SYMBOLIC DYNAMICS

AARON GEELON SO

ABSTRACT. This paper provides an introduction to dynamical systems and topological dynamics: how a system's configurations change over time, and specifically, how similar initial states grow dissimilar. Here, we focus on symbolic dynamics, a type of dynamical system, and how they can model other systems using Markov partitions. We end with a quantitative measure of complexity: topological entropy.

CONTENTS

INTRODUCTION

A *dynamical system* is a system with a set of possible states and a law that determines how the system changes. For example, a deck of cards has 52! possible configurations, and shuffling changes the system to another state. A pendulum can take on different positions and velocities, and Newton's laws determine how the position and velocity change over time. We call the set of all possible states the *phase space* X, and we can think of the law that governs how the system changes as the *time-evolution law* f. Here, we consider a more specific group of systems:

Definition 0.1. A *(discrete-time) dynamical system* is a pair (X, f) where X is a compact metric space, and f is a map from X into itself.

Card shuffling is an example of a discrete-time system; its states change in discrete time steps. In contrast, a swinging pendulum is a continuous-time system. We have required the phase space to be a compact metric so that our space has 'nice' properties, such as the existence of a convergent subsequence.

In this paper, we will first run through some canonical examples of dynamical systems in order to motivate definitions and to help us determine what properties to focus on. Then, we look at a specific type of dynamical system: symbolic dynamics. Finally, we will see how our study in symbolic dynamics help us study a more general class of systems. In order to maintain focus, we either explain unenlightening proofs intuitively or leave references for the curious reader.

Date: September 20, 2014.

1. Dynamics

1.1. **Preliminary examples.** We will use the following dynamical systems to help motivate the definitions such as topological transitivity, mixing, conjugacy, and so on.[1]

1.1.1. *Rotations on the circle.* Consider the dynamical system (S^1, R_α), where S^1 is the unit circle in $\mathbb{C}$ and $R_\alpha : S^1 \longrightarrow S^1$ is the map that rotates the circle by $2\pi\alpha$:

$$z \xmapsto{R_\alpha} e^{i2\pi\alpha} z.$$

As z is in the unit circle, it has the form $z = e^{i2\pi x}$, where $x \in [0, 1)$. Thus, $R_\alpha(z) = e^{i2\pi(x+\alpha)}$. Clearly, we can equivalently define this system on $\mathbb{R}/\mathbb{Z}$ with the map $S_\alpha : \mathbb{R}/\mathbb{Z} \longrightarrow \mathbb{R}/\mathbb{Z}$ by:

$$[x] \xmapsto{S_\alpha} [x + \alpha],$$

where $x + \alpha$ is equivalent to its fractional part, $x + \alpha \pmod 1$. We define a metric d on S^1 as the shortest arclenth between two points. We see that R_α is an *isometry*; it preserves lengths:

$$d(x, y) = d\big(R_\alpha(x), R_\alpha(y)\big).$$

In short, points that are close together stay close together under a rotation map. We can also look at iterations of a single point under R_α. To do this, we introduce some terminology:

Definition 1.1. Let (X, f) be a dynamical system and f invertible. The *orbit* of a point x is the set of images under the iterations of f or f^{-1} of x:

$$\mathcal{O}(x) := \{y \in X \mid y = f^n(x) \text{ where } n \in \mathbb{Z}\}.$$

Definition 1.2. If $f(x) = x$, then x is a *fixed point.* If $f^n(x) = x$ for some $n \in \mathbb{Z}$, then x is a *periodic point* with period n. We define $\mathrm{Fix}(f)$ to be the set of fixed points in X, and $P_n(f)$ the set of periodic points with period n. A subset $Y \subseteq X$ is *invariant* under f if $f(Y) = Y$.

For example, we can look at periodic points of the rotation map R_α. When $\alpha = p/q$ is rational, the set of periodic points is $P_q(R_\alpha) = X$. In fact, we can easily check that periodic points exist if and only if α is rational.

If α is irrational, there are no periodic points, and the orbit of any point is infinite. Because X is compact, the orbit of any x must have a convergent subsequence. So, we can find two points $R_\alpha^n x$ and $R_\alpha^m x$ within an arbitrarily small distance ϵ from each other. Because R_α is an isometry, x and $R_\alpha^{m-n} x$ is also within ϵ from each other. Iterations of R_α^{m-n} will tile S^1 with intervals less than ϵ. Since no point on S^1 is greater than ϵ from a point in the orbit of x, we see that $\mathcal{O}(x)$ is dense in S^1. A dense orbit is related to another concept called topological transitivity:

Definition 1.3. A dynamical system (X, f) is *topologically transitive* if for all nonempty open subsets $U, V \subseteq X$, there exists some integer $n \geq 1$ such that $f^n(U) \cap V \neq \varnothing$.

While the existence of a dense orbit and topological transitivity are in general independent, in the case with R_α, it is easy to see that since the orbit of any point is dense, every open set will contain a point in the orbit of another open set; hence R_α is topologically transitive.

1.1.2. *Dense orbits and topological transitivity.* We can show conditions of the phase space that make these two equivalent.

Proposition 1.4. *If X is perfect (i.e., X has no isolated points), and there exists some x such that $\mathcal{O}(x)$ is dense on X, then f is topologically transitive.*

Proof. Let $U, V \subseteq X$ be open sets. Since $\mathcal{O}(x)$ is dense, there is a $f^n(x) \in U$. And because X is perfect, the set $V \setminus \{x, \ldots, f^n(x)\}$ is nonempty and open; therefore, there exists some $f^m(x) \in V$, where $m > n$. Hence, $f^{m-n}(U) \cap V \neq \varnothing$. $\diamondsuit$

[1]The examples in Section 1.1 are based off of [HK] and [KH]. Section 1.1.2 is a discussion that results from [AD], [SI], and [KS].

Lemma 1.5 (Baire Category Theorem)**.** *Let X be a compact metric space. The intersection of a countable collection of open, dense subsets of X is dense.*

The proof of this theorem can be found [AD, p.17]. We use this theorem to show the following:

Proposition 1.6. *Let X be a compact, second-countable metric space. If f is topologically transitive, then there exists an $x \in X$ with a dense orbit.*

Proof. Let $\{U_n\}_{n\in\mathbb{N}}$ be a basis. Since f is a continuous map, $f^{-k}(U_n)$ is open. Furthermore,

$$V_n := \bigcup_{k\in\mathbb{N}_0} f^{-k}(U_n)$$

is an open, dense set in X, as f is continuous and topologically transitive. If we take $V_n \cap V_m$ for $n \neq m$, we obtain a set of points whose orbit include both U_n and U_m. This set of points is open and dense. Similarly, the set of points whose orbits are dense is the following:

$$\bigcap_{n\in\mathbb{N}} V_n.$$

By the Baire Category theorem, this set is nonempty (in fact, it is dense). ♢

1.1.3. *Expanding maps on the circle.* Define the dynamical system (S^1, F_n) where F_n maps:

$$z \xmapsto{F_n} z^n,$$

for $n \in \mathbb{N}$. As before, we can define the same system on $\mathbb{R}/\mathbb{Z}$ with the map E_n where we have:

$$[x] \xmapsto{E_n} [nx \pmod 1].$$

We require n to be a natural number greater than 1 because we have identified the point 0 with 1; we must have $E_n(0) = E_n(1)$, which holds only if $n \in \mathbb{N}$. For example, let $n = 3$.

