Chaos:
From Seeing to Believing

Tate Jacobson

Whitman Mathematics Department

Spring 2016

Contents

Introduction: Motivating the Study of Chaos

In popular culture, the term "chaos" is most often a generic stand-in for "disorder." It calls to mind countless headlines and over-the-top action movies—at the time of writing, searching for "chaos" on Google produces articles about recent Nascar wrecks, drama among the Kardashians, and "Captain American: Civil War."
These popular renderings of chaos as mere disorder fail to capture the central premise of the mathematical study of chaos: that chaos is not disorder, but rather *order in apparent disorder*. Before cracking into any rigorous definitions, let us take a naïve look at an example of this sort of disorderly order.

A Small Adjustment with Big Implications

Imagine that we are scientists tracking the size of a single population of rabbits over time. Furthermore, suppose that, unbeknownst to us, the size of that population perfectly follows the logistic growth model described by the function $f(x) = 2x(1-x)$, where x denotes the population (in millions) of rabbits at the beginning of a given year and $f(x)$ denotes the population by the end of that year.
Since the end of each year is the beginning of the next we see that we are dealing with a recursive process: for any integer n, if we let x_n denote the population at the end of the nth year, then we see that $x_n = f(x_{n-1})$.
Suppose, then, that $x_0 = 0.01$, where x_0 denotes the initial population. We find that $x_1 = f(x_0) = 2(0.01)(0.99) = 0.0198$ and that $x_2 = f(x_1) = 2(0.0198)(.9802) = 0.03881592$. Repeating this process, we produce the following table and graph which relate n and x_n for the first few values of n:

n	x_n
0	0.01000000
1	0.01980000
2	0.03881592
3	0.07461848
4	0.13810113
5	0.23805842
6	0.36277322
7	0.46233762
8	0.49716309
9	0.49998390
10	0.49999999
11	0.50000000
12	0.50000000

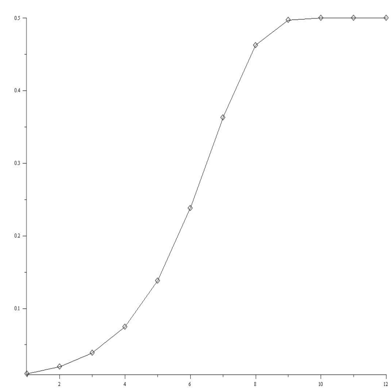


Figure 1: A sequence plot of $\{x_n\}$ under f.

Looking at this raw data we see that the population grows quickly and seems to level off at 0.5 (it is actually just approaching 0.5, but our calculators round off the tiny differences past a certain point). Intuitively, this is how we would expect a population to grow: exploding initially, then reaching a stable value as the scarcity of resources forces members of the population to compete with each other. In our above formulation we might redundantly call this "orderly" order. Since this behavior fits our intuition we take the data at face-value.

Now suppose that the size of the population instead follows the function $g(x) = 4x(1-x)$, and suppose again that $x_0 = .01$. Applying g a few times, we see:

n	x_n
0	0.01000000
1	0.03960000
2	0.15212736
3	0.51593851
4	0.99898386
5	0.00406043
6	0.01617577
7	0.06365646
8	0.23841726
9	0.72629788
10	0.79515708
11	0.65152919
12	0.90815562

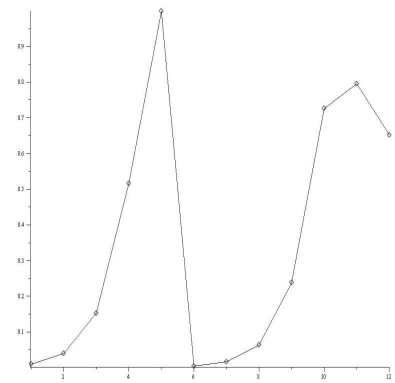


Figure 2: A sequence plot of $\{x_n\}$ under g.

Seeing this data, we will likely assume that some catastrophe occurred during the fifth year or that the population's growth is entirely random.
When we slightly change our input value, we find that g has another bizarre characteristic. Let $x_0 = .01$ as above and let $y_0 = .0099$. The following image shows the sequence plots of $\{x_n\}$ and $\{y_n\}$ up to $n = 25$, with the former plotted in red and the latter in blue:

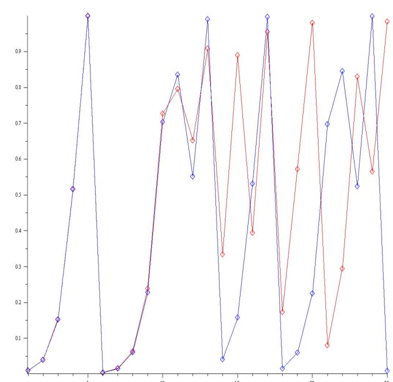


Figure 3: Overlaid sequence plots of $\{x_n\}$ and $\{y_n\}$ under g, where $x_0 = .01$ and $y_0 = .0099$.

While the plots of $\{x_n\}$ and $\{y_n\}$ cleave to each other at first, past $n = 14$ their behavior seems to completely diverge.
As readers, we know that a simple rule is guiding the growth of this population, yet our data seems to defy any predictable patterns. How do we account for this behavior?
Our experience tracking the behavior of $\{x_n\}$ and $\{y_n\}$ under g echoes meteorologist Edward Lorenz's experience in analyzing his primitive weather simulations in the early 1960s. Though he based his model on a handful of equations, he found that even the slightest changes in his initial values led to remarkably different outcomes. He initially assumed that something was wrong with

his computer or with his code, yet found that everything was in working order. As in our own experiment, the models themselves produced the seemingly random behavior. It was with this discovery of such disorderly order that Lorenz first sparked interest in what eventually became *Chaos Theory*. In observing the behavior of our seemingly bizarre logistic map, we too are beginning our journey into chaos in much the same way. (For a more complete and dramatic account of Lorenz's story and the origins of Chaos theory, see Gleick [4].)

Our Goals

In this paper we will characterize the differences between non-chaotic and chaotic dynamics. The first section will introduce *dynamical systems* in general, providing vocabulary and highlighting why f in our naïve example is so predictable. The second section will then provide a general definition of chaos and give an example of a chaotic dynamical system. The third and fourth sections will then explore a powerful method for proving that dynamical systems are chaotic: *topological conjugacy*. We will then close by using this method to prove that many logistic maps, like g, are chaotic.

As I am aiming this paper at a fairly general audience, I have included a lengthy appendix introducing *metric spaces* and *topology*, as a basic understanding of both is essential to some of the deeper results that I cover. If at any point you encounter an unfamiliar term or question an assumption in a proof, check the appendix for details.

1 Dynamical Systems

Before we discuss any particulars, we need to establish a broad definition of the sorts of systems we will be talking about.

Definition 1.0.1. A **dynamical system** consists of a set of possible states along with a rule that determines the present state as a function of past states.

It is important to note at this point that the dynamical systems which we will be exploring have a few restrictions. First, they are *deterministic* rather than *stochastic*. Simply put, this means that our rule will always return the same output for a given input, meaning that it is in no way random. Second, they are *discrete*, as opposed to *continuous*, dynamical systems.
In our first example the set of states was the set of values the population of rabbits could take on and the rule which determined the present state as a function of past states was f. We used the output from one application of our rule as the input value for the next application of our rule. We call each re-application of our rule an *iteration*. As is clear from our example, with each iteration we are composing our rule with itself. To streamline our notation we write the second iterate of f, namely $f(f(x))$, as $f^2(x)$, the third iterate of f, namely $f(f(f(x)))$, as $f^3(x)$, and the nth iterate of f as $f^n(x)$.

1.1 Orbits

Definition 1.1.1. Given a map f and a point x in the domain of f, we call the set of points $\{x, f(x), f^2(x), \ldots\}$ the **forward orbit** of x under f and denote it $O^+(x)$. As x is the starting point of the orbit, we call it the **initial value** or **seed** of the orbit.

Returning to our examples, we see that our first table provides the forward orbit of .01 under f and that our second table provides the forward orbit of .01 under g.
While all of this vocabulary is helpful, a visual representation of orbits helps solidify the concept. We call these diagrams *cobweb plots* and construct them as follows:
Let x_0 be the seed of our orbit. In our plot we graph both our function $f(x)$ and the line $g(x) = x$. With these guidelines, we first trace a line, in red, from $(x_0, 0)$ to $(x_0, f(x_0))$, then from $(x_0, f(x_0))$ to $(f(x_0), f(x_0))$ (this is where plotting $g(x) = x$ is useful). From there we can trace a line to $(f(x_0), f^2(x_0))$, then to $(f^2(x_0), f^2(x_0))$, and so on. With these plots, we can find $f^n(x)$ for any n and, perhaps more importantly, see how the orbit of x got to $f^n(x)$.
Figure 4 provides an easy-to-follow, albeit somewhat bland, example of a cobweb plot.

With a basic understanding of cobweb plots, we can start to visual the behavior of $f(x) = 2x(1-x)$ and $g(x) = 4x(1-x)$ for larger values of n.

Figure 5 and Figure 6 show us that the orbit of $x_0 = .01$ under f continues to approach .5. Figures 7, 8, 9, and 10 reveal that the orbit of $x_0 = .01$ under g seems to travel all over the interval $[0, 1]$.

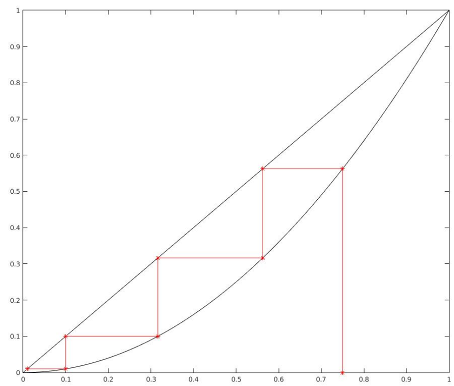


Figure 4: The cobweb plot of $x_0 = .75$ under the map $f(x) = x^2$ up to 5 iterations.

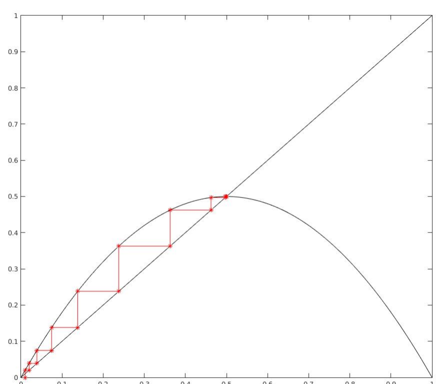


Figure 5: The cobweb plot of $x_0 = .01$ under the map $f(x) = 2x(1-x)$ up to 12 iterations.

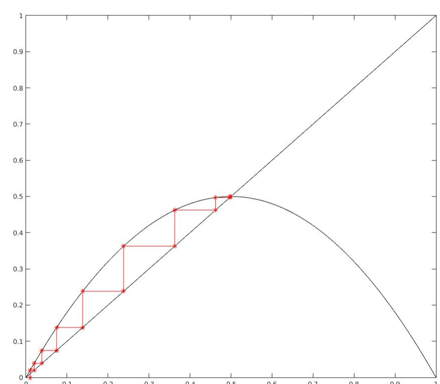


Figure 6: The cobweb plot of $x_0 = .01$ under the map $f(x) = 2x(1-x)$ up to 100 iterations. We see that nothing unexpected happens with the orbit of .01.

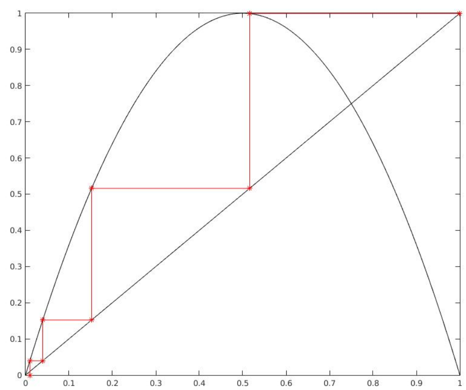


Figure 7: The cobweb plot of $x_0 = .01$ under the map $f(x) = 4x(1-x)$ up to 5 iterations. We see the sudden drop after the fifth iteration as we did in our table.

