n+1} \vee T^{-k}(\xi) \mid_{\lim_{k\to\infty}\alpha_1^n\vee T^{-k}(\xi)}) = E(\chi_A \mid \lim_{k\to\infty} \alpha_1^n \vee T^{-k}(\xi))$$

Also, $E(E(\chi_A \mid \lim_{k\to\infty} \alpha_1^n \vee T^{-k}(\xi))) < \infty$ for all n by Lemma 5.3.3.
By Martingale Convergence Theorem, $g = \lim_{n\to\infty} g_n$ exists a.e. in L^1. ◊

Lemma 5.3.6 We let $\mathcal{B}_\infty = \bigvee_{n=1}^{\infty} \mathcal{B}_n$ if $\{\mathcal{B}_n\}$ is an increasing sequence of sub-σ-algebras of X and let $\mathcal{B}_\infty = \bigcap \mathcal{B}_n$ if $\{\mathcal{B}_n\}$ is a decreasing sequence, and α is a finite partition, then

$$\lim_{n\to\infty} H_\mu(\alpha \mid \mathcal{B}_n) = H_\mu(\alpha \mid \mathcal{B}_\infty).$$

PROOF

We show the decreasing case and a similar discussion for the increasing sequence. [[10], proposition 5.2.11]

Let $A \in \alpha$, because $E(E(\chi_A \mid \mathcal{B}_{n-1}) \mid \mathcal{B}_n) = E(\chi_A \mid \mathcal{B}_n)$,

by reversed martingale theorem and Billingsley [1], Theorem 35.9,

$\lim_{n\to\infty} E(\chi_A \mid \mathcal{B}_n) = E(\chi_A \mid \mathcal{B}_\infty)$.

And $I_{\alpha|\mathcal{B}_n} = -\sum_{A\in\alpha} \log E(\chi_A \mid \mathcal{B}_n) \cdot E(\chi_A \mid \mathcal{B}_n)$ is a bounded continuous function, by the bounded convergence theorem, we can get

$$\lim_{n\to\infty} H(\alpha \mid \mathcal{B}_n) = \lim_{n\to\infty} \int I_{\alpha|\mathcal{B}n}\, d\mu$$
$$= \int \lim_{n\to\infty} I_{\alpha|\mathcal{B}n}\, d\mu = H(\alpha \mid \mathcal{B}_\infty)$$

◊

Lemma 5.3.7 Let α be a finite partition, then

$$h_\mu(\alpha \mid \xi^-) = h_\mu(T \mid \xi^-, \alpha) = \lim_{n\to\infty} \lim_{k\to\infty} H_\mu(\alpha \mid \bigvee_{l=1}^{n-1} T^{-l}\alpha \vee T^{-k}(\xi))$$
$$= H_\mu(\alpha \mid \lim_{n\to\infty} \lim_{k\to\infty} \bigvee_{l=1}^{n-1} T^{-l}\alpha \vee T^{-k}(\xi))$$

PROOF

Since $\lim_{k\to\infty} H_\mu(\alpha \vee T^{-1}\alpha \vee \ldots \vee T^{-(j-1)}\alpha \mid T^{-k}(\xi)) = \lim_{k\to\infty} H_\mu(\alpha \mid \bigvee_{l=1}^{j-1} T^{-l}\alpha \vee$

$T^{-k}(\xi)) + \lim_{k\to\infty} H_\mu(\bigvee_{l=1}^{j-1} T^{-l}\alpha \mid T^{-k}(\xi))$,

Then

$$\begin{aligned}
&\lim_{k\to\infty} H_\mu(\alpha \mid \bigvee_{l=1}^{j-1} T^{-l}\alpha \vee T^{-k}(\xi)) \\
&= \lim_{k\to\infty} H_\mu(\alpha \vee T^{-1}\alpha \vee \ldots \vee T^{-(j-1)}\alpha \mid T^{-k}(\xi)) - \lim_{k\to\infty} H_\mu(\bigvee_{l=1}^{j-1} T^{-l}\alpha \mid \\
&T^{-k}(\xi)) \\
&= \lim_{k\to\infty} H_\mu(\bigvee_{l=0}^{j-1} T^{-l}\alpha \mid T^{-k}(\xi)) - \lim_{k\to\infty} H_\mu(\bigvee_{l=0}^{j-2} T^{-l}\alpha \mid T^{-(k-1)}(\xi))
\end{aligned}$$

