72-5810

ROTHSCHILD, Joseph Gad, 1939-
ON THE COMPUTATION OF TOPOLOGICAL ENTROPY.

The City University of New York, Ph.D., 1971
Mathematics

University Microfilms, A XEROX Company, Ann Arbor, Michigan

ON THE COMPUTATION OF TOPOLOGICAL ENTROPY

by

JOSEPH GAD ROTHSCHILD

A dissertation submitted to the Graduate Faculty in Mathematics in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York.

1971

This manuscript has been read and accepted for the University Committee in Mathematics in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

August 24, 1971 Richard Sacksteder
date Professor Richard Sacksteder
Chairman of Examining Committee

August 24, 1971 Edgar A. Feldman
date Professor Edgar Feldman
Deputy Executive Officer

Professor Burton Randol

Professor Edgar A. Feldman

Supervisory Committee

ACKNOWLEDGMENTS

In connection with this work, it is a pleasure to acknowledge my indebtedness to Professor Richard Sacksteder, my advisor, and Dr. Alan Konheim.

Special thanks to my wife for her patience, love, and understanding.

Finally, I would like to thank the secretaries and The City University Mathematics Department.

TABLE OF CONTENTS

INTRODUCTION

A central problem in mathematics is the determination of conjugacy classes of maps: two continuous maps, f and g, are conjugate if there exists a homeomorphism ψ, such that $g = \psi f \psi^{-1}$. Topological entropy, $h(f)$, was introduced by R.L. Adler, A.G. Konheim, and M.H. McAndrew (1) as a conjugacy invariant, that is, if f and g are conjugate, then $h(f) = h(g)$. In recent years, topological entropy has been employed by various authors (2), (3), (7), (10), who have been concerned with conjugacy classes of maps.

In spite of the progress made in this area, it remains rather difficult to compute the topological entropy of particular maps. In fact, it is only possible to calculate the topological entropy explicitly for a limited class of maps or spaces. Accordingly, some of our work is restricted to spaces homeomorphic to an interval I, and we require the maps to be injective on subintervals in a finite partition of I. The results we obtain for an interval I can also be extended to a circle.

The notation and basic definitions relevant to the concept of topological entropy are contained in (1), and are reviewed in Chapter I. We include only those proofs which are needed for future reference.

In Chapter II, we show that the role of open covers in the definition of topological entropy can be filled by partitions consisting of a finite number of connected elements. We then prove that the role of these partitions can in fact be filled by a unique partition (Thm.II-1). Finally, we are able to represent the topological entropy, h(f), in terms of the iterates of a map f, (Thm.II-2). Thus, we have completely eliminated the role of open covers from the definition of topological entropy.

In the third chapter, we prove that for a map of X into itself, there is no loss of generality in computing the topological entropy for a restriction of the map to a subspace on which it is surjective (Thm.III-1).

An additional characterization of the topological entropy is obtained from an investigation of the multiplicity function $\eta(y)$, which is defined as the number of pre-images of $f^{-1}(y)$; this characterization is in terms of an expression analogous to the total variation (Thm.III-2).

In the last chapter we apply the techniques developed here to compute the topological entropy for certain classes of maps.

CHAPTER I. PRELIMINARIES

This chapter contains a brief summary of some of the results obtained by R.L. Adler, A.G. Konheim, and M.H. McAndrew (1).

Unless otherwise specified, X will denote a compact Hausdorff space, and all maps will be assumed to be continuous.

Definition (I-1): *For any cover* $\mathcal{A}$ *of* X, *denote by* $N(\mathcal{A})$ *the number of elements in a subcover of minimal cardinality. A subcover of a cover is called minimal, if no other subcover contains fewer elements.*

Note that if f is a mapping of X into itself, and $\mathcal{A}$ is a cover of X, then $f^{-1}\mathcal{A} = \{ f^{-1}A : A \in \mathcal{A} \}$ is also a cover.

Definition (I-2): *For any two covers,* $\mathcal{A}$ *and* $\mathcal{B}$ *of* X, *let the join of* $\mathcal{A}$ *and* $\mathcal{B}$ *be defined as the cover*

$$\mathcal{A} \vee \mathcal{B} = \{ A \cap B : A \in \mathcal{A},\ B \in \mathcal{B} \} .$$

Definition (I-3): *A cover* $\mathcal{B}$ *is a refinement of a cover* $\mathcal{A}$, *(denoted by* $\mathcal{A} \prec \mathcal{B}$), *if for every* B *in* $\mathcal{B}$ *there exists an* A *in* $\mathcal{A}$ *containing* B .

The following four results are easily verified; for the proofs, see (1). Let X be a compact Hausdorff space, and f a map from X into itself, then the following assertions are true.

(I-1) If $\mathcal{A} \prec \mathcal{B}$, then $N(\mathcal{A}) \leq N(\mathcal{B})$.

(I-2) $N(f^{-1}\mathcal{A}) \leq N(\mathcal{A})$ for all $\mathcal{A}$.

(I-3) If f is surjective, then $N(\mathcal{A}) = N(f^{-1}\mathcal{A})$.

(I-4) $N(\mathcal{A} \vee \mathcal{B}) \leq N(\mathcal{A})\, N(\mathcal{B})$.

We shall also abreviate the expression $\mathcal{A} \vee f^{-1}\mathcal{A} \vee, \dots, \vee f^{-n+1}\mathcal{A}$ to $\mathcal{A}^n$, where $\mathcal{A}^1 = \mathcal{A}$.

Proposition (I-1): *Let* $\mathcal{A}$ *be a finite cover of* X, *then*

$$\lim_{n\to\infty} 1/n \log N(\mathcal{A}^n)$$ *exists and is finite.*

Proof: We shall first prove that the sequence $\{\log N(\mathcal{A}^n)\}$ is subadditive, that is, $\log N(\mathcal{A}^{n+m}) \leq \log N(\mathcal{A}^n) + \log N(\mathcal{A}^m)$. In fact, $N(\mathcal{A}^{n+m}) = N(\mathcal{A}^m \vee f^{-m}\mathcal{A}^n) \leq N(\mathcal{A}^m)\, N(f^{-m}\mathcal{A}^n) \leq N(\mathcal{A}^m)\, N(\mathcal{A}^n)$. The first inequality follows from property (I-4), and the second inequality follows from property (I-2).

We now prove the general assertion that if $a_{n+m} \leq a_n + a_m$, and $a_n \geq 0$, $(n,m = 1,2,\dots)$, then $\lim_{n\to\infty} (1/n)a_n$ exists and is finite. For any fixed j, we can express n uniquely by, $n = mj+p$ for $p = 1,2,\dots,j-1$,

$$a_n/n \leq [a_{mj}/(mj+p) + a_p/(mj+p)] \leq [a_{mj}/mj + a_p/(mj+p)] \leq$$

$$\leq [a_j/j + a_p/(mj+p)],$$

where the first and last inequalities follow from the subadditive condition. Therefore, since $n \to \infty$ implies $m \to \infty$,

$$\limsup_{n\to\infty} a_n/n \leq \limsup_{n\to\infty} (a_j/j + a_p/(mj+p)) \leq a_j/j .$$

But this is true for all j, thus $\limsup_{n\to\infty} a_n/n \leq \liminf_{n\to\infty} a_n/n$.

This proves the limit exists; it must also be finite, since $(1/n)a_n \leq a_1$ for all n. This completes the proof of the proposition.

Definition (I-4): *Let* X *be a compact Hausdorff space, and* $\mathcal{A}$ *be an open cover of* X. *For a continuous map* f, *from* X *into* X, *the Topological Entropy of* f, *with respect to* $\mathcal{A}$, *is defined by*

$$\lim_{n\to\infty} 1/n \log N(\mathcal{A}^n) = h(f,\mathcal{A}).$$

Proposition (I-2): *If* $\mathcal{A}$ *and* $\mathcal{B}$ *are open covers such that* $\mathcal{A} < \mathcal{B}$, *then* $h(f,\mathcal{A}) \leq h(f,\mathcal{B})$.