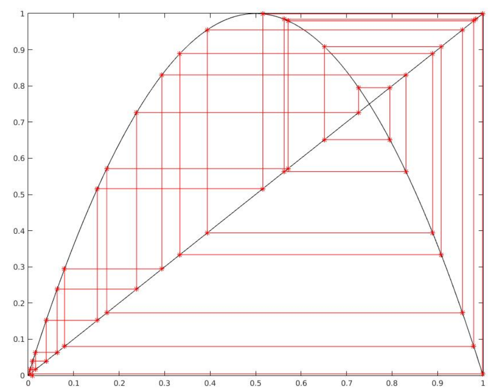


Figure 8: The cobweb plot of $x_0 = .01$ under the map $f(x) = 4x(1-x)$ up to 25 iterations. Clearly the orbit of .01 is covering a fair amount of the interval $[0, 1]$.

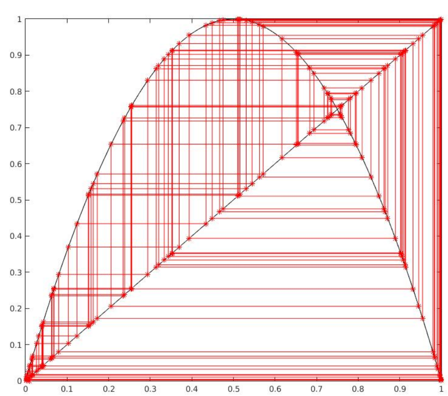


Figure 9: The cobweb plot of $x_0 = .01$ under the map $f(x) = 4x(1-x)$ up to 100 iterations. We see that the orbit of .01 continues to hit new points.

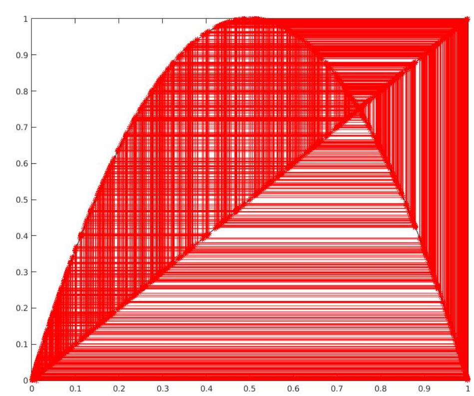


Figure 10: The cobweb plot of $x_0 = .01$ under the map $f(x) = 4x(1-x)$ up to 1000 iterations. The orbits of .01 is still covering new ground.

1.2 Types of Points in a Dynamical System

Now that we are familiar with what dynamical systems are, we might wonder whether we can determine how the orbits of certain points will behave. The most predictable points in a dynamical system are undoubtedly *fixed points*, followed closely by *periodic points*.

Definition 1.2.1. Given a point p in the domain of f, if $f(p) = p$ we call p a **fixed point** of the map f. We denote the set of fixed points of f by $Fix(f)$.

Example 1.2.2. We can rather easily find the fixed points of $f(x) = 2x(1-x)$. We simply need to solve the equation $f(p) = p$ for p:

$$\begin{aligned} 2p(1-p) &= p & (1)\\ 2p - 2p^2 &= p & (2)\\ p - 2p^2 &= 0 & (3)\\ p(1-2p) &= 0\ . & (4) \end{aligned}$$

Clearly $Fix(f) = \{0, \frac{1}{2}\}$.

Example 1.2.3. We can find the fixed points of $f(x) = 4x(1-x)$ in a similar manner. We see:

$$\begin{aligned} 4p(1-p) &= p & (5)\\ 4p - 4p^2 &= p & (6)\\ 3p - 4p^2 &= 0 & (7)\\ p(3-4p) &= 0\ , & (8) \end{aligned}$$

so that $Fix(f) = \{0, \frac{3}{4}\}$.

Pulling together a few concepts from calculus and this definition, we have the following theorem which provides conditions on f and its domain which guarantee that f has a fixed point.

Theorem 1.2.4. *If $f : [a,b] \to [a,b]$ is continuous, then f has at least one fixed point in $[a,b]$.*

Proof. Let $g(x) = f(x) - x$. As g is the difference of continuous functions, it is likewise continuous. We note that $a \leq f(a)$ and that $f(b) \leq b$ as $[a,b]$ is the range of f. If either $f(a) = a$ or $f(b) = b$ then we are already done. Suppose then that $a < f(a)$ and $f(b) < b$. Then $g(a) > 0$ and $g(b) < 0$, so the Intermediate Value Theorem provides that there exists a point c between a and b such that $g(c) = 0$. Therefore $f(c) = c$ and we are done. □

We can show that if f has one additional property, then f has a unique fixed point. Note that in the following discussion, f' denotes the derivative of f.

Theorem 1.2.5. *Let I be a closed interval. If $f : I \to I$ and $|f'(x)| < 1$ for all x in I, then there exists a unique fixed point for f in I.*

Proof. As f is differentiable on I it is continuous on I. Therefore our previous theorem guarantees that f has at least one fixed point on I. Suppose, then, that both x and y are fixed points and $x \neq y$. By the Mean Value Theorem, there exists a point c between x and y such that

$$f'(c) = \frac{f(y) - f(x)}{y - x} = \frac{y - x}{y - x} = 1$$

a clear contradiction of our assumption that $|f'(x)| < 1$ for all x in I. Thus $x = y$. □

While fixed points have a number of unique qualities, they ultimately fall into the much larger category of *periodic points*.

Definition 1.2.6. We say that a point x_0 is **periodic** of period n if $f^n(x_0) = x_0$ for some $n > 0$. We call the orbit of x_0 a **periodic orbit** or **cycle** in this case.
We denote the set of periodic points of period n under f by $Per_n(f)$ and set of all periodic points under f by $Per(f)$.

With this definition, we can see that $Fix(f) = Per_1(x)$.
It is important to note that the set of periodic points of period n might contain points with different periods. For example, if x_1 has period 6 and x_2 has period 4, then both are members of $Per_{12}(f)$.

Example 1.2.7. Let $f(x) = x^2 - 1$. We see that $x = -1$ is a periodic point of period 2 as $f^2(-1) = f(f(-1)) = f(0) = -1$.

Moreover, it is interesting to note that if x is periodic of period n under f, then for every $i \in \{1, \ldots, n-1\}$ $f^i(x)$ is likewise periodic of period n. We see that this is the case as

$$f^n(f^i(x)) = f^{n+i}(x) = f^{i+n}(x) = f^i(x).$$

Therefore, as $x = -1$ is periodic of period 2 under $f(x) = x^2 - 1$ and $f(-1) = 0$, we know that 0 is periodic of period 2 under f as well.
As you can imagine, the cobweb plot of the forward orbit of a fixed point is not too terribly exciting. The cobweb plot of the forward orbit a periodic point, however, can prove a little more informative. Figures 11 and 12 depict cycles of period 2 and 3 respectively.

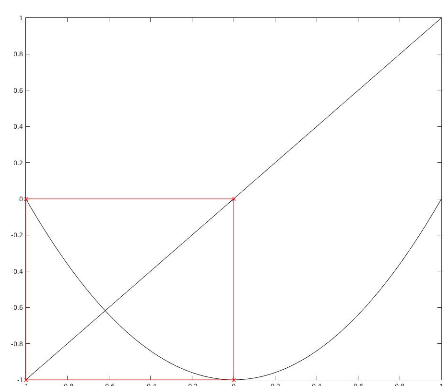


Figure 11: The forward orbit of $x = 0$ under $f(x) = x^2 - 1$. As we can see, 0 is periodic of period 2 under this map.

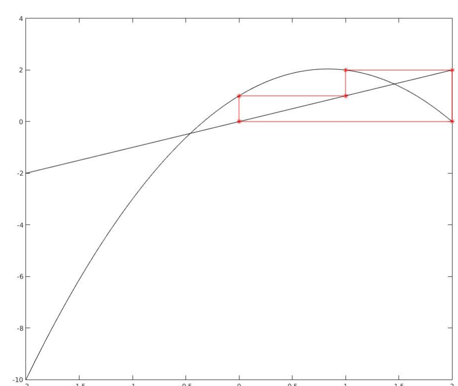


Figure 12: The forward orbit of $x = 0$ under $f(x) = -\frac{3}{2}x^2 + \frac{5}{2}x + 1$. As we can see, 0 is periodic of period 3 under this map.

Note that a general dynamical system, might not have any non-trivial periodic points at all. Consider $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = 2x$. Clearly only $x_0 = 0$ is a periodic point under f.

A point that is not fixed or periodic itself may have points in its forward orbit that are either fixed or periodic.

Definition 1.2.8. We say that a point x is **eventually fixed** if x is not fixed but there exists an $m > 0$ such that $f^{i+1}(x) = f^i(x)$ for all $i \geq m$ (that is, $f^i(x)$ is fixed for $i \geq m$).

As we noted in the introduction, the orbit of $x_0 = .01$ under $f(x) = 2x(1-x)$ seemed to level out at .5, but was actually just approaching 0.5. As such, .01 is *not* eventually fixed under f.

Definition 1.2.9. We say that a point x is **eventually periodic** of period n if x is not periodic but there exists an $m > 0$ such that $f^{n+i}(x) = f^i(x)$ for all $i \geq m$ (that is, $f^i(x)$ is periodic for $i \geq m$).

Example 1.2.10. Let $f(x) = x^2 - 1$ again. We see that $x = 1$ is an eventually periodic point of period 2 as

$$f^2(1) = f(f(1)) = f(0) = -1$$

and

$$f^4(1) = f(f(f^2(1))) = f(f(-1)) = f(0) = -1.$$

1.3 Limits of Orbits

While not all points in a dynamical system have orbits which eventually cycle through a finite set of values, many do have orbits which approach certain points (again, .01 under $f(x) = 2x(1-x)$). Just as we study the limit of a function $f(x)$ as x approaches infinity, we may also study the limit of the orbit of a particular point in our dynamical system as n approaches infinity.

Definition 1.3.1. Let p be periodic of period n. A point x is **forward asymptotic** to p if

$$\lim_{i \to \infty} f^{in}(x) = p \ .$$

The **stable set** of p, denoted by $W^s(p)$, consists of all points forward asymptotic to p.

Note that in this paper we will primarily concern ourselves with points which are forward asymptotic to fixed points.

Definition 1.3.2. Suppose that f is invertible. We say that x is **backward asymptotic** to p if

$$\lim_{i \to -\infty} f^{in}(x) = p \ .$$

We call the set of backward asymptotic points to p the **unstable set** of p and denote it $W^u(p)$.

Example 1.3.3. In Figure 6 we see that even though $x_0 = .01$ is relatively close close 0, the forward orbit of x_0 climbs towards 0.5, suggesting that the stable set of f is the interval $(0, 1]$.

1.4 Hyperbolicity

While we might understand how points can be forward asymptotic or backward asymptotic to a periodic point p, we still need a way to tell whether p has points converging to it. Whether p has this property has to do with whether p is *hyperbolic*.

Definition 1.4.1. Let p be a periodic point of period n. We say p is **hyperbolic** if $|(f^n)'(p)| \neq 1$. We call $(f^n)'(p)$ the **multiplier** of p.

Example 1.4.2. Consider the function $f(x) = x^2$. Clearly f has fixed points at $x = 0$ and $x = 1$. As $f'(0) = 0$ and $f'(1) = 2$ clearly both fixed points are hyperbolic.

Example 1.4.3. As a counterexample, consider the function $f(x) = x$. Clearly every point x is a fixed point. However, none these fixed points is hyperbolic as $f'(x) = 1$ for all x.

With the concept of hyperbolicity established, we can develop a way of telling whether a periodic point p has points converging to it. For the following problems we are assuming that f is differentiable and that its derivative is continuous.