We can get

$$\begin{aligned}
&\sum_{j=2}^{n} \lim_{k\to\infty} H_\mu(\alpha \mid \bigvee_{l=1}^{j-1} T^{-l}\alpha \vee T^{-k}(\xi)) \\
&= \lim_{k\to\infty} H_\mu(\bigvee_{l=0}^{n-1} T^{-l}\alpha \mid T^{-k}(\xi)) - \lim_{k\to\infty} H_\mu(\alpha \mid T^{-k}(\xi))
\end{aligned}$$

By Cesaro theorem and Lemma 5.3.6,

$$\begin{aligned}
h_\mu(\alpha \mid \xi^-) &= \lim_{n\to\infty} \lim_{k\to\infty} H_\mu(\alpha \mid \bigvee_{l=1}^{n-1} T^{-l}\alpha \vee T^{-k}(\xi)) \\
&= H_\mu(\alpha \mid \lim_{n\to\infty} \lim_{k\to\infty} \bigvee_{l=1}^{n-1} T^{-l}\alpha \vee T^{-k}(\xi)).
\end{aligned}$$

◇

Now we are ready to show the following theorem which is similar to Shannon-McMillan-Breiman theory.

Theorem 5.3.2 (preimage S-M-B theorem)

Let $T : X \to X$ be an ergodic m.p.t. on the probability space $(X, \mathcal{B}, \mu)$ and α a finite partition of X.

Then

$$\lim_{n\to\infty} \lim_{k\to\infty} \frac{1}{n+1} I_{\bigvee_{i=0}^{n} T^{-i}\alpha \mid T^{-k}(\xi)}(x) = h_\mu(T \mid \xi^-, \alpha) \text{ a.e.}$$

PROOF

Let $g_n = \lim_{k\to\infty} I_{\alpha \mid \bigvee_{l=1}^{n} T^{-l}(\alpha) \vee T^{-k}(\xi)}$ for n=1,2,3,...

Then we can find that

$$\begin{aligned}
\lim_{k\to\infty} I_{\alpha_0^n|T^{-k}(\xi)} &= \lim_{k\to\infty} I_{T^{-1}\alpha\vee\ldots\vee T^{-n}\alpha|T^{-k}(\xi)} \\
&\quad + \lim_{k\to\infty} I_{\alpha|T^{-1}\alpha\vee\ldots\vee T^{-n}\alpha\vee T^{-k}(\xi)} \\
&= \lim_{k\to\infty} I_{\alpha_0^{n-1}|T^{-(k-1)}(\xi)} \circ T \\
&\quad + \lim_{k\to\infty} I_{\alpha|T^{-1}\alpha\vee\ldots\vee T^{-n}\alpha\vee T^{-k}(\xi)} \\
&= g_n + g_{n-1} \circ T + \ldots + g_1 \circ T^{n-1} + g_0 \circ T^n \\
&= \sum_{s=0}^{n} g_{n-s} \circ T^s \text{ where } g_0 = \lim_{k\to\infty} I_{\alpha|T^{-k}(\xi)}
\end{aligned}$$

Let $g = \lim_{n\to\infty} \lim_{k\to\infty} I_{\alpha|\bigvee_{l=1}^{n-1} T^{-l}\alpha\vee T^{-k}(\xi)}$ By Lemma 5.3.5, g exists a.e. in L^1.

Then we can write

$$\begin{aligned}
\frac{1}{n+1}\lim_{k\to\infty} I_{\alpha_0^n|T^{-k}(\xi)} &= \frac{1}{n+1}\sum_{s=0}^{n} g_{n-s} \circ T^s \\
&= \frac{1}{n+1}\sum_{s=0}^{n} g \circ T^s + \frac{1}{n+1}\sum_{s=0}^{n}(g_{n-s} - g) \circ T^s
\end{aligned}$$

By Birkhoff Ergodic Theorem, Proposition 5.3.6 and 5.3.7,

$$\begin{aligned}
\lim_{n\to\infty}\frac{1}{n+1}\sum_{s=0}^{n} g \circ T^s &= \int_X g\, d\mu \\
&= \int \lim_{n\to\infty}\lim_{k\to\infty} I_{\alpha|\alpha_1^n\vee T^{-k}(\xi)}\, d\mu \\
&= h_\mu(\alpha, T \mid \xi^-) \text{ a.e.}
\end{aligned}$$