Note: If we replace the open covers $\mathcal{A}$ and $\mathcal{B}$, in proposition (I-2), by partitions $\mathcal{A}'$ and $\mathcal{B}'$, consisting of a finite number of elements, we get an analogous result, namely,

$$\lim_{n\to\infty} 1/n \log N(\mathcal{A}'^n) \leq \lim_{n\to\infty} 1/n \log N(\mathcal{B}'^n).$$

Definition (I-5): *Let* X *be a compact Hausdorff space, and* f *be a continuous map from* X *into* X. *Define the Topological Entropy of* f, *as* $\sup_{\mathcal{A}} h(f,\mathcal{A}) = h(f)$, *where the supremum is taken over all open covers of* X. (*Note that* $0 \leq h(f) \leq \infty$.).

Remark 1: The compactness of X allows us to take the supremum in definition (I-5) over all finite open covers of X, as the following argument shows:

Let β be the class of all finite open covers of X, and α be the class of all open covers of X. Then clearly $\sup_{\mathcal{A}\in\alpha} h(f,\mathcal{A}) \geq \sup_{\mathcal{B}\in\beta} h(f,\mathcal{B})$. On the other hand, for every open cover $\mathcal{A}$, there exists finite open subcovers $\mathcal{B}$; this implies $N(\mathcal{A}^n) \leq N(\mathcal{B}^n)$ for all n, and thus $h(f,\mathcal{A}) \leq h(f,\mathcal{B})$. Therefore, $\sup_{\mathcal{A}\in\alpha} h(f,\mathcal{A}) \leq \sup_{\mathcal{B}\in\beta} h(f,\mathcal{B})$.

Definition (I-6): *A sequence of covers*, $\{\mathcal{A}_n\}$, *is a refining sequence if* (1) $\mathcal{A}_n < \mathcal{A}_{n+1}$ $(n=1,2,\ldots)$, *and* (2) *for every finite cover* $\mathcal{B}$, *there exists an* $\mathcal{A}_n$, *such that* $\mathcal{B} < \mathcal{A}_n$.

A refining sequence of open covers, when it exists, simplifies the computation of the topological entropy, as the following results reveal.

Proposition (I-3): *If* $\{\mathcal{A}_n\}$ *is a refining sequence of open covers, then* $h(f) = \lim_{n \to \infty} h(f,\mathcal{A}_n)$.

Proof: The result is an immediate consequence of proposition (I-2) and remark 1.

For metric spaces (X,d), proposition (I-3) can be restates in terms of the following definition and the Lebesgue Covering Lemma.

Definition (I-7): *The diameter* $d(\mathcal{A})$ *of a cover* $\mathcal{A}$ *is defined by* $d(\mathcal{A}) = \sup_{A \in \mathcal{A}} \{ d(A) \}$, *where* $d(A)$ *is the diameter of the set* A.

Lebesgue Covering Lemma: *For every open cover* $\mathcal{A}$ *of a compact metric space* (X,d), *there exists* $\delta > 0$ *such that if* $U \subseteq X$ *is a set for which* $d(U) < \delta$, *then* U *is contained in one of the elements of* $\mathcal{A}$. *The supremum over all such numbers* δ *is called the Lebesgue number of* $\mathcal{A}$.

Rephrasing of Lebesgue's Covering Lemma: *For open covers* $\mathcal{A}$ *and* $\mathcal{B}$ *of a compact metric space* (X,d), *if* $d(\mathcal{B})$ *is less than the Lebesgue number of* $\mathcal{A}$, *then* $\mathcal{A} < \mathcal{B}$.

Proposition (I-3) now takes the form,

Corollary (I-1): *If* $\{\mathcal{A}_n\}$ *is a sequence of finite open covers of the compact metric space* (X,d), *such that* $\mathcal{A}_n < \mathcal{A}_{n+1}$ *and* $d(\mathcal{A}_n) \to 0$ *as* $n \to \infty$, *then* $\{\mathcal{A}_n\}$ *is a refining sequence and* $h(f) = \lim_{n \to \infty} h(f,\mathcal{A}_n)$.

A consequence of the above corollary is the existence of refining sequences for compact metric spaces.

The remaining propositions in this chapter exhibit some of the properties of the topological entropy. The proofs of these propositions can be found in (1) and will be omitted here except for those needed for reference in subsequent chapters.

Proposition (I-4): [(1); Theorem 2, p.311] Let X and Y be compact Hausdorff spaces, f be a continuous map from X into X, and ψ be a homeomorphism of X onto Y, then $h(f) = h(\psi f \psi^{-1})$.

The original proof of the next assertion [(1); Theorem 3, p.312] is incorrect, and a correct version can be found in (5). We offer another proof in chapter IV for $X = [a,b]$.

Proposition (I-5): Let f and g be continuous maps of X into itself and Y into itself, respectively. Then $h(f \times g) = h(f) + h(g)$, where $f \times g$ maps $X \times Y$ into itself, and is defined by $(f \times g)(x,y) = (f(x),g(y))$ for x in X, and y in Y .

Lemma (I-1): Let $\{a_n\}$ and $\{b_n\}$ be two sequences such that $a_n \geq 0$ and $b_n \geq 0$ for all n. Suppose $a = \lim_{n \to \infty} 1/n \log a_n$, and $b = \lim_{n \to \infty} 1/n \log b_n$ exists, then $\lim_{n \to \infty} 1/n \log(a_n + b_n) = \max.\{a,b\}$.

Proof: For any $c > \max.\{a,b\}$ there exists an integer p, such that $\log a_n < nc$ and $\log b_n < nc$ whenever $n \geq p$. Thus

$\log(a_n + b_n) < nc + \log 2$ and consequently,

$$\max.\{a,b\} \leq \liminf_{n \to \infty} 1/n \log(a_n + b_n) \leq \limsup_{n \to \infty} 1/n \log(a_n + b_n) < c .$$

Therefore, $\lim_{n \to \infty} 1/n \log(a_n + b_n) = \max.\{a,b\}$.

Proposition (I-6): [(1); Theorem 4, p.313] Let X_1 and X_2 be two closed subsets of X such that $X_1 \cup X_2 = X$ and $f(X_i) \subseteq X_i$, $(i=1,2)$.

Then $h(f) = \max.\{\, h(f|_{X_1}),\ h(f|_{X_2})\}$.

Proposition (I-7): $h(f^k) = k\, h(f)$, $(k = 1,2,\ldots)$.

Proof:

$$h(f^k) \geq h(f^k,\mathcal{A}^k) = \lim_{n\to\infty} 1/n \log N(\mathcal{A}^k \vee f^{-k}\mathcal{A}^k \vee \ldots \vee f^{(-n+1)k}\mathcal{A}^k)$$

$$= k \lim_{n\to\infty} 1/nk \log N(\mathcal{A}^{-nk+1}) = k\, h(f,\mathcal{A})$$

for any open cover $\mathcal{A}$. Therefore, $h(f^k) \geq k\, h(f)$. On the other hand,

$$\mathcal{A} \vee (f^k)^{-1}\mathcal{A} \vee \ldots \vee (f^k)^{-n+1}\mathcal{A} \prec \mathcal{A} \vee f^{-1}\mathcal{A} \vee \ldots \vee f^{-nk+1}\mathcal{A} = \mathcal{A}^{nk} .$$

Therefore, by property (I-1), $h(f,\mathcal{A}) = \lim_{n\to\infty} 1/nk \log N(\mathcal{A}^{nk}) \geq 1/k\, h(f^k,\mathcal{A})$, where the inequality follows from proposition (I-2). As this result is true for all open covers $\mathcal{A}$, we get the inequality $h(f) \geq 1/k\, h(f^k)$.

The above propositions were mainly concerned with the properties of the topological entropy, but are of little help in its actual computation. One appreciates the difficulty in calculating the topological entropy by considering the paper by M.H. McAndrew and R.L. Adler (8), which is devotes entirely to the computation of the topological entropy of the Chebyshev Polynomial; in chapter IV, we give a simple computation based on the results of this paper. The difficulty lies in evaluating the quantity $N(\mathcal{A}^n)$, which depends on the given map and the choice of the open cover $\mathcal{A}$. We shall show in the following chapter how the dependence on the cover can be neglected, for $X = [a,b]$.

CHAPTER II. TOPOLOGICAL ENTROPY WITH RESPECT TO PARTITIONS.