Theorem 1.4.4. *Let p be a hyperbolic fixed point. If $|f'(p)| < 1$, then there is an open interval U about p such that if $x \in U$, then*

$$\lim_{n\to\infty} f^n(x) = p \ .$$

Proof. We know that there exists a real number A such that $|f'(p)| < A < 1$. We will show by induction that as p is a fixed point and $|f'(p)| < A$, $|(f^n)'(p)| < A^n$ for all n. For our base case we see that the statement is obviously true for $n = 1$. Suppose it is true for $n - 1$. We see that

$$\begin{aligned} |(f^n)'(p)| &= |(f(f^{n-1}))'(p)| = |f'(f^{n-1}(p))||(f^{n-1})'(p)| && (9)\\ &= |f'(p)||(f^{n-1})'(p)| < AA^{n-1} = A^n && (10) \end{aligned}$$

as $f^{n-1}(p) = p$ since p is a fixed point under f. We know that the derivative of f^n is continuous. Let $\epsilon = A^n - |(f^n)'(p)|$. As $(f^n)'$ is continuous at p there exists a $\delta > 0$ such that $|(f^n)'(x) - (f^n)'(p)| < \epsilon$ if $x \in (p - \delta, p + \delta)$. We then see that

$$\begin{aligned} |(f^n)'(x)| &= |(f^n)'(x) - (f^n)'(p) + (f^n)'(p)| && (11)\\ &\leq |(f^n)'(x) - (f^n)'(p)| + |(f^n)'(p)| && (12)\\ &< A^n - |(f^n)'(p)| + |(f^n)'(p)| = A^n && (13) \end{aligned}$$

if $x \in (p - \delta, p + \delta)$.
Let $x \in (p - \delta, p + \delta)$. By the Mean Value Theorem

$$\frac{|f^n(x) - f^n(p)|}{|x - p|} < A^n$$

so that

$$|f^n(x) - p| = |f^n(x) - f^n(p)| < A^n|x - p| < A^n\delta \ .$$

As δ is fixed by ϵ and $|A| < 1$, clearly $\lim_{n\to\infty} A^n\delta = 0$. Thus

$$\lim_{n\to\infty} f^n(x) = p \ .$$

□

A direct consequence of this is that the interval $(p - \delta, p + \delta)$ is a subset of the stable set of p.

Definition 1.4.5. If p is a hyperbolic periodic point of period n with $|(f^n)'(p)| < 1$ we say that p is an **attracting periodic point** or a **sink**.

Example 1.4.6. Returning to $f(x) = 2x(1-x)$, we note that $f'(x) = 2-4x$. As such $f'(1/2) = 0$. Therefore $\frac{1}{2}$ is a sink, which helps explain the behavior we saw in Figure 6.

Definition 1.4.7. If p is a fixed point with $|f'(p)| > 1$ we call p a **repelling fixed point** or **source**.

Theorem 1.4.8. *Let f be a function which is infinitely differentiable where all of its derivatives are continuous and let p be a hyperbolic fixed point with $|f'(p)| > 1$. Then there is an open interval U around p such that, if $x \in U$ and $x \neq p$, then there exists an integer k such that $f^k(x) \notin U$.*

The proof of this theorem is similar enough to our last proof that we will not bother with it here.

Example 1.4.9. Returning to $f(x) = 2x(1-x)$, we note that $f'(x) = 2-4x$. As such $f'(0) = 2$. Therefore 0 is a source, which again helps explain the behavior we saw in Figure 6.

As our analysis has suggested so far, the map $f(x) = 2x(1-x)$ seems incredibly predictable. We can generalize this behavior to large class of logistic maps. Recall that logistic maps are functions of the form $F_\mu(x) = \mu x(1-x)$.

Theorem 1.4.10. *Let $1 < \mu < 3$.*

1. *F_μ has a sink at $p_\mu = \frac{\mu-1}{\mu}$ and a source at 0.*
2. *If $0 < x < 1$, then*
$$\lim_{n\to\infty} F_\mu^n(x) = p_\mu.$$

Unfortunately, this theorem is quite difficult to prove. Given that non-chaotic systems are not our primary focus, we will leave it without proof here.

1.5 Dynamical Systems: A Wide Field of Study

Having now established a firm understanding of dynamical systems we can see what a wide variety of forms they can take. As such, we cannot make too many general statements about the behavior of dynamical systems. We did find, however, that any logistic map F_μ with $1 < \mu < 3$ behaves quite predictably. As we delve into chaos and, ultimately, chaotic logistic maps, we will see how diverse this seemingly mundane family of functions really is.

2 Chaos

Note that the following discussion of chaos presumes a general knowledge of *metric spaces* and *topology.* If you are unfamiliar with these concepts or would like a brief refresher, consult the Appendix.
Chaotic dynamics are defined by a few key properties which we must explore in isolation before pulling together a complete definition of chaos. Note that I have adapted the following definitions from Devaney [3], generalizing his results in $(\mathbb{R}, d_1)$ to any metric space (X, d).

2.1 Our Definition of Chaos

Definition 2.1.1. Let (X, d) be a metric space and let $J \subseteq X$.
$f : J \to J$ has **sensitive dependence on initial conditions** if there exists a $\delta > 0$ such that, for any $x \in J$ and any open set N containing x, there exists a $y \in N$ and an $n \geq 0$ such that $d(f^n(x), f^n(y)) > \delta$.

Simply put, f has sensitive dependence if even the slightest change in initial conditions eventually leads to a substantially different outcome. This is the very behavior which we saw with $g(x) = 4x(1-x)$ and which first alarmed Lorenz (see Introduction). Moreover, this property is frequently referred to as *The Butterfly Effect.* Among the properties of a chaotic dynamical system, this is perhaps the most well-known. Some popular depictions of chaos even equate it solely with the Butterfly Effect, ignoring the other two key properties of a chaotic dynamical system entirely.

Example 2.1.2. Consider the map $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = 2x$.
Let $s \in \mathbb{R}$ and let N be an open set containing s. Since N is open, we know that there exists a real number $\epsilon > 0$ such that $(s - \epsilon, s + \epsilon) = B_{d_1}(s, \epsilon) \subseteq N$. Let $y = s + \frac{\epsilon}{2}$. Clearly $y \in (s - \epsilon, s + \epsilon)$.
We know that there exists a positive integer M such that $2^M \epsilon > 1$. We see that for any positive integer n and any $t \in \mathbb{R}$, $f^n(t) = 2^n t$. Therefore

$$|f^{M+1}(s) - f^{M+1}(y)| = |2^{M+1}s - 2^{M+1}y| = 2^{M+1}|s - y| = 2^M \epsilon > 1 \ .$$

Therefore f has sensitive dependence on initial conditions.

Definition 2.1.3. Let (X, d) be a metric space and let $J \subseteq X$.
$f : J \to J$ is said to be **topologically transitive** if for any pair of open sets $U, V \subseteq J$ there exists a k >0 such that

$$f^k(U) \cap V \neq \emptyset \ .$$

This is to say that if you take two open sets U and V in J and look far enough in the orbits of all of the elements of U under f you will eventually find some element in V. In even more general terms: wherever you look in J you will be able to find a point which has an orbit which travels all over J. We saw this behavior in the orbit of .01 under $g(x) = 4x(1 - x)$ in Figures 8, 9, and 10.

Example 2.1.4. As a counterexample note that the map $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = 2x$ is certainly not topologically transitive. Let $V = (0, 1)$ and $U = (1, 2)$. Clearly if $x \in U$, then $f^k(x) > 1$ for any k. As such, for every k we see that $f^k(U) \cap V = \emptyset$.

Our third property requires no new definition. It is simply that *periodic points are dense under* f. If you are unfamiliar with the concept of density, see the Appendix. Essentially, this means that wherever you look in J, you will find a periodic point under f.

Example 2.1.5. We will once again use the map $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = 2x$ as a counterexample. Clearly f does not have any periodic points other than 0 as $f^{k+1}(x) > f^k(x)$ for every k.

With all of the pieces in place we can finally formalize our definition of chaos:

Definition 2.1.6. Let (X, d) be a metric space and let $V \subseteq X$.
We say that $f : V \to V$ is **chaotic** on V if

1. f has sensitive dependence on initial conditions,
2. f is topologically transitive,
3. periodic points under f are dense in V.

2.2 A Straightforward Example of a Chaotic Dynamical System

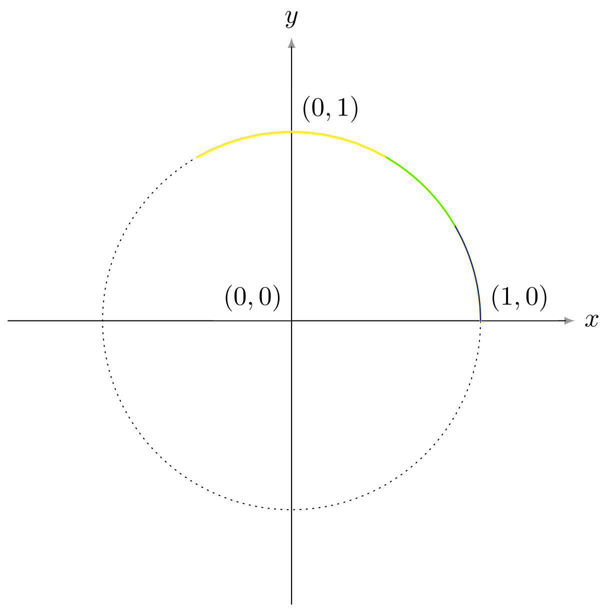
With this definitions of chaos we can show that the following system is chaotic: Consider the function $f(\theta) = 2\theta$ defined on the unit circle, denoted S^1, under the metric $|x - y|$. We note that $\theta = \theta + 2k\pi$ for any integer k.
To see that f has sensitive dependence on initial conditions, let $\delta = 1$ and consider two points $\theta_1, \theta_2 \in S^1$ where $\theta_1 < \theta_2$. We recognize that $f^n(\theta) = 2^n\theta$.
Therefore

$$|f^n(\theta_2) - f^n(\theta_1)| = |2^n\theta_2 - 2^n\theta_1| = 2^n(\theta_2 - \theta_1) \ .$$

By the Archimedean property there exists a real number r such that $1 < r(\theta_2 - \theta_1)$. As the sequence $\{2^n\}$ is unbounded above, there exists a positive integer N such that $2^N > r$. Thus

$$|f^N(\theta_2) - f^N(\theta_1)| = 2^N(\theta_2 - \theta_1) > r(\theta_2 - \theta_1) > 1 \ .$$

We will now show that f is topologically transitive on S^1. We note that any open subset of S^1 must contain an open interval, say (θ_a, θ_b). We see that the codomain of (θ_a, θ_b) under $f^n(\theta)$ will be $(2^n\theta_a, 2^n\theta_b)$. It is not hard to see, using the same argument we used in our discussion of the sensitive dependence of f, that there exists an integer M such that $2^M\theta_a - 2^M\theta_b > 2\pi$. Thus eventually the orbits of elements from (θ_a, θ_b) will cover S^1 and, by extension, any subset of S^1. To further elucidate this concept, I have included the following image. (θ_a, θ_b) in the argument is the blue arc below, $(f(\theta_a), f(\theta_b))$ is the green arc, and $(f^2(\theta_a), f^2(\theta_b))$ is the yellow arc.



The density of periodic points under f is perhaps the most difficult to prove.
We note that $f^n(\theta) = 2^n\theta$ so θ is a periodic point of period n if and only if

$$2^n\theta = \theta + 2k\pi$$

for some integer n, that is, if and only if

$$\theta = \frac{2k\pi}{2^n - 1} \ .$$

From here we only need to see that numbers of the form $\frac{k}{2^n-1}$ where k and n are integers are dense in the interval $[0, 1]$.
Consider any open subset of $[0, 1]$. We know that it will contain some open interval (a, b). By the Archimedean property there exists a positive integer m such that $1 < m(b - a)$. Moreover, since $\{2^n - 1\}$ is unbounded above, there exists a positive integer p such that $2^p - 1 > m$. Therefore

$$1 < m(b - a) < (2^p - 1)(b - a)$$

so that

$$\frac{1}{2^p - 1} < b - a$$

and

$$a < \frac{1}{2^p - 1} + a < b \ .$$

As

$$\frac{1}{2^p - 1} + a = \frac{1 + a(2^p - 1)}{2^p - 1}$$

and p and $1 + a(2^p - 1)$ are both integers, we have found a number of the form $\frac{k}{2^n-1}$, where k and n are integers, in a generic open subset of $[0, 1]$. Thus numbers of this form are dense in $[0, 1]$ and, by extension, numbers of the form $\frac{2k\pi}{2^n-1}$ are dense in S^1.
Therefore f is chaotic.

3 Topological Conjugacy: A Powerful Tool

All told, directly using the definition of chaos to prove that the doubling map on the unit circle is chaotic is not too difficult: The orbits of two distinct points diverge predictably, the images of subsets of S^1 clearly move around S^1, and periodic points are relatively easy to find.

Unfortunately, there are many chaotic maps whose chaotic properties are not so immediately apparent. In order to prove that such maps are chaotic, we show that they are *topologically conjugate* to other maps which we know are chaotic.

First, however, we must unpack what it means for two maps to be topologically conjugate. To this end, we define a *homeomorphism*.