Working with $\mathbb{R}/\mathbb{Z}$, it is easy to see that E_3 has two fixed points: 0 and 1/2. Suppose x is a rational number, p/q. E_3 simply maps p/q to the equivalence class containing $3p/q$. Since there are a finite number of equivalence classes with denominator q, all rational numbers are periodic points. On the other hand, if x is irrational, $E_3^n(x)$ can never equal x; otherwise, we can express x as a fraction. So, x is periodic if and only if x is rational.

It also turns out that E_3 is topologically transitive; however, we can state a stronger property:

Definition 1.7. A dynamical system (X, f) is *topologically mixing* if for all nonempty open subsets $U, V \subseteq X$, there exists some $N \in \mathbb{N}$ such that $f^n(U) \cap V \neq \varnothing$ for all $n \geq N$.

To distinguish transitivity and mixing, we can think of the two in the following way. If f is topologically transitive, in some sense, it will take a group of points 'close' together and these points will visit every neighborhood of the space. If f is topologically mixing, it will eventually 'spread' these points across the whole space. The following proposition shows that mixing is a stronger criterion than transitivity:

Proposition 1.8. *If f is topologically mixing, then f is topologically transitive.*

The proof is a direct application of the definitions. More interestingly, we can consider why E_3 is topologically mixing. It is telling that E_3 is called an expanding map: for any open interval $(a, b) \subseteq [0, 1)$, the length of the interval is $b - a$. Every time E_3 is applied, the length of the resulting interval is tripled, until the length exceeds 1. However, when the length reaches 1, the interval covers the whole space. Then, any subsequent iterations of E_3 will continue to cover the whole space. Since every open set contains an open interval, over time, every open set will come to cover the whole space. In fact, all expanding maps E_n are topologically mixing.

Under E_n, points that are originally close together will not remain close together. We say that the orbits of points under E_n exhibit a *sensitive dependency on initial conditions*, and up to some distance δ, two orbits will diverge exponentially [KH, p.41]. In contrast, the map R_α is an isometry, so the distance between two orbits remains constant; R_α is not topologically mixing.

Proposition 1.9. *Isometries are not topologically mixing.*

Proof. Intuitively, we let U be some small open interval, and V_1 and V_2 also be open intervals sufficiently small and far apart so that an interval of U's length cannot intersect both V_1 and V_2 at the same time. Because f is an isometry, for every n, either $f^n(U) \cap V_1 = \varnothing$ or $f^n(U) \cap V_2 = \varnothing$. Therefore, isometries cannot be topologically mixing. ♢

1.1.4. *Hyperbolic toral automorphisms.* Expanding maps are noninvertible. Now, we look at a system similar to expanding maps, except that these are invertible.[2] Let $A \in \mathrm{SL}_n\mathbb{Z}$ be a matrix with integer entries with $\det(A) = 1$. Define the map $F_A : \mathbb{T}^n \longrightarrow \mathbb{T}^n$ by:

$$F_A(x) = Ax \pmod 1.$$

Since the determinant of A is 1, F_A is invertible (in fact, the matrix $A^{-1} \in \mathrm{SL}_n\mathbb{Z}$). We call F_A a *toral automorphism* since it is an automorphism of the group $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$. We say that F_A is *hyperbolic* if the magnitude of the eigenvalues of A are different from 1.

A common example of a hyperbolic toral automorphism on $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ is defined by:

$$L = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

That is, we have the following map: $(x, y) \overset{F_L}{\longmapsto} (2x + y, x + y) \pmod 1$, represented in Figure 1. We can calculate the eigenvalues and associated eigenvectors, which turn out to be:

$$\lambda_\pm = \frac{3 \pm \sqrt{5}}{2} \quad \text{and} \quad v_\pm = \begin{bmatrix} 2 \\ -1 \pm \sqrt{5} \end{bmatrix}.$$

The eigenlines defined by translations of the span of each eigenvector are therefore invariant under the map F_L. Notice that $\lambda_- < 1 < \lambda_+$. The points on the line spanned by v_+ are *expanding* away from the origin (in an analogous fashion to the expanding maps). Likewise, the points on the line spanned by v_- are *contracting* toward the origin. Additionally, eigenlines are not the only F_L-invariant set; *any* line parallel to an eigenline is also invariant under F_L. Of course, since these lines are projected from the real plane to the 2-torus, the orbits are ostensibly much more complicated on the torus than on the plane. We can think of these lines as *linear flow*, the path of an object moving in a straight line on the torus.

Lemma 1.10. *If a linear flow is invariant under F_L, then it is dense on $\mathbb{T}^2$.*

Proof. On the torus, consider only the points (x, y) where $x \sim 0$; that is, points on the segment from the origin to the point $(0, 1)$ in Figure 1. Geometrically, this is a circle, S^1. A linear flow with slope λ_+ beginning at the origin will cross S^1 infinitely many times. In Figure 1, the first time this occurs is at the point a. The map that sends a point on S^1 to the point where this linear flow will next intersect S^1 is exactly a rotation map with $\alpha = \lambda_+$. Since the orbit of R_α is dense on S^1, translates of the eigenlines are dense on $\mathbb{T}^2$. ♢

Lemma 1.11. *Periodic points of F_L are dense.*

Proof. The proof is very similar to the above explanation to why rational numbers are periodic points under expanding maps. We look at points $(x, y) \in \mathbb{T}^2$ where $x, y \in \mathbb{Q}$. We can always express x and y with the same denominator q. Notice that the components of $F_L(x, y)$ may also be expressed with the denominator q. Because there are at most q^2 points with shared denominator q, points with rational coordinates are periodic. These points are dense in $\mathbb{T}^2$. ♢

Proposition 1.12. *F_L is topologically transitive.*

Proof. Let U and V be two open sets. By Lemma 1.11, periodic points are dense; there are periodic points $p \in U$ and $q \in V$. Consider the linear flow through p parallel to v_+. By Lemma 1.10, this line is dense; therefore, this line will have a non-empty intersection with V. Similarly, consider the linear flow through q parallel to v_-. This line will intersect the first line

[2] Section 1.1.4 follows the presentation of hyperbolic toral automorphisms in [KH].