One of the original conjectures posed by the authors of (1) was that the topological entropy of a map could be calculated by considering partitions of X rather then covers consisting of open sets. This conjecture suggests itself since, as we have seen in the previous chapter, the results leading to the definition of $h(f,\mathcal{R})$ are true for arbitrary finite covers $\mathcal{R}$, in particular for finite partitions. Furthermore, by remark 1, the topological entropy of a map can be defined by taking the supremum over all finite open covers of X.

In this chapter we shall prove that the above conjecture is indeed true, that is, the topological entropy can be evaluated by considering finite partitions of X, where X is an interval. The interval $[a,b]$ will be denoted by I.

Lemma(II-1): *Let $\mathcal{R}$ be a partition of I consisting of connected sets and let f be injective on each R in $\mathcal{R}$. Then for every n, all elements of $\mathcal{R}^n$ are connected.*

Proof: The proof is by induction on n.

By assumption the assertion is true for $n = 1$. Assume the assertion is true for $n = k$, that is, for all P in $\mathcal{R}^k$, P is connected.

Let $S = R \cap f^{-1}P$ be a typical element of $\mathcal{R}^{k+1}$ where R is in $\mathcal{R}$ and P is in $\mathcal{R}^k$. P and $f(R)$ are connected by the induction hypothesis and the definition of f. But the restriction $f|_R$ is a homeomorphism, and thus $S = (f|_R)^{-1}(P \cap f(R))$ is connected.

Lemma (II-2): *Let $\mathcal{R}$ be a partition of a compact Hausdorff space X, and let f be injective on each element of $\mathcal{R}$. Then f^k is injective on each element of $\mathcal{R}^n$, $k = 1,2,\ldots,n$.*

Proof: The proof is by induction on n.

For $n = 1$ the assertion is true by hypothesis. Assume that f^k is injective on P in $\mathcal{R}^n$, $(k = 1,2,\ldots,n)$, and let $S = P \cap f^{-n}R$, be a typical element of $\mathcal{R}^{n+1}$, where R is in $\mathcal{R}$ and P is in $\mathcal{R}^n$. Since S is contained in P, f^k is injective on S for $k = 1,2\ldots n$. Furthermore, for x,y in S, $x \neq y$, the induction hypothesis implies, $f^n(x) \neq f^n(y)$. Since $f^n(x)$ and $f^n(y)$ are in R, $f(f^n(x)) = f^{n+1}(x) \neq f^{n+1}(y)$, and the lemma is proved.

It will be useful, in what follows, to calculate the topological entropy of a map f with respect to a special class of partitions, which we shall call f - injective. We shall now define these partitions and prove their existance.

Definition (II-1): *Let f map X into itself. A partition of X consisting of the family of disjoint subsets $\{ \mathcal{U}_\alpha : \alpha \in A \}$, and $F = X - \bigcup_{\alpha \in A} \mathcal{U}_\alpha$, is called f-injective if,*

(i) *Each $\mathcal{U}_\alpha$ is open and connected.*

(ii) *$f|_{\mathcal{U}_\alpha}$ is injective.*

(Either the family $\{ \mathcal{U}_\alpha : \alpha \in A \}$ or F may be empty.)

Since the collection $\{ \mathcal{U}_\alpha : \alpha \in A \}$, specifies the partition, we shall sometimes abuse language and refer to $\{ \mathcal{U}_\alpha : \alpha \in A \}$ as the injective partition of X.

We partially order the injective partitions of X by defining a partial order relation, $\{ \mathcal{U}_\alpha : \alpha \in A \} < \{ \mathcal{V}_\beta : \beta \in B \}$ if for every $\mathcal{V}_\beta$, β in B, there exists $\mathcal{U}_\alpha$, α in A such that $\mathcal{V}_\beta \subseteq \mathcal{U}_\alpha$.

Definition (II-2): An f-injective partition is minimal if it is minimal with respect to the partial ordering defined above. A minimal f-injective partition will be denoted by $\mathcal{P}_o$.

Definition (II-3): Let $\mathcal{P}_o$ *be a minimal f-injective partition of* X. *An open connected element of* $\mathcal{P}_o$ *is called a branch of* f *with respect to* $\mathcal{P}_o$, *and* $N(f,\mathcal{P}_o)$ *will denote the number of these branches. One further defines the number of branches of* f *by* $N(f) = \min_{\mathcal{P}_o} N(f,\mathcal{P}_o)$, *where the minimum is taken over all minimal* f - *injective partitions.*

Proposition (II-1): Let X *be a compact Hausdorff space, and* f *a continuous map of* X *into itself. Then a minimal* f-*injective partition exists.*

Proof: If f is a constant map, the assertion is trivially true; therefore, we may assume that f is not identically constant. Let $\mathcal{A}$ be a collection of open connected disjoint sets, on which f is injective. Let $\Pi = \{ \mathcal{A}_\alpha : \alpha \in \Gamma \}$ be the family of all such collections with the partial order relation defined above. Let $\{ \mathcal{A}_\alpha : \alpha \in T \subseteq \Gamma \}$ be a linearly ordered subset of Π. Define $A_T(x) = \bigcup_{\alpha \in T} A(x,\alpha)$ where $A(x,\alpha)$ is the element of $\mathcal{A}_\alpha$ containing x when such an element exists; and $A(x,\alpha)$ is empty otherwise. $A_T(x)$ is an open connected (and possibly empty) set on which f is injective because, f is injective on each element of the nested collection of open connected sets $A(x,\alpha)$. It is easy to check that if y is in $A_T(x)$, then $A_T(y)$ is contained in $A_T(x)$; hence, $A_T(x) = A_T(y)$.

The disjoint collection $\{ A_T(x) \}$ is a lower bound for the linearly ordered set $\{ \mathcal{A}_\alpha : \alpha \in T \}$. By Zorn's lemma, there exists a minimal element for the family Π. This completes the proof.

Remark 2: If $\mathcal{P}_o$ is a minimal f-injective partition of X, and f is not constant on any open subset of X, then the set $F = X - \bigcup_{\alpha \in A} \mathcal{U}_\alpha$, as defined in definition (II-1), has an empty interior. If, in addition, X = I, then $\mathcal{P}_o$ is unique and we denote $N(f, \mathcal{P}_o)$ by N(f); we call N(f) the number of branches of f. In fact, if N(f) = M, then F is a discrete set consisting of M - 1 points.

For the remainder of this chapter whenever X = I we shall make the following three assumptions:

(i) Partitions of I will have a finite number of elements.

(ii) Maps from I into I will have a finite number of branches.

(iii) Maps from I into I are not constant on any open subset of I.

Note that with the above three assumptions, propositions (I-1) and (I-2) remain true.

Lemma (II-3): If $\mathcal{R}$ and S are finite partitions of I consisting of connected sets, then the partition $\mathcal{R} \vee S$ consists of at most $N(\mathcal{R}) + N(S)$ connected sets.

Proof: Let $\mathcal{R}$ be a partition containing $N(\mathcal{R})$ connected elements, and S be a partition consisting of one connected element, I. Then $\mathcal{R} \vee S = \mathcal{R}$ and $N(\mathcal{R} \vee S) \le N(\mathcal{R}) + N(S)$. We proceed by induction; assume that for any two partitions $\mathcal{R}$ and S, for which N(S) = n, we have $N(\mathcal{R} \vee S) \le N(\mathcal{R}) + N(S)$.

Consider a partition S for which N(S) = n + 1, where $S = \{ S_1, S_2, \ldots, S_n, S_{n+1} \}$. Without essential loss of generality, assume S_n and S_{n+1} abut. Then $S_n \cup S_{n+1}$ is connected and we define the partition $S' = \{ S_1, S_2, \ldots, (S_n \cup S_{n+1}) \}$, where N(S') = n .

