A BRIEF OVERVIEW OF CHAOS

MARIYA PERSHYNA

Abstract. In this paper we will discuss the basics of chaos theory, using examples to highlight various nuances. We begin by providing some of the fundamental definitions and concepts in dynamical systems. We then define chaos and provide examples of chaotic functions as well as non-chaotic functions to develop the intuition of what it means for a function to be chaotic. We focus in particular on the quadratic map and Horseshoe Map, using topological conjugacy and symbolic dynamics to analyze these systems. We conclude with a brief discussion of alternative definitions of chaos and quantitative measures of chaos.

Contents

1. Preliminaries

We will begin by defining a dynamical system. Dynamical systems are systems that evolve based on some fixed rule or function. A differential equation, for example, is a dynamical system. They can be either continuous or discrete. Continuous dynamical systems are defined for every point in time, while discrete systems are only considered at discrete time points. Discrete dynamical systems, on the other hand, are formed by iterating a function ($f(x), f(f(x)), f(f(f(x)))$, and so on). The k^{th} iterate of a function is denoted as $f^k(x)$ (not to be confused with taking $f(x)$ to the power of k or the k^{th} derivative of f).

In this paper, we will always be working in a metric space X (no assumptions about completeness or compactness are necessary unless otherwise specified).

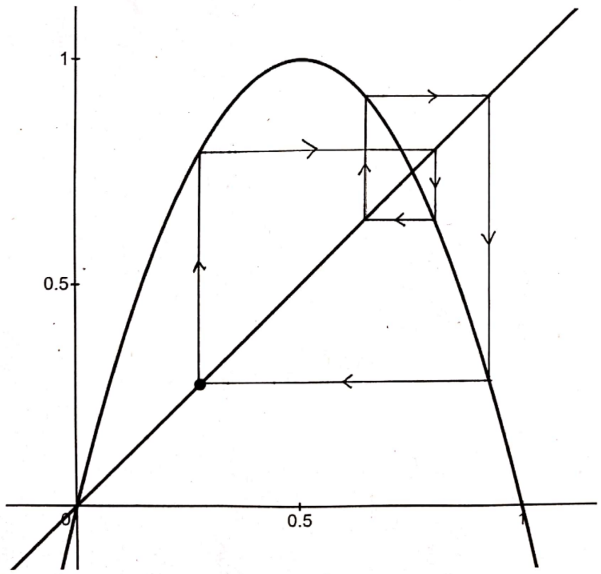

FIGURE 1. A point of period 4 for $f(x) = 4x(1-x)$.

Definition 1.1. A *fixed point* is a point $x \in X$ such that $f(x) = x$. Fixed points can be either *attracting*, where nearby points go towards the fixed point under iterations of f, or *repelling*, where nearby points move away under iterations of f.

One can easily find fixed points graphically by looking at the intersection of f with the line $y = x$. See Figure 1 for an example.

Definition 1.2. A *periodic point* is a point $x \in X$ such that for some $k > 0$, $f^k(x) = x$.

Definition 1.3. The *orbit* of a point is the set $\mathcal{O}(x) = \{f^k(x) \mid k \geq 0\}$. If f is invertible, the orbit of a point is $\mathcal{O}(x) = \{f^k(x) \mid k \in \mathbb{Z}\}$.

A common goal in dynamical systems is to understand the orbits of all points in the domain of the function. One can graphically examine the orbits of points and find periodic points by starting at the point $(x, f(x))$ and moving horizontally until you reach the line $y = x$. Moving vertically back to the graph of $f(x)$ at this point will land you on the point $(f(x), f(f(x)))$. Continuing this process allows you to quickly see the long-term behavior of a point. An example of this can be seen in Figure 1, where the graphs of $f(x) = 4x(1-x)$ and $y = x$ are shown. We see that f has two fixed points where f intersects $y = x$, at $x = 0$ and $x = 0.75$. Using the aforementioned graphical analysis process, we see that $x = 0.2771$ (rounded to four decimal places) is a point of period 4.

The following definition is for a commonly-studied aspect of dynamical systems.

Definition 1.4. An *invariant set* is a subset $Y \subset X$ such that $f(Y) = Y$.

Invariant sets often exhibit the most interesting behavior of a system, such as chaos. One common "shape" of an invariant set is the Cantor set.

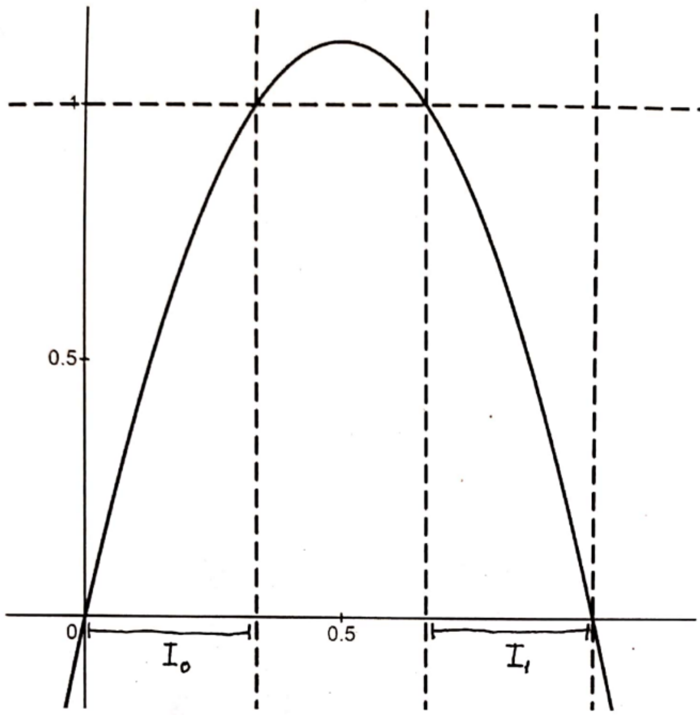

FIGURE 2. I_0 and I_1 for the quadratic map when $\mu > 2+\sqrt{5}$

Definition 1.5. A subset $\Lambda \subset X$ is called a *Cantor set* if it is closed, totally disconnected, and perfect. A set P is totally disconnected if between any two points $x, y \in P$ there exists a point $q \notin P$. A set P is perfect if $P = P'$ (where P' is the set of limit points of P).

Example 1.6. Consider the quadratic map on $\mathbb{R}$, $f = \mu x(1-x)$, where μ is a constant in $\mathbb{R}$. In this case, let $\mu > 2+\sqrt{5}$. We will find the invariant set of f on the interval $I = [0,1]$ when $\mu > 2+\sqrt{5}$ and show that it is a Cantor set. For any $x \notin I$, $f(x) \notin I$ and $f^k(x)$ goes to negative infinity as k goes to infinity. Thus once the orbit of a point exits I, it will not re-enter I. Let A_0 be the set of points on I such that $f(x) \notin I$. As can be seen in the diagram, A_0 is an open interval centered around $\frac{1}{2}$. Since $f(A_0) \cap I = \emptyset$, A_0 is not in the invariant set of f. Now let I_0 denote the closed interval in I to the left of A_0 and I_1 denote the closed interval in I to the right of A_0 (see Figure 2).