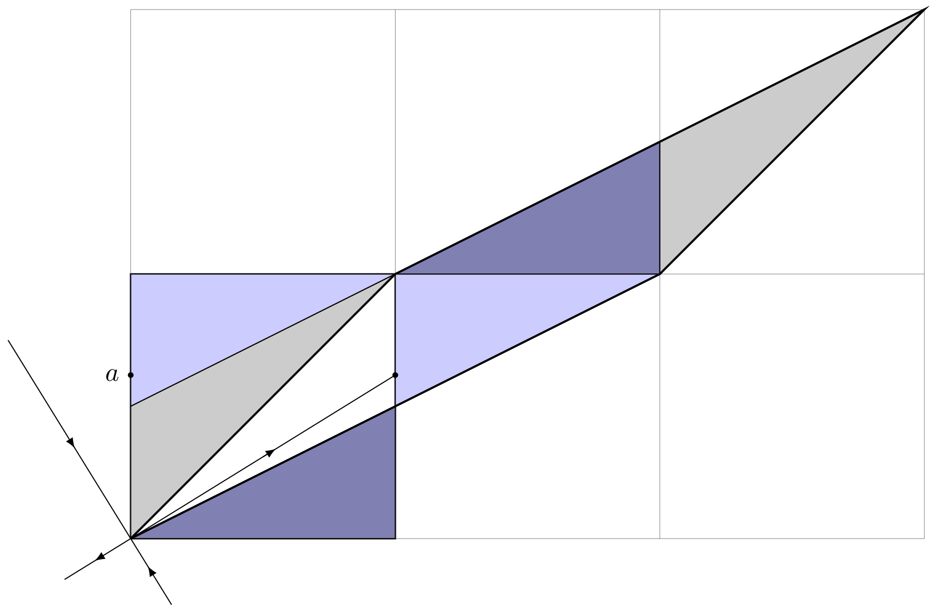

FIGURE 1. The action of F_L on $\mathbb{T}^2$. We also show the expanding and contracting eigenlines. *Adapted from* [KH, p.43].

at some point $r \in V$. If n is a common period of p and q, then under F_L^n, p and q are fixed points. The two lines are also invariant under F_L^n, where F_L^n expands and contracts the lines by λ_+^n and λ_-^n, respectively. Iterates of points on the first line converge to p as time goes to $-\infty$. Likewise, points on the second line converge to q as time approaches ∞.

More specifically, we can consider the point r on both lines:

$$\lim_{k\to\infty} F_L^{-kn}(r) = p.$$

And because r is also on the second line, we also have:

$$\lim_{k\to\infty} F_L^{kn}(r) = q.$$

Thus, for a sufficiently large k, $F_L^{2kn}(U) \cap V$ is nonempty. $\diamondsuit$

In fact, F_L is topologically mixing. Consider a line parallel to the expanding eigenline. Since this line is dense on $\mathbb{T}^2$, for all ϵ, there is a long enough segment of the line such that every point is within a distance ϵ from the line. Let U and V be open sets. Since U is open, it contains a segment of the eigenline. Since V is open, it contains some open ϵ-ball. Since the segment of the eigenline expands under every iteration, after some N, it will be sufficiently long so that it will intersect any ϵ-ball. Thus, $f^n(U) \cap V \neq \varnothing$ for all $n \geq N$.

Furthermore, since L is symmetric, by the spectral theorem, the eigenvectors are orthogonal. This is obvious in Figure 1. We can also use our orthogonal eigenbasis to change the fundamental region of our geometric picture. That is, instead of thinking of F_L acting on the unit square, we can construct a more 'natural' fundamental region. We have done just this in Figure 2. Then, the action of F_L on the two rectangles R_1 and R_2 below is to expand along their lengths and contract along their widths by a factor of λ_+ and λ_-, respectively. This geometric property is one we will later come back to exploit.

1.2. **Symbolic dynamics.** On these systems, the so-called *shift spaces* act as phase spaces and the shift maps as time-evolution laws. First, we will provide some necessary terminology before discussing how symbolics can be seen as dynamical systems, and how they can code other systems.

Definition 1.13. We call a finite set $\mathcal{A}$ an *alphabet* and its elements *symbols* or *letters*. A finite sequence of letters $x_0 \dots x_k$ is called a *block* or *word*. Infinite sequences $x : \mathbb{N}_0 \longrightarrow \mathcal{A}$ built

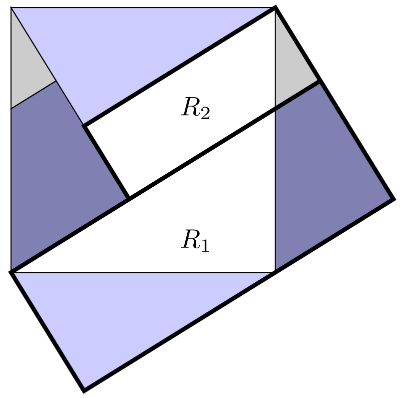

FIGURE 2. A different fundamental region for F_L. *Adapted from* [KH, p.84].

over these letters are denoted by $x = (x_i)_{i\in\mathbb{N}_0}$. Similarly, bi-infinite sequences $x : \mathbb{Z} \longrightarrow \mathcal{A}$ by $x = (x_i)_{i\in\mathbb{Z}}$. Often, for compactness of notation, we represent x by:

$$\dots x_{-2}x_{-1}.x_0x_1x_2\dots,$$

where the point separates the negative indices from the others. The collection of all bi-infinite (or infinite) sequences is $\mathcal{A}^{\mathbb{Z}}$, the *full $\mathcal{A}$-shift*. As $\mathcal{A}$ has N letters, it is isomorphic to:

$$\{0, 1, \dots, N-1\}.$$

We call the full shift over this set a *full N-shift* or a *Bernoulli shift*, which we denote by Ω_N. Similarly, we let Ω_N^R be the *one-sided Bernoulli shift*, the set of all infinite sequences. A Bernoulli shift is an example of a shift space, which we will define rigorously later.

We endow $\Omega_N = \{0, \dots, N-1\}^{\mathbb{Z}}$ with the product topology, where $\{0, \dots, N-1\}$ has the discrete topology. We call sets of sequences with a finite number of fixed coordinates *cylinders*:

$$C_n^\alpha := C_{n_1,\dots,n_k}^{\alpha_1,\dots,\alpha_k} := \{x \in \Omega_N \mid x_{n_i} = \alpha_i\}.$$

We can verify that cylinders form a basis for the topology on Ω_N; that is, open sets are in general unions of cylinders. We can also define a metric d_λ on Ω_N that induces the same topology as the product topology for any $|\lambda| > 1$ by:

$$d_\lambda(x, y) := \sum_{n\in\mathbb{Z}} \frac{|x_n - y_n|}{\lambda^{|n|}}.$$

Definition 1.14. The *shift map* on Ω_N is the invertible map $\sigma_N : \Omega_N \longrightarrow \Omega_N$ such that if $y = \sigma_N(x)$, then $y_i = x_{i+1}$. The shift map σ_N^R on Ω_N^R is defined similarly, but it is not invertible.

We consider the pair (Ω_N, σ_N) a dynamical system (similarly for the one-sided shift). For example, we can analyze the following dynamical system: (Ω_3^R, σ_3^R). The fixed points are the sequences with equal entries, such as $.000\dots$. Periodic points are sequences in which there is a block that is repeated infinitely.

Proposition 1.15. *Periodic points are dense.*

Proof. As cylinders form a basis on Ω_3^R, any open set will contain at least one cylinder C_n^α. Since a cylinder fixes a finite number of coordinates, pick a block $w = w_0 \dots w_k$ such that if the cylinder fixes the ith letter to be α_i, then $w_i = \alpha_i$. The infinite concatenation of w with itself is an infinite sequence in C_n^α. Hence, the periodic point:

$$.w_0 \dots w_k w_0 \dots w_k \dots$$

is in the original open set, proving that periodic points are dense. $\diamondsuit$

1.2.1. *The relation between σ_3^R and E_3.* Consider the one-sided Bernoulli shift, Ω_3^R. Again, we will denote the infinite sequences in Ω_3^R by:

$$x = .x_0x_1x_2\dots .$$

The one-sided shift transformation acts on x in the following way:

$$\sigma_3^R(x) = .x_1x_2\dots .$$

We now look at properties of (Ω_3^R, σ_3^R).

