AN INTRODUCTION TO TOPOLOGICAL ENTROPY

KAREN BUTT

Abstract. We develop the basic notions necessary to define topological entropy for topological dynamical systems. We prove a number of results which are then used to compute the entropy of some standard dynamical systems. Finally, we show topological entropy is an invariant of topological conjugacy.

Contents

1. Introduction

A discrete-time dynamical system is a nonempty set X and a map $f : X \to X$, together with some additional structure, such as measure-theoretic or topological structure. The study of dynamical systems as a whole is primarily concerned with the asymptotic behavior of such systems, that is how the system evolves after repeated applications of f. In this paper, we are concerned with the study of topological dynamics, where we restrict X to be a metric space and f to be a continuous transformation.

Topological entropy measures the evolution of distinguishable orbits over time, thereby providing an idea of how complex the orbit structure of a system is. Entropy distinguishes a dynamical system where points that are close together remain close from a dynamical system in which groups of points move farther.

In the first section, we develop the notion of entropy by considering coverings of X. Next, we explain the notions of spanning and separated sets and prove that entropy can also be defined in these two ways. With a more flexible definition of entropy, we are able to compute it for specific examples in the subsequent section. As guiding examples, we consider the doubling map, isometric transformations and hyperbolic toral automorphisms. In the final section, we explain topological conjugacy, which is the notion of equivalence in topological dynamical systems. We then prove that topological entropy is an invariant.

Date: 2014.

2. Definition with Covers

We will consider the exponential growth of distinguishable orbits with respect to the number of applications of our map f. First we will consider finite orbit segments of length n distinguishable at some finite resolution ϵ. Given a compact metric space (X, d) and map $f : X \to X$, we define the function $d_n : X \times X \to \mathbb{R}$ by

$$d_n(x, y) = \max_{0 \le k < n} d(f^k(x), f^k(y)).$$

For each n, d_n is a metric on X, since d is already defined to be a metric on X. With this new notion of distance, points are ϵ-close if they remain ϵ-close for n iterates of f. We can think of ϵ as a resolution, the smallest distance at which we can distinguish two points from one another.

Next we define a quantity that counts these distinguishable orbits. We want to collect indistinguishable orbits – that is, points that are ϵ-close with respect to the d_n-metric – and count how many of these collections we have. However, we want to avoid the possibility of the collections themselves being ϵ-close to one another as this would over estimate the distinguishable orbits. So we consider a covering of X that contains the smallest possible amount of collections of ϵ-close points.

Definition 2.1. Fix $\epsilon > 0$. Let $\text{cov}(n, \epsilon, f)$ denote the minimal cardinality of a covering of X by sets of d_n-diameter less than ϵ, where the diameter of a set is given by the supremum of distances between pairs of points in the set.

Remark 2.2. Since compactness means every open cover of X contains a finite subcover, it follows that $\text{cov}(n, \epsilon, f)$ is a finite quantity.

Definition 2.3. Let

$$h_\epsilon(f) = \limsup_{n\to\infty} \frac{1}{n} \log(\text{cov}(n, \epsilon, f)).$$

This limit gives the exponential growth of $\text{cov}(n, \epsilon, f)$ with a fixed resolution as the length of orbits we consider tends to infinity. We are considering the limit superior because we do not know if the limit actually exists. To prove it does, we require the following lemma from calculus.

Lemma 2.4. *Let $\{a_n\}_{n\ge 1}$ be a subadditive sequence, that is $a_{m+n} \le a_m + a_n$ $\forall m, n$. Then $\lim_{n\to\infty} a_n/n$ exists and is equal to $\inf_n a_n/n$.*

The proof uses basic computation to show the limit inferior is equal to the limit superior. For the complete proof, see [4, Theorem 4.9].

Proposition 2.5. *The limit $\lim_{n\to\infty} \frac{1}{n} \log(\text{cov}(n, \epsilon, f)) = h_\epsilon(f)$ exists and is finite.*

Proof. Given the previous lemma, we need to show that the sequence $\log(\text{cov}(n, \epsilon, f))$ is subadditive. Then, since $\log(\text{cov}(n, \epsilon, f))$ is bounded below by 0, $\inf \log \text{cov}(n, \epsilon, f)/n$ must be finite.