Let $A_1 = \{x \in I \mid f(x) \in A_0\}$. For any $x \in A_1$, $x \in I$ and $f(x) \in A_0 \subset I$, but $f^2(x) \notin I$. Since $f(I_0) = f(I_1) = I$, we have that A_1 consists of two open intervals, one in the middle of I_0 and one in the middle of I_1. So far we have determined that A_0 and A_1 are not in the invariant set of f since points in these intervals eventually leave I. We can continue in the same manner by "removing" the middle of each remaining interval to get the invariant set of f. Note that at the n^{th} iteration of this process (starting with $n = 0$), 2^n intervals are removed. Let

$A_{k+1} = \{x \in I \mid f(x) \in A_k\}$ for $k \geq 0$. Then the invariant set Λ of f on I is

$$\Lambda = I - \bigcup_{k \geq 0} A_k.$$

We will now show that Λ is a Cantor set. Because Λ is an infinite intersection of closed intervals, Λ itself is also closed. Next we will show that Λ is totally disconnected. When $\mu > 2 + \sqrt{5}$, $|f'(x)| > 1$ for any $x \in I_0 \cup I_1$. Let $\lambda > 1$ be such that $|f'(x)| > \lambda$ for any $x \in I_0 \cup I_1$. It follows that $|(f^n)'(x)| > \lambda^n$ for any $x \in \Lambda \subset I_0 \cup I_1$. Suppose for the sake of contradiction that there exist $x, y \in \Lambda$ such that $[x, y] \subset \Lambda$. We showed that $|(f^n)'(\alpha)| > \lambda^n$ for all $\alpha \in [x, y]$. Take n such that $\lambda^n |x - y| > 1$. By the Mean Value Theorem, there exists some $\alpha \in [x, y]$ such that

$$\frac{|f^n(x) - f^n(y)|}{|x - y|} = |(f^n)'(\alpha)| > \lambda^n,$$

so $|f^n(x) - f^n(y)| > \lambda^n |x - y| > 1$. However, this is a contradiction since $f^n(x), f^n(y) \in I$ so $|f^n(x) - f^n(y)| < 1$. It follows that our assumption was incorrect and between any $x, y \in \Lambda$ there exists a point $q \notin \Lambda$, so Λ is totally disconnected.

We will now show that Λ is perfect. Since Λ is closed, it must contain all of its limit points, so $\Lambda' \subset \Lambda$. For any $x \in I$ such that x is the endpoint of some A_k, x must be in Λ because either $f(x) = 0$ or $f(x) = 1$ and $f^2(x) = 0$, so x is eventually fixed in I. Consider some $p \in \Lambda$. As shown earlier, for any y arbitrarily close to x, there must be some $q \in [x, y]$ such that $q \notin \Lambda$, which means that $q \in A_k$ for some k. This implies that the intervals A_k get arbitrarily close to x, so there is a sequence of endpoints of the A_k that approach x. Since all of these endpoints are in Λ, any $x \in \Lambda$ is the limit of a sequence of points in Λ, so $\Lambda \subset \Lambda'$. We have now shown that Λ is perfect, closed, and totally disconnected, so Λ is a Cantor set.

Definition 1.7. Two functions $f : X \to X$, $g : Y \to Y$ are *topologically conjugate* if there exists a homeomorphism $h : X \to Y$ such that $h \circ f = g \circ h$.

Intuitively, topological conjugacy can be thought of as a topological change of coordinates, similar to conjugacy in linear algebra. The importance of topological conjugacy is that it preserves the behavior of dynamical systems. Some dynamical systems can be difficult to study directly but are topologically conjugate to a map that is easier to understand. Such a topological conjugacy can provide a significant amount of insight into the behavior of the original system.

2. Defining Chaos

Before we can define chaos, we will need to define a few more properties of dynamical systems.

Definition 2.1. A map f is *(topologically) transitive* on an invariant set Y if there exists some point $x \in Y$ such that the forward orbit of x is dense in Y.

See Example 2.5 for an example of a transitive map. Note that a map can be transitive on its entire domain or simply on an invariant subset of the domain. While this is a commonly used definition for transitivity, a more intuitive version is provided in the following theorem.

Theorem 2.2. *A map f is transitive on a set Y if and only if for any two open sets U, V in Y, there exists some $k \in \mathbb{Z}$, $k > 0$, such that $f^k(U) \cap V \neq \emptyset$.*

This result is proved by the Birkhoff Transitivity Theorem. For a proof of the theorem, see Section 7.2 of [4]. Intuitively, transitivity means that a map "mixes" all of the points in a set Y.

We will also need to define the notion of sensitive dependence on initial conditions.

Definition 2.3. A map $f : X \to X$ has *sensitive dependence on initial conditions* if there exists some $r > 0$ such that for any $x \in X$ and $\epsilon > 0$ there exists some $y \in X$ with $d(x, y) < \epsilon$ and a $k \geq 0$ such that $d(f^k(x), f^k(y)) \geq r$.

See Example 2.6 for an example of a map that has sensitive dependence on initial conditions. Throughout this paper we will refer to sensitive dependence on initial conditions as sensitivity for the sake of conciseness.

We can now state a definition for a chaotic map.

Definition 2.4. A map $f : X \to X$ is *chaotic* on a set Y if f is transitive on Y, periodic points are dense in Y, and f has sensitive dependence on initial conditions.

This definition of chaos was proposed by Devaney in 1989, but other, subtly different, definitions exist as well. For a discussion of other notions of chaos see Sections 2.4 and 3.

In order to fully understand what it means for a map to be chaotic, it is helpful to first consider examples of maps that are not chaotic. First, we consider a map that is transitive on its entire domain but lacks sensitivity and dense periodic orbits.

Example 2.5. Let $f : S^1 \to S^1$ be a rotation of the unit circle by an irrational number λ. An equivalent map is $F : \mathbb{R} \to \mathbb{R}$ such that $F(x) = x + \lambda \mod 1$. This map is transitive on the unit circle but is not sensitive and does not have dense periodic points (in fact, f has no periodic points at all).

We will first show that f has no periodic points. Suppose that there exists some $x \in S^1$ such that $f^k(x) = x$ for $k > 0$. This means that $F^k(x) = x + k\lambda \mod 1 = x$, so $k\lambda = n$ for some integer n. Thus, $\lambda = \frac{n}{k}$ for $n, k \in \mathbb{Z}$, $k > 0$, so λ is a rational number. By the contrapositive, we get that if λ is irrational, f has no periodic points.

We will now show that f is transitive by showing that the orbit of every point in S^1 is dense in S^1. Take $x \in S^1$ and $\varepsilon > 0$. Since x is not periodic, the points $f^k(x)$ are distinct for every k. Because S^1 is compact and therefore every sequence has a convergent subsequence, $f^k(x)$ must have a limit point in S^1. Thus, there must exist $n, m \in \mathbb{N}$ such that $d(f^n(x), f^m(x)) < \varepsilon$, where d is the arc length metric. Let $q = n - m$. Then $d(f^q(x), x) < \varepsilon$ because f preserves the distance between points. We also get that $d(f^{2q}(x), f^q(x)) < \varepsilon$, $d(f^{3q}(x), f^{2q}(x)) < \varepsilon$, and so on. These intervals must eventually cover all of S^1. Thus, for any $\varepsilon > 0$ and $y \in S^1$, there exists some $k > 0$ such that $d(f^k(x), y) < \varepsilon$. This shows the orbit of any point is dense in S^1, which means that f is transitive on S^1.

Finally, we will show that f does not have sensitive dependence on initial conditions. A rotation of the circle preserves distance between points, so $d(x, y) = d(f(x), f(y))$ for any $x, y \in S^1$. Thus points $x, y \in S^1$ that begin a distance $\varepsilon > 0$ apart will remain a distance ε apart under iterations of f, so f does not have sensitive dependence on initial conditions.

Why is this a good non-example of a chaotic map? Despite being transitive on its domain, f is incredibly orderly in the sense that it preserves the order and distance of points on the circle (for any $x, y \in S^1$, we have that $d(x, y) = d(f(x), f(y))$).

We will now consider an example of a function that has sensitive dependence on initial conditions but is not chaotic due to its lack of transitivity.

Example 2.6. Let $f : \mathbb{R} \to \mathbb{R}$ be the function $f(x) = 2x$. This map has sensitive dependence on initial conditions but is not transitive on any non-finite subset of $\mathbb{R}$ and does not have dense periodic points.