Proposition 1.16. *The map σ_3^R is topologically transitive and mixing.*

Proof. Let U and V be open sets, containing the cylinders C_n^α and C_m^β, respectively. Find blocks $p_0\dots p_{k-1}$ and $q_0\dots q_{m-1}$ in U and V as done in Proposition 1.15. Then, points x where $x_i = p_i$ for $0 \le i \le k-1$ and $x_i = q_i$ for $n \le i \le n+m-1$ satisfy $x \in U$ and $\sigma_3^n(x) \in V$ for any $n > k$. So, σ_3^R is topologically mixing, which also implies topological transitivity. $\diamondsuit$

And since σ_3^R is topologically transitive, there exists a point with a dense orbit in Ω_3^R. In fact, we can construct such a point by writing a sequence that contains all possible finite blocks.

Interestingly, our notation utilizing the decimal point is evocative of decimal representation of numbers. In Ω_3^R, the sequences correspond to points in the unit interval in base 3 representation. (In fact, except for rationals of the form $p/3^n$, there is a one-to-one correspondence from the sequence to the real number). The shift map σ_3^R is then equivalent to multiplying a point by 3 then taking the fractional part. In other words, (Ω_3^R, σ_3^R) seems equivalent to $(\mathbb{R}/\mathbb{Z}, E_3)$.

We can more precisely define equivalence between dynamical systems, called *conjugacy*, in the following way:

Definition 1.17. A *homomorphism* $h : (X, f) \longrightarrow (Y, g)$ from one dynamical system to another is a continuous function $h : X \longrightarrow Y$ such that $h\circ f = g\circ h$ (i.e., the following diagram commutes):

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[d, "h"'] & X \arrow[d, "h"] \\ Y \arrow[r, "g"] & Y \end{tikzcd}$$

If h is onto, then h is a *factor map*, and we say that there is a *topological semiconjugacy* from X to Y. If h is one-to-one, then h is an *embedding.* If h is both onto and one-to-one, then it is a *topological conjugacy.*

We can define $h : \Omega_3^R \longrightarrow \mathbb{R}/\mathbb{Z}$ to take a sequence to the real number it represents in base 3. Since h is onto and satisfies $h \circ E_3 = \sigma_3^R \circ h$, it is a factor map; it forms a *semiconjugacy* from Ω_3^R to $\mathbb{R}/\mathbb{Z}$. Figure 3 gives a geometric representation of the semiconjugacy. For example, in Figure 3, the symbolic representation x of the point r would have $x_0 = 0$ and $x_1 = 2$. We might think of the partitions of $\mathbb{R}/\mathbb{Z}$ as a way to classify points in terms of three different states, and the symbolic representation maps out the path of a point through these three states.

We often use symbolic dynamics to analyze properties of other dynamical systems. For example, consider a set in Ω_3^R that is homeomorphic to the ternary Cantor set:[3]

$$K = \{x \in \Omega_3^R \mid x_i = 0, 2 \text{ for all } i \in \mathbb{N}_0\}.$$

Clearly, K is invariant under σ_3^R. This implies that there is a set $h(K)$ in $\mathbb{R}/\mathbb{Z}$ that is homeomorphic to the Cantor set that is invariant under E_3. We have:

$$\sigma_3^R(K) = K \implies E_3 \circ h(K) = h \circ \sigma_3^R(K) = h(K).$$

Furthermore, notice that K is isomorphic to Ω_2^R by the conjugacy map $g : \Omega_2^R \longrightarrow K$ where g is defined coordinate-wise such that $g(0) = 0$ and $g(1) = 2$.

Since Ω_2^R is a compact, second-countable metric space, topological transitivity implies the existence of a dense orbit. So, there is a point whose orbit is dense in K, which further implies that there exists a point in $h(K) \in \mathbb{R}/\mathbb{Z}$ whose orbit is dense in $h(K)$.

[3]This example with the ternary Cantor set is based off of [HK].

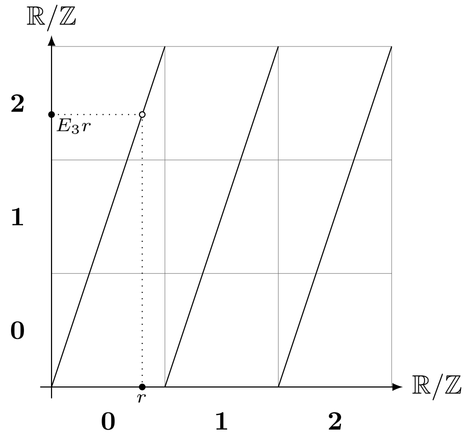

FIGURE 3. Partition $\mathbb{R}/\mathbb{Z}$ into the three sections **0**, **1**, and **2** above. If $x \in \Omega_3^R$ maps to $r \in \mathbb{R}/\mathbb{Z}$ under h, then x_k tells us which partition $E_3^k r$ is in.

We have shown that there exists an orbit of a point under E_3 with a closure homeomorphic to the Cantor set. We proved this easily because (semi)conjugacies preserve certain properties we are interested in dynamical systems:

Proposition 1.18. *Let h be a semiconjugacy from (X, f) to (Y, g). Fixed and periodic points are preserved under h. Topological transitivity and mixing in (X, f) imply topological transitivity and mixing in (Y, g), respectively.*

Proof. The proof is a simple application of continuity of h and the fact that h commutes:

$$h \circ g = f \circ h.$$

Clearly, if h is a conjugacy, then properties in (Y, g) extend to (X, f) as well. ♢

We will now formalize the symbolics a little more before moving on to discuss entropy.

2. Symbolics

2.1. **Topological Markov shifts.** We often use symbolic systems to model other dynamical systems, as we did in the above example with Ω_3^R and $\mathbb{R}/\mathbb{Z}$, where sequences represent the path of a point through different states. However, we might want to work with a more restricted space than the full Bernoulli shift.[4]

For example, if we want to model a system where state 1 is never followed by another state 1, we want to exclude all sequences containing the block 11. For dynamical systems, the type of restricted shifts we want to look at are called *topological Markov shifts.*

Definition 2.1. We can define a set of forbidden words $\mathcal{F}$ that determines whether certain sequences are *admissible* or not. That is, a sequence that contains a forbidden block is not admissible. We denote the set of all admissible sequences by $X_{\mathcal{F}}$. A *shift space* (or *shift*) is a subset $X \subseteq \mathcal{A}^{\mathbb{Z}}$ such that $X = X_{\mathcal{F}}$ for some set of forbidden blocks $\mathcal{F}$. If $X \subseteq Y$ where Y is a shift, then X is a *subshift.* If $\mathcal{F}$ is finite, then $X_{\mathcal{F}}$ is called a *subshift of finite type.*

An example of a subshift of finite type is the *golden-mean shift*, a subshift of Ω_2 with $\mathcal{F} = \{11\}$. Often, we look at a graphical representation of the shift, as in Figure 4. There are two ways we can represent a subshift of finite type: the *edge shift* and the *vertex shift.* In both, admissible sequences are bi-infinite walks on the graph. Based on Figure 4, a walk on the edge graph might look like:

$$\begin{bmatrix}0\\1\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix}\begin{bmatrix}0\\0\end{bmatrix}\begin{bmatrix}0\\0\end{bmatrix},$$

[4]Section 2.1 is is based off of [LM].