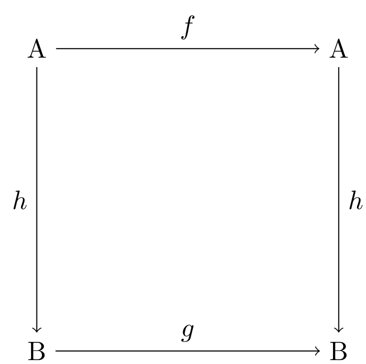
Definition 3.0.1. Let $f : I \to J$.
We say that f is a **homeomorphism** if f is one-to-one, onto, and continuous, and f^{-1} is also continuous.

Example 3.0.2. The function $f(x) = \sqrt{x}$, where $x \in \mathbb{R}^+$ is a homeomorphism.

With that definition in place, we can define topological conjugacy:

Definition 3.0.3. Let $f : A \to A$ and $g : B \to B$ be two maps. We say that f and g are **topologically conjugate** if there exists a homeomorphism $h : A \to B$ such that $h \circ f = g \circ h$. We call h a **topological conjugacy**.

As a visual reference, we have the following diagram:



3.1 The Properties Shared by Topologically Conjugate Maps

It is not immediately apparent that two topologically conjugate maps, such as $f : A \to A$ and $g : B \to B$, have the same dynamics. Thus before we can use topological conjugacy to prove that a dynamical system is chaotic, we need to explore the implications of topological conjugacy at some length.

3.1.1 Fixed and Periodic Points in Topologically Conjugate Maps

We will starting by exploring the many relationships between the dynamics of $x \in A$ under f and the dynamics of $h(x) \in B$ under g.

Theorem 3.1.1. *Suppose that $f : A \to A$ and $g : B \to B$ are topologically conjugate under the homeomorphism $h : A \to B$. Then a point p in A is a fixed point under f if and only if $h(p)$ is a fixed point under g.*

Proof. We will start with the forward direction.
Let $p \in A$ be a fixed point under f. Then $f(p) = p$. Therefore

$$h(p) = (h \circ f)(p) = (g \circ h)(p) ,$$

making $h(p)$ a fixed point under g by definition.
Moving in the opposite direction, let $p \in A$ and suppose that $h(p)$ is a fixed point under g. Then $(g \circ h)(p) = h(p)$. Therefore

$$p = (h^{-1} \circ h)(p) = (h^{-1} \circ g \circ h)(p) = (h^{-1} \circ h \circ f)(p) = f(p) ,$$

making p a fixed point under f by definition. □

We can use the same logic to find a similar correspondence between periodic points under f and periodic points under g.

Theorem 3.1.2. *Suppose that $f : A \to A$ and $g : B \to B$ are topologically conjugate under the homeomorphism $h : A \to B$. Then a point x in A is a periodic point of period n under f if and only if $h(x)$ is a periodic point of period n under g.*

Proof. We will start with the forward direction.
Let $x \in A$ be a periodic point of period n under f. Then $f^n(x) = x$. We see that

$$h(x) = (h \circ f^n)(x) = (g \circ h \circ f^{n-1})(x) = (g^2 \circ h \circ f^{n-2})(x) .$$

Continuing the process of substituting $g \circ h$ for $h \circ f$, we find that

$$h(x) = (g^n \circ h)(x) .$$

Therefore $h(x)$ is a periodic point of period n under g by definition.
Moving in the opposite direction, let $x \in A$ and suppose that $h(x)$ is a periodic point of period n under g. Then $(g^n \circ h)(x) = h(x)$. Therefore

$$\begin{aligned} x &= (h^{-1} \circ h)(x) = (h^{-1} \circ g^n \circ h)(x) \\ &= (h^{-1} \circ g^{n-1} \circ h \circ f)(x) \\ &= (h^{-1} \circ g^{n-2} \circ h \circ f^2)(x) . \end{aligned}$$

Continuing the process of substituting $h \circ f$ for $g \circ h$, we find that

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

Therefore x is a periodic point of period n under f by definition. □

Therefore h gives a one-to-one correspondence between $Per_n(f)$ and $Per_n(g)$.

Recall that a point x is eventually fixed if x is not fixed but there exists a positive integer $m > 0$ such that $f^{i+1}(x) = f^i(x)$ for all $i \geq m$. Also recall that a point x is eventually periodic of period n

if x is not periodic but there exists a positive integer $m > 0$ such that $f^{n+i}(x) = f^i(x)$ for all $i \geq m$. It does not take too much imagination to adapt the argument used in the proof of the previous theorem to show that a point $x \in A$ is eventually fixed under f if and only if $h(x)$ is eventually fixed under g, and that x is eventually periodic under f if and only if $h(x)$ is eventually periodic under g.
Thus there are direct correspondences between the most predictable types of points in two topologically conjugate maps.

3.1.2 Asymptotic Orbits in Topologically Conjugate Maps

As in our general discussion of dynamical systems, we might be curious about the end-behavior of points under f and g.
Let p be periodic of period n. Recall that a point x is *forward asymptotic* to p if $\lim_{i\to\infty} f^{in}(x) = p$.

Theorem 3.1.3. *Let (A, d_1) and (B, d_2) be metric spaces and let $H : A \to B$ be a homeomorphism. Suppose that p is a period point of period n under f. Then a point $x \in A$ is forward asymptotic to p if and only if $h(x)$ is forward asymptotic to $h(p)$.*

Proof. We will prove the forward direction here and leave the other direction without proof as it requires essentially the same argument. Let $\epsilon > 0$. We note that as h is continuous, there exists a $\delta > 0$ such that

$$d_2(h(s), h(t)) < \epsilon$$

if $d_1(s, t) < \delta$. As x is forward asymptotic to p, we know that there exists a positive integer M such that

$$d_1(f^{in}(x), p) < \delta$$

if $i \geq M$.
Therefore

$$d_2((h \circ f^{in})(x), h(p)) < \epsilon$$

if $i \geq M$. Moreover, we note by the argument employed in the proof of Theorem 2 that $(h \circ f^{in})(x) = (g^{in} \circ h)(x)$, so that

$$d_2((g^{in} \circ h)(x), h(p)) < \epsilon$$

if $i \geq M$. Therefore $h(x)$ is forward asymptotic to $h(p)$ under g by definition. ☐

Provided f and g are invertible, we can apply this same argument to show an equivalence between backward asymptotic orbits as well.

3.1.3 How a Topological Conjugacy Maps Sets

While we have discussed how a topological conjugacy acts on individual points, we have yet to discuss how it acts on subsets of a generic metric space (X, d). We need to explore these concepts before discussing topological conjugacy and chaos because both topological transitivity and density are properties which describe how a map acts on subsets of its domain.

Theorem 3.1.4. *Let (X, d_1) and (Y, d_2) be metric spaces and let $h : X \to Y$ be a continuous function. If $M \subseteq Y$ is an open set, then $\{x \in X : f(x) \in M\}$, the preimage of M under f, is also an open set.*

Proof. Let $x_0 \in \{x \in X : f(x) \in M\}$. We know that $f(x_0) \in M$. Since M is open we know that $f(x_0)$ is an interior point of M. As such there exists a real number $r > 0$ such that $B_{d_2}(f(x_0), r) \subseteq M$.
Let $0 < \epsilon \leq r$. Because $f(x)$ is continuous at x_0 we know that there exists a $\delta > 0$ such that $f(x) \in B_{d_2}(f(x_0), \epsilon)$ if $x \in B_{d_1}(x_0, \delta)$.
Let $x \in B_{d_1}(x_0, \delta)$. Then $f(x) \in B_{d_2}(f(x_0), \epsilon) \subseteq M$, meaning that $x \in \{x \in X : f(x) \in M\}$.
Thus $B_{d_1}(x_0, \delta) \subseteq \{x \in X : f(x) \in M\}$, making x_0 is an interior point of $\{x \in X : f(x) \in M\}$.
Therefore $\{x \in X : f(x) \in M\}$ is an open set. □

With this theorem about continuous maps established, we can discuss how *homeomorphisms* map sets. Recall that both a homeomorphism and its inverse are continuous.

Theorem 3.1.5. *Let (X, d_1) and (Y, d_2) be metric spaces, let $E \subseteq X$, and let $h : X \to Y$ be a homeomorphism. Then E is an open subset of (X, d_1) if and only if $h(E)$ is an open subset of (Y, d_2).*

Proof. Let E be an open subset of X. Since h^{-1} is a continuous map we know from our last theorem that $\{y \in Y : h^{-1}(y) \in E\} = h(E)$ is open in Y.
Let E be a subset of X with the property that $h(E)$ is an open subset of y. Then $\{x \in X : h(x) \in W\} = (h^{-1} \circ h)(E) = E$ is open in X. □

We know that if E is open in X, then $X \setminus E$ is closed (see Appendix). Since a homeomorphism maps open sets to open sets, we can conclude that *a homeomorphism also maps closed sets to closed sets.*

3.2 Topological Conjugacy and Chaos

Having discussed how a topological conjugacy acts on sets we can show that if two maps f and g are topologically conjugate, then f is chaotic if and only if g is chaotic.

Theorem 3.2.1. *Let (X, d_1) and (Y, d_2) be metric spaces and let $f : X \to X$ and $g : Y \to Y$ be two maps with a topological conjugacy $h : X \to Y$ between them. Then f is topologically transitive if and only if g is topologically transitive.*

Proof. Suppose that f is topologically transitive and let U and V be two open sets in Y. We know that $h^{-1}(U)$ and $h^{-1}(V)$ are both open sets in X. Since f is topologically transitive, we know that there exists a positive integer k such that

$$(f^k \circ h^{-1})(U) \cap h^{-1}(V) \neq \emptyset \ .$$

We see that $(h \circ f^k \circ h^{-1})(U) = (g^k \circ h \circ h^{-1})(U) = g^k(U)$. Therefore

$$g^k(U) \cap V \neq \emptyset \ .$$

Hence g is topologically transitive.
Suppose that g is topologically transitive and let U and V be two open sets in X. We know that $h(U)$ and $h(v)$ are both open sets in Y. Since g is topologically transitive, we know that there exists a positive integer j such that

$$(g^j \circ h)(U) \cap h(V) \neq \emptyset \ .$$

We see that $(h^{-1} \circ g^j \circ h)(U) = (f^j \circ h^{-1} \circ h)(U) = f^j(U)$. Therefore

$$f^j(U) \cap V \neq \emptyset \ .$$

Hence f is topologically transitive.

□

Theorem 3.2.2. *Let (X, d_1) and (Y, d_2) be metric spaces and let $f : X \to X$ and $g : Y \to Y$ be two maps with a topological conjugacy $h : X \to Y$ between them. Then periodic points are dense in X under f if and only if periodic points are dense in Y under g.*

Proof. Suppose that periodic points are dense in X under f. Let U be an open set in Y. We then know that $h^{-1}(U)$ is an open set in X. Since periodic points are dense in X under f we know that there exists a periodic point $p \in h^{-1}(U)$ under f. We know that since p is a periodic point under f, $h(p)$ is a periodic point under g. Since $h(p)$ is an element of U, we can see that every open set in Y contains a periodic point under g.
Suppose that periodic points are dense in Y under g. Let U be an open set in X. We then know that $h(U)$ is an open set in Y. Since periodic points are dense in Y under g we know that there exists a periodic point $p \in h(U)$ under g. We know that since p is a periodic point under f, $h^{-1}(p)$ is a periodic point under f. Since $h^{-1}(p)$ is an element of U, we can see that every open set in X contains a periodic point under f. □

Theorem 3.2.3. *Let (X, d_1) and (Y, d_2) be metric spaces and let $f : X \to X$ and $g : Y \to Y$ be two maps with a topological conjugacy $h : X \to Y$ between them. Then f has sensitive dependence on initial conditions if and only if g has sensitive dependence on initial conditions.*

Proof. Suppose that f has sensitive dependence on initial conditions and let $\epsilon > 0$ be the real number guaranteed to exist by that property. Let $x \in Y$ and let N be a neighborhood of x. We know that there exists some point $y \in h^{-1}(N) \subseteq X$ and some positive integer n such that

$$d_1((f^n \circ h^{-1})(x), f^n(y)) > \epsilon \ .$$

Since h^{-1} is continuous we know that there exists some $\delta > 0$ such that $d_1(h^{-1}(s), h^{-1}(t)) < \epsilon$ if $d_2(s,t) \leq \delta$. Note that since ϵ is fixed, δ is as well.
Suppose that $d_2(g^n(x), (g^n \circ h)(y)) \leq \delta$. Then

$$d_1((h^{-1} \circ g^n)(x), (h^{-1} \circ g^n \circ h)(y)) = d_1((f^n \circ h^{-1})(x), f^n(y)) < \epsilon \ ,$$

a contradiction. Therefore $d_2(g^n(x), (g^n \circ h)(y)) > \delta$. Thus we have shown that given any point $x \in Y$ and any neighborhood N around it, we can find a point $w \in N$ (the $h(y)$ we found using the sensitive dependence of f) and a positive integer n such that $d_2(g^n(x), g^n(w)) > \delta$. As such, g has sensitive dependence on initial conditions.
Given that the proof of the converse is essentially the same (the only difference being that it relies on the fact that h is continuous instead of the fact that h^{-1} is continuous) we will omit it here. □

With all of these pieces in place we have our powerful tool for proving that sets are chaotic:

Theorem 3.2.4. *Suppose that f and g are topologically conjugate. Then f is chaotic if and only if g is chaotic.*

4 Proving a Dynamical System is Chaotic Using Topological Conjugacy

As we noted at the beginning of our discussion of topological conjugacy, given two topologically conjugate chaotic maps, $f : A \to A$ and $g : N \to B$, it can often be far easier to prove that f is chaotic than that g is chaotic. As such, we choose to prove directly that f is chaotic in order to show that g is chaotic. In this section we will be doing exactly that, with the *shift map*, $\sigma : \Sigma_2 \to \Sigma_2$, playing the role of f in our scheme and any logistic map, $F_\mu = \mu x(1-x)$ where $\mu > 2+\sqrt{5}$, playing the role of g (we will cover F_4, from the introduction, later as it is a special case).