Fix $\epsilon > 0$. Suppose A is a cover of X with sets of d_n-diameter less than ϵ and B is a cover of X with sets of d_m-diameter less than ϵ. Then $\text{cov}(n, \epsilon, f) \le |A|$ and $\text{cov}(m, \epsilon, f) \le |B|$, because cov refers to the minimum cardinality.

Let $U \in A$ and $V \in B$. If $U \cap f^{-n}(V) \ne \emptyset$, consider $x, y \in U \cap f^{-n}(V)$. We have $x, y \in U$, so

$$\max_{0\le i<n} d(f^i(x), f^i(y)) < \epsilon.$$

Also $x, y \in f^{-n}(V)$ which means $f^n(x), f^n(y) \in V$. Since V has d_m-diameter less than ϵ, we have

$$\max_{0 \leq i < m} d(f^i(f^n(x)), f^i(f^n(y))) < \epsilon.$$

The above two inequalities give

$$\max_{0 \leq i < m+n} d(f^i(x), f^i(y)) < \epsilon.$$

So the set $U \cap f^{-n}(V)$ has d_{n+m}-diameter less than ϵ.

Let $C = \{U \cap f^{-n}(V) : U \in A, V \in B\}$. Then $|C| \leq |A||B|$; equality holds when each intersection is non-empty. Additionally, we have already shown that C has d_{n+m}-diameter less than ϵ, so $\text{cov}(m+n, \epsilon, f) \leq |C|$. Recalling the initial bounding of $\text{cov}(n, \epsilon, f)$ and $\text{cov}(m, \epsilon, f)$ we have

$$\text{cov}(m+n, \epsilon, f) \leq |C| \leq |A||B| \leq \text{cov}(n, \epsilon, f)\text{cov}(m, \epsilon, f).$$

Recalling basic properties of logarithms, we have that the sequence $\log(\text{cov}(n, \epsilon, f)$ is subadditive, which completes the proof. ☐

Next, we want to decrease the resolution. Clearly if we decrease ϵ, the amount of sets of diameter less than ϵ needed to cover X can only increase. Therefore, $\text{cov}(n, \epsilon, f)$ is monotonic, and so is $h_\epsilon(f)$, which means it has a limit as ϵ goes to 0.

Definition 2.6. The topological entropy of f is given by $h(f) = \lim_{\epsilon \to 0^+} h_\epsilon(f)$.

This definition can be used to prove a number of properties about topological entropy, but before we proceed to examples, we develop more useful definitions in the next section.

3. Spanning and Separated Sets

We will again consider the metric d_n to count distinguishable orbit segments at a fixed resolution.

Definition 3.1. Fix $\epsilon > 0$. Let $n \in \mathbb{N}$. A set $A \subset X$ is an (n, ϵ)-spanning set if $\forall x \in X$, $\exists y \in A$ such that $d_n(x, y) < \epsilon$.

Remark 3.2. Recall that the definition of d_n depends on f.

This is a useful characterization because we can easily construct such sets for known dynamical systems. The first dynamical system we will consider is the doubling map $f : S^1 \to S^1$ where $f(x) = 2x \mod 1$. Here we are taking S^1 to be $[0, 1]$ with endpoints identified. The natural choice of metric for the unit circle is

$$d(x, y) = \min(|x - y|, 1 - |x - y|)$$

with respect to which f is continuous. We will always consider this metric when discussing the doubling map.

Before we can construct an (n, ϵ)-spanning set for the doubling map, we require the following lemma.