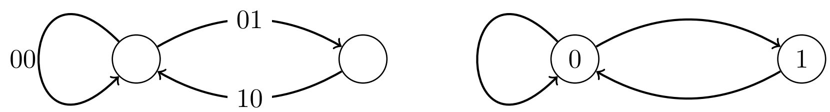

FIGURE 4. The edge shift (*left*) and vertex shift (*right*) representations of $X_{\mathcal{F}}$.

where we have indicated each edge by a column vector. The analogous walk on the vertex graph looks like 01000. Here, it is obvious how we can convert from one shift to the other.

The two types of graphs have different advantages. In an edge-labeled graph, multiple edges may connect two vertices, whereas having multiple edges is redundant on a vertex-labeled graph. On the other hand, vertex shifts are in some sense simpler:

If we walk on the vertex graph, our next step depends only on where we are now. Therefore, we have a *memory* of 1 step; we can only 'remember' our present state when we move to the next state. Suppose we have a shift where the block 111 is forbidden while the block 11 is not. Then we could not immediately represent the shift as a vertex shift, since this new shift requires a memory of 2 steps. So, vertex shifts are simpler in the sense that we do not need to keep track of very much information to know where we can go.

However, we can find a way to represent the shift with the forbidden block 111 as a vertex shift in the following way:

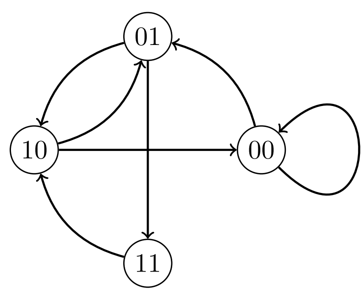

FIGURE 5. Higher block presentation of $X_{\mathcal{F}'}$ where $\mathcal{F}' = \{111\}$.

For a subshift of finite type with set of forbidden blocks $\mathcal{F}$, let the length of the longest block be M. There is a finite set of admissible blocks of length $M-1$. Let these blocks of length $M-1$ form a new alphabet. For example, let $y_1 = x_1 \dots x_{M-1}$ and $y_2 = x_2 \dots x_M$. We admit the block $y_1 y_2$ if the block $x_1 \dots x_M$ is not forbidden under $\mathcal{F}$. In this way, we let the block $y_1 y_2$ represent the analogous block $x_1 \dots x_M$. The new shift produced in this way is called the *higher block presentation.*

This way of turning any subshift of finite type into an equivalent shift of 1-step memory allows us to develop a theory of subshifts that are also vertex shifts. These shifts are closely related to Markov processes, where the probability of the successive event depends only on the prior event.

We also think of these subshifts as topological spaces (the topology is inherited from the full shift). Thus, we call these subshifts *topological Markov shifts*, and these are the shifts we are mainly interested for symbolic dynamics here.

While we can define a topological Markov shift by a set of forbidden blocks, we can also define it by the transition matrix:

Definition 2.2. Let X be a topological Markov shift. The corresponding *transition matrix* T is the 0-1 matrix where $T_{ij} = 1$ if the block ij is admissible, and $T_{ij} = 0$ if ij is forbidden.

For example, the transition matrix associated with the golden mean shift is:

$$T = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}.$$

We can also look at the powers of the transition matrix.

Proposition 2.3. *Let T be the transition matrix for a topological Markov shift. The ijth entry of T^k gives the number of admissible blocks of length $k+1$ that begins with i and ends with j.*

Proof. Consider the block $x_1 \dots x_k$. The following value can take values of 0 or 1 depending on whether the block forbidden or not:

$$T_{x_1} \cdot \dots \cdot T_{x_k} = \begin{cases} 0 & \text{if the block is forbidden} \\ 1 & \text{if the block is admissible.} \end{cases}$$

The ijth entry of T^k is the following:

$$\left[T^k\right]_{ij} = \sum_{x_2} \cdots \sum_{x_k} T_{ix_2} \cdot \dots \cdot T_{x_k j}.$$

Therefore, every admissible block that begins with i and ends with j contributes 1 to the sum, while forbidden blocks contribute nothing. $\diamondsuit$

Corollary 2.4. *The number of periodic points with period n is $P_n(\sigma) = \operatorname{tr} T^n$.*

Proof. This follows because the diagonal entries of T^n give the number of ways that a block of length $n+1$ can begin and end on the same number. We can use this to build periodic sequences of period n. $\diamondsuit$

Definition 2.5. A 0-1 matrix T is called *transitive* if there exists an n such that T^n is positive (i.e. T^n has all positive entries).

This is not a case of unfortunate nomenclature; we will soon see that if the transition matrix T is transitive, then the shift defined by T is topologically transitive. First, we need the following:

Lemma 2.6. *Let T be a 0-1 matrix. If T^n is positive, then for all $m \geq n$, T^m is positive.*

Proof. Notice that if the entries of the ith row of T are all zero, then the ith entry of any sensible matrix product TA also has an ith row of all zeroes. Thus, every row of a transitive matrix T has at least one nonzero entry. Suppose that T^n is positive. Then we have:

$$T^{n+1}_{ik} = \sum_j T_{ij} T^n_{jk} > 0$$

since there exists at least one $T_{ij} = 1$ and $T^n_{jk} > 0$. By induction, T^m is positive for $m \geq n$. $\diamondsuit$

Proposition 2.7. *If T is transitive, then the topological Markov shift Ω_T is topologically mixing.*

Proof. Recall that the basis for our topology on Ω_T are cylinder sets intersected with Ω_T. In fact, we can restrict the basis to sets of the following form:

$$C^\alpha_{k,T} := \Omega_T \cap \{x \in \Omega \mid x_i = \alpha_i \text{ for all } -k \leq i \leq k\}.$$

That is, we have restricted our cylinder sets to be centered at the 0th coordinate. Notice that for T transitive, $C^\alpha_{k,T}$ is nonempty if and only if α is an admissible sequence.

Let $U, V \subseteq \Omega_T$ be open sets. Then, there are cylinder sets contained in U and V:

$$C^\alpha_{k,T} \subseteq U \quad \text{and} \quad C^\beta_{\ell,T} \subseteq V,$$

where $\alpha = \alpha_{-k} \dots \alpha_k$ and $\beta = \beta_{-\ell} \dots \beta_\ell$. Let T^n be positive. Then, there is an admissible block of length $m+1$ for all $m \geq n$ that begins with α_k and ends with $\beta_{-\ell}$. Thus, we have:

$$\sigma_T^{k+m+\ell}(C^\alpha_{k,T}) \cap C^\beta_{\ell,T} \neq \varnothing$$

for all $m \geq n$. So, T is topologically mixing, and it is topologically transitive. $\diamondsuit$

In a similar way, we can also show that periodic points are dense if T is transitive.

Proposition 2.8. *If T is transitive, then periodic points in Ω_T are dense.*

Proof. Consider the cylinder $C^{\alpha}_{k,T}$. For some n, T^n is positive, so there is an admissible block of length $n+1$ that begins with α_k and ends with α_{-k}. We can clearly construct a periodic point in all basis sets for Ω_T. $\diamondsuit$

We alluded to the possibility of topological Markov shifts as modeling other dynamical systems, as in the above example with $(\mathbb{R}/\mathbb{Z}, E_3)$ and (Ω_3^R, σ_3^R). Specifically, we showed how (semi)conjugacies preserved periodic points, or topological transitivity or mixing, and so on. Now, we will be more explicit in how we might produce these homeomorphisms of systems.