Consider $R \cap (S_n \cup S_{n+1}) = (R \cap S_n) \cup (R \cap S_{n+1})$ for some R in $\mathcal{R}$, where $R \cap (S_n \cup S_{n+1}) \in \mathcal{R} \vee S'$ and $R \cap S_i \in \mathcal{R} \vee S$ ($i = n, n+1$). Since $R \cap (S_n \cup S_{n+1})$ is connected, $R \cap S_n$ and $R \cap S_{n+1}$ abut, then there is at most one set R in $\mathcal{R}$ for which $R \cap S_n \neq \emptyset$, and $R \cap S_{n+1} \neq \emptyset$. Therefore

$$N(\mathcal{R} \vee S) = N(\mathcal{R} \vee S') + 1 \leq N(\mathcal{R}) + N(S') + 1 = N(\mathcal{R}) + N(S).$$

Note that for higher dimensional spaces, lemma (II-3) is false; in general, by assertion (I-4) $N(\mathcal{R} \vee S) \leq N(\mathcal{R})\, N(S)$.

Proposition (II-2): *Let* $\mathcal{P}$ *and* $\mathcal{R}$ *be finite partitions of* I *consisting of connected sets, such that* $\mathcal{P} < \mathcal{R}$ *and* f *is injective on each element of* $\mathcal{P}$. *Then* $N(\mathcal{R}^n) \leq n\, N(\mathcal{R})\, N(\mathcal{P}^n)$ *for all* n.

Proof: The proof is by induction on n. For $n = 1$, $N(\mathcal{R}) \leq N(\mathcal{R}) N(\mathcal{P})$ is trivially true. Since $\mathcal{P} < \mathcal{R}$, we have $\mathcal{P}^k < \mathcal{R}^k$. For a P in $\mathcal{P}^k$ let $R_1, R_2, \ldots, R_{n(P)}$ be such that $\bigcup_{j=1}^{n(P)} R_j = P$, where R_j belongs to $\mathcal{R}^k$ for $j = 1,2,\ldots,n(P)$. By the induction hypothesis

$$\text{(II-1)} \qquad N(\mathcal{R}^k) = \sum_{P \in \mathcal{P}^k} n(P) \leq k\, N(\mathcal{R})\, N(\mathcal{P}^k).$$

Now consider all sets S in $\mathcal{R}^{k+1}$, for which $S \subseteq P$. A typical element S must be of the form $S_{ij} = R_j \cap f^{-k} R'_i$, where R'_i is in $\mathcal{R}$ for $i = 1,2,\ldots,N(\mathcal{R})$. By lemma (II-2), f^k is injective on P, and by lemma (II-1), P, R_j, R'_i are connected; thus $\{ f^{-k}R' \cap P\colon i=1,2,\ldots,N(\mathcal{R}) \}$ partitions P into at most $N(\mathcal{R})$ connected sets. By lemma (II-3), $\{ S_{ij}\colon i=1,2,\ldots,N(\mathcal{R});\ j=1,2,\ldots,n(P) \}$ partitions P into at most $N(\mathcal{R}) + n(P)$ connected sets; hence,

$$N(\mathcal{R}^{k+1}) \leq \sum_{P \in \mathcal{P}^k} [N(\mathcal{R}) + n(P)] \leq N(\mathcal{R})N(\mathcal{P}^k) + kN(\mathcal{R})N(\mathcal{P}^k) \leq (k+1)N(\mathcal{R})N(\mathcal{P}^k)$$

where the second inequality follows from equation (II-1).

But $\mathcal{P}^{k} < \mathcal{P}^{k+1}$ so $N(\mathcal{P}^{k}) \leq N(\mathcal{P}^{k+1})$; therefore

$$N(\mathcal{R}^{k+1}) \leq (k+1)\, N(\mathcal{R})\, N(\mathcal{P}^{k+1}).$$

Definition (II-4): *For a finite partition* $\mathcal{R}$ *of a compact Hausdorff space* X *we define* $\overline{h}(f,\mathcal{R}) = \lim_{n\to\infty} 1/n \log N(\mathcal{R}^{n})$ *and* $\overline{h}(f) = \sup_{\mathcal{R}} \overline{h}(f,\mathcal{R})$, *where the supremum is taken over all finite partitions consisting of connected elements.*

Note: By proposition (I-1), $\overline{h}(f,\mathcal{R})$ exists and is finite.

Corollary (II-1): *If* $\mathcal{P}$ *and* $\mathcal{R}$ *are partitions of* I *which satisfy the hypothesis of proposition (II-2), then* $\overline{h}(f,\mathcal{R}) = \overline{h}(f,\mathcal{P})$.

Proof: Since $\mathcal{P} < \mathcal{R}$, $N(\mathcal{P}^{n}) \leq N(\mathcal{R}^{n})$ and $\overline{h}(f,\mathcal{P}) \leq \overline{h}(f,\mathcal{R})$. From proposition (II-2) $1/n \log N(\mathcal{R}^{n}) \leq 1/n \log[n\, N(\mathcal{R})\, N(\mathcal{P}^{n})]$. Passing to the limit, we get $\overline{h}(f,\mathcal{R}) \leq \overline{h}(f,\mathcal{P})$ and therefore $\overline{h}(f,\mathcal{R}) = \overline{h}(f,\mathcal{P})$.

Remark 3: In order to apply Corollary (II-1), the partitions considered must consist of connected elements. Accordingly, let $\mathcal{M}_{o}$ be the partition consisting of the open connected elements P of $\mathcal{P}_{o}$ (on which f is injective) and the collection $\{\{x\} : x \in F\}$, where $F = X - \bigcup_{P\in\mathcal{P}_{o}} P$.

Corollary (II-2): *Let* $\mathcal{M}_{o}$ *be a partition of* I, *as described in remark 3, and suppose* N(f) *is finite, then* $\overline{h}(f) = \overline{h}(f,\mathcal{M}_{o})$.

Proof: By remark 2 $N(\mathcal{M}_{o}) = 2N(f) - 1$; hence, $N(\mathcal{M}_{o})$ is finite. Therefore, for every finite partition $\mathcal{R}$ for which $\mathcal{M}_{o} < \mathcal{R}$ we have $\overline{h}(f,\mathcal{M}_{o}) = \overline{h}(f,\mathcal{R})$ by corollary (II-1). And so by proposition (I-2), $\overline{h}(f) = \overline{h}(f,\mathcal{M}_{o})$.

We shall now prove that $h(f) = \overline{h}(f)$. We need the following three results.

Lemma (II-4): Let $\mathcal{R}$ *be a finite cover of the compact Hausdorff space* X. *Then every element* R *of* $\mathcal{R}^{n+m}$ *is in* $(\mathcal{R}^n)^{m+1}$.

Proof: By definition, $R = R_o \cap f^{-1}R_1 \cap \ldots \cap f^{-n-m+1}R_{n+m-1}$ for R_i in $\mathcal{R}$. But

$$R = (R_o \cap f^{-1}R_1 \cap \ldots \cap f^{-n+1}R_{n-1}) \cap f^{-1}(R_1 \cap f^{-1}R_2 \cap \ldots \cap f^{-n+1}R_n) \cap, \ldots,$$
$$\cap f^{-m}(R_m \cap f^{-1}R_{m+1} \cap \ldots \cap f^{-n+1}R_{n+m-1}),$$

that is R is in $(\mathcal{R}^n)^{m+1}$.

Lemma (II-5): Let $\mathcal{R}$ *be a finite partition of the compact Hausdorff space* X, *then* $\bar{h}(f,\mathcal{R}) = \bar{h}(f,\mathcal{R}^n)$ *for all* n.

Proof: Since $\mathcal{R} < \mathcal{R}^n$, it follows that $\bar{h}(f,\mathcal{R}) \le \bar{h}(f,\mathcal{R}^n)$ for all n. By lemma (II-4), $N(\mathcal{R}^n)^{m+1} \le N(\mathcal{R}^{n+m})$. Therefore,

$$\bar{h}(f,\mathcal{R}^n) = \lim_{m\to\infty} 1/(m+1) \log N[(\mathcal{R}^n)^{m+1}] \le \lim_{m\to\infty} 1/(m+1) \log N(\mathcal{R}^{n+m}) =$$

$$= \lim_{m\to\infty} 1/(n+m) \log N(\mathcal{R}^{n+m}) = \bar{h}(f,\mathcal{R}) \quad \text{for all } n.$$

Proposition (II-3): Let $\mathcal{R}$ *be a finite partition of the compact Hausdorff space* X, *then* $\bar{h}(f^n,\mathcal{R}) = n\,\bar{h}(f,\mathcal{R})$ *and* $\bar{h}(f^n) = n\,\bar{h}(f)$.