Let $r = 1$. Take $x \in \mathbb{R}$ and $\varepsilon > 0$. Take $y \in \mathbb{R}$ such that $|x - y| = \delta < \varepsilon$. Since $f^k(x) = 2^k x$, if $|x - y| = \delta$, then $|f^k(x) - f^k(y)| = 2^k\delta$. Take $k > 0$ such that $2^k\delta \geq 1$, so $|f^k(x) - f^k(y)| \geq r$. We have now shown that f has sensitive dependence on initial conditions.

We can see that f has a single fixed point at $x = 0$ since f intersects the line $y = x$ only at $x = 0$. Using graphical analysis, we see that any point $x > 0$ goes to infinity under iterations of f and any point $x < 0$ goes to negative infinity under iterations of f. Thus f has a single periodic point, the fixed point at $x = 0$. This shows that f does not have dense periodic points.

A function is defined as transitive on an invariant set if there exists a point with a dense forward orbit. The only invariant sets of f are the set $\{0\}$ and all of $\mathbb{R}$ since for any proper subset $X \subset R$ such that $X \neq \{0\}$, any $x \in X$ will leave X under iterations of f. Sets consisting of single points are not very interesting in the study of chaos and do not fulfill the conditions for sensitive dependence on initial conditions. Thus we will only consider whether f is transitive on $\mathbb{R}$. If $x = 0$, the forward orbit of x is $\{0\}$, which is not dense in $\mathbb{R}$. If $x \neq 0$, the forward orbit of x is $\{x, 2x, 4x \ldots\}$, which is not dense in $\mathbb{R}$. Thus f is not transitive on $\mathbb{R}$.

Although $f(x) = 2x$ has sensitive dependence on initial conditions, it doesn't "feel" chaotic due to its lack of transitivity. On the other hand, the irrational rotation map on the unit circle "mixes" together all points of the circle under iterations of the map, yet it doesn't "feel" chaotic because of its distance-preserving nature.

After considering maps that are not chaotic, we can now begin to think about what it means for a map to be chaotic. As we've seen in the examples, one important aspect of chaotic maps is transitivity, which means that all neighborhoods of the chaotic region get mixed together under iterations of the map. Sensitive dependence on initial conditions is important because without it, the behavior of the map would be significantly more predictable with computations. As for dense period orbits, Devaney's inclusion of this property in the definition of a chaotic map reveals an interesting aspect of chaos. According to Devaney, even a seemingly random chaotic map has an element of regularity, as shown by the dense periodic orbits. However, this regular behavior is "hidden" since even points infinitesimally close to a periodic orbit will have drastically different behavior.

As mentioned previously, topological conjugacy preserves behavior of dynamical systems. Thus topological conjugacy is an excellent tool for studying chaotic maps. The following theorem shows that conjugacy preserves transitivity and dense periodic orbits, two of the components of chaos.

Theorem 2.7. *Suppose $f : X \to X$ and $g : Y \to Y$ are topologically conjugate. If f is topologically transitive and has dense periodic points, then g is topologically transitive and has dense periodic points.*

Proof. Let $h : X \to Y$ be the conjugacy between f and g. We will first show that conjugacy preserves transitivity. Take two open sets $U, V \subset Y$. Because h is continuous, $h^{-1}(U), h^{-1}(V)$ are open sets in X. By transitivity of f and Theorem 2.2, there exists $x \in h^{-1}(U)$ and $k \geq 0$ such that $f^k(x) \in h^{-1}(V)$. Since $g \circ h = h \circ f$, we have $g = h \circ f \circ h^{-1}$ and $g^k = h \circ f^k \circ h^{-1}$. Consider $h(x) \in U$:

$$g^k(h(x)) = h \circ f^k \circ h^{-1}(h(x)) = h \circ f^k(x).$$

Because h is invertible and $f^k(x) \in h^{-1}(V)$, we have $g^k(h(x)) = h(f^k(x)) \in V$. Thus for any open sets $U, V \subset Y$, there exists a $k \geq 0$ such that $g^k(U) \cap V \neq \emptyset$. By Theorem 2.2, g is transitive.

We will now show that conjugacy preserves dense periodic orbits. Take $y \in Y$ and $\varepsilon > 0$. Let U be an open ball of radius ε around y. Consider $h^{-1}(U)$, which is an open set containing $h^{-1}(y)$. Take $B \subset h^{-1}(U)$, a ball of radius $\varepsilon' > 0$ around $h^{-1}(y)$. Because periodic points of f are dense in X, there exist $x \in B$ and $k \geq 0$ such that $f^k(x) = x$. Consider $h(x) \in U \subset Y$:

$$g^k(h(x)) = h \circ f^k \circ h^{-1}(h(x)) = h(f^k(x)) = h(x).$$

Thus $h(x)$ is a periodic point for g with period k such that $d(y, h(x)) < \varepsilon$. We have now shown that g has dense periodic orbits. □

It may be easier to prove that a map is chaotic by showing topological conjugacy to a simpler chaotic map rather than by directly showing that the original map is chaotic. However, while topological conjugacy necessarily preserves dense periodic orbits and transitivity due to these being topological properties, conjugacy doesn't necessarily preserve sensitivity, which is a metric property. It is important to know when sensitivity is preserved to avoid making incorrect assumptions about maps. The following theorem describes when conjugacy preserves sensitivity.

Theorem 2.8. *Suppose $f : X \to X$ is topologically conjugate to $g : Y \to Y$ where X, Y are compact metric spaces. If g has sensitive dependence on initial conditions on Y, then f has sensitive dependence on initial conditions on X.*

Proof. Let $r > 0$ be the constant guaranteed by the sensitivity of g, so for any $x \in Y$ and $\varepsilon > 0$ there exists some $y \in Y$ with $d(x, y) < \varepsilon$ and $k \geq 0$ such that $d(f^k(x), f^k(y)) \geq r$. Let $h : X \to Y$ be the conjugacy between f and g. Because h is continuous and X is compact, h must be uniformly continuous on X. Thus there exists a $\delta > 0$ such that if $d(p, q) < \delta$ for $p, q \in X$, then $d(h(p), h(q)) < r$. This implies that if $d(p, q) \geq r$ in Y then $d(h^{-1}(p), h^{-1}(q)) \geq \delta$ in X. Take some $x \in X$ and $\varepsilon > 0$. By the continuity of h^{-1}, there exists $\varepsilon' > 0$ such that if $q \in Y$ and $d(q, h(x)) < \varepsilon'$, then $d(h^{-1}(q), x) < \varepsilon$. Fix some $q \in Y$ such that $d(q, h(x)) < \varepsilon'$. From the previous sentence, we know that $d(h^{-1}(q), x) < \varepsilon$ in X. Due to the sensitivity of g there must be some $k \geq 0$ such that $d(g^k(q), g^k(h(x))) \geq r$, which implies that $d(h^{-1}(g^k(q)), h^{-1}(g^k(h(x)))) \geq \delta$. Because $g = h \circ f \circ h^{-1}$, we have $g^k = h \circ f^k \circ h^{-1}$. Thus, $h^{-1}(g^k(q)) = f^k(h^{-1}(q))$ and $h^{-1}(g^k(h(x))) = f^k(x)$. We have found a constant $\delta > 0$ such that there exists a $h^{-1}(q) \in X$ where $d(h^{-1}(q), x) < \varepsilon$ and a $k \geq 0$ such that $d(f^k(h^{-1}(q)), f^k(x)) \geq \delta$, which means that f has sensitive dependence on initial conditions on X. □

2.1. **The quadratic map.** We will now return to our earlier example of the quadratic map. The quadratic map is a surprising example of how chaotic behavior can occur in a relatively simple map on $\mathbb{R}$.

The quadratic map (also sometimes called the logistic map) is a family of functions of the form

$$f(x) = \mu x(1-x).$$

Different values of μ will lead to drastically different behavior of iterates of f. For $\mu = 4$, the logistic map exhibits chaotic behavior on its invariant set, the interval $[0,1]$. To prove this, we will first need to define symbolic dynamics as a tool for studying more complicated functions.