Therefore, we must show the following to prove the Theorem

$$\lim_{n\to\infty}\frac{1}{n+1}\sum_{s=0}^{n} |g_{n-s} - g| \circ T^s = 0 \text{ a.e.} \tag{5.4}$$

For each N=1,2,3,...,let $G_N = \sup_{s\geq N} |g_s - g|$. Then

$$\begin{aligned}
\frac{1}{n+1}\sum_{s=0}^{n} |g_{n-s} - g| \circ T^s &= \frac{1}{n+1}\sum_{s=0}^{n-N} |g_{n-s} - g| \circ T^s + \frac{1}{n+1}\sum_{s=n-N+1}^{n} |g_{n-s} - g| \circ T^s \\
&\leq \frac{1}{n+1}\sum_{s=0}^{n-N} G_N \circ T^s + \frac{1}{n+1}\sum_{s=n-N+1}^{n} |g_{n-s} - g| \circ T^s.
\end{aligned}$$

We fix N and let n go to infinity. Since $|g_{n-s} - g| \leq g^* + g \in L^1$, the second term above tends to 0 a.e. Similarly, $G_N \leq g^* + g \in L^1$,so we may apply Birkhoff Ergodic Theorem to the first term:

$$\lim_{n\to\infty} \frac{1}{n+1} \sum_{s=0}^{n-N} |g_{n-s} - g| \circ T^s \leq \lim_{n\to\infty} \frac{1}{n+1} \sum_{s=0}^{n-N} G_N T^s$$
$$= \int_X G_N \, d\mu \to 0$$

by the dominated convergence theorem and that $G_N \to 0$ a.e. then we can get (5.4). ◇

Lemma 5.3.8 $h_\mu(T \mid \xi^-) \leq h_\mu(T)$.

PROOF

For any finite partition α, $H_\mu(\alpha \mid \bigvee_{i=1}^{n-1} T^{-i}\alpha \vee T^{-k}(\xi)) \leq H_\mu(\alpha \mid \bigvee_{i=1}^{n-1} T^{-i}\alpha)$ and $h_\mu(T, \alpha) = \lim_{n\to\infty} H_\mu(\alpha \mid \bigvee_{i=1}^{n-1} T^{-i}\alpha)$

By Lemma 5.3.7, $h_\mu(T \mid \xi^-) \leq h_\mu(T)$. ◇

5.4 Ergodic Decomposition of Metric Preimage Entropy

Lemma 5.4.1:([14], Lemma 4.15) Let $r \geq 1$ be a fixed integer. For each $\epsilon > 0$ there exists $\delta > 0$ such that if $\zeta = \{A_1, ..., A_r\}, \eta = \{C_1, ..., C_r\}$ are any two partitions of X into r sets with $\Sigma_{i=1}^{r} \mu(A_i \Delta C_i) < \delta$, then $H_\mu(\zeta \mid \eta) + H_\mu(\eta \mid \zeta) < \epsilon$.

Lemma 5.4.2:(cf.[14], Theorem 8.3) Let $T : X \to X$ be a continuous map of a compact metric space. Let $(\zeta_n)_{n=1}^{\infty}$ be a sequence of partitions such that $\mathrm{diam}(\zeta_n) \to 0$. Then

$$h_\mu(T \mid \xi^-) = \lim_{n\to\infty} h_\mu(T \mid \xi^-, \zeta_n)$$

Proof:

Let $\epsilon > 0$. Choose a finite partition $\zeta = \{A_1, A_2, ..., A_r\}$ such that $h_\mu(T \mid \xi^-, \zeta) > h_\mu(T \mid \xi^-) - \epsilon$ if $h_\mu(T \mid \xi^-) < \infty$, or $h_\mu(T \mid \xi^-, \zeta) > 1/\epsilon$ if $h_\mu(T \mid \xi^-) = \infty$. Choose $\delta > 0$ to correspond to ϵ and r in Lemma 5.4.1. Choose compact sets $K_i \subset A_i, 1 \leq i \leq r$, with $\mu(A_i \setminus K_i) < \delta/(r+1)$. Let $\acute{\delta} = \inf_{i\neq j} d(K_i, K_j)$ and choose n with $\mathrm{diam}(\xi_n)$

$< \acute{\delta}/2$.