4.1 The Shift Map

4.1.1 The Shift Map's Domain: Sequence Space

Before we can discuss the shift map, we need to discuss the set on which it acts, the *sequence space* on the symbols 0 and 1.

Definition 4.1.1. The **sequence space** on the symbols 0 and 1 is the set

$$\Sigma_2 = \{\mathbf{s} = (s_0 s_1 s_2 ...) | s_j = 0 \text{ or } 1\} \ .$$

Elements of Σ_2 are infinite strings of 0's and 1's, such as

$$(001101\ldots) \quad \text{or} \quad (101001\ldots) \ .$$

As our previous discussions have indicated, we need Σ_2 to be a metric space in order to talk about chaos and topological conjugacy. To this end, we develop the following definition of the distance between elements in the sequence space.

Definition 4.1.2. We define the distance between two sequences $\mathbf{s} = (s_0 s_1 s_2 \ldots)$ and $\mathbf{t} = (t_0 t_1 t_2 \ldots)$ to be

$$d(\mathbf{s}, \mathbf{t}) = \sum_{i=0}^{\infty} \frac{|s_i - t_i|}{2^i} \ .$$

Since $\frac{|s_i - t_i|}{2^i} \leq \frac{1}{2^i}$ for every positive integer i and since $\sum_{i=0}^{\infty} \frac{1}{2^i}$, a geometric series, converges, we know that $\sum_{i=0}^{\infty} \frac{|s_i - t_i|}{2^i}$ converges. As such $d(\mathbf{s}, \mathbf{t}) \in \mathbb{R}$.

Example 4.1.3. We see that if $\mathbf{r} = (000000\ldots)$, $\mathbf{s} = (010101\ldots)$, and $\mathbf{t} = (111111\ldots)$, then

$$d(\mathbf{r}, \mathbf{t}) = \sum_{i=0}^{\infty} \frac{1}{2^i} = 2$$

and

$$d(\mathbf{r}, \mathbf{s}) = d(\mathbf{s}, \mathbf{t}) = \sum_{i=0}^{\infty} \frac{1}{2^{2i}} = \frac{4}{3} \ .$$

It is not too difficult to verify that $d(\mathbf{s}, \mathbf{t})$ is a metric:

Theorem 4.1.4. *$d(\mathbf{s}, \mathbf{t})$ is a metric.*

Proof. We will work through the definition of a metric systematically to show that $d(\mathbf{s}, \mathbf{t})$ is a metric.

1. It is clear that $|a - b| \geq 0$ for any points $a, b \in \{0, 1\}$. As such $d(\mathbf{s}, \mathbf{t}) \geq 0$ for any $\mathbf{s}, \mathbf{t} \in \Sigma_2$.
2. We note that $\mathbf{s} = \mathbf{t}$ if and only if $s_i = t_i$ for each i, that $s_i = t_i$ if and only if $|s_i - t_i| = 0$, and that $|s_i - t_i| = 0$ for each i if and only if $d(\mathbf{s}, \mathbf{t}) = \sum_{i=0}^{\infty} \frac{|s_i - t_i|}{2^i} = 0$.
3. As $|s_i - t_i| = |t_i - s_i|$ for each i, it is clear that $d(\mathbf{s}, \mathbf{t}) = d(\mathbf{t}, \mathbf{s})$.
4. Let $\mathbf{r}, \mathbf{s}, \mathbf{t} \in \Sigma_2$. We know that $|r_i - s_i| \leq |r_i - t_i| + |t_i - s_i|$ for each i. As such $d(\mathbf{r}, \mathbf{s}) \leq d(\mathbf{r}, \mathbf{t}) + d(\mathbf{t}, \mathbf{s})$.

□

As may be clear from our discussions of chaos and topological conjugacy, we will often want to find an upper bound for $d(\mathbf{s}, \mathbf{t})$. The following theorem helps us to find such an upper bound fairly easily.

Theorem 4.1.5. *Let $\mathbf{s}, \mathbf{t} \in \Sigma_2$. If $s_i = t_i$ for all $i \in \{0, 1, \ldots, n\}$, then $d(\mathbf{s}, \mathbf{t}) \leq \frac{1}{2^n}$. Conversely, if $d(\mathbf{s}, \mathbf{t}) < \frac{1}{2^n}$, then $s_i = t_i$ for all $i \in \{0, 1, \ldots, n\}$.*

Proof. Suppose $s_i = t_i$ for all $i \in \{0, 1, \ldots, n\}$. We see

$$\begin{aligned} d(\mathbf{s}, \mathbf{t}) &= \sum_{i=0}^{n} \frac{|s_i - t_i|}{2^i} + \sum_{i=n+1}^{\infty} \frac{|s_i - t_i|}{2^i} \\ &\leq \sum_{i=n+1}^{\infty} \frac{|s_i - t_i|}{2^i} = \frac{1}{2^{n+1}} \sum_{i=0}^{\infty} \frac{|s_i - t_i|}{2^i} = \frac{1}{2^n} \ . \end{aligned}$$

We will prove the converse by contrapositive. Suppose that $s_j \neq t_j$ for some $j \leq n$. Then

$$d(\mathbf{s}, \mathbf{t}) \geq \frac{1}{2^j} \geq \frac{1}{2^n} \ .$$

Thus if $d(\mathbf{s}, \mathbf{t}) < \frac{1}{2^n}$, then $s_i = t_i$ for all $i \in \{0, 1, \ldots, n\}$. □

4.1.2 The Shift Map

With all of this background information established, we can finally talk about the shift map.

Definition 4.1.6. The **shift map** $\sigma : \Sigma_2 \to \Sigma_2$ is defined by

$$\sigma(s_0 s_1 s_2 \ldots) = (s_1 s_2 s_3 \ldots) \ .$$

Intuitively, the shift map just "shifts" every entry in the sequence one entry to the left and cuts off the first entry.

4.1.3 The Shift Map is Chaotic

By design, it is relatively easy to see that the shift map is chaotic even though we hardly know anything about it.

Theorem 4.1.7. *The map* $\sigma : \Sigma_2 \to \Sigma_2$ *is chaotic.*

Proof. We will show that the set of periodic points under σ, $Per(\sigma)$, is dense in Σ_2 by showing that any point $\mathbf{s} \in \Sigma_2$ is a limit point of $Per(\sigma)$, and therefore $\Sigma_2 \subseteq \overline{Per(\sigma)}$.
Let $\mathbf{s} = (s_0 s_1 s_2 \ldots) \in \Sigma_2$ and let $\epsilon > 0$. We know that there exists a positive integer N such that $\frac{1}{2^N} < \epsilon$.
Let $\tau_n = (s_0 s_1 \ldots s_n s_0 s_1 \ldots s_n s_0 \ldots)$ for each n. Clearly each τ_n is periodic of period n.
From our previous theorem, since for a given n τ_n and $\mathbf{s}$ agree in their first $n+1$ entries, we know that

$$d(\mathbf{s}, \tau_n) \leq \frac{1}{2^n} \ .$$

Thus we see that if $m \geq N$, then

$$d(\mathbf{s}, \tau_m) \leq \frac{1}{2^m} \leq \frac{1}{2^N} < \epsilon \ .$$

As such $\{\tau_n\}$ converges to $\mathbf{s}$.
Turning our focus towards sensitive dependence, we consider a sequence $\mathbf{s}$ in Σ_2 and an open set N containing it. We know that there exists some $\epsilon > 0$ such that $B_d(\mathbf{s}, \epsilon) \subseteq N$. Moreover, we know that there exists some integer M such that $\frac{1}{2^M} < \epsilon$. Let $\mathbf{r} \in B_d(\mathbf{s}, \frac{1}{2^M})$ with $r_m \neq s_m$ for some $m > M+1$ (our previous theorem provides that so long as $r_i = s_i$ for $i \in \{0, \ldots, M+1\}$, we are guaranteed that $d(\mathbf{s}, \mathbf{r}) < \frac{1}{2^{M+1}} < \frac{1}{2^M}$).
We note that $\sigma^m(\mathbf{s}) = (s_m s_{m+1} \ldots)$ and $\sigma^m(\mathbf{r}) = (r_m r_{m+1} \ldots)$.
As $s_m \neq r_m$, we know that

$$d(\sigma^m(\mathbf{s}), \sigma^m(\mathbf{r})) \geq \frac{1}{2} \ .$$

Finally, we will show that σ is topologically transitive by finding an element $\mathbf{s}'$ of Σ_2 such that given two open sets U and V in Σ_2, $\sigma^n(\mathbf{s}') \in U$ and $\sigma^m(\mathbf{s}') \in V$ where $m \geq n$.
Consider the sequence

$$\mathbf{s}' = (0\,1|00\,01\,10\,11|000\,001\ldots|\ldots) \ .$$

We see that $\mathbf{s}'$ consists of all possible strings of 0's and 1's of length 1, followed by all strings of length 2, and so on. Let U and V be open sets containing points $\mathbf{u}$ and $\mathbf{v}$, respectively. Since U and V are open, there exist real numbers $\epsilon_u > 0$ and $\epsilon_v > 0$ such that

$$B_d(\mathbf{u}, \epsilon_u) \subseteq U \text{ and } B_d(\mathbf{v}, \epsilon_v) \subseteq V \ .$$

Moreover, we know that there exist positive integers N_u and N_v such that $\frac{1}{2^{N_u}} < \epsilon_u$ and $\frac{1}{2^{N_v}} < \epsilon_v$. Given how we constructed $\mathbf{s}'$, we know that there exists an integer M_u such that $\sigma^{M_u}(\mathbf{s}')$ agrees with $\mathbf{u}$ in its first N_{u+1} entries. As such

$$d(\sigma^{M_u}(\mathbf{s}'), \mathbf{u}) \leq \frac{1}{2^{N_u}} < \epsilon_u,$$

so that $\sigma^{M_u}(\mathbf{s}') \in B_d(\mathbf{u}, \epsilon_u) \subseteq U$.
Likewise, we know that there exists integer $M_v > M_u$ such that $\sigma^{M_v}(\mathbf{s}')$ agrees with $\mathbf{v}$ in its first N_{v+1} entries. As such

$$d(\sigma^{M_v}(\mathbf{s}'), \mathbf{v}) \leq \frac{1}{2^{N_v}} < \epsilon_v,$$

so that $\sigma^{M_v}(\mathbf{s}') \in B_d(\mathbf{v}, \epsilon_v) \subseteq V$.
Therefore

$$\sigma^{M_v}(\mathbf{s}') \in \sigma^{M_v - M_u}(U) \cap V \ .$$

□

As in our discussion of the doubling map on the unit circle, we see that chaotic properties come to the shift map fairly easily. It is for this very reason that we use it to show that other maps are chaotic.