Proof: From lemma (II-5)

$$\bar{h}(f^n,\mathcal{R}) = \bar{h}(f^n,\mathcal{R}^n) = \bar{h}(f^n,\mathcal{R} \vee f^{-1}\mathcal{R} \vee \ldots \vee f^{-n+1}\mathcal{R}) =$$

$$= n \lim_{m\to\infty} 1/nm \log N(\mathcal{R} \vee f^{-1}\mathcal{R} \vee \ldots \vee f^{-nm+1}\mathcal{R}) = n\,\bar{h}(f,\mathcal{R})$$

for all $\mathcal{R}$.

The second assertion follows by taking the supremum over all finite partitions with connected elements.

We shall now construct a finite open cover $\mathcal{A}_o$, of I and compare $h(f,\mathcal{M}_o)$ with $h(f,\mathcal{A}_o)$. Let $\mathcal{A}_o$ be a finite open cover of I consisting of the open connected elements of $\mathcal{M}_o$ and disjoint open connected intervals $\{ S_j \}$, $j=1,2,\ldots,N(f)-1$, such that each element of $\{ S_j \}$ contains exactly one of the remaining elements of $\mathcal{M}_o$.

Lemma (II-6): *Let* $X = I$, *and let* $\mathcal{M}_o$ *and* $\mathcal{A}_o$ *be as above, then at most* 3^n *elements of* $\mathcal{M}_o^n$ *are required to cover any element of* $\mathcal{A}_o^n$.

Proof: From the definition of $\mathcal{A}_o$ and $\mathcal{M}_o$ we observe that at most three elements of $\mathcal{M}_o$ are required to cover any set in $\mathcal{A}_o$. Therefore if $A = A_o \cap f^{-1}A_1 \cap \ldots \cap f^{-n+1}A_{n-1}$ (A_i in $\mathcal{A}_o$), then at most 3^n elements of $\mathcal{M}_o^n$ are required to cover A in $\mathcal{A}_o^n$.

Lemma (II-7): $h(f) + \log 3 \geq \bar{h}(f)$.

Proof: It suffices to prove that $h(f,\mathcal{A}_o) + \log 3 \geq \bar{h}(f,\mathcal{M}_o)$ since $h(f) \geq h(f,\mathcal{A}_o)$ and by corollary (II-2) $\bar{h}(f) = \bar{h}(f,\mathcal{M}_o)$. By lemma (II-6) $3^n N(\mathcal{A}_o^n) \geq N(\mathcal{M}_o^n)$, which implies the result.

We are now in a position to prove the main theorem of this chapter, that $h(f) = \bar{h}(f,\mathcal{M}_o)$ provided $N(f)$ is finite. This allows us to calculate the topological entropy by restricting our attention to a particular partition $\mathcal{M}_o$, and calculating $N(\mathcal{M}_o^n)$ in the same manner as for covers $\mathcal{A}$, consisting of open sets. This is a surprising result since for covers $\mathcal{A}$ consisting of open sets, $h(f) = h(f,\mathcal{A})$ only for separating covers (7), that is, covers for which $\bigcap_{-\infty}^{\infty} f^{-n}\bar{A}_n$ is at most a point, where A_n is in $\mathcal{A}$ and $\bar{A}$ is the closure of A. What will make the theorem work is lemma (II-7) which is true only because $\bar{h}(f) = \bar{h}(f,\mathcal{M}_o)$ in contrast to the case of topological entropy for which we only have $h(f) \geq h(f,\mathcal{A})$. The other interesting aspect of the theorem is that since $\bar{h}(f,\mathcal{M}_o)$ exists and is finite, $h(f)$ is always finite provided f has a finite number of branches.

Theorem (II-1): Let $I = [a,b]$ and suppose $N(f)$ is finite, then $h(f) = \overline{h}(f)$ exists and is finite.

Proof: By lemma (II-7) $h(f^n) + \log 3 \geq \overline{h}(f^n)$; applying proposition (I-7) and proposition (II-3), $n\,h(f) + \log 3 \geq n\,\overline{h}(f)$ for all n. Therefore, $h(f) \geq \overline{h}(f)$.

On the other hand, let $\{ \mathcal{A}_o < \mathcal{A}_i < \dots \}$ be a sequence of open covers, where $\mathcal{A}_o$ is an open cover defined as above, and $\mathcal{A}_i$ is a finite open cover consisting of connected sets. This is a refining sequence by corollary (I-1) provided we let $d(\mathcal{A}_i) \to 0$ as $i \to \infty$, where $d(\mathcal{A}_i) = \sup\limits_{A \in \mathcal{A}_i} \{ d(A) \}$. For each $\mathcal{A}_i$ there exists a finite partition $\mathcal{R}_i$, such that $\mathcal{A}_i < \mathcal{R}_i < \mathcal{R}_i \vee \mathcal{M}_o$, where $\mathcal{M}_o$ is as in remark 2. From assertion (I-1), $N(\mathcal{A}_i^n) \leq N(\mathcal{R}_i^n) \leq N[(\mathcal{R}_i \vee \mathcal{M}_o)^n]$ and thus $h(f,\mathcal{A}_i) \leq \overline{h}(f,\mathcal{R}_i) \leq \overline{h}(f,\mathcal{R}_i \vee \mathcal{M}_o) = \overline{h}(f)$, where the last equality follows from corollary (II-2). Therefore, by corollary (I-1), $h(f) = \lim\limits_{i \to \infty} h(f,\mathcal{A}_i) \leq \overline{h}(f)$.

It is worthwhile to examine the partitions $\mathcal{M}_o$ and $\mathcal{P}_o$ more closely since they play a major role in the revious results. The following propositions will illustrate the unique properties of $\mathcal{M}_o^n$ and $\mathcal{P}_o^n$.

Proposition (II-4): $\overline{h}(f,\mathcal{P}_o) = \overline{h}(f,\mathcal{M}_o) = h(f)$.

Proof: Since $\mathcal{P}_o < \mathcal{M}_o$ we have $\overline{h}(f,\mathcal{P}_o) \leq \overline{h}(f,\mathcal{M}_o)$ by proposition (I-2). It now suffices to observe that the open connected elements of $\mathcal{M}_o^n$ and $\mathcal{P}_o^n$ are identical, and if M_n denotes the number of these elements, then $N(\mathcal{M}_o^n) \leq 2M_n \leq 2\,N(\mathcal{P}_o^n)$, which implies $\overline{h}(f,\mathcal{M}_o) \leq \overline{h}(f,\mathcal{P}_o)$.

Therefore by theorem (II-1) $\overline{h}(f,\mathcal{M}_o) = \overline{h}(f,\mathcal{P}_o) = h(f)$.

Proposition (II-5): *Let* $\mathcal{P}_o$ *be a minimal* f-*injective partition of* I, *then* $\mathcal{P}_o^n$ *is a minimal* f^n-*injective partition of* I.

Proof: We prove that the open connected elements of $\mathcal{P}_o^n$ are the branches of f^n. By definition, the branches of f^n consist of maximal open sets on which f^n is injective. Therefore, since f^k, ($k=1,2,\ldots,n$), is injective on each open connected set $Q \in \mathcal{P}_o^n$, we see by lemma (II-1) and lemma (II-2) that Q must be contained in some branch of f^n. On the other hand, if P and Q are open connected elements in $\mathcal{P}_o^n$ such that $P \cup Q$ is contained in a branch of f^n, then there exists ξ such that without loss of generality $x < \xi < y$, for all x in P and for all y in Q, and such that $\alpha = f^k(\xi)$ for some α in F and for some k, $k=0,1,\ldots,n-1$. But then f^{k+1} is not injective in some neighborhood of ξ and therefore neither is f^n. This contradicts the assumption that $P \cup Q$ is contained in a branch of f^n. Therefore, $P \cup Q$ is a connected open element contained in $\mathcal{P}_o^n$.

Theorem (II-1) can now be restated in terms of $N(f^n)$.

