M375T: Dynamical Systems

Reese Lance

Spring 2020

Abstract

Class taught by Lewis Bowen, notes taken by Reese Lance. The notes are not live texed. They are post-mortem-texed, that is, taken by hand during class and typed later. Some of my own thoughts are interjected, but quite rarely. I initially thought to try to separate my thoughts from the professor's but it becomes too difficult. As such I will also try to expand on examples which are mentioned in passing in class, spell out proofs which are glossed over, and add insight where I think it is helpful. This helps to justify the existence of this set of notes, as opposed to live-texed notes. Especially because some of my own content is interspersed throughout these notes, any corrections, questions, comments, suggestions, etc., can be sent via email (reese.lance@utexas.edu) or if you can find any other way to communicate with me, that is also fine. At the moment I'm trying to get the notes written, and worrying about making the format not look like trash later. I'm also not going to track theorem and lemma numbers, as I think that's mostly useless. If a proof somewhere says "Applying Theorem X", it can usually be determined from context what theorems need to be invoked, and if the reader doesn't find it readily apparent, then searching for the theorem in question will be a valuable experience. Also I always forget to write down the numbers. Also as I revisit and add in more stuff, the numbering becomes involved and I'd have to actually figure out how to number properly instead of just manually putting numbers, which is what would have been the plan. Thanks to Arun Debray whose formatting choices inspired my own.

Table of Contents

CHAPTER 1

Overview and a little bit of history

Lecture 1, Jan 21. Given a set Ω and a map $f : \Omega \to \Omega$, there is not much we can say about this scenario. If we embed $\Omega \subset \mathbb{R}^n$, we gain a lot more structure, in particular there is a smooth structure, so that we can speak about smoothness of f and do calculus if Ω is a nice enough submanifold. In this class, we will often consider $\Omega = I = [0, 1]$. Often in real world practice, studying dynamical systems necessarily involve large dimensions[1], but to really understand the mathematics of dynamical systems, we consider one dimension often. One of the main questions of dynamical systems is this:

$$\text{Given } f : \Omega \to \Omega, \text{ what can we say about } f^n(x) \text{ when } n \text{ becomes large}$$

Here $f^n(x)$ denotes the process of applying f n times. At least my first instinct was "nothing, that is too general". It seems like dynamical systems is broadly the study of 'systems' $f : \Omega \to \Omega$ such that the above question has an interesting answer. Here is one such example:

Example: Define $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = x^2$. Then we know

$$\lim_{n\to\infty} f^n(x) = \begin{cases} \infty & |x| > 1 \\ 1 & |x| = 1 \\ 0 & |x| < 1 \end{cases}$$

Looking at the positive half of the real line, points greater than 1 diverge to infinity, the point 1 itself is a fixed point of f, and points less than 1 converge to 0, and the negative half mirrors the positive half.

Example: Define $S^1 = \{e^{i\theta} \mid 0 \le \theta \le 2\pi\}$, the unit circle in the complex plane. Then for any $\alpha \in \mathbb{R}$, we define R_α, the rotation map by an angle α:

$$\begin{aligned} R_\alpha &: S^1 \to S^1 \\ e^{i\theta} &\mapsto e^{i(\theta+\alpha)} \end{aligned}$$

this has the effect of rotating the entire unit circle by an angle θ. This map is an isometry, that is, it preserves the distance between any two points, taking the usual metric by embedding $S^1 \subset \mathbb{R}^2$. But what can we say about its long term behavior? We should probably

[1] For example in classical mechanics, the phase space of a system generally has $6n$ dimensions for n particles, three position and three velocity degrees of freedom.

guess that it depends on α. For example, if $\alpha = \frac{2\pi p}{q}$, for some $p, q \in \mathbb{Z}$, then $R_\alpha^q = Id$, because

$$\begin{aligned}
R_\alpha \circ \cdots \circ R_\alpha(e^{i\theta}) &= R_\alpha \circ \cdots \circ R_\alpha(e^{i(\theta+\alpha)}) \\
&= R_\alpha \circ \cdots \circ R_\alpha(e^{i(\theta+2\alpha)}) \\
&\vdots \\
&= e^{i(\theta+q\alpha)} \\
&= e^{i(\theta+q\frac{2\pi p}{q})} = e^{i\theta}e^{2\pi i p} = e^{i\theta}
\end{aligned}$$

If α is not a rational multiple of 2π, i.e. it is irrational, then there will not be a fixed point. We denote the orbit[2] of an element $x = e^{i\theta}$ as

$$O_x = \{f^n(e^{i\theta}) \mid n \in \mathbb{Z}\}$$

The orbit of any point under this action is dense, in the following sense: Take any point in S^1. Then any neighborhood of this point[3] has infinite intersection with O_x. This is a little more difficult to see, but eventually, the distribution becomes equidistributed, meaning that a point in the orbit, given large enough n, is equally likely to be in a certain arc of the circle as any other, provided they have the same arc length.