Lemma 3.3. *Let f be the doubling map on S^1. Then we have*

$$d(x, y) \leq \frac{1}{4} \Rightarrow d(f(x), f(y)) = 2d(x, y).$$

Proof. Clearly, $d(x, y) = |x - y|$ when $|x - y| \leq 1/2$. Let x, y be such that $d(x, y) \leq 1/4$. This means $|x - y| \leq 1/4$. Using the definition of f, we have

$$\begin{aligned} d(f(x), f(y)) &= d(2x \bmod 1, 2y \bmod 1) \\ &= \min(|2x - 2y \bmod 1|, 1 - |2x - 2y \bmod 1|). \end{aligned}$$

Note that $|2x - 2y| \leq 1/2$, so $2x - 2y \bmod 1 = 2x - 2y$.
Therefore,

$$\begin{aligned} d(f(x), f(y)) &= 2|x - y| \\ &= 2d(x, y). \end{aligned}$$

□

Notation 3.4. Let S_k denote the set of dyadic rationals with denominator 2^k, that is

$$S_k = \big\{\frac{i}{2^k}, 0 \leq i < 2^k - 1\big\}.$$

Proposition 3.5. *The set of fractions S_{n+k} is an (n, ϵ)-spanning set for the doubling map.*

Proof. Fix $\epsilon > 0$. Choose $k \geq 2$ such that $1/2^{k+1} \leq \epsilon < 1/2^k$. Note that for any $x \in S^1$, there exists $i \in \{0, ..., 2^{n+k} - 1\}$ such that

$$x \in \Big[\frac{i}{2^{n+k}}, \frac{i+1}{2^{n+k}}\Big).$$

Then, choose $y \in S_{n+k}$ to be either of the endpoints of this dyadic interval. Then, $d(x, y) \leq 1/2^{n+k} < 1/4$.

Using Lemma 3.3, this implies

$$\begin{aligned} d(f(x), f(y)) &= 2d(x, y) \\ &\leq 2/2^{n+k} \\ &< 1/4. \end{aligned}$$

So we can apply the lemma again to get

$$\begin{aligned} d(f^2(x), f^2(y)) &= 2d(f(x), f(y)) \\ &\leq 2^2/2^{n+k} \end{aligned}$$

Applying the lemma j consecutive times, for any j satisfying $0 \leq j < n$, gives

$$d(f^j(x), f^j(y)) = 2^j d(x, y) \leq \frac{2^j}{2^{n-k}} \leq \frac{2^n - 1}{2^{n-k}} < \frac{1}{2^{k+1}} \leq \epsilon.$$

So for any $x \in S^1$ we have,

$$\max_{0 \leq j < n} (d(f^j(x), f^j(y)) = d_n(x, y) < \epsilon,$$

for some $y \in S_{n+k}$, which concludes the proof. □

Before we can proceed to calculating entropy, we need to introduce a final way of counting distinguishable orbits.

Definition 3.6. Fix $\epsilon > 0$. Let $n \in \mathbb{N}$. A set $A \subset X$ is an (n, ϵ)-separated set if for all $x, y \in A$ with $x \neq y$, we have $d_n(x, y) \geq \epsilon$.

Proposition 3.7. *Let f be the doubling map on S^1. Then S_{n-1+k} is an (n,ϵ)-separated set.*

Proof. Fix $\epsilon > 0$. Choose $k \geq 2$ such that $1/2^{k+1} \leq \epsilon < 1/2^k$. Let x, y be two distinct points in S_{n-1+k}. We want to show that $d_n(x,y) \geq \epsilon$. This means showing that there exists j satisfying $0 \leq j \leq n-1$ such that $d(f^j(x), f^j(y)) \geq \epsilon$. Suppose there exists j such that $d(f^j(x), f^j(y)) \geq 1/4$, then we are done because $\epsilon < 1/4$ by assumption.

If this is not the case, then for every j we have that $d(f^j(x), f^j(y)) < 1/4$, so we can apply Lemma 3.3 $n-1$ times to get

$$d(f^{n-1}(x), f^{n-1}(y)) = 2^{n-1}d(x,y).$$

Next note that for all distinct $x, y \in S_{n-1+k}$ we have $d(x,y) \geq 1/2^{n-1+k}$, giving

$$2^{n-1}d(x,y) \geq \frac{2^{n-1}}{2^{n-1+k}} = \frac{1}{2^k} \geq \epsilon,$$

which completes the proof. □

In Section 2, we had to consider the minimal cardinality of a covering of X to count distinguishable orbits. The same applies to a spanning set.