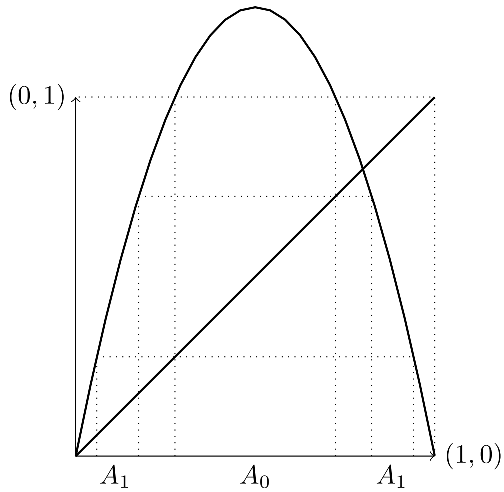
4.2 Logistic Maps

Recall that logistic maps are functions of the form $F_\mu(x) = \mu x(1-x)$. In the following few sections, we will show that if $\mu > 2 + \sqrt{5}$, then F_μ defined on a subset of $[0, 1]$ and σ are topologically conjugate, and, by extension, F_μ is chaotic.
In order to see this we have to discuss the subset of $[0, 1]$ in question.
We will be exploring the set of elements whose orbits under F_μ do not leave $[0, 1]$. To start, we have the following theorem:

Theorem 4.2.1. *Suppose $\mu > 1$. If $x < 0$, then $\lim_{n\to\infty} F_\mu^n(x) = -\infty$, and if $x > 1$, then $\lim_{n\to\infty} F_\mu^n(x) = -\infty$.*

Proof. We see that if $x < 0$, then $\mu(1 - x) > 1$, so that $F_\mu(x) = \mu x(1 - x) < x$. Therefore the sequence $\{F_\mu^n(x)\}$ is strictly decreasing. Clearly the sequence does not converge to a point y, for if it did then the sequence $\{F_\mu^{n+1}(x)\}$ would converge to $F_\mu(y) < y$. Therefore $\{F_\mu^n(x)\}$ diverges to $-\infty$.
If $x > 1$, then $F_\mu(x) < 0$. The argument that we just used thus applies to $F_\mu(x)$ and $\{F_\mu^{n+1}(x)\}$ diverges to $-\infty$. □

We find some interesting behavior when $\mu > 4$. We will denote $I = [0, 1]$ for convenience. We find that $F'_\mu(x) = \mu - 2\mu x$, so that F_μ attains its maximum at $\frac{1}{2}$. We note that $F_\mu(\frac{1}{2}) = \mu\frac{1}{2}(1 - \frac{1}{2}) = \frac{\mu}{4}$ is that maximum. As such, the maximum value of F_μ on I is greater than 1 when $\mu > 4$.
Since F_μ is continuous at $\frac{1}{2}$, we know that there exists some open interval A_0 centered at $\frac{1}{2}$, such that if $x \in A_0$, then $F_\mu(x) > 1$. This means that for any point $x \in A_0$, the orbit of x leaves I after a single iteration of F_μ and from there tends inexorably towards $-\infty$.
Let $A_1 = \{x \in I : F_\mu(x) \in A_0\}$. We see that if $x \in A_1$, then $F_\mu^2(x) > 1$, so that the orbit of x tends towards $-\infty$ from there. The following diagram depicts A_0 and A_1 for F_5:



With this basic idea in place, we can define

$$A_n = \{x \in I : F_\mu^n(x) \in A_0\}$$

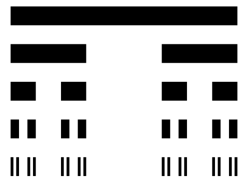
for any positive integer n. We recognize that A_n consists of all points in I for which $F_\mu^i(x) \in I$ if $i \leq n$, but $F_\mu^{n+1}(x) \notin I$. Clearly $\bigcup_{n=0}^{\infty} A_n$ consists of every point in I which eventually escapes I and tends towards $-\infty$.
We will thus focus on the set $I \setminus \bigcup_{n=0}^{\infty} A_n$ which we will denote Λ.
The set Λ has one key quality which helps us show that F_μ is topologically conjugate to the shift map.

Definition 4.2.2. A set is **totally disconnected** if it contains no intervals.

Theorem 4.2.3. *If $\mu > 2 + \sqrt{5}$, then Λ is totally disconnected.*

Proof. We need $|F'_\mu(x)| > 1$ for all $x \in I \setminus A_0$. We will not bother to prove it here, but F_μ has this property if $\mu > 2 + \sqrt{5}$. Nevertheless, as this is the case, we know that there exists some real number r such that $|F'_\mu(x)| > r > 1$ for all $x \in \Lambda$ as $\Lambda \subset I \setminus A_0$. As we showed in our discussion of hyperbolicity, this implies that $|(F_\mu^n)'(x)| > r^n$. Now suppose that Λ contains some interval $[x, y]$. Then $|(F_\mu^n)'(t)| > r^n$ for all $t \in [x, y]$. We know that there exists a positive integer N such that $r^N|x - y| > 1$. By the Mean Value Theorem, we know that $|F^N(x) - F^N(y)| \geq r^N|x - y| > 1$, which implies that at least one of $F^N(x)$ or $F^N(y)$ is not contained in I, contradicting the fact that x and y are in Λ. Thus Λ does not contain any intervals. $\square$

To further understand the concept of a totally disconnected set, we look at perhaps the most familiar totally disconnected set: the Cantor Middle Third set. The Cantor set is constructed as follows: Start with the interval $[0, 1]$; remove the open interval $(1/3, 2/3)$, that being the middle third; remove the middle thirds of two intervals that remain; repeat. The following diagram illustrates how we construct the set:

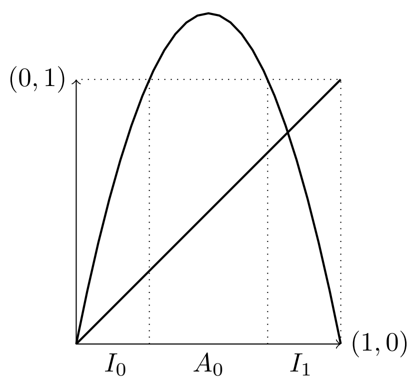


Note that we construct Λ in much the same way that we construct the Cantor set: by removing open intervals from the middles of larger intervals.

4.3 Building our Topological Conjugacy: The Itinerary

With strong understandings of σ and F_μ we can now connect the two maps with a topological conjugacy: the *itinerary.*

Definition 4.3.1. The **itinerary** of a point $x \in \Lambda$ is a sequence $S(x) = s_0 s_1 s_2 \ldots$ where $s_j = 0$ if $F_\mu^j(x) \in I_0$ and $s_j = 1$ if $F_\mu j(x) \in I_1$.



We recall that points in Λ never escape I. As $\Lambda \subseteq I_0 \cup I_1$ this means that we can follow the orbit of any point $x \in \Lambda$ as it moves between I_0 and I_1. As the name suggests, the itinerary could thus be thought of as a way to record where the orbit of $x \in \Lambda$ goes using a sequence $S(x) \in \Sigma_2$.

4.4 $S : \Lambda \to \Sigma_2$ is a Topological Conjugacy

Theorem 4.4.1. $S : \Lambda \to \Sigma_2$ *is a homeomorphism.*

Proof. As the task of showing that a function is a homeomorphism entails many smaller arguments, we will break this section up into more manageable pieces.

1. We will start by using a contradiction to prove that S is one-to-one.
   Let $x, y \in \Lambda$, where $x \neq y$ and $S(x) = S(y)$. If this is the case, then for each integer n, $F_\mu^n(x)$ and $F_\mu^n(y)$ are both in either I_0 or I_1. This means that both are on the same side of $\frac{1}{2}$. As such F_μ is monotonic on the interval between $F_\mu^n(x)$ and $F_\mu^n(y)$.
   Therefore, if z is between x and y, then $F_\mu(z)$ is between $F\mu(x)$ and $F\mu(y)$, $F_\mu^2(z)$ is between $F^2\mu(x)$ and $F^2\mu(y)$, and so on. Thus the orbit of z remains in I for all n, making z an element of Λ. Since z is a generic element of the interval between x and y, this implies that

Λ contains the interval between x and y. This contradicts that Λ is totally disconnected.

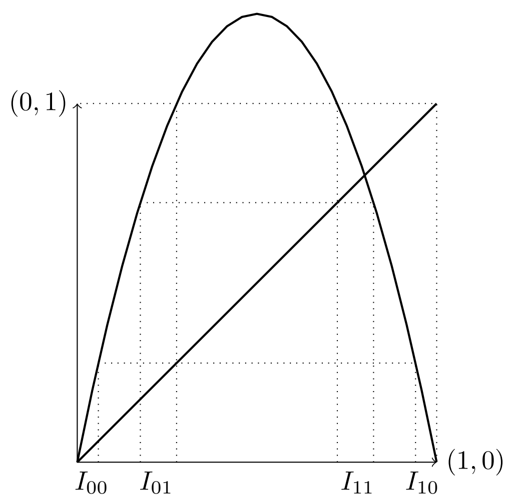
2. We will now show that S is onto.
   Let $\mathbf{s} = (s_0 s_1 s_2 \ldots)$. We want to find an point $x \in \Lambda$ such that $S(x) = \mathbf{s}$.
   As one might expect, given a closed interval $J \subseteq I$ we might want to talk about the preimage of J and more generally about the set

$$F_\mu^{-n}(J) = \{x \in I : F_\mu^n(x) \in J\},$$

   for each integer n. We note that as J is a closed interval, its preimage $F_\mu^{-1}(J)$ must have components in I_0 and I_1 which are themselves closed intervals as F_μ is symmetrical. For ease of notation, we define

$$\begin{aligned} I_{(s_0 s_1 \ldots s_n)} &= \{x \in I : x \in I_{s_0}, F_\mu(x) \in I_{s_1}, \ldots, F_\mu^n(x) \in I_{s_n}\} \\ &= I_{s_0} \cap F_\mu^{-1}(I_{s_1}) \cap \ldots \cap F_\mu^{-n}(I_{s_n}) \ . \end{aligned}$$

   The following graph depicts the four possibilities for $I_{(s_0 s_1)}$:



   We will use proof by induction to show that the $I_{(s_0 s_1 s_2 \ldots s_n)}$ form a nested sequence of nonempty closed intervals. We note that

$$I_{(s_0 s_1 s_2 \ldots s_n)} = I_{s_0} \cap F_\mu^{-1}(I_{(s_1 s_2 \ldots s_n)}) \ .$$

   By our induction hypothesis, $I_{(s_1 s_2 \ldots s_n)}$ is nonempty and closed. As we showed above, $F_\mu^{-1}(I_{(s_1 s_2 \ldots s_n)})$ consists of two closed intervals, one in each of I_0 and I_1. Thus $I_{s_0} \cap F_\mu^{-1}(I_{(s_1 s_2 \ldots s_n)})$ is a single closed interval, as I_{s_0} is really one of I_0 or I_1. These intervals are nested because

$$I_{(s_0 s_1 s_2 \ldots s_n)} = I_{(s_0 s_1 s_2 \ldots s_{n-1})} \cap F_\mu^{-n}(I_{(s_1 s_2 \ldots s_n)}) \subset I_{(s_0 s_1 s_2 \ldots s_{n-1})} \ .$$

Thus we find that $\bigcap_{n\geq 0} I_{(s_0 s_1 s_2 \ldots s_n)}$ is nonempty.
From this we can conclude that if $x \in \bigcap_{n\geq 0} I_{(s_0 s_1 s_2 \ldots s_n)}$ then $x \in I_{s_0}$, $F_\mu(x) \in I_{s_1}$, and so on for any integer n. As such, $S(x) = (s_0 s_1 \ldots)$. Since $\mathbf{s}$ is a generic point in Σ_2, this implies that S is onto.
It is important to note here that since S is one to one, $\bigcap_{n\geq 0} I_{(s_0 s_1 s_2 \ldots s_n)} = \{x\}$.