Example: Define

$$\begin{aligned}
D : S^1 &\to S^1 \\
e^{i\theta} &\mapsto e^{2i\theta}
\end{aligned}$$

This is known as the doubling map, as it doubles the angle of any given point in the circle. What does this map actually do to the unit circle? It maps the upper half to all of S^1, for example, and the first quartile to the upper half, and so on. We will now show some elementary properties of this map:

Lemma: *For any $p \in \mathbb{R}$, there is a point in S^1 which has period p.*

Proof: The point 1 trivially satisfies this condition, but the point is that there is a non-trivial point which does this. In fact, there will be infinitely many points in the circle which satisfy this condition: If $p \in \mathbb{R}$ is the desired period, then consider the point, for any $k \in \mathbb{Z}$,

$$e^{i\theta}, \quad \theta = \frac{\pi k}{2^{p-1} - \frac{1}{2}}$$

Compute

$$D^p(e^{i\theta}) = e^{i\frac{2^p \pi k}{2^{p-1}-\frac{1}{2}}} = e^{\frac{i\pi k}{2^{p-1}-\frac{1}{2}}} = e^{i\theta}$$

[2]Bowen states that this nomenclature comes from studying the orbits of celestial objects. It made me think of group actions. Indeed it is used in the same spirit as in group actions.

[3]Which can be thought of as an open arc on the circle containing the point in question.

as required.

$\square$

For example, if we set $p = 5$, then $\theta = \frac{2\pi}{31}$, and

$$D^5(e^{i\theta}) = e^{\frac{i64\pi}{31}} = e^{\frac{i62\pi}{31}} e^{\frac{i2\pi}{31}} = e^{\frac{i2\pi}{31}}$$

There are also points which may not be periodic, but after applying D a couple times, they become periodic. For a trivial example, the point -1 is not periodic, but

$$D(-1) = 1$$

which is periodic. There are also pairs of distinct points whose distance goes to 0 as $n \to \infty$.

Example: Let $\Sigma_2 = \left\{(x_i)_{i=1}^{\infty} \mid x_i \in \{0,1\}\right\}$ be the set of sequences of points in $\{0,1\}$, and define the shift map $\sigma : \Sigma_2 \to \Sigma_2$ by

$$\sigma(x_1, x_2, \dots) = (x_2, x_3, \dots)$$

Also define

$$\Phi : \Sigma_2 \to S^1$$

via

$$\Phi(x_i) = e^{2\pi i \sum_i x_i 2^{-i}}$$

Lemma: *There is a commutative diagram*

$$\begin{array}{ccc} \Sigma_2 & \xrightarrow{\sigma} & \Sigma_2 \\ \Phi\downarrow & & \downarrow\Phi \\ S^1 & \xrightarrow[D]{} & S^1 \end{array}$$

Proof: We check

$$(D \circ \Phi)(x_i) = D(e^{2\pi i \sum_i x_i 2^{-i}}) = (e^{2\pi i \sum_i x_i 2^{-i+1}})$$
$$e^{2\pi i (x_1 + \frac{x_2}{2} + \frac{x_3}{4} + \dots)}$$
$$= e^{2\pi i (\frac{x_2}{2} + \frac{x_3}{4} + \dots)} e^{2\pi i x_1} = e^{2\pi i (\frac{x_2}{2} + \frac{x_3}{4} + \dots)}$$

Similarly,

$$(\Phi \circ \sigma)(x_i) = \Phi(x_2, x_3, \dots)$$
$$= e^{2\pi i (\frac{x_2}{2} + \frac{x_3}{4} + \dots)}$$