Definition 3.8. Let $\mathrm{span}(n,\epsilon,f)$ be the minimum cardinality of an (n,ϵ)-spanning set.

Separated sets are essentially the opposite of spanning sets. It is possible to have a collection of points that is too sparse, missing many distinguishable orbits. So it makes sense to consider the maximal cardinality of such a set.

Definition 3.9. Let $\mathrm{sep}(n,\epsilon,f)$ be the maximum cardinality of an (n,ϵ)-separated set.

Next, we prove a lemma that relates our three orbit-counting quantities. From this result, we immediately get two additional equivalent definitions for entropy.

Lemma 3.10. $\mathrm{cov}(n,2\epsilon,f) \leq \mathrm{span}(n,\epsilon,f) \leq \mathrm{sep}(n,\epsilon,f) \leq \mathrm{cov}(n,\epsilon,f)$.

Proof. Let A be an (n,ϵ)-spanning set of minimum cardinality. Then,

$$\bigcup_{y \in A} B(y,\epsilon) \supset X,$$

where $B(y,\epsilon)$ denotes balls centered at y with d_n-radius less than ϵ. Then these balls have d_n-diameter less than 2ϵ, and since they cover X,

$$\mathrm{cov}(n,2\epsilon,f) \leq |A| = \mathrm{span}(n,\epsilon,f),$$

proving the first inequality.

Next, let B be an (n,ϵ)-separated set of maximum cardinality, meaning we cannot add any more points to B such that it still has the separated property. Then for every $x \in X \setminus B$ and every $y \in B$, the inequality $d_n(x,y) \geq \epsilon$ cannot hold. Therefore, for all $x \in X$ we can choose $y \in B$, such that $d_n(x,y) < \epsilon$. Hence B is also an (n,ϵ)-spanning set. Then

$$\mathrm{span}(n,\epsilon) \leq |B| = \mathrm{sep}(n,\epsilon,f),$$

proving the second inequality.

Finally, consider B again and let C be a covering of X with sets of d_n-diameter less than ϵ. Also, let C be such that $|C| = \text{cov}(n, \epsilon, f)$ holds. Suppose that $|B| > |C|$ holds. Then there exists an element of C which contains more than one point of B. In other words, there exists a set of d_n-diameter less than ϵ which contains more than one point of B. Thus, there are two points in B which are less than ϵ apart in the d_n-metric, contradicting the fact that B is (n, ϵ)-separated. Therefore we must have $|B| \leq |C|$, which proves the final inequality. □

Corollary 3.11. $h(f) = \lim_{\epsilon \to 0^+} \lim_{n \to \infty} \frac{1}{n} \log(\text{sep}(n, \epsilon, f) = \lim_{\epsilon \to 0^+} \lim_{n \to \infty} \frac{1}{n} \log \text{span}(n, \epsilon, f)$.

Proof. We consider the inequalities in lemma 3.10 and apply the appropriate logarithms and limit operations to everything so the leftmost and rightmost term have the form $h(f)$. Then, by squeeze theorem for limits, both the span and sep terms' limits exist and are equal to $h(f)$. □

4. Topological Entropy for Some Examples

Proposition 4.1. *Let f be the doubling map on S^1. Then $h(f) = \log 2$.*

Proof. We know from Proposition 3.5 that S_{n+k} is an (n, ϵ)-spanning set. It has cardinality 2^{n+k}, since the numerators of the fractions range from 0 to $2^{n+k} - 1$. Therefore, $\text{span}(n, \epsilon, f) \leq 2^{n+k}$.

$$\begin{aligned} h_\epsilon(f) &= \lim_{n \to \infty} \frac{1}{n} \log(\text{span}(n, \epsilon, f)) \\ &\leq \lim_{n \to \infty} \frac{(n+k)\log 2}{n} \\ &= \log 2. \end{aligned}$$

Similarly, since by Proposition 3.7 S_{n-1+k} is an (n, ϵ)-separated set, we have

$$\begin{aligned} h_\epsilon(f) &\geq \lim_{n \to \infty} \frac{(n-1+k)\log 2}{n} \\ &= \log 2. \end{aligned}$$

By the squeeze theorem for limits, $\lim_{\epsilon \to 0^+} h_\epsilon(f)$ exists and is equal to $\log 2$. □

This result is true in greater generality for maps $f(x) = \alpha x \mod 1$ for $\alpha \in \mathbb{N}$, where we find $h(f) = \log \alpha$. The proof is very similar: we just need to consider fractions with denominators α^k instead of 2^k.