Theorem (II-2): *Let* f *map* I *into* I *and suppose* N(f) *is finite*, *then* $h(f) = \bar{h}(f) = \lim_{n\to\infty} 1/n \log N(f^n)$ *exists and is finite*.

Proof: The proof follows easily from proposition (II-4) and proposition (II-5)., since $N(\mathcal{P}_o^n) \le 2M_n \le 2N(\mathcal{P}_o^n)$

For future reference, we state the following inequalities in equations (II-2) and (II-3).

(II-2) $N(f^{n+k}) \le N(f^n)\, N(f^k)$

(II-3) If A and B are subintervals of I, such that $A \subseteq B$, then $N(f|_A) \le N(f|_B)$.

CHAPTER III. TOTAL VARIATION

The final result of the last chapter, expressed the topological entropy in terms of the number of branches of the iterates of a map. In this chapter we relate the number of branches, and thus the topological entropy, to the "Multiplicity function", defined below.

The following notation will be used throughout this chapter. The interval $[a,b]$ will be denoted by I, and dx will denote Lebesgue measure on I.

Definition (III-1): Let f map X into itself. Define the multiplicity function of f, $\eta : X \to Z \cup \{\infty\}$, $Z = \{0,1,\dots\}$, by

$$\eta(y) = \text{cardinality } \{ x : f(x) = y \}.$$

Remark 4: If $N(f)$ is finite, it follows from eq. (II-2), that $N(f^n)$ is finite, and therefore we can define

$$\eta^n(y) = \text{cardinality } \{ x : f^n(x) = y \}.$$

The multiplicity function is Lebesgue integrable on I, if and only if, f has finite total variation, and we can express the total variation, $V(f)$, of a map, in terms of the following integral (9).

$$V(f) = \int \eta(y)dy$$

The integration is over the range of f.

Definition (III-2): Let f be a continuous map from I into I for which $N(f)$ is finite. Let μ denote probability measure, that is, a measure for which $\mu(I) = 1$. Define

$$1) \quad V(f,\mu) = \int \eta(y)\mu(dy).$$

2) $\mathcal{V}(f,\mu) = \lim_{n\to\infty} 1/n \log V(f^n,\mu)$

3) $\mathcal{V}(f) = \sup_{\mu} \mathcal{V}(f,\mu)$, where the supremum is taken over all probability measures.

(Note that $\mathcal{V}(f)$ can be negative unless f is surjective.)

The main result of this chapter will be to show that if f has a finite number of branches, then $h(f) = \mathcal{V}(f)$, and in special cases $h(f) = \mathcal{V}(f,dy)$. For differentiable maps we also have the useful relation $\mathcal{V}(f,dx) = \lim_{n\to\infty} 1/n \log \int |(Df^n)(x)|dx$, where Df^n is the derivative of f^n, and the integration is over the domain of f.

Before proceeding to the main theorem, we prove a preliminary result which will simplify many of the proofs, but is also of interest independent of its role in that theorem.

We shall prove there is no loss of generality in assuming that f is surjective when one calculates the topological entropy of f.

Let $X_n = f^n(X)$, $X_\infty = \bigcap_{n=0}^{\infty} X_n$ ($X_0 = X$), and $g = f|_{X_\infty}$.

Our aim is to prove $h(f) = h(g)$.

Proposition (III-1): Let f be a continuous map from the compact Hausdorff space X into itself, then g maps X_∞ onto itself.

Proof: First we show that g maps X_∞ into itself; this follows since $X_\infty \subseteq f(X_n) = X_{n+1}$ for all n, hence $g(X_\infty) \subseteq X_\infty$.

To prove g is onto, we first note that since f is continuous, $f^n(X) = X_n$ and X_∞ are closed non-empty subsets of X. Suppose y is in X_∞, we shall construct a point x in X_∞ for which $g(x) = y$. Since y is in X_∞ we have y in X_n for all n, which implies there exists x_{n-1} in X_{n-1} such that $f(x_{n-1}) = y$. Consider a sequence $\{x_n\}$, (x_n in X_n), such that $f(x_n) = y$ for all n .Since

X is compact, there exists a subsequence, $\{x_{n_i}\}$, such that x_{n_i} approaches x in X, as $i \to \infty$. But since x_{n_i} is in X_{n_i} and X_∞ is closed, x is in X_∞.

Lemma (III-1): *Let* $\mathcal{A}$ *be a finite cover of* X, *then for every* m, $h(f,\mathcal{A}) \leq 1/m \log N(\mathcal{A}^m)$; *in particular*, $h(f,\mathcal{A}) \leq \log N(\mathcal{A})$.

Proof: Consider

$$\begin{aligned}\mathcal{A}^{nm} &= \mathcal{A} \vee f^{-1}\mathcal{A} \vee,\dots,\vee f^{-nm+1}\mathcal{A} \\ &= \mathcal{A}^m \vee f^{-m}\mathcal{A}^m \vee,\dots,\vee f^{(-n+1)m}\mathcal{A}^m.\end{aligned}$$

Using properties (I-2) and (I-4), we obtain $N(\mathcal{A}^{nm}) \leq [N(\mathcal{A}^m)]^n$, and since $h(f,\mathcal{A})$ exists,

$$h(f,\mathcal{A}) = \lim_{n\to\infty} 1/nm \log N(\mathcal{A}^{nm}) \leq \lim_{n\to\infty} 1/nm \log[N(\mathcal{A}^m)]^n = 1/m \log N(\mathcal{A}^m).$$

The following additional notation will be used below. For a finite open cover $\mathcal{A}$ of X, let $(\mathcal{A})_n = \mathcal{A} \vee X_n = \{ A \cap X_n : A \in \mathcal{A},\ 0 \leq n \leq \infty \}$, $(X_0 = X)$, so that $(\mathcal{A})_n$ is an open cover of X_n, open in the subspace topology; and let $f_n = f|_{X_n}$, $n = 1,2,\dots$.

The following results are easily verified,

(III-1) $\qquad (\mathcal{A} \vee \mathcal{B})_n = (\mathcal{A})_n \vee (\mathcal{B})_n$,

(III-2) $\qquad f_n^{-1}(\mathcal{A})_n = (f^{-1}\mathcal{A})_n$.

Lemma (III-2): *Let* X *be a compact Hausdorff space, and* f *be a map from* X *into itself, then* $h(f_n,(\mathcal{A})_n) = h(f,\mathcal{A})$ $(n = 1,2,\dots)$.

Proof: Given an open cover $\mathcal{A}$ of X, form the open cover $(\mathcal{A})_n$ of X_n and let $N((\mathcal{A})_n)$ be the minimal cardinality of a subcover of $(\mathcal{A})_n$. Thus by properties (III-1) and (III-2),

$$\begin{aligned}N[(\mathcal{A})_n \vee f_n^{-1}(\mathcal{A})_n \vee,\dots,\vee f_n^{-k+1}(\mathcal{A})_n] &= N[(\mathcal{A} \vee f^{-1}\mathcal{A} \vee,\dots,\vee f^{-k+1}\mathcal{A})_n] \\ &= N[(\mathcal{A}^k)_n] .\end{aligned}$$

Clearly $N[(\mathcal{A})_n] \leq N(\mathcal{A})$, therefore, $h(f_n,(\mathcal{A})_n) \leq h(f,\mathcal{A})$.

On the other hand, since $f^{-n}X_n = X$, it follows that $f^{-n}(\mathcal{A}^k)_n$ is a cover of X, and $f^{-n}(\mathcal{A}^{k})_n \vee \mathcal{A}^{n} \subseteq \mathcal{A}^{n+k}$. Therefore,

$$N(\mathcal{A}^{n+k}) \le N(f^{-n}(\mathcal{A}^{k})_n)N(\mathcal{A}^{n}) \le N((\mathcal{A}^{k})_n)N(\mathcal{A}^{n}) ,$$

where the inequalities follow from properties (I-4) and (I-2), respectively.