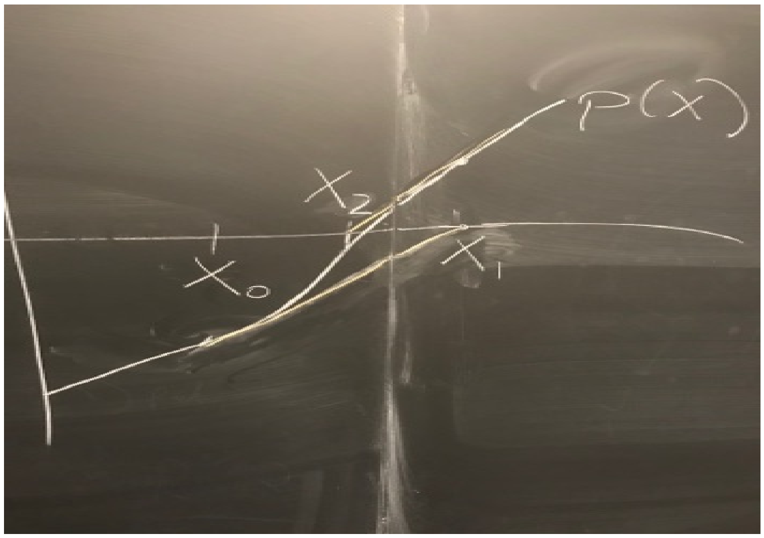
as required.

$\square$

Example: Newton's Method: Let

$$p(x) = \sum_{n=0}^{N} a_n x^n$$

Suppose you want to find a root of this polynomial. Then Newton's method says to make an initial guess for the root, x_0. Then we want to make better and better guesses, until they finally converge to a root. How do we make a better guess? Take the tangent line at x_0. If x_0 is not a critical point, then the tangent line will intersect the x-axis at a point, call it x_1. Then take the tangent line at x_1, and its intersection with the x-axis will be x_2, and so on.



How can we be sure this process converges? The picture should give intuition: Each successive choice is closer to the desired root. For an actual proof, we will set

$$x_{i+1} = x_i - \frac{p(x_i)}{p'(x_i)}$$

and $N(x) = x - \frac{p(x)}{p'(x)}$. Then want to show

$$\lim_{n\to\infty} N^n(x) = p \quad \text{and} \quad P(p) = 0$$

Theorem: *There exists a root p of P and $\epsilon > 0$, such that*

$$|x_0 - p| < \epsilon \Rightarrow \lim_{n\to\infty} N^n(x_0) = p$$

□

So Newton's Method works! And can be used to find roots of polynomials[4].

Now for a little bit of history: We all know the famous equation $F = ma$. This is a differential equation, so if you write down all the forces in a given system, and manage to solve the differential equation, if you have suitable initial conditions,[5] you can describe the time evolution of your system. For example, if one were to write down force diagrams

[4]This is in general an impossible task, as we learn via Galois theory for degree 5 or higher polynomials.
[5]In this case, you need 2 initial conditions to uniquely specify a solution.

for our solar system, there would need to be 54 initial conditions to describe the evolution of our system, since there are 3 position coordinates, 3 velocity coordinates, and 1 sun + 8 planets[6]. In general, the phase space of an n-body system in 3D will be $\mathbb{R}^{6n}$.
Poincare's viewpoint: We should study $\Omega \subset \mathbb{R}^n$ and all possible trajectories in the phase space. In classical physics, we view dynamics as deterministic: If we have the initial state of the system, and the differential equation describing its dynamics is known, then the future state of the system is completely determined. Poincare held that determinism in this sense is wrong because dynamics are inherently chaotic: If there is a small error in the measurement of the initial conditions, as there must be because there is no such thing as a perfect measurement, then that error can propagate exponentially when predicting the future states of the system, i.e. a small error can become large in a chaotic system. For example, there are many animations online of the double pendulum which show close initial conditions that become chaotic very quickly.
Historical aside over.
Here are a series of definitions of objects we will see frequently in the course:
Definition: Let I, J be intervals in $\mathbb{R}$. A <u>homeomorphism</u> from I to J is a continuous, invertible map whose inverse is also continuous.
Definition: Given a set $M \subset \mathbb{R}$, then

$$C^r(M) := \{f : M \to \mathbb{R} \mid f \text{ is } r \text{ times continuously differentiable}\}$$