Next, we compute the entropy of an isometry, which should clearly be 0 since distance between points is preserved so the number of distinguishable orbits is constant under applications of f.

Theorem 4.2. *Let (X, d) be a compact metric space and let $f : X \to X$ be a continuous isometry, that is $d(x, y) = d(f(x), f(y))$. Then $h(f) = 0$.*

Proof. Since f is an isometry, we have $d(x, y) = d(f^n(x), f^n(y))$ for every n. Thus, $\text{cov}(n, \epsilon, f)$ does not change with n. So

$$h_\epsilon(f) = \lim_{n \to \infty} \frac{1}{n} \log(\text{cov}(n, \epsilon, f)) = 0,$$

which concludes the proof. □

Corollary 4.3. *The entropy of a rotation is 0.*

Next we consider a class of transformations called hyperbolic toral automorphisms. Consider a 2×2 matrix A with integer entries. If the matrix has determinant 1, then its inverse is also an integer matrix, and this notion of invertibility is why we use the term automorphism. We then apply this matrix to $\mathbb{R}^2$. Since the matrix preserves the integer lattice, the map A descends to a map on the quotient $\mathbb{T}^2$ which we denote by f_A.

Also note that since the determinant, which is also the product of the eigenvalues, is equal to 1 then the eigenvalues are λ and λ^{-1}.

Proposition 4.4. *Let $A \in M_2(\mathbb{Z})$ with determinant 1 and eigenvalues λ and λ^{-1}, with $|\lambda| > 1$. Let $f_A : \mathbb{T}^2 \to \mathbb{T}^2$ be the associated toral automorphism. Then $h(f_A) = \log |\lambda|$.*

Proof. We approach this by considering both spanning and separated sets to bound the entropy from both sides. Denote the eigenvectors of A, normalized to unit length for reasons that will be apparent later on, by v_1 and v_2, such that $Av_1 = \lambda v_1$ and $Av_2 = \lambda^{-1} v_2$. Note that since $|\lambda| > 1$, we have $|\lambda^{-1}| < 1$, so the eigenvalues are distinct. A known result from linear algebra tells us that their corresponding eigenvectors are linearly independent. So clearly, v_1 and v_2 form a basis for $\mathbb{T}^2$. Hence for any $x, y \in \mathbb{T}^2$ we can write

$$x - y = \alpha v_1 + \beta v_2 \text{ for some } \alpha, \beta \in \mathbb{R}.$$

Now fix $\epsilon > 0$ and choose N such that $1/N \leq \epsilon/2$. Since $\mathbb{T}^2$ is the unit square with opposite sides identified, we construct our spanning set on the unit square because it is easier to visualize. Consider lines on the unit square in the direction of v_1. Additionally, we require that there are enough lines such that the distance between any two lines in the direction of v_2 is less than $1/N$.

Next, consider the set of points on these lines where the points are chosen such that their spacing along the lines is less than $\epsilon/2|\lambda|^{n-1}$. Denote this set by D. Let $x \in \mathbb{T}^2$ and let y be the point in D which is closest to x. We wish to bound the coefficients of the linear combination

$$x - y = \alpha v_1 + \beta v_2.$$

Considering the predetermined spacing between the points along the direction of v_1, we have $|\alpha| \leq \epsilon/2|\lambda|^{n-1}$. Due to the aforementioned restriction on the spacing between the lines, we have $|\beta| < 1/N \leq \epsilon/2$.