Definition 2.9. Let Σ_p^+ be the space of sequences where each term is an integer from the set $\{1,2,\ldots,p\}$. The space Σ_p^+ can also be viewed as the space of functions from $\mathbb{N}$ into the set $\{1,2,\ldots,p\}$. We can define a metric on Σ_p^+ with

$$d(\mathbf{s},\mathbf{t}) = \sum_{k=0}^{\infty} \frac{\delta(s_k,t_k)}{2^k}$$

where $\mathbf{s} = (s_0 s_1 \ldots)$ and $\mathbf{t} = (t_0 t_1 \ldots)$. The value $\delta(s_k,t_k)$ is defined as

$$\delta(s_k,t_k) = \begin{cases} 0 & s_k = t_k \\ 1 & s_k \neq t_k. \end{cases}$$

Note that the distance between two points $\mathbf{s},\mathbf{t} \in \Sigma_p^+$ depends on how many corresponding terms they have. Since a difference in s_k, t_k for a smaller k affects the distance more than a difference in s_k, t_k for larger k, a smaller distance between two points means that they correspond in a greater number of initial terms.

We also define a shift map σ on Σ_p^+ with $\sigma(\mathbf{s}) = \mathbf{t}$, where $t_k = s_{k+1}$. This map is called the shift map because it shifts the sequence over by one term.

One can check that the shift map is continuous. Take $\mathbf{s} \in \Sigma_p^+$ and $\varepsilon > 0$. Take $n \geq 0$ such that $\frac{p-1}{2^n} < \varepsilon$ and let $\delta = \frac{p-1}{2^{n+1}}$. If $\mathbf{t} \in \Sigma_p^+$ is such that $d(\mathbf{s},\mathbf{t}) < \delta$, then the first $n+2$ terms of $\mathbf{s}$ and $\mathbf{t}$ must be the same. This is because if $s_k - t_k \neq 0$ for any $k \leq n+1$, then $d(\mathbf{s},\mathbf{t}) = \sum_{k=0}^{\infty} \frac{\delta(s_k,t_k)}{2^k} \geq \frac{1}{2^k} \geq \frac{1}{2^{n+1}}$. Thus if the first $n+2$ terms of $\mathbf{s}$ and $\mathbf{t}$ align, then the first $n+1$ terms of $\sigma(\mathbf{s})$ and $\sigma(\mathbf{t})$ will align, which means that

$$\begin{aligned} d(\sigma(\mathbf{s}),\sigma(\mathbf{t})) &= \sum_{k=0}^{\infty} \frac{\delta(\sigma(s_k),\sigma(t_k))}{2^k} \\ &= \sum_{k=0}^{n} \frac{\delta(\sigma(s_k),\sigma(t_k))}{2^k} + \sum_{k=n+1}^{\infty} \frac{\delta(\sigma(s_k),\sigma(t_k))}{2^k} \\ &\leq 0 + \sum_{k=n+1}^{\infty} \frac{p-1}{2^k} \\ &= \frac{p-1}{2^n} < \varepsilon. \end{aligned}$$

This shows that σ is continuous. The space Σ_p^+ together with the shift map σ is called the *symbol space on p symbols*.

In order to use the shift map to show that the quadratic map is chaotic, we need to establish topological conjugacy between the quadratic map and the shift map. To do so, we define the following map.

Definition 2.10. The *itinerary map* $S : \Lambda \to \Sigma_2^+$ on the invariant set described in Example 1.6 for the quadratic map f_μ for $\mu > 2+\sqrt{5}$ is defined as $S(x) = (s_0 s_1 \ldots)$, where $s_k = 1$ if $f_\mu^k \in I_0$ and $s_k = 2$ if $f_\mu^k \in I_1$. The sequence $S(x)$ is called the *itinerary* of x.

For a definition of I_0 and I_1 see Example 1.6. Thus the itinerary map provides a correspondence between points in the invariant set of the quadratic map and points in σ_2^+ by tracking the orbit of points. Note that $S \circ f_\mu = \sigma \circ S$.

Theorem 2.11. *For $\mu > 2 + \sqrt{5}$, the itinerary map $S : \Lambda \to \Sigma_2^+$ is a homeomorphism. Thus f_μ is topologically conjugate to σ.*

Proof. We will first show that S is injective. Suppose $S(x) = S(y)$ for $x, y \in \Lambda$. This means that $f_\mu^n(x)$ and $f_\mu^n(y)$ are both in either I_0 or I_1 for all n. Without loss of generality suppose $f_\mu^n(x) \leq f_\mu^n(y)$ for some n. Since f_μ is injective on both I_0 and I_1, this means that for all $q \in [f_\mu^n(x), f_\mu^n(y)]$, $f_\mu(q)$ is also on the same side of $\frac{1}{2}$ as $f_\mu^{n+1}(x), f_\mu^{n+1}(y)$. This means that $q \in \Lambda$ for all $q \in [f_\mu^n(x), f_\mu^n(y)]$. However, if $x \neq y$, this contradicts the fact that Λ is totally disconnected. Thus, $x = y$ and S is injective.

We will now show that S is surjective. Take $\mathbf{s} = (s_0 s_1 s_2 \ldots) \in \Sigma_2^+$. We want to find some $x \in \Lambda$ such that $S(x) = \mathbf{s}$. Define

$$I_{s_0 s_1 \ldots s_n} = \{x \in I \mid x \in I_{s_0}, f_\mu(x) \in I_{s_1}, \ldots, f_\mu^n(x) \in I_{s_n}\}.$$

Note that I_{s_0} is a non-empty closed interval, and $I_{s_0 s_1}$ is a single non-empty closed interval such that $I_{s_0 s_1} \subset I_{s_0}$, since each I_j has one closed interval that maps to I_0 and one closed interval that maps to I_1. By induction, $I_{s_0 s_1 \ldots s_n}$ is a series of non-empty nested closed intervals as n goes to infinity. Thus

$$\bigcap_{n \geq 0} I_{s_0 s_1 \ldots s_n}$$

is non-empty, so there exists some $x \in \Lambda$ such that $S(x) = \mathbf{s}$. By injectivity of S, this x is unique.

We will now show that S is continuous. Take $x \in \Lambda$ and let $S(x) = (s_0 s_1 s_2 \ldots)$. Take $\varepsilon > 0$. Take $n \geq 0$ such that $\frac{1}{2^n} < \varepsilon$. Consider the closed interval $I_{s_0 s_1 \ldots s_n}$, as defined above. Take δ to be the shorter distance between x and the edge of $I_{s_0 s_1 \ldots s_n}$, so if $|x - y| < \delta$ for some $y \in \Lambda$, then $y \in I_{s_0 s_1 \ldots s_n}$. This implies that $S(y)$ agrees with $S(x)$ in the first $n+1$ terms, so $d(S(x), S(y)) \leq \frac{1}{2^n} < \varepsilon$. We have now shown that S is continuous.

Lastly, we want to show that $S^{-1} : \Sigma_2^+ \to \Lambda$ is continuous. We know the inverse exists because S is injective and surjective, as shown earlier. Take $\mathbf{s} = (s_0 s_1 s_2 \ldots) \in \Sigma_2^+$ and $\varepsilon > 0$. Find $n \geq 0$ such that the length of $I_{s_0 s_1 \ldots s_n}$ is less than ε (as $n \to \infty$, the length of $I_{s_0 s_1 \ldots s_n}$ goes to 0). Let $\delta = \frac{1}{2^n}$. For any $\mathbf{t} \in \Sigma_2^+$ such that $d(\mathbf{s}, \mathbf{t}) < \delta$, the terms s_k, t_k are equal for $k \leq n$. This means that $S^{-1}(\mathbf{t}) \in I_{s_0 s_1 \ldots s_n}$, so $d(S^{-1}(\mathbf{s}), S^{-1}(\mathbf{t})) < \varepsilon$. Thus S^{-1} is continuous, concluding the proof that S is a homeomorphism. □

We will now prove that the quadratic map is chaotic.