ie the $f'^{(r)}(x)$ exists and is continuous. A map $f : M \to \mathbb{R}$ is said to be C^r smooth, or simply C^r if it lies in $C^r(M)$.
Definition: A C^r diffeomorphism is an invertible C^r map such that its inverse is also C^r.
Example: $f(x) = x^3$ is invertible, with inverse $f^{-1}(x) = x^{\frac{1}{3}}$. As such, it is a homeomorphism, since both x^3 and $x^{\frac{1}{3}}$ are continuous. However, it is not a C^1 diffeomorphism because $(f^{-1}(x))' = \frac{1}{3}x^{\frac{-2}{3}}$, which is not even defined at the origin, and is thus not differentiable at the origin. Away from the origin, however, this function is a C^∞ diffeomorphism.
Theorem: *Let $I = [a, b] \subset \mathbb{R}$, and $f : I \to I$ is continuous. Then f has a fixed point, i.e. a point $p \in I$ such that $f(p) = p$.*
Proof: We offered a visual proof which relies on showing the plot $y = x$ intersects the plot of f. It can be formalized using the intermediate value theorem:
Theorem: *If $f : [a, b] \to \mathbb{R}$ is continuous and $z \in [f(a), f(b)]$, then $\exists c \in [a, b]$ such that $f(c) = z$.*
In this setting, IVT can be thought of as the statement "The image of a connected set through a continuous function is a connected set."
If we apply IVT to the function $g(x) := f(x) - x$, then $g(a) < 0, g(b) > 0$, so by IVT, there must be a point c such that $g(c) = 0$, which shows the result. c is the fixed point.

$\square$

[6] RIP Pluto.

CHAPTER 2

MVT, CMT, and Applications

Lecture 2, Jan 23. **Theorem:** *(Mean Value Theorem) If $f : [a,b] \to \mathbb{R}$ is C^1, then $\exists\, c \in [a,b]$ such that*

$$\frac{f(b)-f(a)}{b-a} = f'(c)$$

Theorem: *(Contraction Mapping Theorem) Let $f : [a,b] \to [a,b]$, and assume $|f'(x)| < 1 \forall\, x \in [a,b]$. Then f has a fixed point, and*

$$\lim_{n\to\infty} f^n(x) = x_0 \qquad \forall\, x \in [a,b]$$

Proof: We already know the interval $[a,b]$ has the fixed point property[1]. Then by MVT,

$$\left|\frac{f(x)-f(x_0)}{x-x_0}\right| = |f'(c)|$$
$$\Rightarrow |(f(x)-f(x_0)| = f'(c)|x-x_0|$$

Now $f([a,b])$ is compact, so it contains its supremum. So let $\max_{x\in[a,b]} f(x) = \lambda$. Then

$$|f(x)-f(x_0)| \le \lambda$$
$$\Rightarrow (f^n)'(c) \le \lambda^n$$

So iterating shows

$$|f^n(x) - f^n(x_0)| = |(f^n)'(c)|x-x_0| \le \lambda^n \to 0$$

But x_0 is a fixed point, so

$$|f^n(x) - f^n(x_0)| \to 0$$

so all points are attracted to the fixed point, which is unique.

□

We can apply this to Newton's method: Let p be a polynomial of degree n, and

$$N(x) = x - \frac{p(x)}{p'(x)}$$

[1] You can get this from Brouwer fixed point theorem for D^0.

We want to apply CMT here to show that this process converges to the root. We also observe x_0 is a fixed point of $N(x)$, and $N'(x_0) = 0$, so that CMT applies, and we can be sure that the process converges to the root.
CMT can be used to prove the existence of solutions to PDE. In this setting, we often take Ω to be a complete metric space[2]. Suppose

$$f : \Omega \to \Omega$$

satisfies

$$d(f(x), f(y)) \leq \lambda d(x,y) \quad \forall x, y \in \Omega$$

where $\lambda \in [0,1]$.
Definition: A limit point of a topological space S is a point $x \in S$ such that $\exists\, (x_n)_{n\in\mathbb{N}}, x_n \in S/\{x\}$ with

$$\lim_{n\to\infty} x_n = x$$

So essentially, a limit point is a point which is the limit of some sequence. Not we say that the sequence can't contain the point x, otherwise every point would be a limit via the constant sequence $x_i = x_0$.
Example: $S = \left\{\frac{1}{n}\right\}_{n=1}^{\infty} \cup \{0\}$. 0 is a limit point, but $\frac{1}{2}$ is not.
Definition: The closure[3] of a topological space, S, is $\overline{S} = S \cup \{\text{limit points}\}$.
=**Definition:** The Cantor set is defined as

$$C = \left\{ \sum_{i=1}^{\infty} \epsilon_i 3^{-1} \mid \epsilon_i \in \{0,2\} \right\}$$