2.2. **Markov partitions.** The main idea of Markov partitions[5] on a dynamical system (X, f) is to divide the phase space into a finite number of closed subsets, $X_0, X_1, \ldots, X_N$. Then, we can associate each point $x \in X$ with its orbit through the partitions:

$$x \longmapsto \omega,$$

where $f^n(x) \in X_{\omega_n}$. Then, we might be able to produce a conjugacy between X and the shift space. However, there are two difficulties: (1) a point might be coded by more than one sequence, and (2) a sequence might code more than one point.

Often, the phase space X is a connected manifold; if the set $\{X_i\}$ covers X, and if X_i is closed, overlap of the 'partition' is unavoidable. However, if we require the interior of each X_i to be disjoint from each other, then the set of overlapping points has measure zero, and it is unimportant to much of what we want to study on (X, f). In short, we generally disregard the first difficulty.

However, we would like for each sequence to code for at most one point. That is, the set:

$$\bigcap_{n\in\mathbb{Z}} f^n\left(X_{\omega_n}\right) = \{x\}.$$

If we can find a partition of the phase space for which this is always true, we can define a continuous map $h : \Lambda \subseteq \Omega_N \longrightarrow X$ such that h is a factor map.

As an example of how we might do this, we will construct a Markov partition of the torus for our toral automorphism example in section 1.1.4. Recall that the dynamical system we looked at was defined by $(\mathbb{T}^2, F_L)$, where F_L maps (x, y) to $(2x + y, x + y) \pmod 1$.

Furthermore, recall that while the unit square can function as the fundamental region on which F_L acts, we also found another fundamental region defined by the eigenvectors of F_L in Figure 2. In Figure 2, we have two rectangles: R_1 and R_2. We can see the action of F_L on these two rectangles in Figure 6.
Since F_L is invertible, it is clear that the projection of the image of the linear map applied to the rectangles onto its fundamental region is one-to-one. So, we can partition the torus based on this fact, as we have done in Figure 6.

Based on this figure, we can write the transition matrix associated with F_L:

$$T = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix}$$

It is simple to check either by calculating T^2 or by examining Figure 6 to see that T is transitive. Thus, we know that $\Omega_T \subset \Omega_5$ is topologically mixing and transitive. The question is whether we can use this information to determine that $(\mathbb{T}^2, F_L)$ is topologically mixing and transitive. In other words, we want to know whether there exists a factor map $h : \Omega_T \longrightarrow \mathbb{T}^2$.

To show that every sequence encodes at most one point, we look to the hyperbolicity of F_L. As we apply F_L to a partition Δ_i, the region expands in one direction and contracts in another.

[5]This section on Markov partitions combines elements from [AD] and [KH], the latter especially for the toral automorphism Markov partition example.

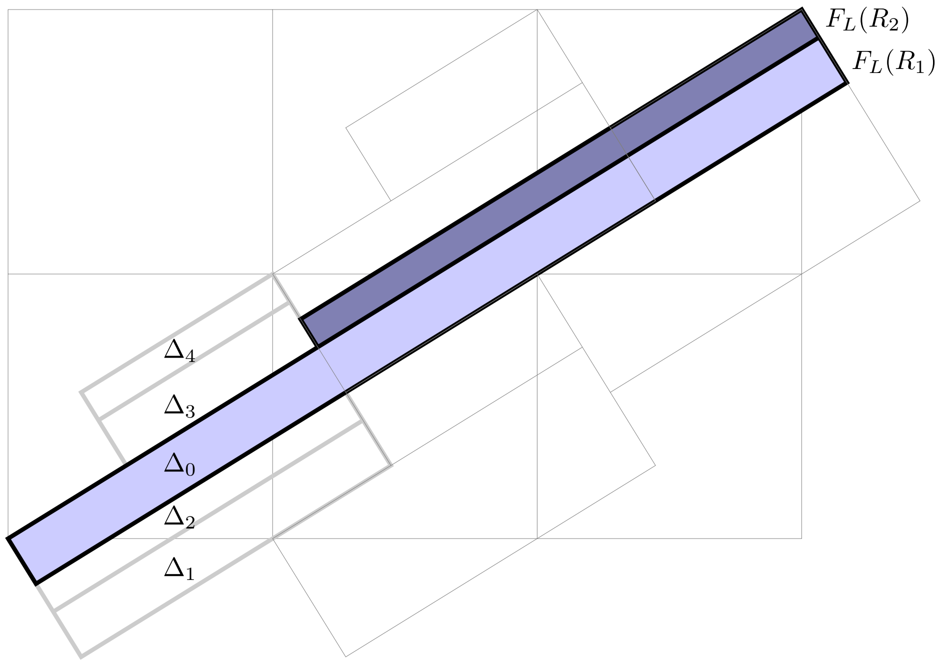

FIGURE 6. Markov partition of $\mathbb{T}^2$ under F_L. *Adapted from* [KH, p.85].

Similarly, F_L^{-1} contracts and expands Δ_i in the other way. Consider, for some $\omega \in \Omega_T$, the following set:

$$\bigcap_{n\in\mathbb{Z}} F_L^n\left(\Delta_{\omega_n}\right), \tag{2.1}$$

which is equivalently the intersection of these two sets:

$$\bigcap_{n\in\mathbb{N}_0} F_L^n\left(\Delta_{\omega_n}\right) \quad \text{and} \quad \bigcap_{n\in\mathbb{N}_0} F_L^{-n}\left(\Delta_{\omega_{-n}}\right). \tag{2.2}$$

These two sets are perpendicular line segments—a 'vertical' and 'horizontal' line with respect to the eigendirections—through the partition Δ_{ω_0}.

To clarify, consider the first set. Any point in this set must be within the partition Δ_{ω_0}. Since we know that this point must end up in Δ_{ω_1} after one iteration of F_L, we can determine the smaller region of Δ_{ω_0} that maps into Δ_{ω_1}. By examining Figure 6, we can see that these successively smaller regions continue to extend along the the whole width of Δ_{ω_0}, but not its length. The whole intersection is just a single line segment running along the width of Δ_{ω_0}. Similarly, the second set in Equation 2.2 is a single line segment running along the length of Δ_{ω_0}. Thus, the intersection of these two sets, equal to the set in Equation 2.1, is a single point.

So, we do have a factor map h from the shift space to the torus. We can use this to determine fixed points, for example. Based on T, there are three fixed points: the sequence of all 0's, 1's, and 4's. It is not difficult to see from Figure 6 that these three points in fact correspond to different corners of the unit square; that is, they are all identified with the origin.

Now, we will discuss topological entropy on shifts, which is a measure of the complexity of the shift space.

3. TOPOLOGICAL ENTROPY

Given that we have a notion of equivalency of dynamical systems, we want to determine what properties remain unchanged under conjugacy. Such properties are called *invariants*. They provide a way to determine quickly whether two dynamical systems could possibly be conjugate. Topological transitivity is an example of an invariant, as we showed in Proposition 1.18.