3. To show that S is continuous let $x \in \Lambda$, let $S(x) = (s_0 s_1 s_2 \ldots)$, and let $\epsilon > 0$. We know that there exists a positive integer N such that $\frac{1}{2^N} < \epsilon$. Consider all possible sequences $(t_0 t_1 \ldots t_N)$ and the associated intervals $I_{(t_0 t_1 \ldots t_N)}$. We note that all of these intervals are disjoint and Λ is a subset of their union. We find combinatorially that there are 2^{N+1} such intervals and note that $I_{(s_0 s_1 s_2 \ldots s_N)}$ is among them. Thus we may find a $\delta > 0$ such that if $|x - y| < \delta$ and $y \in \Lambda$, then $y \in I_{(s_0 s_1 s_2 \ldots s_N)}$. As such, $S(x)$ and $S(y)$ agree in their first $N+1$ terms, so that, by our earlier theorem,
$$d(S(x), S(y)) \leq \frac{1}{2^N} < \epsilon \ .$$

4. Finally, we want to show that S^{-1} is continuous.
Let $\epsilon > 0$. Consider a sequence $\mathbf{s} = (s_0 s_1 \ldots)$. Since the $I_{(s_0 s_1 \ldots s_n)}$ form a nested sequence of intervals that converges to $\{S^{-1}(\mathbf{s})\}$, we know that there exists some integer N such that $I_{(s_0 s_1 \ldots s_N)} \subset (S^{-1}(\mathbf{s}) - \epsilon, S^{-1}(\mathbf{s}) + \epsilon)$.
Let $\mathbf{r}$ be a sequence where $d(\mathbf{s}, \mathbf{r}) < \frac{1}{2^N}$. We know that $\mathbf{r}$ agrees with $\mathbf{s}$ in its first $N+1$ terms. As such $S^{-1}(\mathbf{r}) \in I_{(s_0 s_1 \ldots s_N)}$. Therefore $|S^{-1}(\mathbf{s}) - S^{-1}(\mathbf{r})| < \epsilon$.

□

Theorem 4.4.2. *$S \circ F_\mu = \sigma \circ S$.*

Proof. Recall that
$$F_\mu^{-n}(J) = \{x \in I : F_\mu^n(x) \in J\},$$
and that
$$I_{s_0 s_1 \ldots s_n} = I_{s_0} \cap F_\mu^{-1}(I_{s_1}) \cap \ldots \cap F_\mu^{-n}(I_{s_n}) \ .$$
Since S is one-to-one we know that $\bigcap_{n\geq 0} I_{s_0 s_1 \ldots s_n}$ consists of a single point in Λ which we will call x.
Since
$$I_{s_0 s_1 \ldots s_n} = I_{s_0} \cap F_\mu^{-1}(I_{s_1}) \cap \ldots \cap F_\mu^{-n}(I_{s_n})$$
and $F_\mu(I_{s_0}) = I$, we see that
$$F_\mu(I_{s_0 s_1 \ldots s_n}) = I_{s_1} \cap F_\mu^{-1}(I_{s_2}) \cap \ldots \cap F_\mu^{-n+1}(I_{s_n}) = I_{s_1 s_2 \ldots s_n} \ .$$
Thus
$$\begin{aligned} S \circ F_\mu(x) &= S \circ F_\mu \left(\bigcap_{n=0}^{\infty} I_{s_0 s_1 \ldots s_n} \right) \\ &= S \left(\bigcap_{n=1}^{\infty} I_{s_1 \ldots s_n} \right) \\ &= s_1 s_2 s_3 \ldots \\ &= \sigma \circ S(x) \ . \end{aligned}$$

□

4.5 Chaotic Logistic Maps

Thus our big result falls out:

Theorem 4.5.1. *A logisitic map F_μ with $\mu > 2+\sqrt{5}$ is chaotic.*

Proof. The shift map σ is chaotic. Any logistic map F_μ with $\mu > 2+\sqrt{5}$ is topologically conjugate with σ. As σ is chaotic, so too is F_μ. □

4.6 F_4 : A Special Case

Since we started by looking at the chaotic behavior of $F_4(x) = 4x(1-x)$(simply called g at the time), it behooves us to prove that F_4 is actually chaotic. Unfortunately, the itinerary does not serve as a topological conjugacy between F_4 and the shift map, as points do not escape I under F_4 (it is for this very reason that I used F_4 for the pictures in the first place). Instead, we must prove this one from the ground up, albeit with a few mappings very similar to conjugacies. Given the similarities of the following proof to several previous proofs, we will not dwell on the details for too long.

Theorem 4.6.1. *F_4 is chaotic.*

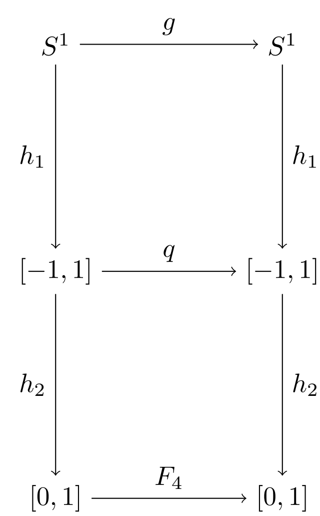
Proof. Recall that the function $g(\theta) = 2\theta$ defined on the unit circle, denoted S^1, is chaotic. Define $h_1 : S^1 \to [-1,1]$ by $h_1(\theta) = cos(\theta)$ and define $q : [-1,1] \to [-1,1]$ by $q(x) = 2x^2 - 1$. We see that

$$\begin{aligned}(h_1 \circ g)(\theta) &= \cos(2\theta) \\ &= 2\cos(\theta)^2 - 1 \\ &= (q \circ h_1)(\theta)\ .\end{aligned}$$

Define $h_2 : [-1,1] \to [0,1]$ by $h_2(t) = \frac{1}{2}(1-t)$, we see that

$$\begin{aligned}(F_4 \circ h_2)(t) &= 4\left(\frac{1}{2}(1-t)\right)\left(1-\left(\frac{1}{2}(1-t)\right)\right) \\ &= 4\left(\frac{1}{2}(1-t)\right)\left(\frac{1}{2}(1+t)\right) \\ &= (1-t)(1+t) \\ &= 1-t^2 \\ &= \frac{1}{2}(2-2t^2) \\ &= \frac{1}{2}(1-(2t^2-1)) \\ &= (h_2 \circ q)(t)\ .\end{aligned}$$

This gives us:



We will start by showing that F_4 is topologically transitive.
We note that as h_1 and h_2 are both surjective and continuous, $(h_2 \circ h_1)$ is also surjective and continuous. As such, if U and V are two open intervals in $[0,1]$, we can find open arcs U' and V' in S^1 such that $(h_2 \circ h_1)(U') = U$ and $(h_2 \circ h_1)(V') = V$. As g is topologically transitive there exists a positive integer N such that $g^N(U') \cap V' \neq \emptyset$. Therefore we see that

$$\begin{aligned} F_4^N(U) &= (F_4^N \circ h_2 \circ h_1)(U') \\ &= (h_2 \circ q^N \circ h_1)(U') \\ &= (h_2 \circ h_1 \circ g^N)(U') \ . \end{aligned}$$

Thus $F_4^N(U) \cap V = (h_2 \circ h_1)(g^N(U') \cap V')$, so that $F_4^N(U) \cap V \neq \emptyset$.
To see that f has sensitive dependence on initial conditions, let $x \in [0,1]$ and let U be an open set containing x. We know that the preimage of U under $(h_2 \circ h_1)$, let's call it U', is an open set. Recall from our proof that g is chaotic that since U' is an open set there exists a positive integer M such that $g^M(U')$ covers S^1. Therefore $F_4^M(U) = (h_2 \circ h_1 \circ g^M)(U')$ covers $[0,1]$. Thus we know that there certainly exists some point $y \in U$ such that $|F_4^M(x) - F_4^M(y)| \geq \frac{1}{2}$.
To show that periodic points are dense under F_4, we will first show that $(h_2 \circ h_1)$ maps periodic points to periodic points. Let $p \in$ be a periodic point of period n under g. Then $g^n(x) = x$. We see that

$$(h_2 \circ h_1)(p) = (h_2 \circ h_1 \circ g^n)(p) = (F_4^n \circ h_2 \circ h_1)(p) \ .$$

Therefore $(h_2 \circ h_1)(p)$ is a periodic point of period n under F_4 by definition.
Therefore if we pick an open interval U in $[0,1]$, we know that its preimage U', also an open set, will contain some periodic point p. Thus U contains the periodic point $(h_2 \circ h_1)(p)$.
As such, periodic points are dense under F_4. □

Note that $(h_2 \circ h_1) : S^1 \to [0,1]$ is not a topological conjugacy, as h_1 is a two-to-one function at most points. Nevertheless, the method we used to show that F_4 is chaotic relies on the same basic principles as our method of using a topological conjugacy.

Conclusion: Seeing Order

Thus we have come full circle, showing why our two original functions F_2 and F_4 have such drastically different behavior. Like Edward Lorenz, we started by looking at some alarming data and tried to unravel why it was occurring. While we returned to the behavior of logistic maps a number of times, along the way we also developed a vocabulary for describing discrete dynamical systems in general, rigorously defined chaos, and constructed a powerful tool for proving that dynamical systems are chaotic. In the end, we saw that chaotic systems have a very particular kind of order underlying their seemingly disorderly behavior: their very nature makes any two sets of distinct initial conditions lead to drastically different outcomes, and causes the orbits of points travel all over their domains.

As an application-driven field, chaos theory invites us to seek new patterns to unravel. Weather forecasting, for instance, has improved dramatically thanks to developments in chaos theory. Chaos might be hiding where we never expect to find it; perhaps we just need to run a few simulations to start our search.

Further Research

It is my hope that a motivated calculus student could read this paper, albeit with frequent visits to the Appendix. Perhaps it will spark their interest for their own eventual senior project. I feel that using the itinerary as a topological conjugacy for a wider variety of systems could be a particularly ripe topic. Additionally, given how much work I put into generalizing this paper's results to metric spaces (unlike any of the sources I was drawing on), I would hope that future students could find other chaotic maps defined on strange metric spaces.

Acknowledgments

I would like to thank Alec Foote for carefully reading and editing this paper every single week this semester, Professor Doug Hundley for guiding me through a topic that was entirely foreign to me before I started this project, and Professor Barry Balof for teaching the Senior Project course this semester.

Appendix: Necessary Background for Chaos

This appendix is intended as a primer on some basic concepts about *metric spaces* and *topology* which we require to discuss chaos. As it is intended for readers entirely unfamiliar with these ideas, this appendix features many examples explained at length. Many of the following definitions have been adapted from Gordon [5].

Metric Spaces

Before we can discuss whether a map over a space is chaotic we need to have a means of measuring the distance between two elements in that space. To that end we have the concept of a *metric space*:

Definition 4.6.2. (Gordon [5])
A **metric space** (X, d) consists of a set X and a function $d : X \times X \to \mathbb{R}$, called a **metric**, that satisfies the following four properties.

1. $d(x, y) \geq 0$ for all $x, y \in X$.
2. $d(x, y) = 0$ if and only if $x = y$.
3. $d(x, y) = d(y, x)$ for all $x, y \in X$.
4. $d(x, y) \leq d(x, z) + d(y, z)$ for all $x, y, z \in X$.

Example 4.6.3. Perhaps the most familiar metric space is $(\mathbb{R}, d_1)$ where $\mathbb{R}$ denotes the set of real numbers and $d_1(x, y) = |x - y|$, where x and y are real numbers. Let us verify that $d_1(x, y)$ is a metric by checking each property:

1. We know that $|t| \geq 0$ for any real number t, so clearly $d_1(x, y) = |x - y| \geq 0$ for all $x, y \in \mathbb{R}$.
2. We recall that $|0| = 0$, so that $d_1(x, x) = |x - x| = 0$. Likewise, we recall that $|t| > 0$ when $t \neq 0$, so that if $x \neq y$ then $d_1(x, y) = |x - y| > 0$.
3. We know that $|x - y| = |y - x|$ for all $x, y \in \mathbb{R}$.
4. We know that $|a + b| \leq |a| + |b|$ for all $a, b \in \mathbb{R}$. As such, $|x - y| \leq |x - z| + |z - y|$ for all $x, y, z \in \mathbb{R}$.