C contains $\frac{1}{3}$, but not under the expansion $\frac{1}{3} + 0 + 0 + \dots$, because you can't have a 1 coefficient. Instead it is included under the expansion $\frac{1}{3} = 0 + 2\frac{1}{3} + 2\frac{1}{3^2} + \dots$. Some properties of the cantor set:
i) C is closed in I.
ii) $[0,1]/C$ is dense in $[0,1]$.
Here we go into an aside about bump functions, but I made a note about this a while ago based on the chapter in Lee, so I'm just going to insert that here. The spirit is the same, and I went through and proved some things. There's also a bit of background on manifolds and partitions of unity: The copy paste starts here

In topology, we have a notion of combining locally smooth functions on a topological space into a global continuous function:

[2]Not clear at the moment why we need completeness, but maybe that has more to do with PDE than with CMT.

[3]There are a couple of equivalent definitions of closure, but I think this one is most amenable to a dynamical systems perspective, where we are interested in sequences.

Gluing Lemma for Continuous Maps: *Given X, Y topological spaces, and either one of the conditions:*
a) B_i is a finite collection of closed subsets of X such that $\bigcup B_i = X$
b) B_i is a collection of open subsets of X such that $\bigcup B_i = X$ and suppose there exists maps $f_i : B_i \to Y$ which agree on all intersections $B_i \cap B_j$. Then there exists a unique continuous map $f : X \to Y$ such that $f|_{B_i} = f_i$ for all i.

We have a generalization to smooth manifolds for smooth maps, which only works on open sets.

Gluing Lemma for Smooth Maps: *Let M, N smooth manifolds with or without boundary, and let $\{U_\alpha\}$ be an open cover of M. Suppose for each α, there exists a smooth map $f_\alpha : U_\alpha \to N$ such that the maps agree on all intersections $U_\alpha \cap U_{\alpha'}$. Then there exists a unique smooth map $f : M \to N$ such that $f|_{U_\alpha} = f_\alpha$ for all α.*

An immediate example of this lemma not being applicable given a closed cover is the case of the absolute value function on $\mathbb{R}$. It is smooth on the closed sets $(-\infty, 0], [0, \infty)$, but gluing it together, i.e. $f(x) = |x|$ is obviously not smooth at $x = 0$, it is not even differentiable. However, it is continuous, so we can use the gluing lemma for continuous maps. In practice, the gluing lemma on smooth maps is cumbersome to use, since it requires that we construct maps which agree on large open subsets of a manifold. We will now discuss *partitions of unity*, which will allow us to glue together locally smooth maps into global smooth maps without checking this condition. To do so, we will construct what is called a *bump function* through a rather lengthy process, in gory detail:
Lemma: *The function $f : \mathbb{R} \to \mathbb{R}$ defined by*

$$f(x) = \begin{cases} e^{-1/t} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

is smooth.
Proof: We want to show $f \in C^\infty(\mathbb{R})$, which we will do by demonstrating that the kth derivative is continuous for all k. First, it is clear that the function is smooth on $\mathbb{R}/0$, so we need to verify smoothness at 0. The left handed derivative is clearly 0, so we will show that the right hand derivative is also 0 at $x = 0$ and continuous. We will prove by induction that for $x > 0$, the kth derivative is given by

$$f^{(k)}(x) = p_k(x) \frac{e^{-1/x}}{x^{2k}}$$

The 0th case is clear, as it reduces to the definition of f we gave. Assume the inductive hypothesis. Then

$$\begin{aligned}
f^{(k+1)}(x) &= \frac{d}{dx}f^{(k)}(x) = \frac{d}{dx}p_k(x)\frac{e^{-1/x}}{x^{2k}} \\
&= p_k'(x)\frac{e^{-1/x}}{x^{2k}} + p_k(x)\Big[\frac{e^{-1/x}x^{-2}}{x^{2k}} - \frac{2xe^{-1/x}}{x^{2k+1}}\Big] \\
&= \left(x^2p_k'(x) + p_k(x) - 2kxp_k(x)\right)\frac{e^{-1/x}}{x^{2(k+1)}}
\end{aligned}$$