Finally, $h(f,\mathcal{A}) = \lim_{k\to\infty} 1/(n+k) \log N(\mathcal{A}^{n+k})$

$$\le \lim_{k\to\infty} (k/n+k)1/k \log[N((\mathcal{A}^{k})_n)N(\mathcal{A}^{n})] = h(f_n,(\mathcal{A})_n) .$$

Lemma (III-3): *Let* X *be a compact Hausdorff space, and* f *be a continuous map from* X *into itself.* Let $g = f|_{X_\infty}$. *Then*

$$h(f,\mathcal{A}) \le h(g,(\mathcal{A})_\infty) , \text{ where } (\mathcal{A})_\infty = \mathcal{A} \vee X_\infty , \text{ and } X_\infty = X \cap \bigcap_{n=1}^{\infty} f^{n}X .$$

Proof: Given $\epsilon > 0$, fix n so that

$$h(g,(\mathcal{A})_\infty) + \epsilon > 1/n \log N[((\mathcal{A})_\infty)^{n}] = 1/n \log N((\mathcal{A}^{n})_\infty)$$

such an n exists by proposition (I-1). Let $\mathcal{B}_k = \{B_i \in \mathcal{A}^k : i = 1,2,\dots,M_k\}$ be a cover of X_∞ of minimal cardinality, consisting of M_k elements. Clearly, $M_k = N((\mathcal{A}^{k})_\infty)$, and one easily sees that $h(g,(\mathcal{A})_\infty) = \lim_{k\to\infty} 1/k \log M_k$.

Choose m sufficiently large so that $\bigcup_{i=1}^{M_n} B_i \supseteq X_m \supseteq X_\infty$, and as $\mathcal{B}_n \subseteq \mathcal{A}^n$, we have (III-3)

$$h(f_m^n,(\mathcal{A}^{n})_m) \le h(f_m^n,\mathcal{B}_n) \le \log M_n ,$$

where the second inequality follows from lemma (III-1). A re-examination of the proof of lemma (II-5) shows that it is true for arbitrary finite covers; therefore, using lemma (III-3) and (II-5), proposition (I-7) with properties (III-1) and III-2), we have $h(f,\mathcal{A}) = h(f_m,(\mathcal{A})_m) = h(f_m,(\mathcal{A}^n)_m) = 1/n\, h(f_m^n,(\mathcal{A}^{n})_m)$, which, if substituted into (III-3), yields

$$h(f,\mathcal{A}) \le 1/n \log M_n < h(g,(\mathcal{A})_\infty) + \epsilon .$$

Therefore, $h(f,\mathcal{A}) \leq h(g,(\mathcal{A})_\infty)$.

Theorem (III-1): Let X be a compact Hausdorff space, and f be a map from X into itself. If $g = f|X_\infty$, where $X_\infty = X \cap \bigcap_{n=1}^{\infty} f^n X$, then $h(f) = h(g)$.

Proof: Let $\mathcal{A}_\infty$ be an arbitrary open cover of X_∞ , open in the subspace topology of X_∞ . There exists an open cover $\mathcal{A}$ of X , such that $(\mathcal{A})_\infty = \mathcal{A}_\infty$, namely, $\mathcal{A} = \{A \cup (X - X_\infty) : A \in \mathcal{A}_\infty\}$. Thus by lemma (III-3), we have $h(f) \leq h(g)$.

On the other hand, by proposition (I-6),

$$h(f) = \max.\{h(f),h(g)\} ,$$

which implies $h(g) \leq h(f)$.

This last theorem enables us to restrict our attention to surjective maps; for if f maps X into X , we shall restrict our attention to g and X_∞ . Any measure considered in this chapter will be assumed to satisfy $\mu(X_\infty) = 1$. The fact that f is surjective enables us to assume that $\eta^n(y) \geq 1$ for all y in X , and $\{\eta^n(y)\}$ is an increasing sequence, which implies

$$V(f^n,\mu) \geq 1 \quad \text{and} \quad \mathcal{V}(f,\mu) \geq 0 \quad \text{for all } \mu .$$

It is also apparent that if X_∞ is a point, then $h(f) = 0$. This would be the case, for example, if f is a map having the absolute value of its derivative almost everywhere less than one.

We proceed to the proof of the main result of this section, that $h(f) = \mathcal{V}(f)$. The idea will be to estimate $N(f^n)$ in terms of the multiplicity function $\eta^n(y)$ evaluated at fixed points. The following lemma gives us this estimate.

Lemma (III-4): *Let* f *map* I *onto* I , $(I = [\alpha_0, \alpha_{N(f)}])$, *and suppose* $N(f)$ *is finite*. *Let* $\{\alpha_i : i = 1,2,\ldots,N(f)-1\}$ *determine the branches of* f , *so that* f *is injective on* (α_i, α_{i+1}) . *Then*

$$N(f^n) \le n \, \Sigma_{i=1}^{N(f)-1} \, \eta^n(\alpha_i) \, .$$

Proof: f^n has $N(f^n)$ branches determined by $N(f^n)-1$ points $\{x_j : j = 1,2,\ldots,N(f^n)-1\}$ such that $f^k(x_j) = \alpha_i$ for some k , $(k = 0,1,\ldots,n-1)$. Therefore,

$$N(f^n)-1 \le \text{Cardinality } \{\bigcup_{\substack{0 \le k \le n-1 \\ 1 \le i \le N(f)-1}} \{f^{-k}\alpha_i\}\}$$

$$\le \Sigma_{k=0}^{n-1} \, [\eta^k(\alpha_1) + \eta^k(\alpha_2) + \ldots + \eta^k(\alpha_{N(f)-1})] \, .$$

Strict inequality in the first instance occurs if some x_j is in $\{f^{-k}\{\alpha_0, \alpha_{N(f)}\}\}$, $k = 0,1,\ldots,n-1$, otherwise we have equality; and $\eta^0(\alpha_i) = 1$ $(i = 1,2,\ldots,N(f)-1)$. But since f is injective and $\eta^k(y) \ge 1$ for all k , $\{\eta^k(y)\}$ is a non-decreasing sequence; and in particular, $\eta^n(\alpha_i) \ge \eta^k(\alpha_i)$ $(k = 0,1,\ldots,n)$. Therefore,

$$N(f^n) \le n[\eta^n(\alpha_1) + \eta^n(\alpha_2) + \ldots + \eta^n(\alpha_{N(f)-1})] \, .$$

Theorem (III-2): *Let* f *map* I *onto* I , $I = [\alpha_0, \alpha_{N(f)}]$, *and suppose* $N(f)$ *is finite*, *then* $h(f) = \mathcal{V}(f)$.

Proof: Let $F = \bigcup_{i=1}^{N(f)-1} \{\alpha_i\}$ and define a measure on I by $\mu_0(\alpha_i) = 1/[N(f)-1]$, $i = 1,2,\ldots,N(f)-1$, , and $\mu_0(I-F) = 0$.

Using lemma (III-4),

$$V(f^n, \mu_0) = \Sigma_{i=1}^{N(f)-1} \, \eta^n(\alpha_i)\mu_0(\alpha_i) \ge N(f^n)\left[\frac{1}{n[N(f)-1]}\right] ;$$

passing to the limit, and using theorems (II-2) and (II-1)

$$\mathcal{V}(f) \geq \mathcal{V}(f,\mu_0) \geq \bar{h}(f) = h(f) .$$

The other inequality follows easily since $\eta^n(y) \leq N(f^n)$ for all y in I, and for all n. Thus

$$V(f^n,\mu) = \int_I \eta^n(y)\mu(dy) \leq N(f^n) ,$$

which implies

$$\mathcal{V}(f,\mu) \leq \bar{h}(f) = h(f) \quad \text{for all } \mu .$$

Therefore, $\mathcal{V}(f) = h(f)$.

A simple consequence of lemma (III-4) is the following proposition.

Proposition (III-3): If $\overline{\eta^n} = \max_{y\in I} \eta^n(y)$, then $h(f) = \lim_{n\to\infty} 1/n \log \overline{\eta^n}$.

Proof: From lemma (III-4)

$$\overline{\eta^n} \leq N(f^n) \leq n \sum_{i=1}^{N(f)-1} \eta^n(\alpha_i) \leq n\, N(f)\, \overline{\eta^n} .$$

Hence, by theorem (II-2) ,

$$h(f) = \lim_{n\to\infty} 1/n \log \overline{\eta^n} .$$

CHAPTER IV. EXAMPLES AND APPLICATIONS

The following examples will illustrate some applications of the results obtained in the previous chapters.