Theorem 2.12. *The quadratic map f_μ is chaotic on Λ for $\mu > 2 + \sqrt{5}$.*

Proof. We first show that f_μ has dense periodic orbits by showing that the shift map σ has dense periodic orbits. Take a point $\mathbf{s} = (s_0 s_1 \ldots) \in \Sigma_2^+$. Define the sequence

$a_n = (s_0 s_1 \dots s_n s_0 s_1 \dots s_n s_0 s_1 \dots s_n \dots)$, where each element a_n is a sequence that repeats the first $n+1$ terms of $\mathbf{s}$. Each a_n is periodic with respect to σ because the removing the first n terms of a_n will return us to a_n. Because the first n terms of a_n and $\mathbf{s}$ are the same, we have that $d(a_n, \mathbf{s}) \leq \frac{1}{2^n}$. Thus as n approaches infinity, the periodic points a_n get arbitrarily close to $\mathbf{s}$. We've now shown that σ has dense periodic orbits. By topological conjugacy to σ, f_μ must also have dense periodic orbits.

Next we show that f_μ is transitive by showing that σ has a point whose orbit is dense in Σ_2^+. Consider the point

$$\mathbf{s} = (0\ 1\ 00\ 01\ 10\ 11\ 000\ 001\ \dots).$$

This point is constructed by first listing all "blocks" of length 1 (0 and 1), then all "blocks" of length 2 (00, 01, 10, 11), and so on. The order of blocks of length k doesn't matter, as long as we first start with blocks of length 1, then blocks of length 2, and so on. For any $\mathbf{t} \in \Sigma_2^+$, some iterate $\sigma^k(\mathbf{s})$ agrees with $\mathbf{t}$ in the first n terms. Thus $\sigma^k(\mathbf{s})$ gets arbitrarily close to any $t \in \Sigma_2^+$. In order to get closer to $\mathbf{t}$, we just need an iterate $\sigma^k(\mathbf{s})$ that agrees with $\mathbf{t}$ in a greater number of initial terms. As discussed earlier, if $s_k = t_k$ for $k \leq n$ for some $\mathbf{s}, \mathbf{t} \in \Sigma_2^+$, then $d(\mathbf{s}, \mathbf{t}) \leq \frac{1}{2^n}$. Thus, σ is transitive on Σ_2^+, which means that f_μ is transitive on Λ.

Finally, we will show that f_μ has sensitive dependence on initial conditions on Λ. Let $r > 0$ be less than the width of the gap between I_0 and I_1. Take any $x, y \in \Lambda$ such that $x \neq y$ (we know Λ contains at least two points because $0, 1 \in \Lambda$). Since $x \neq y$, we must have that $S(x) \neq S(y)$ (the itineraries of x and y differ in at least one spot). Suppose $S(x)$ and $S(y)$ first differ in the nth term. Because of how S was defined, this means that either $S(x) \in I_0$ and $S(y) \in I_1$ or $S(x) \in I_1$ and $S(y) \in I_0$. In either case, $|f_\mu^n(x) - f_\mu^n(y)| > r$. This concludes our proof that f_μ is chaotic on the invariant set on $[0, 1]$ for $\mu > 2 + \sqrt{5}$. $\square$

2.2. **The Horseshoe Map.** So far we have limited our discussion of dynamics and chaos to one dimension. The following example is a relatively simple example of chaos in two dimensions that has a lot of similarities to the quadratic map. Once again, symbolic dynamics will be used to understand the behavior of this map.

To consider the dynamics of the Horseshoe Map, we will first need to construct the region on which it acts. Let D be a region consisting of a square D_1 of side length 1 with two semicircles D_2, D_3 at either end, as shown in Figure 3.

The map $F : D \to D$ acts by contracting D horizontally by a factor of $\delta < \frac{1}{2}$, then stretching D vertically by a factor of $1/\delta$. The region D is then folded in half, causing it to resemble a horseshoe. As such, $F(D) \subset D$ and F is one-to-one, but not onto.

Let us now consider the effects of this map. The two semicircles D_2 and D_3 are mapped back into D_2. The preimage of the square D_1 consists of two horizontal rectangles, V_0 and V_1, which have a width of δ and a length of 1. The two rectangles V_0 and V_1 are mapped linearly onto two vertical rectangles, which also have a width of δ and a length of 1. The rectangles $F(V_0)$ and $F(V_1)$ are equal to the region $F(D_1) \cap D_1$ (see Figure 4).

We will need the following theorem to consider the dynamics of F on the semicircle D_2.

Theorem 2.13 (Contraction Mapping Theorem)**.** *Let (X, d) be a complete metric space and let $T : X \to X$ be a contraction, which means that there exists a $c \in \mathbb{R}$,*

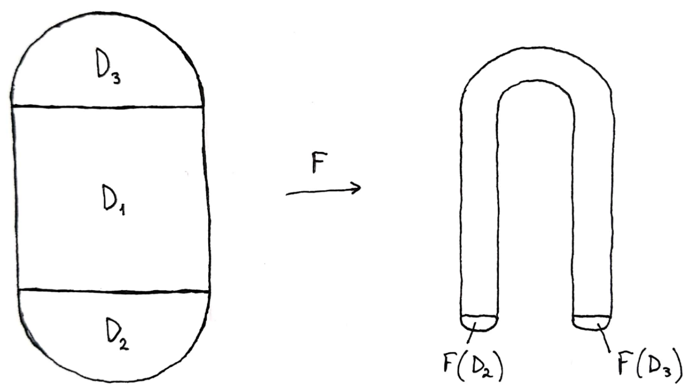

FIGURE 3. The first iteration of the Horseshoe Map

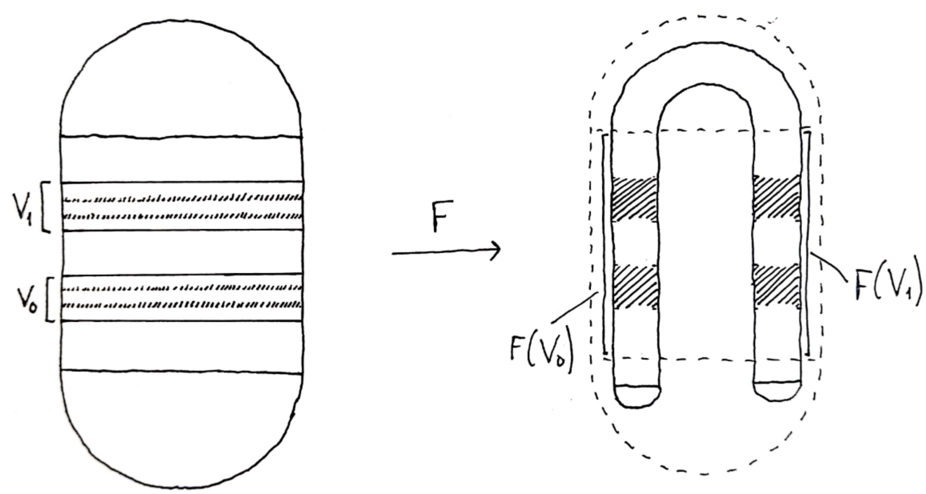

FIGURE 4. The pre-image of D_1 and the set $\{x \in D_1 \mid F^2(x) \in D_1\}$

$0 \leq c < 1$, *such that* $d(T(x), T(y)) \leq cd(x, y)$ *for all* $x, y \in X$. *Then there exists a unique fixed point of* T *in* X.

For a proof of this theorem, see Section 5.2 of [4].