Now we are ready to show D is an (n, ϵ)-spanning set, which means showing that the distance between x and y is still small even after we apply the map f_A. So consider

$$A^k(x - y) = \alpha A^k(v_1) - \beta A^k(v_2) = \alpha \lambda^k(v_1) - \beta \lambda^{-k}(v_2).$$

By the triangle inequality,

$$d(f_A^k(x), f_A^k(y)) \leq d(f_A^k(x), f_A^k(x + \alpha v_1)) + d(f_A^k(x + \alpha v_1), f_A^k(y)).$$

Additionally,

$$d(f_A^k(x), f_A^k(x + \alpha v_1)) = d(A^k(x), A^k(x) + \alpha \lambda^k v_1) \leq \alpha |\lambda|^k$$

because the norm of v_1 is 1. By a similar argument for the second term $d(f_A^k(x+\alpha v_1), f_A^k(y))$, we conclude

$$d(f_A^k(x), f_A^k(y)) \le |\alpha||\lambda|^k + |\beta||\lambda|^{-k} \le |\lambda|^k \frac{\epsilon}{2|\lambda|^{n-1}} + |\lambda|^{-k}\frac{\epsilon}{2}.$$

Recalling that $|\lambda| > 1$ and $k \le n-1$ we have

$$\frac{\epsilon}{2|\lambda|^{n-1}} + |\lambda|^{-k}\frac{\epsilon}{2} < \epsilon,$$

which shows D is (n,ϵ)-spanning. Thus, $\mathrm{span}(n,\epsilon,f) \le |D|$.

The cardinality of D is the number of points on each line multiplied by the number of lines. Thus, it is bounded above by

$$\frac{L}{2\epsilon/|\lambda|^{n-1}} 2NC,$$

where L is the length of the longest line and $C \ge 1$ is a constant that depends on the angle between v_1 and v_2. C determines how many lines we need to satisfy the $1/N$ spacing between lines in the direction of v_2, with $C = 1$ when v_1 and v_2 are perpendicular. Computing the appropriate limits gives $h(f_A) \le \log|\lambda|$.

Next, we construct an (n,ϵ)-separated set E. Draw a line on the unit square in the direction of v_1 such that it has maximum length; call this length L. Let E be the set of points with spacing $\epsilon/|\lambda|^{n-1}$ on this line. To prove this is a separated set, we show that two consecutive points are farther than ϵ apart under the first n applications of f. Let x, y be two consecutive points on the line. Then,

$$x - y = v_1\epsilon/|\lambda|^{n-1}.$$

As before, we apply A k times to get

$$A^k(x-y) = \frac{\epsilon}{\lambda^{n-1}} A^k(v_1) = \frac{\epsilon}{\lambda^{n-1}} \lambda^k(v_1).$$

The absolute value of the above quantity reaches its maximum when $k = n-1$, which gives

$$d(f_A^{n-1}(x), f_A^{n-1}(x)) = A^{n-1}(x-y) = |\lambda|^{n-1}\frac{\epsilon}{|\lambda|^{n-1}} = \epsilon.$$

Therefore, E is (n,ϵ)-separated. Thus, $\mathrm{sep}(n,\epsilon) \ge |E|$. Additionally,

$$|E| \ge \frac{L}{\epsilon/|\lambda|^{n-1}}.$$

Taking the appropriate limits, we obtain that $h(f_A) \ge \log|\lambda|$.

Recalling the inequality obtained from considering the spanning set D and the squeeze theorem for limits, we have $h(f_A) = \log|\lambda|$. □

It is important to note that the entropy depends only on the expanding direction (the largest eigenvalue) of the transformation. The largest eigenvalue was used to bound $|\alpha|$ from above, whereas the upper bound for $|\beta|$ was determined based on the initial position of x and y. This is due to the fact that points get farther apart in the direction of v_1 after repeated applications of f, whereas they get closer together in the direction of v_2. Thus, the eigenvalue to v_2 does not affect the distance $d_n(x,y)$ and consequently does not affect the entropy either.