The term in the parenthesis is clearly a polynomial of order $k+1$ at most, with the first and last term containing the highest power terms. Now returning to the f defined on the whole real line, we want to show that $f^{(k)}(0) = 0$, which we also do by induction. It is clear by definition in the base case. Assume the inductive hypothesis. Clearly, the left hand derivative is 0. We need to show the right hand derivative is also 0. To do so, we know that for $x > 0$,

$$\begin{aligned}
f^{(k+1)}(0) &= \lim_{x\to 0}\frac{f^{(k)}(0+x) - f^{(k)}(0)}{x} \\
&= \lim_{x\to 0}\frac{f^{(k)}(x)}{x} = \lim_{x\to 0}\frac{p_k(x)\frac{e^{-1/x}}{x^{2k}}}{x} \\
&= p_k(0)\lim_{x\to 0}\frac{e^{-1/x}}{x^{2k+1}} = 0
\end{aligned}$$

Thus the kth left and right derivatives of f exist and are equal, so we have $f \in C^\infty(\mathbb{R})$. $\square$

Now we will define the notion of a cutoff function. Intuitively, we want this to have some behavior in a certain region and not do much else outside of that region. With that in mind, we show

Lemma: *Given $r_1, r_2 \in \mathbb{R}$ such that $r_1 < r_2$, there exists a smooth function $h : \mathbb{R} \to \mathbb{R}$ such that*

$$h(t) = \begin{cases} 1 & x \leq r_1 \\ 0 < h(t) < 1 & r_1 < x < r_2 \\ 0 & x \geq r_2e \end{cases}$$

Proof: Using the f from the previous lemma, let

$$h(t) = \frac{f(r_2 - x)}{f(r_2 - x) + f(x - r_1)}$$

You can see that this works by just drawing the graph. For $x < r_1$, the second term in the denominator is 0, and we have 1. For x between r_1 and r_2, the function is always positive, and through a couple of simple estimates, you can show its norm is less than 1, using the fact that $|f| < 1$. Then for $x > r_2$, the numerator is 0, the first term in the denominator is 0, and the second term in the denominator is non-zero, thus h is 0. We have constructed

a cutoff function. Finally, we can clearly extend this notion to arbitrary dimension by just treating every component the same. That is, there exists smooth $H : \mathbb{R}^n \to \mathbb{R}$ which is 1 in a disk around 0, $0 < H(x) < 1$ in the annulus around 0, and 0 everywhere else, for any radius/disk size. This is made precise by the definition $H(\vec{x}) = h(|\vec{x}|)$. Having constructed smooth bump functions, we can now see an important application of paracompactness of manifolds. Given a manifold M and an open cover A_α, we define a partition of unity subordinate to A_α as a family of continuous functions $\psi_\alpha : M \to \mathbb{R}$ such that:
a) $0 \leq \psi_\alpha(x) \leq 1$ for all α and x
b) $\text{supp}(\psi_\alpha) \subset A_\alpha$, for all α
c) The family of supports indexed by alpha, $\text{supp}(\psi_\alpha)$ is locally finite. That is, every point of M has a neighborhood which intersects only finitely many $\text{supp}(\psi_\alpha)$.
d) For all $x \in M$,

$$\sum_\alpha \psi_\alpha(x) = 1$$

Then a smooth partition of unity is one in which the ψ_α are all smooth, as functions from $\mathbb{R}^n \to \mathbb{R}$.
Theorem: (Existence of Partitions of Unity)
Proof: For simplicity, assume M is a manifold without boundary. Then it has a basis of regular coordinate balls. Given an open cover, X_α, we acquire a basis on M, $\{B_\alpha\}$. By paracompactness, we know there exists a refinement of X_α, B_i, which is countable and locally finite. For each i, B_i is a regular coordinate ball for some X_α, which guarantees the existence of B_i' smooth coordinate ball, along with the associated smooth chart φ_i, which maps the balls in the manifold into balls in $\mathbb{R}^n$. Then our proposed smooth partition of unity subordinate to X_α is