We see that T is a contraction on D_2 because the image of D_2 is contained entirely in D_2. By the Contraction Mapping Theorem, there exists a unique fixed point p of F in D_2. It follows that for all $q \in D_2$,

$$\lim_{n\to\infty} F^n(q) = p.$$

Similarly, because $F(D_3) \subset D_2$, the orbits of all points in D_3 also approach p. Points in the square D_1 which are not in the invariant set of D behave similarly. If for some $q \in D_1$ there exists a $k > 0$ such that $F^k(q) \notin D_1$, then $F^k(q)$ must be in either D_2 or D_3. Then, as before, $\lim_{n\to\infty} F^n(q) = p$.

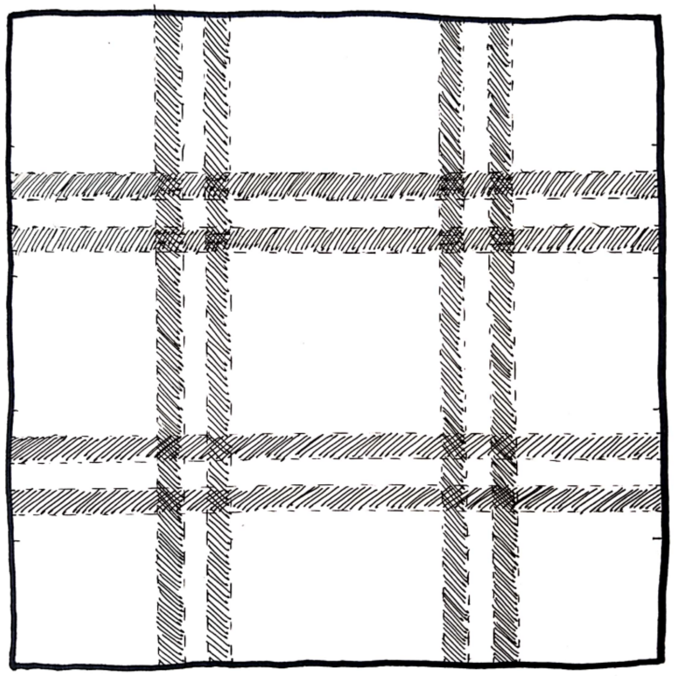

FIGURE 5. The invariant set of the Horseshoe Map on D_1

Thus to understand the behavior of all points in D under F it remains for us to understand the points whose orbits remain in D_1 for all iterations of F (the invariant set of D_1). Let $\Lambda = \{q \in D_1 \mid F^n(q) \in D_1 \text{ for all } k \in \mathbb{Z}\}$. Here we consider both the forward and backward orbits of points in the invariant set because F is invertible, unlike the quadratic map. To describe the invariant set Λ, we will first consider the structure of the following two sets:

$$\Lambda_+ = \{q \in D_1 \mid F^n(q) \in D_1 \text{ for all } k \geq 0\},$$

$$\Lambda_- = \{q \in D_1 \mid F^{-n}(q) \in D_1 \text{ for all } k < 0\}.$$

As we mentioned previously, the preimage of D_1 is the two horizontal rectangles V_0 and V_1. Thus $V_0 \cup V_1 = \{q \in D_1 \mid F(q) \in D_1\}$. If we apply another iteration of F, we see that there are two intervals in each of V_0 and V_1 such that $F^2(q) \in D_1$. Thus the shaded portion of Figure 4 is equal to the set $\{q \in D_1 \mid F^2(q) \in D_1\}$. Continuing this process, we see that Λ_+ is formed by taking three "slices" out of the set with each iteration of F. Thus in each V_i, Λ_+ is the product of a Cantor set and a horizontal interval. Analogously, Λ_- is the product of a Cantor set and a vertical interval. The invariant set Λ is then formed by taking the intersection of Λ_+ and Λ_-. See Figure 5 for a representation of Λ.

Just as with our analysis of the quadratic map, we can use symbolic dynamics to analyze the Horseshoe Map. Because the Horseshoe Map is now two-dimensional instead of one-dimensional, we will need a two-sided itinerary for each point. However, the basic premise for the itinerary map for the horseshoe map is the same as that for the quadratic map. The itinerary map for the horseshoe map is also defined by which region the orbit of each point lies in (this map is two-sided because it considers both the forward and backward orbits).

Definition 2.14. Let Σ_2 be the sequence space

$$\Sigma_2 = \{\mathbf{s} = (\dots s_{-2}s_{-1}.s_0s_1s_2\dots) \mid s_i = 1 \text{ or } 2\}.$$

The metric on this space is the same as that on Σ_2^+: $d(\mathbf{s},\mathbf{t}) = \sum_{k=0}^{\infty} \frac{|s_k - t_k|}{2^{|k|}}$. Let the itinerary map $S : \Lambda \to \Sigma_2$ be given by $S(p) = (\dots s_{-2}s_{-1}.s_0s_1s_2\dots)$ for $p \in \Lambda$, where $s_k = 1$ if $F^k(p) \in V_0$ and $s_k = 2$ if $F^k(p) \in V_1$ for $k \geq 0$. For $k < 0$, $s_{-k} = 1$ if $F^{-k}(p) \in V_0$ and $s_{-k} = 2$ if $F^{-k}(p) \in V_1$. The shift map on Σ_2 is also defined analogously to that on Σ_2^+: $\sigma(\dots s_{-2}s_{-1}.s_0s_1s_2\dots) = (\dots s_{-2}s_{-1}s_0.s_1s_2\dots)$.

Just as with the metric on Σ_2^+, with the metric on Σ_2 a smaller distance between two points $\mathbf{s},\mathbf{t}$ corresponds to a greater number of corresponding terms s_k, t_k for lower $|k|$ (the terms that are closer to the center of the sequence). Unlike the shift map on Σ_2^+, the shift map σ on $\sum_2$ has an inverse (shifting the sequence to the right instead of the left). Thus one can check that σ is a homeomorphism on Σ_2. The proof of this is nearly identical to the proof that σ is continuous on Σ_2^+, so it will be omitted. The itinerary map $S : \Lambda \to \Sigma_2$ gives a topological conjugacy between $F : \Lambda \to \Lambda$ and $\sigma : \Sigma_2 \to \Sigma_2$ (the proof of which will be omitted because it is nearly identical to the proof in Theorem 2.11).

We are now ready to show that the Horseshoe Map is chaotic. This proof is very similar to the proof that the quadratic map is chaotic.

Theorem 2.15. *The Horseshoe Map $F : D \to D$ is chaotic on Λ.*

Proof. We will use topological conjugacy to σ to show that F has dense periodic orbits. Fix some $\mathbf{q} \in \Sigma_2$ and $\delta > 0$. Find a value of $k \in \mathbb{N}$ such that $1/2^{k-1} < \delta$. Let $\mathbf{s} = (\dots s_{-2}s_{-1}.s_0s_1s_2\dots)$, where $s_i = q_i$ for $|i| \leq k$. This sequence, s_i for $|i| \leq k$, then repeats to the left of s_{-k} and to the right of s_k to make $\mathbf{s}$ a periodic point. Because the k terms to the left and right of s_0 and q_0 correspond, we have $d(\mathbf{s},\mathbf{q}) \leq \frac{1}{2^{k-1}} < \delta$. Thus there is a periodic point arbitrarily close to any $\mathbf{q} \in \Sigma_2$. We have now shown that F has dense periodic orbits.

In proving transitivity of the quadratic map, we constructed a point whose orbit was dense in Σ_2^+ under the shift map. In fact, the same point works to show transitivity of $\sigma : \Sigma_2 \to \Sigma_2$. Let $\mathbf{s} \in \Sigma_2$ be such that

$$\mathbf{s} = (\dots s_{-2}s_{-1}.\ 0\ 1\ 00\ 01\ 10\ 11\ 000\ 001\ \dots).$$

The terms s_i for $i < 0$ are arbitrary. Under iterations of σ, any finite sequence of 0s and 1s will eventually be centered around s_0, bringing us arbitrarily close to any point in Σ_2. Thus σ is transitive on Σ_2, showing transitivity of F on Λ.