Example 1: We shall adopt the results of chapter III to calculate the entropy for a class of functions for which the multiplicity function is constant except at finitely many points.

Proposition (IV-1): *Let* f *map* I *onto itself, where* f *has a finite number of branches. Suppose* $\eta(y) = \upsilon$, (y *in* I), *except at finitely many points* $y = \alpha_i$, ($i=1,2,\ldots,p$). *Then* $h(f) = \log \upsilon$.

Proof : The points at which $\eta(y) < \upsilon$ are the relative maxima and minima of f . Let $\overline{\eta^n} = \sup_{y\in I} \eta^n(y)$ and define $\Gamma_n' = \{ y: \eta^n(y) < \overline{\eta^n} \}$, and $\Gamma_n = \{ y: f^k(\alpha_i) = y,\ 1 \leqq i \leqq p,\ 1 \leqq k \leqq n \}$, then clearly $\Gamma_n' \subseteq \Gamma_n$. As $\bigcup_{i=1}^{\infty} \Gamma_i$ is a countable set, $J = I - \bigcup_{i=1}^{\infty} \Gamma_i \neq \emptyset$, and $\eta^n(y) = \overline{\eta^n}$ for all n , provided y is in J . We prove inductively that $\overline{\eta^n} = \upsilon^n$, in fact, for only $y \in J$, $\eta^n(y) = \upsilon^n$.

For y in J , $\eta(y) = \upsilon$. Assume that $\eta^n(y) = \upsilon^n$ for y in J, then $\eta^{n+1}(y) = \Sigma_{i=1}^{\eta(y)} \eta^n(x_i)$ where $f(x_i) = y$, $i = 1,2,\ldots,\eta(y)$. We claim that x_i is in J, since otherwise, $x_i = f^k(\alpha_j)$, for some $j = 1,2,\ldots,p$ and some $k = 1,2,\ldots,n$, which implies $y = f(x_i) = f^{k+1}(\alpha_j)$ and thereby contradicts the assumption that y is in J. Therefore, $\eta^n(x_i) = \upsilon^n$ by the induction hypothesis, and similarly $\eta(y) = \upsilon$.

Hence, $\eta^{n+1}(y) = \Sigma_{i=1}^{\upsilon} \upsilon^n = \upsilon^{n+1} = \overline{\eta^{n+1}}$.

By proposition (III-3), $h(f) = \lim_{n\to\infty} 1/n \log \overline{\eta^n} = \lim_{n\to\infty} 1/n \log \upsilon^n = \log \upsilon$.

Definition (IV-1) : *Let* $I = [0,1]$, *and* f *be a map from* I *onto* I, *for which* $N(f)$ *is finite*. *If the image of* $f|_{[\alpha_i, \alpha_{i+1}]}$, $(i = 0,1,\ldots,\upsilon - 1;\ \alpha_0 = 0;\ \alpha_\upsilon = 1)$, *is homeomorphic to* I, *we call* f *a full function* (4).

Example 2 : If f is a full function with υ branches, then proposition (IV-1), or a direct application of theorem (III-2), implies $h(f) = \log \upsilon$.

Example 3 : The topological entropy of the Chebyshev Polynomial was calculated by M. H. McAndrew and R. L. Adler (8). It is easily verified that the Chebyshev Polynomial of degree υ is conjugate to a map defined on $[0,\pi]$ which is a piecewise linear full function with υ branches. Application of theorem (III-2), or the previous two examples, give the entropy as $\log \upsilon$.

Example 4 : We prove proposition (I-5) for the special case where X and Y are intervals, and f and g have a finite number of branches. The error in the original proof ((1); Theorem 3,p 312) was in the claim that open covers of $X \times Y$ of the form $\mathcal{A} \times \mathcal{B} = \{A \times B;\ A \in \mathcal{A},\ B \in \mathcal{B}\}$ have the property that $N_{X \times Y}(\mathcal{A} \times \mathcal{B}) = N_X(\mathcal{A})\, N_Y(\mathcal{B})$. For open covers we only have the inequality $N_{X \times Y}(\mathcal{A} \times \mathcal{B}) \leqq N_X(\mathcal{A})\, N_Y(\mathcal{B})$; however, for partitions $\mathcal{A}$ and $\mathcal{B}$, of X and Y respectively, we have equality since $\mathcal{A} \times \mathcal{B}$ is a partition of $X \times Y$. Therefore, if we restrict our attention to partitions rather than open covers, the original proof (1) that $h(f \times g) = h(f) + h(g)$ holds.

A reexamination of the proofs in chapters II and III shows that the essential criteria are the restrictions of X to a one dimensional space, and the maps to those with a finite number of branches. It is easily verified that all the results obtained are true for the circle S^1.

Example 5 : Let $X = S^1$ and f be a homeomorphism, then $h(f) = 0$. This result is an immediate consequence of proposition (III-3), since $\pi^n(y) = 1$ for all y in S^1 and therefore, $\overline{\pi^n} = 1$.

CONCLUDING REMARKS

The results of Chapter II show that the computation of the topological entropy can be simplified for continuous maps having a finite number of branches on an interval. This suggests the investigation of the following problems.

Let $X = [a_1,b_1] \times [a_2,b_2] \times \ldots \times [a_n,b_n]$ and let f, mapping X into X, be injective on elements of a partition consisting of a finite number of convex sets. Is it still possible to compute the topological entropy by restricting one's attention to minimal f-injective partitions? Proposition (I-5) gives a favorable answer for the case of product maps.

Consider a sequence f_n converging to f ; for what notion of convergence does $h(f_n)$ converge to $h(f)$. The sequence

$$f_n(x) = \begin{cases} -2x + 2^{-n} & 0 \le x < 2^{-n-1} \\ 2x - 2^{-n} & 2^{-n-1} \le x < 2^{-n} \\ x & 2^{-n} \le x \le 1 \, . \end{cases}$$

together with propositions (I-6) and (IV-1) demonstrate that uniform convergence is not sufficient. In fact, the results in Chapter II suggest that the convergence needed should be such, that $N(f_n)$ approaches $N(f)$.

Additional results may be obtained if Theorem (III-2) remains true for a modified definition of $\mathcal{V}(f)$ (Def. (III-2)). Thus, if we define $\mathcal{V}(f) = \sup_\mu \mathcal{V}(f,\mu)$, where the supremum is taken over all invariant measures rather than all measures, it is reasonable to conjecture that the topological entropy for a large class of functions may be computed from $\mathcal{V}(f,\mu)$ (Def. (III-2)).

Lastly, proposition (III-3) may be true for arbitrary compact spaces, and functions for which $\bar{\eta}$ is finite.

BIBLIOGRAPHY

(1) R. L. Adler, A. G. Konheim, and M. H. McAndrew,"Topological Entropy",*Trans. Amer. Math. Soc.* *114*(1965), pp.309-319.

(2) R. L. Adler and B. Weiss,"Entropy, a complete metric invariant for automorphisms of the torus", *Proc. Natl. Acad. Sci.* *57*(1967), pp.1573-1576.

(3) R. Bowen, *Topological entropy and axiom A*, Global Analysis Symposium, Vol. XIV.

(4) H. Cohen,"On fixed points of commuting functions", *Proc. Amer. Math. Soc.* *15*(1964), pp.293-296.

(5) L. W. Goodwyn,"The product theorem for topological entropy", to appear.

(6) L. W. Goodwyn,"Topological entropy bounds measure theoretic entropy", *Proc. Amer. Math. Soc.* *23*(1969), pp.679-688.

(7) H. B. Keynes and J. B. Robertson,"Generators for topological entropy and expansiveness", to appear.

(8) M. H. McAndrew and R. L. Adler,"Topological entropy of Chebyshev polynomials", *Trans. Amer. Math. Soc.* *121*(1966), pp.236-241.

(9) I. Natanson, *Theory of Functions of a Real Variable, Vol I*, p.225, Fredrich Unger Publishing Company.

(10) W. Parry,"Symbolic Dynamics and Transformations of the Unit Interval", *Trans. Amer. Math. Soc.* *122*(1966), pp.368-378.