This means that we can immediately tell that two dynamical systems are not conjugate if one is topologically transitive and the other is not. However, just because two systems are both transitive does not mean that they *are* conjugate. An invariant of a dynamical system that determines precisely when two systems are conjugate is a *complete invariant*.

While not a complete invariance, topological entropy is an important invariant:

3.1. **Entropy on symbolic systems.** The complexity of a dynamical system should be preserved by conjugacy. Intuitively, the expanding map is more complex than the rotation map because the orbits of two close points do not remain close for the former, whereas the distance in fact remains constant in the latter. Therefore, suppose we knew the orbit of a point of the rotation map, it is a simple matter to determine the orbit of any other point. However, small errors in expanding maps grow exponentially, as though there are more paths in this system.

In symbolic dynamics, a way to measure the complexity is to count the number of admissible words of length n, and see how the number changes as n approaches infinity.[6] For example, the one-sided N-shift has 1 zero-length block, N one-length blocks, and N^n n-length blocks. However, a subshift Ω will have fewer. If w_n is the number of admissible n-letter blocks of a subshift, then we have:

$$w_{n+m} \leq w_n \cdot w_m.$$

We can take the logarithm to find the following nondecreasing subadditive sequence:

$$\log(w_{n+m}) \leq \log(w_n) + \log(w_m).$$

Lemma 3.1. *If (a_n) is a subadditive sequence (that is, $a_{n+m} \leq a_n + a_m$), then the following limit exists in $\mathbb{R} \cup \{\infty\}$:*

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

The proof may be found in [KH, p.374]. This lemma implies the following limit exists:

$$h(\Omega) := \lim_{n\to\infty} \frac{1}{n} \log(w_n).$$

This is the *topological entropy* of a shift.

For example, the topological entropy of a full N-shift is $\log N$. However, the calculation of the topological entropy is not so clear in general, so we introduce the following lemma:

Lemma 3.2 (Perron-Frobenius Theorem). *Let T be a transitive $r \times r$ matrix. Then T has one eigenvector v (up to scalar) with positive coordinates. The eigenvalue λ corresponding to v is positive and is greater than the magnitude of all other eigenvalues.*

The proof may be found in [KH, p.52]. We are more interested in its applications:

Proposition 3.3. *Let T be a transitive $r \times r$ matrix for the shift $\Omega_T \subseteq \Omega_r$. Let λ be the greatest eigenvalue of T. Then, the topological entropy is $h(\Omega_T) = \log \lambda$.*

Proof. Based on Proposition 2.3, it is clear that the number of admissible $(n+1)$-letter words of the shift Ω_T is the sum of all the entries of T^n. That is,

$$w_{n+1} = \sum_{i=0}^{r-1} \sum_{j=0}^{r-1} T_{ij}^n.$$

From the Perron-Frobenius theorem, let v be an eigenvector corresponding to λ, with positive coordinates bounded below by m and above by M. Therefore, we have:

$$m \sum_j T_{ij}^n \leq \sum_j T_{ij}^n v_j = \lambda^n v_i \leq \lambda^n M. \tag{3.1}$$

So, we further have:

$$m \sum_i \sum_j T_{ij}^n \leq r\lambda^n M,$$

[6] This way of presenting topological entropy on shifts follows [ST].

which allows us to give an upper bound to $h(\Omega_T)$:

$$h(\Omega_T) = \lim_{n\to\infty} \frac{1}{n} \log(w_n) \le \lim_{n\to\infty} \frac{1}{n} \log\left(\lambda^n \cdot \frac{rM}{m}\right) = \log\lambda. \tag{3.2}$$

Similarly, we can bound $h(\Omega_T)$ below in the same manner:

$$\lambda^n m \le \lambda^n v_i = \sum_j T_{ij}^n v_j \ge M \sum_j T_{ij}^n. \tag{3.3}$$

Equation 3.3 is analogous to Equation 3.1. And, we bound $h(\Omega_T)$ below also by $\log\lambda$. Thus, $h(\Omega_T) = \log\lambda$. $\diamondsuit$

For example, we can calculate the topological entropy of Ω_T for the symbolic representation of the hyperbolic toral automorphism in section 1.1.4. The maximal eigenvalue for the 5×5 transition matrix turns out to be:

$$\lambda = \frac{3+\sqrt{5}}{2}. \tag{3.4}$$

The entropy is $\log\lambda$. To connect the topological entropy on symbolic systems, we look at how topological entropy is defined in general.

3.2. **Topological entropy.** Imagine we are looking at a dynamical system with finite resolution; points within some $\epsilon > 0$ from each other are indistinguishable to us. However, two close points may move away from each other after some time; thus, if we look at the orbits of points, we can tell more points apart. In general, the longer we look, the more orbits we can discern (at least, the number of orbits is nondecreasing). The topological entropy of the system is a measurement of the growth of the number of orbits that we can distinguish given our finite resolution.[7]

Here, we consider dynamical systems with compact, second-countable metric space (with metric d). Suppose we can distinguish points with arbitrary, but finite, precision $\epsilon > 0$. That is, we cannot tell points less than a distance ϵ apart.

We now define the metric d_n^f on (X, f) to be the following:

$$d_n^f(x,y) := \max_{0\le k\le n-1} d\big(f^k(x), f^k(y)\big).$$

That is, we track the orbits of two points x and y for $n-1$ steps. The distance between two points under d_n^f is the maximum distance between their first n iterations under f. If we can look at the system for n time steps, we can distinguish points that are ϵ apart on the metric d_n^f.

We can now define topological entropy in an analogous way to that on shift spaces. On a topological Markov shift, if our resolution is so coarse that we can determine only the first coordinate of a sequence, then we can partition the space so that indistinguishable points are grouped together. The number of partitions is the number of admissible words of length 1. If we can track the system for $(n-1)$ iterations of the shift map, we could refine the partition (now, the number of partitions is w_n). To perform the same analysis on a general dynamical system, we define the following:

Definition 3.4. Given (X, f), the *d_n^f-diameter* of a subset $Y \subseteq X$ is the supremum of the d_n^f-distance between two points in Y.

Therefore, if a subset of X has a d_n^f-diameter less than ϵ, its points are indistinguishable to us for at least $(n-1)$ iterations of the map f. We can form covers of X using sets with d_n^f-diameter less than ϵ. This is related to the idea of partitioning the shift space based on the n coordinates. On the shift space, we counted the number of partitions. But, we cannot just count the number of sets in a cover. However, we assumed that X is compact; therefore, finite covers exist. So, we may define the following quantity:

Definition 3.5. Let $\mathrm{cov}(n, \epsilon)$ be the cardinality of a minimal covering of X by sets with d_n^f-diameter less than ϵ.

[7]This presentation of topological entropy on a general dynamical system mostly uses [BS]; however, proofs for Proposition 3.6 and Proposition 3.8 are adapted from [KH].