Topology in a Metric Space

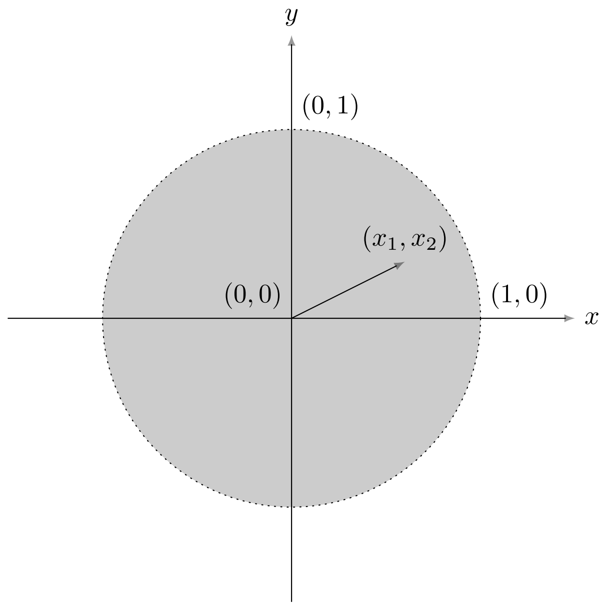
While thus far we have focused on how our maps act on individual points in a dynamical system, we also need a firm understanding of how our maps act on sets of points. Before we can do this, however, we need to develop a vocabulary for discussing different types of sets.

Definition 4.6.4. (Gordon [5])
Let (X, d) be a metric space.
Let $v \in X$ and let $r > 0$. The **open ball** centered at v with radius r is defined by

$$B_d(v, r) = \{x \in X : d(x, v) < r\} .$$

Example 4.6.5. Consider the familiar metric space $(\mathbb{R}, d_1)$ where $d_1(x, y) = |x-y|$ for all $x, y \in \mathbb{R}$. We see that $B_{d_1}(0, 1) = (-1, 1)$.

Example 4.6.6. Consider the metric space $(\mathbb{R}^2, d_2)$ where $d_2(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$ for all points $x = (x_1, x_2)$ and $y = (y_1, y_2)$ in $\mathbb{R}^2$. We can see that $B_{d_2}((0, 0), 1)$ is the interior of the circle of radius 1 centered at the origin. In the following image we can see that $d_2((0, 0), (x_1, x_2)) < 1$, so that $(x_1, x_2) \in B_{d_1}((0, 0), 1)$.



Definition 4.6.7. (Gordon [5])
Let (X, d) be a metric space, let $E \subseteq X$, and let $x \in X$.

1. The point x is an **interior point** of E if there exists an $r > 0$ such that $B_d(x, r) \subseteq E$.
2. The point x is a **limit point** of E if for each $r > 0$, the set $E \cap B_d(x, r)$ contains a point of E other than x.
   An equivalent definition is that x is a limit point of E if there exists a sequence in $E \setminus \{x\}$ that converges to x.
3. The set E is **open** in (X, d) if each point of E is an interior point of E.
4. The set E is **closed** in (X, d) if E contains all of its limit points.

Example 4.6.8. Consider the metric space $(\mathbb{R}, d_1)$. Let $E = (0, 1)$. We see that .5 is an interior point of E as $B_{d_1}(.5, .25) = (.5 - .25, .5 + .25) \subseteq (0, 1)$.

Example 4.6.9. Consider the metric space $(\mathbb{R}, d_1)$. Let $G = \{\frac{1}{n} : n \in \mathbb{Z}^+\}$. We will show that 0 is a limit point of G.
Let $\epsilon > 0$. We note that there exists a positive integer N such that $\frac{1}{N} < \epsilon$. Therefore $\frac{1}{N} \in B_{d_1}(0, \epsilon)$. Thus for every $\epsilon > 0$, $B_{d_1}(0, \epsilon)$ contains a point in E other than 0.

Example 4.6.10. We will show that (a, b), where $a, b \in \mathbb{R}$, is an open set in $(\mathbb{R}, d_1)$.
Let $x \in (a, b)$. Since $a < x < b$, we know that $b - x > 0$ and $x - a > 0$. Let $r = \min\{b - x, x - a\}$.
We see that $B_{d_1}(x, r) = (x - r, x + r) \subseteq (a, b)$.

Example 4.6.11. As a counterexample, we will show that $(0, 1]$ is not an open as 1 is not an interior point of $(0, 1]$
Let $\epsilon > 0$. Clearly $1 + \frac{\epsilon}{2} \in B_{d_1}(x, \epsilon)$. We note, however, that $1 + \frac{\epsilon}{2} > 1$, so that $1 + \frac{\epsilon}{2} \notin (0, 1]$.

Example 4.6.12. We will show that $[a, b]$, where $a, b \in \mathbb{R}$, is a closed set in $(\mathbb{R}, d_1)$.
Let $x \notin [a, b]$. We want to show that x is not a limit point of $[a, b]$.
Since $x \notin [a, b]$, either $x < a$ or $b < x$. Suppose that $x < a$. Let $\epsilon = a - x$ and consider the open ball $B_{d_1}(x, \epsilon) = (x - \epsilon, x + \epsilon)$. Let $y \in B_{d_1}(x, \epsilon)$. We know that $y < x + \epsilon = x + (a - x) = a$. Therefore $y \notin [a, b]$. As such, the open ball $B_{d_1}(x, \epsilon)$ does not contain any point in $[a, b]$, meaning that y is clearly not a limit point of $[a, b]$.
We can see rather easily that a similar argument applies if $b < x$, so we will not bother with that case here.
Thus $[a, b]$ contains all of its limit points, and is therefore closed.

Example 4.6.13. As a counterexample, we will show that the set $G = \{\frac{1}{n} : n \in \mathbb{Z}^+\}$ is not closed.
As we showed above, 0 is a limit point of G. We see, however, that $0 \notin G$.

In order to test your comfort with these terms, attempt to prove the following helpful theorem.

Theorem 4.6.14. *Any open ball is an open set.*

As we will be discussing mappings of open and closed sets, it is helpful to know the relationship between the two in a generic metric space.

Theorem 4.6.15. *Let (X, d) be a metric space and let $E \subseteq X$.*

1. *The set E is open in (X, d) if and only if $X \setminus E$ is closed in (X, d).*
2. *The set E is closed in (X, d) if and only if $X \setminus E$ is open in (X, d).*

Proof. We will start by showing that if E is open in (X, d), then $X \setminus E$ is closed in (X, d).
Let $x \in E$. As x is necessarily an interior point of E we know that there exists some $r > 0$ such that $B_d(x, r) \subseteq E$. This implies that $B_d(x, r) \cap X \setminus E = \emptyset$. Therefore x is clearly not a limit point of $X \setminus E$. Thus $X \setminus E$ must contain all of its limit points and be closed.
Moving to showing that if $X \setminus E$ is closed, then E is open, suppose that x is a point in E that is not an interior point of E. This means that for every $r > 0$, $B_d(x, r) \cap X \setminus E \neq \emptyset$. Since $x \notin X \setminus E$, this means that for every $r > 0$, $B_d(x, r)$ contains a point in $X \setminus E$ other than x. Thus x is a limit point of $X \setminus E$. As $X \setminus E$ is closed, however, this means that $x \in X \setminus E$, a contradiction. Thus every point in E is an interior point of E, making E open.
The proof of the second statement follows immediately from the first statement and the fact that $X \setminus (X \setminus E) = E$. □

Density

Definition 4.6.16. (Gordon [5])
Let (X, d) be a metric space and let $E \subseteq X$.

1. The **derived set** of E, denoted E', is the set of all limit points of E.

2. The **closure** of E, denoted $\overline{E}$, is the set $E \cup E'$.

Example 4.6.17. Consider the metric space $(\mathbb{R}, d_1)$.
If $E = (a, b)$, then $E' = [a, b]$ and $\overline{E} = [a, b]$.

Example 4.6.18. Consider the metric space $(\mathbb{R}, d_1)$.
If $G = \{\frac{1}{n} : n \in \mathbb{Z}^+\}$, then $G' = \{0\}$ and $\overline{G} = \{\frac{1}{n} : n \in \mathbb{Z}^+\} \cup \{0\}$.

Definition 4.6.19. (Trench [6]) A set U is **dense** in a set S if $U \subseteq S \subseteq \overline{U}$.

While this definition is certainly intuitive, there are easier ways of showing that a set U is dense in a set S. To this end we have the following theorem:

Theorem 4.6.20. *Let (X, d) be a metric space and let $S \subseteq X$.
If each open ball that intersects S contains a point in U, then U is dense in S.*

Proof. Suppose that every open ball that intersects S contains a point in U. We want to show that every point in S is either a point in U or a limit point of U.
Let $s \in S$. If s is a point in U, then we're finished, so suppose that s is not a point in U. Consider a sequence $\{\epsilon_n\}$, where each $\epsilon_n > 0$, that converges to 0. As every open which intersects S contains a point in U, for each integer n there exists an $x_n \in U$ such that $x_n \in B_d(s, \epsilon_n) \cap S$. Consider, then, the sequence $\{x_n\}$. Let $\epsilon > 0$. Since $\{\epsilon_n\}$ converges to 0 there exists a positive integer N such that $\epsilon_n < \epsilon$ if $n \geq N$. Therefore $d(x_n, s) < \epsilon_n < \epsilon$ if $n \geq N$. As such, s is the limit point of a sequence of points in U. Therefore every point in S is either a point in U or a limit point of U, that is $U \subseteq S \subseteq \overline{U}$. □

As an example, we have the following theorem and proof adapted from Trench [6].

Theorem 4.6.21 (Trench [6])**.** *The rational numbers are dense in the reals.*

Proof. Let a and b be two real numbers. We want to show that there exists a rational number $p/q \in (a, b)$.
By the Archimedean Property, we know that there exists a positive integer q such that $q(b-a) > 1$.
Let p be the smallest integer such that $p > qa$. Then $p - 1 \leq qa$, so that

$$qa < p \leq qa + 1.$$

As $q(b - a) > 1$, clearly

$$qa < p < qa + q(b - a) = qb.$$

Thus

$$a < p/q < b\ .$$

□

Continuity

Finally, it is helpful to develop a vocabulary for describing how functions map points which are close together to other points which are close together. To this end, we formulate a generic concept of *continuity* in a metric space.

Definition 4.6.22. Let (X, d_1) and (Y, d_2) be metric spaces.

1. A function $f : X \to Y$ is **continuous** at $x_0 \in X$ if for each $\epsilon > 0$ there exists a $\delta > 0$ such that $f(x) \in B_{d_2}(f(x_0), \epsilon)$ for all $x \in B_{d_1}(x_0, \delta)$.

2. A function $f : X \to Y$ is continuous on X if it is continuous at each point of X.

We can see fairly easily how this concept carries over to $(\mathbb{R}, d_1)$. We may, however, want a more interesting example. As such, we will show that the *shift map* $\sigma : \Sigma_2 \to \Sigma_2$ (see Section 4.1.2) is continuous.

Theorem 4.6.23. *The shift map* $\sigma : \Sigma_2 \to \Sigma_2$ *is continuous.*

Proof. Let $\epsilon > 0$ and $\mathbf{s} = (s_0 s_1 s_2 \ldots)$. We know that there exists a positive integer N such that $\frac{1}{2^N} < \epsilon$. Let $\delta = \frac{1}{2^{N+1}}$. If $\mathbf{t} = (t_0 t_1 t_2 \ldots)$ is a sequence in Σ_2. such that $d(\mathbf{s}, \mathbf{t}) < \delta$, then $s_i = t_i$ for $i \leq N + 1$. Therefore, the ith entries of $\sigma(\mathbf{s})$ and $\sigma(\mathbf{t})$ agree for $i \leq N$, so that

$$d(\sigma(\mathbf{s}), \sigma(\mathbf{t})) \leq \frac{1}{2^N} < \epsilon.$$

□

References

[1] Kathleen T. Alligood, Tim D. Sauer, and James A. Yorke, *Chaos: An Introduction to Dynamical Systems*, Springer-Verlag, New York, 1997.

[2] Robert L. Devaney, *A First Course in Chaotic Dynamical Systems: Theory and Experiment*, Westview Press, Boulder, Co., 1992.

[3] Robert L. Devaney, *An Introduction to Chaotic Dynamical Systems*, Addison-Wesley Publishing Company, Reading, Ma., 1989.

[4] James Gleick, *Chaos: Making a New Science*, Penguin Books, New York, 1987.

[5] Russell A. Gordon, *Real Analysis: A First Course*, Addison-Wesley Publishing Company, 2002.

[6] William F. Trench, *Introduction to Real Analysis*, The American Institute of Mathematics' Open Textbook Initiative, August 2015.

Alphabetical Index