Finally, we will show that F has sensitive dependence on initial conditions. Let $r < 1$. Take $\mathbf{p} \in \Sigma_2$ and $\varepsilon > 0$. Take $k \in \mathbb{N}$ such that $1/2^{k-1} < \varepsilon$. Let $\mathbf{q} \in \Sigma_2$ be such that $q_i = p_i$ for $|i| \leq k$, and $q_i \neq p_i$ for $i = k+1$. Then $d(\mathbf{p},\mathbf{q}) < 1/2^{k-1} < \varepsilon$ and $d(\sigma^{k+1}(\mathbf{p}), \sigma^{k+1}(\mathbf{q})) \geq 1 > r$. Thus σ has sensitive dependence on initial conditions on Σ_2. By Theorem 2.8, because both Σ_2 and Λ are compact, F has sensitive dependence on initial conditions on Λ. Thus F is chaotic on Λ. $\square$

Because the Horseshoe Map is two-dimensional as opposed to one-dimensional, some new interesting behaviors arise in this map. For more information, see Chapter 28 of [3].

2.3. **Chaos on non-compact sets.** So far we have only considered chaotic systems on compact sets. As can be seen in Theorem 2.8, sensitive dependence on initial conditions is necessarily conserved by topological conjugacy only on compact sets, not non-compact sets. In addition, many of the chaotic systems that are studied arise from natural phenomena. As such, they are often only defined on compact sets or only have interesting behavior on compact sets. Because of the finite energy of such systems, their behavior often becomes more predictable and less interesting outside of some compact set. The following is an example of a function that is chaotic on a non-compact set.

Theorem 2.16. *The function* $T(x) = \tan(x)$ *is chaotic on a non-compact subset of* $\mathbb{R}$*.*

Proof. We will first show that any interval I is eventually mapped to the entire real line. Take some interval $I \subset \mathbb{R}$ and choose an interval $[x, y] \subset I$ such that $x, y \in [\frac{\pi}{2} + k\pi, \frac{\pi}{2} + (k+1)\pi]$ for some $k \in \mathbb{Z}$. In other words, both x and y lie in one period of tangent. Since $\tan'(x) \geq 1$ for all x where it is defined and is only equal to 1 at $k\pi$, we have that

$$\frac{\tan(x) - \tan(y)}{x - y} > 1.$$

Choose some $\lambda > 1$ such that

$$\frac{\tan(x) - \tan(y)}{x - y} > \lambda.$$

Thus we have $|T(x) - T(y)| > \lambda|x - y|$. This also implies that $|T^2(x) - T^2(y)| > \lambda|T(x) - T(y)| > \lambda^2|x - y|$. Inductively, we get

$$|T^n(x) - T^n(y)| > \lambda^n|x - y|.$$

Choose $n \geq 0$ such that $\lambda^n|x - y| > \pi$ (suppose for now that $T^k(x), T^k(y)$ are defined for all $k \leq n + 1$). Thus we get $|T^n(x) - T^n(y)| > \lambda^n|x - y| > \pi$, which means that the interval $J = [T^n(x), T^n(y)]$ covers an entire period of tangent. Iterating the function once more, we get that $T(J) = \mathbb{R}$. Thus, $T^{n+1}(I) = \mathbb{R}$.

What if $T^k(x)$ or $T^k(y)$ is not defined for some $k \leq n + 1$? Without loss of generality, suppose $T^k(x)$ is not defined for some $k \leq n+1$. This means that $T^k(q)$ approaches positive or negative infinity as q approaches x. It follows that $T^k(q)$ moves through infinitely many intervals of the form $[\frac{\pi}{2} + m\pi, \frac{\pi}{2} + (m+1)\pi]$ for $m \in \mathbb{Z}$. After another iteration of T, each of these intervals is mapped to $\mathbb{R}$, so T^{k+1} covers $\mathbb{R}$ infinitely many times when close to x. Thus, $T^{k+1}(x) = \mathbb{R}$ in this case as well.

Transitivity follows closely from the fact that any interval is eventually mapped to $\mathbb{R}$. Take two open sets $U, V \subset \mathbb{R}$. The set U cannot be a union of isolated points, otherwise it would be closed, so it must contain some interval I. We showed that there must be some n such that $T^n(I) = \mathbb{R}$, so $T^n(U) = \mathbb{R}$, which means that $T^n(U) \cap V \neq \emptyset$. By Theorem 2.2, T is transitive on $\mathbb{R}$. Dense periodic orbits also follow from the fact that any interval eventually maps to $\mathbb{R}$. Fix some $x \in \mathbb{R}$ and $\varepsilon > 0$. Let $I = [x - \varepsilon/2, x + \varepsilon/2]$. As proved earlier, there exists some $n \geq 0$ such that $T^n(I) = \mathbb{R}$, which means there exists some $q \in I$ such that $T^n(q) = q$. Thus, q is a periodic point for T with period n such that $|x - q| < \varepsilon$. This proves the existence of dense periodic orbits on $\mathbb{R}$.

We will now show sensitivity of T. Let $r = 1$. Fix some $\varepsilon > 0$ and $x \in \mathbb{R}$ such that $T^n(x)$ exists for all n. Take some $y \in \mathbb{R}$ such that $|x - y| < \varepsilon$. As shown earlier, we can fix $\lambda > 1$ and $n \geq 0$ such that $|T^n(x) - T^n(y)| > \lambda^n |x - y| > 1$. If $T^k(y)$ does not exist for some $k \leq n$ (take k to be the smallest such value where $T^k(y)$ does not exist), then $T^k(q)$ approaches positive or negative infinity as q approaches y. Without loss of generality, assume $y > x$. Then there must exist q such that $x < q < y$ and $|T^k(x) - T^k(q)| > 1$. This shows sensitive dependence on initial conditions.

This concludes our proof that $T = \tan(x)$ is chaotic wherever $T^n(x)$ exists for all n. Since tangent is periodic, this set is unbounded and thus non-compact. $\square$

2.4. **Other definitions of chaos.** The definition given in Definition 2.4 was introduced by Devaney in 1989 [2]. However, Banks et al. proved in 1992 in a surprisingly short proof that transitivity and dense periodic orbits imply sensitive dependence on initial conditions [1]. It is interesting to note that they did not use any assumption of compactness, so transitivity and dense periodic orbits imply sensitivity even on a non-compact set. Does this contradict our earlier discussion and Theorem 2.8, where we stated that topological conjugacy always preserves transitivity and dense periodic orbits, but only preserves sensitivity on compact sets? No, the theorem proved by Banks et al. simply shows that sensitivity can be preserved by topological conjugacy on a non-compact set as long as both transitivity and dense periodic orbits are also present.

In addition to Devaney's definition of chaos being slightly redundant, other mathematicians have proposed different definitions. For example, the definition given by Robinson in [4] states that a system is chaotic as long as it is transitive and has sensitive dependence on initial conditions (but has no mention of dense periodic orbits). For an extended discussion of other definitions of chaos, see [2].

3. Quantitative measures of chaos

The definition of chaos given in the previous section is not quantitative. A dynamical system either is transitive or is not transitive on some invariant set, either is sensitive or is not sensitive, and either has dense periodic points or does not have dense periodic points. Thus, by the previously given definition, a dynamical system is either chaotic or non-chaotic, with no in-between. For two chaotic maps, is it possible to compare their chaoticity and determine that one map is "more" chaotic than the other? The following notions allow us to do that by providing quantitative measures of chaos. These measures can also serve as alternative definitions of chaos to the one given in the previous section.

3.1. **Liapunov Exponents.** One measure which is associated with sensitivity is the Liapunov exponent. More specifically, Liapunov exponents measure the rate at which nearby orbits move apart in a dynamical system.