5. Topological Conjugacy

Topological conjugacy is an important concept for determining when two systems are dynamically equivalent. Many dynamical properties, including topological entropy, are invariant under conjugacy. Thus we can use entropy to more easily determine when two systems are not the same. For example, this invariance and the results from the previous section let us conclude that a rotation and the doubling map cannot be equivalent, since a rotation is an isometry.

Definition 5.1. Let (X, f) and (Y, g) be two topological dynamical systems. A semiconjugacy from g to f is a map $\pi : Y \to X$ that is onto and satisfies

$$f \circ \pi = \pi \circ g.$$

Definition 5.2. A topological conjugacy is an invertible semiconjugacy, that is $\pi : Y \to X$ is both onto and one-to-one.

Entropy being an invariant under conjugacy is an intuitive result, since a topological conjugacy can be thought of as putting the orbits of two dynamical systems in one-to-one correspondence. To prove this theorem, we first require the following lemma.

Lemma 5.3. *For a compact metric space (X, d) and a continuous map $f : X \to X$ the topological entropy of f does not depend on the choice of the metric d.*

Proof. Let d and d' be two metrics on X. For all $\epsilon > 0$ let

$$\delta(\epsilon) = \sup\{d'(x, y) : d(x, y) \leq \epsilon\}.$$

This means if a set U has d_n-diameter less than ϵ then U has d'_n-diameter at most $\delta(\epsilon)$. Since the sets in a cover might be bigger in the d'_n metric, it is possible we require fewer of them to cover X giving

$$\mathrm{cov}'(n, \delta(\epsilon), f) \leq cov(n, \epsilon, f),$$

where cov corresponds to d and cov' corresponds to d'. Also, since X is a compact metric space, $\lim_{\epsilon \to 0} \delta(\epsilon) = 0$. So we have

$$\lim_{\delta \to 0^+} \lim_{n \to \infty} \frac{1}{n} \log(\mathrm{cov}'(n, \delta, f) \leq \lim_{\epsilon \to 0^+} \lim_{n \to \infty} \frac{1}{n} \log(\mathrm{cov}(n, \epsilon, f).$$

By interchanging d and d' in the definition of $\delta(\epsilon)$ we get the opposite inequality, which concludes the proof. ☐

Theorem 5.4. *If (X, f) and (Y, g) are two topologically conjugate dynamical systems with conjugacy $\phi : Y \to X$ then $h(f) = h(g)$.*

Proof. Let d be a metric on X. Let d' be a metric on Y defined as

$$d'(y_1, y_2) = d(\phi(y_1), \phi(y_2)).$$

By the previous lemma, the value of $h(g)$ does not depend on the definition of the metric d'. Now consider

$$\begin{aligned}
d'_n(y_1, y_2) &= \max_{0 \le k < n-1} d'(g^k(y_1), g^k(y_2)) \\
&= \max_{0 \le k < n-1} d(\phi(g^k(y_1)), \phi(g^k(y_2)) \text{ by definition of } d' \\
&= \max_{0 \le k < n-1} d(f^k(\phi(y_1)), f^k(\phi(y_2))) \text{ because } \phi \text{ is a conjugacy} \\
&= d_n(\phi(y_1), \phi(y_2))
\end{aligned}$$

Hence distances in both X and Y depend solely on the d_n metric. Given this along with the fact that ϕ is a bijection, coverings as well as spanning and separated sets must have the same cardinality for X and Y. It follows that $h(g) = h(f)$. □

Acknowledgements. I would like to thank my mentor, Rachel Vishnepolsky, for her help with learning about dynamical systems, choosing an interesting paper topic, and reviewing many drafts. I would also like to thank Peter May for organizing the REU and reviewing this paper.

References

[1] Michael Brin, Garrett Stuck. *Introduction to Dynamical Systems.* Cambridge University Press. 2003.
[2] C. E. Silva. *Invitation to Ergodic Theory.* American Mathematical Society. 2007.
[3] Corrina Ulcigrai. *Topological Entropy.* http://www.maths.bristol.ac.uk/ maxcu/TopologicalEntropy14.pdf.
[4] Peter Walters. *An Introduction to Ergodic Theory.* Springer-Verlag. 1992.