Notice that $\text{cov}(n,\epsilon)$ is nondecreasing with respect to n because d_n^f is a nondecreasing sequence of metrics. That is, if $n > m$, then $d_n^f(x,y) \geq d_m^f(x,y)$. Furthermore, $\text{cov}(n,\epsilon)$ satisfies:

$$\text{cov}(n+m,\epsilon) \leq \text{cov}(n,\epsilon)\cdot\text{cov}(m,\epsilon). \tag{3.5}$$

To see this, let $\mathcal{U}$ and $\mathcal{V}$ be minimal covers of X with sets of d_n^f and d_m^f-diameters less than ϵ, respectively. Let $U \in \mathcal{U}$ and $V \in \mathcal{V}$. Since $\mathcal{U}$ and $\mathcal{V}$ are both covers of X, sets of the form $U \cap f^{-n}(V)$ also cover X. Furthermore, these sets have d_{n+m}^f-diameters less than ϵ. This new cover has cardinality $\text{cov}(n,\epsilon)\cdot\text{cov}(m,\epsilon)$, proving the inequality above. As before, we can take the logarithm to find the following nondecreasing subadditive sequence:

$$\log\text{cov}(n+m,\epsilon) \leq \log\text{cov}(n,\epsilon) + \log\text{cov}(m,\epsilon).$$

By the same logic, we define the following quantity, which we call the *topological entropy*:

$$h(f) := \lim_{\epsilon\to 0}\limsup_{n\to\infty}\frac{1}{n}\log\text{cov}(n,\epsilon). \tag{3.6}$$

It is not obvious that the topological entropy does not depend on the metric on X. However, it is aptly named the topological entropy because two metrics inducing the same topology on X will yield the same topological entropy.

Proposition 3.6. *Let d and d' be equivalent metrics on X. The topological entropies $h_d(f)$ and $h_{d'}(f)$ arising from the two metrics are equal.*

Proof. Consider the subset $Y \subseteq X \times X$ of points (x,y) such that $d(x,y) \geq \epsilon$. This is a compact subset of $X \times X$, with the inherited product topology. Since d' is continuous, by the extreme value theorem, d' reaches a minimum $\delta(\epsilon)$ on some point $(x,y) \in Y$. Furthermore, $\delta(\epsilon) > 0$ since $x \neq y$. Thus, points within ϵ of each other by d are within $\delta(\epsilon)$ by d'. Therefore,

$$\text{cov}_d(n,\epsilon) \geq \text{cov}_{d'}\big(n,\delta(\epsilon)\big).$$

This implies that $h_d(f) \geq h_{d'}(f)$. We get the reverse inequality by switching d and d'. Thus, topological entropy is independent of metric $h_d(f) = h_{d'}(f)$. $\diamondsuit$

Corollary 3.7. *Topological entropy is an invariant of a topological conjugacy.*

Proof. Let (X,f) and (Y,g) be conjugate with the homeomorphism $\phi : X \longrightarrow Y$. Let d be a metric on X. We can define d' a metric on Y such that:

$$d'(y_1,y_2) = d\left(\phi^{-1}(y_1),\phi^{-1}(y_2)\right).$$

Therefore, ϕ is an isometry. Since topological entropy is independent of metric, $h(f) = h(g)$. $\diamondsuit$

Proposition 3.8. *Let (X,f) and (Y,g) be semiconjugates by the factor map $\phi : X \longrightarrow Y$. Then, $h(f) \geq h(g)$.*

Proof. The proof is very similar to that for Proposition 3.6. Since X is compact and ϕ is continuous, ϕ is uniformly continuous. Hence, there exists some $\delta(\epsilon)$ such that

$$d_X(x,y) < \delta(\epsilon) \implies d_Y\big(\phi(x),\phi(y)\big) < \epsilon.$$

Therefore, if two points are within $\delta(\epsilon)$ of each other on X, their images under ϕ are within ϵ on Y. This implies:

$$\text{cov}_X\big(n,\delta(\epsilon)\big) \geq \text{cov}_Y(n,\epsilon).$$

This further implies $h(f) \geq h(g)$. $\diamondsuit$

From this proposition, we see that the topological entropy of our hyperbolic toral example $h(F_L)$ must be less than $\log\lambda$, where λ is the same as in Equation 3.4. Not coincidentally, λ is also the value of the greatest eigenvalue of the matrix L defining the toral automorphism. In fact, the entropy of the toral map and the Markov shift map are equal:

Proposition 3.9. *If F_A is a hyperbolic toral automorphism defined by the 2×2 matrix A with eigenvalues λ, λ^{-1}, where $|\lambda| > 1$, then the topological entropy is $h(F_A) = \log|\lambda|$.*

The proof involves finding a minimal cover for $\mathbb{T}^2$; the details may be found in [BS, p.41]. Similarly, the topological entropy of a full n-shift tells us that $h(E_n) \leq \log n$. And in fact, the topological entropy of the expanding maps E_n are equal to $\log n$.

On one hand, that the topological entropy of the dynamical systems above and their associated topological Markov shifts are equal is a result of the statistical insignificance of points with multiple shift representations. On the other hand, that the entropy are equal should not be surprising: topological entropy effectively measures local expansions by a map; the map E_n locally expands the circle by n. Thus, error terms ϵ are essentially expanded exponentially. And so, we should expect the growth of orbits to be on the order of $\log n$.

Similarly, the map F_L expands by λ on certain lines and contracts by λ^{-1} on other lines. The contraction does not add to the topological entropy. However, locally on an expanding line, the system looks like an expanding map. Thus, the topological entropy $\log |\lambda|$ is not unreasonable.

Epilogue

It was the goal of this paper to provide a brief introduction to dynamical systems, with a focus on symbolic dynamics. Of course, there is so much more to be studied.

For more examples of dynamical systems, I suggest [HK], which gives an accessible introduction to topics discussed in this paper, as well as to topics such as hyperbolic dynamics and variational methods. For a more advanced introduction, [BS] is concise, yet very readable. [KH] is even more comprehensive. The latter two cover topics such as ergodic theory or measure-theoretic entropy that are omitted in the first.

For more theory and applications of symbolics, I suggest [LM], which, for example, goes into much more detail about conjugacy and codes.

Acknowledgments. It is my pleasure to thank Peter May for making possible the 2014 REU program, in which I have learned quite a bit of math. Also, I would like to thank all those, including the lecturers and mentors, who took their time during the summer to make the program run. In particular, I want to thank my mentor Clark Butler. He not only introduced me to dynamical systems with great lessons and answers, but also made this paper much more accurate, focused, and readable.

Thank you to all of my teachers, friends, and family for having giant shoulders.

References

[AD] Adler, R. *Symbolic Dynamics and Markov Partitions*, Bulletin of the American Mathematical Society, 35.1 (1998): 1-56.

[BS] Brin, M., Stuck, G. *Introduction to Dynamical Systems*, Cambridge University Press, New York (2002).

[HK] Hasselblatt, B., Katok, A. *A First Course in Dynamics with a Panorama of Recent Developments*, Cambridge University Press, New York (2003).

[KH] Katok, A., Hasselblatt, B. *Introduction to the Modern Theory of Dynamical Systems*, Cambridge University Press, New York (1995).

[KS] Kolyada, S., Snoha, L. *Topological transitivity*, Scholarpedia, (2009): 4(2):5802.

[LM] Lind, D., Marcus, B. *An Introduction to Symbolic Dynamics and Coding*, Cambridge University Press, New York (1995).

[SI] Silverman, S. *On maps with dense orbits and the definition of chaos*, Rocky Mountain Journal of Mathematics, 22.1 (1992): 353-75.

[ST] Sternberg, S. *Dynamical Systems*, Dover Publications, Inc., New York (2010).