Definition 3.1. Let $f : \mathbb{R} \to \mathbb{R}$ be a C^1 function. The Liapunov exponent of a point $x_0 \in \mathbb{R}$ is defined as

$$\begin{aligned}\lambda(x_0) &= \limsup_{n\to\infty} \frac{1}{n} \log(|(f^n)'(x_0)|)\\ &= \limsup_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \log(|f'(x_j)|)\end{aligned}$$

where $x_j = f^j(x_0)$.

Because the definition of a Liapunov exponent is based on the derivative of f^n, the Liapunov exponent is defined only when $(f^n)'$ is defined for all n. While this is the definition for only one dimension, there is an analagous (but more complicated) definition for higher dimensions. A positive Liapunov exponent at a point indicates that nearby orbits will eventually diverge exponentially, which means the system has sensitive dependence on initial conditions around that point. A negative Liapunov exponent, on the other hand, indicates that the orbit of the point (as well as orbits of points nearby) approach a fixed point or periodic orbit. This behavior is stable and, thus, non-chaotic.

In the following example, we calculate the Liapunov exponent for a function which we already proved to be sensitive.

Example 3.2. Let $f(x) = 2x$. We want to calculate the Liapunov exponent of f. We know that $f^k(x) = 2^k x$, so $(f^k)'(x) = 2^k$. Thus

$$|(f^k)'(x)|^{1/k} = (2^k)^{1/k} = 2.$$

This means that $\lambda(x) = \log(2) > 0$ for all $x \in \mathbb{R}$. As we can see, the Liapunov exponent is positive and the function has sensitive dependence on initial conditions, as proved earlier.

3.2. **Topological Entropy.** Another quantitative measure of chaos is topological entropy. Conceptually, topological entropy is a measure of how many "different orbits" there are for a map, where orbits are considered "different" if they are at least $\varepsilon > 0$ apart at some point.

Definition 3.3. Let $f : X \to X$ be a continuous map. Consider the distance

$$d_{n,f}(x, y) = \sup_{0\le j<n} d(f^j(x), f^j(y)).$$

A set $S \subset X$ is considered *(n, ε)-separated for f* if $d_{n,f} > \varepsilon$ for all points $x, y \in S$ such that $x \neq y$.

Define *the number of different orbits of length n* (as measured by ε) by

$$r(n, \varepsilon, f) = \max\{\#(S) \mid S \subset X \text{ is a } (n, \varepsilon)\text{-separated set for } f\},$$

where $\#(S)$ is the cardinality of S.

Next define

$$h(\varepsilon, f) = \limsup_{n\to\infty} \frac{\log(r(n, \varepsilon, f))}{n},$$

which measures the growth rate of $r(n, \varepsilon, f)$ as n increases.

Now, define the *topological entropy* of f as

$$h(f) = \lim_{\varepsilon\to 0} h(\varepsilon, f).$$

The following is an example of a calculation of topological entropy for a map.

Example 3.4. Let $f : S^1 \to S^1$ be the doubling map on the unit circle. Consider the unit circle as starting at 0 and ending at 1 after one rotation, so a point q on the unit circle is equal to $q = \frac{\theta}{2\pi}$, where θ is the corresponding angle in radians. An equivalent map is $F : \mathbb{R} \to \mathbb{R}$ where $F(x) = 2x \mod 1$. Let the distance between two points on S^1 be equal to the arc length between them divided by 2π. This is equivalent to the standard metric on $\mathbb{R}$ applied to F. We want to show that $h(f) = \log 2$. Take $\varepsilon > 0$. If we consider two points $x, y \in S^1$ as points in $\mathbb{R}$, $d(f^{n-1}(x), f^{n-1}(y)) \le \varepsilon$ if and only if $d(x, y) \le \frac{\varepsilon}{2^{n-1}}$. We can place at most $[\frac{2^{n-1}}{\varepsilon}]$ points spaced exactly $\frac{\varepsilon}{2^{n-1}}$ apart in $[0, 1)$ (where $[\frac{2^{n-1}}{\varepsilon}]$ is the integer part of $\frac{2^{n-1}}{\varepsilon}$). If these points are considered on the unit circle, the point closest to 1 is close to the point nearest to 0, so we can place at most $[\frac{2^{n-1}}{\varepsilon}] - 1$ points spaced exactly $\frac{\varepsilon}{2^{n-1}}$ apart in S^1. There must be at least one neighboring pair of points that are slightly more than $\frac{\varepsilon}{2^{n-1}}$ apart in S^1. Thus all of the points can be spaced apart slightly so that all neighboring points are more than $\frac{\varepsilon}{2^{n-1}}$ apart. Thus, for all neighboring points x, y we have $d(f^{n-1}(x), f^{n-1}(y)) > \varepsilon$, so all of the points are (n, ε)-separated. Thus,

$$r(n, \varepsilon, f) = [\frac{2^{n-1}}{\varepsilon}] - 1.$$

We can now consider $h(\varepsilon, f)$:

$$\begin{aligned} h(\varepsilon, f) &= \limsup_{n\to\infty} \frac{\log([\frac{2^{n-1}}{\varepsilon}] - 1)}{n} \\ &= \limsup_{n\to\infty} \frac{\log(\varepsilon^{-1}) + (n-1)\log(2)}{n} \\ &= \log(2). \end{aligned}$$

Since $h(\varepsilon, f) = \log(2)$ for any $\varepsilon > 0$, we have $h(f) = \log(2)$.

The following theorem sheds some light on what it means for a map to have an entropy of zero.

Theorem 3.5. *Let $f : X \to X$ be a continuous function on a compact metric space. If the set of non-wandering points of f, denoted as $\Omega(f)$, is a finite number of periodic orbits, then the topological entropy of f is zero, $h(f) = 0$.*

For a proof, see Section 8.1 of [4].

Topological entropy is a good quantitative measure of chaos because having a higher number of “different orbits” is correlated with chaotic behavior. In particular, sensitive dependence on initial conditions implies a greater entropy value because it guarantees that any point has an arbitrarily close point whose orbit diverges by some $r > 0$ at some point.

We have now considered two quantitative measures of chaos, topological entropy and the Liapunov exponent. This paper provides merely a brief overview of these concepts. For a more in-depth consideration of Liapunov exponents, topological entropy, and other measures of chaos, see Chapter VIII of [4].

4. Final Remarks

We have now considered multiple examples of chaotic functions that are chaotic and a few examples of functions that are not chaotic to develop an intuitive understanding of chaotic functions. We also discussed a few quantitative measures of chaos. But why study chaotic dynamical systems at all? As alluded to earlier, many chaotic dynamical systems arise from natural phenomena. The Horseshoe Map, for example, bears great similarity to the way that bakers knead dough by stretching it and folding it over itself. The weather is another example of a natural chaotic system. Even with current technological and mathematical tools, the weather cannot be accurately predicted more than two weeks in advance. The difficulty of studying chaotic systems is that despite them being entirely deterministic (their future state depends in a predictable way on the initial state), an inexact approximation of the initial state will not give you even a remotely accurate prediction of a future state. Because of computer rounding error and an inevitable degree of inaccuracy in measurements, we cannot have the exact initial conditions for a natural dynamical system. Thus it is important to know when our models can give us an accurate prediction of a system and when they cannot.

Acknowledgments

First and foremost, I'd like to thank my mentor, Meg Doucette, for providing guidance during the course of my research and helping me work through the material. I'd also like to thank my family for always supporting and believing in me. Finally, I'd like to thank Peter May for running the REU at UChicago.

References

[1] J. Banks, J. Brooks, G. Cairns, G. Davis and P. Stacey. On Devaney's Definition of Chaos. Mathematical Association of America. 1992.
[2] François Blanchard. Topological chaos: what may this mean? HAL. 2008.
[3] Robert Devaney. An Introduction to Chaotic Dynamical Systems. CRC Press. 2022.
[4] Clark Robinson. Dynamical systems: Stability, Symbolic Dynamics, and Chaos. CRC Press. 1995.