For $1 \leq i < r$ let $E_n^{(i)}$ be the union of all the elements of ζ_n that intersect K_i and let $E_n^{(r)}$ be the union of the remaining elements of ζ_n. Since $diam(\zeta_n) < \acute{\delta}/2$ each $C \in \zeta_n$ can intersect at most one K_i. Then $\acute{\zeta_n} = \{E_n^{(1)}, ..., E_n^{(r)}\}$ is so that $\acute{\zeta_n} \leq \zeta_n$ and $\mu(E_n^{(i)} \triangle A_i) < \delta$. By Lemma 1 we have $H_\mu(\zeta \mid \acute{\zeta_n}) < \epsilon$. Therefore if n is such that $diam(\zeta_n) < \acute{\delta}/2$, then

$$\begin{aligned} h_\mu(T \mid \xi^-, \zeta) &\leq h_\mu(T \mid \xi^-, \acute{\zeta_n}) + \epsilon \text{ by Lemma 5.2.3} \\ &\leq h_\mu(T \mid \xi^-, \zeta_n) + \epsilon \end{aligned}$$

Then $diam(\zeta_n) < \acute{\delta}/2$ implies $h_\mu(T \mid \xi^-, \zeta_n) > h_\mu(T \mid \xi^-) - 2\epsilon$ if $h_\mu(T \mid \xi^-) < \infty$ or $h_\mu(T, \zeta_n) > (1/\epsilon) - \epsilon$ if $h_\mu(T \mid \xi^-) = \infty$. Therefore we show that $\lim_{n\to\infty} h_\mu(T \mid \xi^-, \zeta_n)$ exists and equals $h_\mu(T \mid \xi^-)$.

◇

Ergodic Decomposition of Invariant Measures

Let $T : X \to X$ be a measurable map. We define $\sum_0(T)$ as the set of points $x \in X$ such that, for every continuous $f : X \to R$, the limit

$$\tilde{f} = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} f(T^j(x))$$

exists. Further, let $C^0(X)$ be the space of continuous functions $f : X \to R$ endowed with the norm $\|f\| = \sup_{x\in X} \|f(x)\|$. For $x \in \sum_0(T)$ we define $L_x : C^0(X) \to R$ by

$$L_x(f) = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} f(T^j(x))$$

Then L_x is a positive linear functional and $L_x(f) = 1$, so that by Riesz's representation theorem there exists a unique probability measure μ_x on X such that

$$\int_X f \, d\mu_x = L_x(f)$$

We define $\sum_1(T)$ as the set of $x \in \sum_0(T)$ such that μ_x is T-invariant. Then we define $\sum_2(T)$ to be the set of $x \in \sum_1(T)$ for which μ_x is ergodic, and $\sum(T)$ to be the set of $x \in \sum_2(T)$ for which x belongs to the support of μ_x. Now let μ be an invariant measure. Then every integrable function f is μ_x-integrable for μ-almost every $x \in \sum(T)$ and

$$\int_X \left(\int_X f \, d\mu_x \right) d\mu = \int_X f \, d\mu$$

Theorem 5.4.1: Ergodic Decomposition of measure-theoretic entropy

Let (X,T) be a compact dynamical system and α a finite partition. Let $\mu \in M(X,T)$ and $\{\mu_x : x \in E\}$ with $\mu(E) = 1$. Then $x \to h_{\mu_x}(T,\alpha)$ and $x \to h_{\mu_x}(T)$ are measurable functions with

$$h_\mu(T,\alpha) = \int_X h_{\mu_x}(T,\alpha) \, d\mu$$

and

$$h_\mu(T) = \int_X h_{\mu_x}(T) \, d\mu$$

Theorem 5.4.2: Ergodic Decomposition of Metric Preimage Entropy

Let (X,T) be a compact dynamical system and α a finite partition. Let $\mu \in M(X,T)$ and $\{\mu_x : x \in E\}$ with $\mu(E) = 1$. Then $x \to h_{\mu_x}(T \mid \xi^-,\alpha)$ and $x \to h_{\mu_x}(T \mid \xi^-)$ are measurable functions with

$$h_\mu(T \mid \xi^-,\alpha) = \int_X h_{\mu_x}(T \mid \xi^-,\alpha) \, d\mu$$

and

$$h_\mu(T \mid \xi^-) = \int_X h_{\mu_x}(T \mid \xi^-) \, d\mu$$

Proof:

By Lemma 5.3.7, we have

$$\begin{aligned}
h_\mu(T \mid \xi^-,\alpha) &= H_\mu(\alpha \mid \lim_{n\to\infty} \lim_{k\to\infty} \alpha_1^{n-1} \vee T^{-k}(\xi)) \\
&= \int_X \lim_{n\to\infty} \lim_{k\to\infty} I_{\alpha|\alpha_1^{n-1} \vee T^{-k}(\xi)} \, d\mu \\
&= \int_X \left(\int_X \lim_{n\to\infty} \lim_{k\to\infty} I_{\alpha|\alpha_1^{n-1} \vee T^{-k}(\xi)} \, d\mu_x \right) d\mu \\
&= \int_X h_{\mu_x}(T \mid \xi^-,\alpha) \, d\mu
\end{aligned}$$

By Lemma 5.4.2, we choose a sequence of finite partitions (ζ_m) such that $diam(\zeta_m) \to 0$. Then

$$\begin{aligned}
h_\mu(T \mid \xi^-) &= \lim_{m\to\infty} h_\mu(T \mid \xi^-, \zeta_m) \\
&= \lim_{m\to\infty} \lim_{n\to\infty} \lim_{k\to\infty} H_\mu(\zeta_m \mid (\zeta_m)_1^{n-1} \vee T^{-k}(\xi)) \\
&= \lim_{m\to\infty} H_\mu(\lim_{n\to\infty} \lim_{k\to\infty} (\zeta_m \mid (\zeta_m)_1^{n-1} \vee T^{-k}(\xi)) \\
&= \lim_{m\to\infty} (\int_X \lim_{n\to\infty} \lim_{k\to\infty} I_{\zeta_m \mid (\zeta_m)_1^{n-1} \vee T^{-k}(\xi)}) \, d\mu \\
&= \lim_{m\to\infty} \int_X (\int_X \lim_{n\to\infty} \lim_{k\to\infty} I_{\zeta_m \mid (\zeta_m)_1^{n-1} \vee T^{-k}(\xi)} \, d\mu_x) \, d\mu \\
&= \lim_{m\to\infty} \int_X h_{\mu_x}(T \mid \xi^-, \zeta_m) \, d\mu \\
&= \int_X \lim_{m\to\infty} h_{\mu_x}(T \mid \xi^-, \zeta_m) \, d\mu \\
&= \int_X h_{\mu_x}(T \mid \xi^-) \, d\mu
\end{aligned}$$

◇

Bibliography

[1] P.Billingsley;*Probability and Measure*; Wiley-Interscience Publication,(1994)

[2] R.Bowen; *Entropy-expansive maps*, Trans. Amer. Math. Soc. 164,323-331,(1972)

[3] M.Denker,C.Grillenberger and K.Sigmund; *Ergodic Theory on Compact Spaces*, Spring Lecture Notes in Math. 527, Spring: New York, (1976)

[4] D.Fiebig;U.Fiebig and Z.Nitecki; *Entropy and Preimage Sets*, preprint

[5] M.Hurley; *On topological entropy of maps*, Ergodic Theory and Dynamical Systems,15,557-568,(1995)

[6] J.L.Kelley; *General Topology*, Van Nostrand, Princeton, NJ,(1955)

[7] R.Langevin and P.Walczaki; *Entropie de l'image in inverse d'une application*. Bull.Soc.Math.France.120,237-250,(1992)

[8] R.Langevin and P.Walczak; *Entropie d'une dynamique*. C.R. Acad. Sci. Paris. 312. 141-144 ,(1991)

[9] Z.Nitecki and F.Przytycki; *Preimage entropy of mappings*, International Journal of Birfurcation and Chaos, Vol.9, No.9, 1815-1843, (1999)

[10] K. Petersen;*Ergodic Theory*; Cambridge University Press,(1981)

[11] C.Robinson; *Dynamical Systems; stability,symbolics and chaos*,CRC Press Inc., (1995)

[12] V.A.Rohlin; *On the Fundamental Ideals of Measure Theory,* American Mathematical Society, Translation Number 71, (1952).

[13] T.Romeo; *some fundamental properties of continuous functions and topological entropy,* Pacific Journal of Mathematics,Vol.141,No.2,(1990)

[14] P.Walters; *An Introduction to ergodic theory* Springer Lecture Notes, Vol.458, (1982)