$$f_i := \begin{cases} H_i \circ \varphi_i & x \in B_i' \\ 0 & x \in M/\{\overline{B_i}\} \end{cases}$$

where H_i is the smooth bump function from the previous lemma, defined to be potentially non-zero on the ball B_i', and 0 elsewhere. We need to check the the 4 axioms:
a) Follows directly from the properties of H_i. Note that it doesn't matter what φ does (which is good, because we don't really know anything about it), in the end everything is mapped according to H_i, which is bounded by 0 and 1.
b) By definition, f_i is only non-zero on $\overline{B_i}$, thus $\text{supp}(f_i) = \overline{B_i} \subset X_\alpha$ for some α. This containment comes because the B_i' emerged as a refinement of X_α, the open cover, so each B_i and B_i' is contained in some X_α.
c) As we showed previously, the supports of f_i are closures of basis elements, $\overline{B_i}$. However, by paracompactness, the collection $\{B_i\}$ is locally finite, which implies the collection $\{\overline{B_i}\}$ is locally finite.
d) Exercise.
End copy paste. Also end of lecture.

CHAPTER 3

Properties of Fixed Points

Lecture 3, Jan 28. The Collatz conjecture, to enhance your mathematical culture: Given $f : \mathbb{N} \to \mathbb{N}$, defined by

$$f = \begin{cases} \frac{x}{2} & x \text{ even} \\ 3x+1 & x \text{ odd} \end{cases}$$

If we track some orbits, we see

$$1 \mapsto 4 \mapsto 2 \mapsto 1$$

For something a bit more involved,

$$11 \mapsto 34 \mapsto 17 \mapsto 52 \mapsto 26 \mapsto 13 \mapsto 40 \mapsto 20 \mapsto 10 \mapsto 5 \mapsto 16 \mapsto 4 \mapsto 2 \mapsto 1$$

Theorem/Conjecture: (Collatz Conjecture) $\forall\, x \in \mathbb{N}, \exists\, n \textit{ such that } f^n(x) = 1.$
This is a famous conjecture, on which Terrence Tao recently made a contribution. This is interesting because you could show this question to a high schooler, yet the mathematics used to even make progress in it is extremely technical.
Definition: The forward orbit of x is

$$O_x^+ = \{x, f(x), f^2(x), \dots\}$$

Similarly, the backward orbit is

$$O_x^- = \{x, f^{-1}(x), f^{-2}(x), \dots\}$$

Clearly f must be invertible to define the backwards orbit. Finally, the orbit of x is

$$O_x = O_x^+ \cup O_x^-$$

Definition: x is called periodic, with period n, if $f^n(x) = x$. The smallest period is called the prime period[1].
Definition: x is called eventually periodic if $\exists\, n$ such that $f^n(x)$ is periodic.

[1] We make this definition because "period" is inherently ambiguous: If a point has period 4, it may also have period 2. We are generally interested in the smallest period.

We know that $0.\overline{999} = 1$, and $0.123\overline{999} = .0124$. There is nothing special about base 10 in this case. If we were to work in base 2, we have

$$.0\overline{111} = 1 \quad \text{and} \quad .010010\overline{111} = .01001011$$

So $\Phi : \Sigma_2 \to S^1$ is not a bijection: It is 2 to 1 in countably many places. Returning to the doubling map on S^1, we saw in lecture 1 that for any $n \in \mathbb{N}$, there is a point $p \in S^1$ such that p has period n. But how many such points are there? If we compute

$$\begin{aligned} D^n(p) &= p \\ p^{2^n} &= p \\ \Rightarrow 1 &= p^{2^n - 1} \end{aligned}$$

The final equation shows that p is a $2^n - 1$ root of unity, of which there are $2^n - 1$, so

$$|Per_n(D)| = 2^n - 1$$

What about the shift map? To see this, consider a periodic sequence: If a sequence is periodic, with period n, then the sequence itself is just that first n bit string, repeated infinitely many times. In other words, we need only an n bit string to determine a periodic sequence with period n. To make a choice of 0 or 1 in each spot makes 2^n choices, so

$$|Per_n(\sigma)| = 2^n$$

Example: Consider the map $f(x) = 5 + \alpha(x - 5)$. This map has a fixed point, 5, and we want to learn about its long term behavior, i.e. what is $\lim_{n\to\infty} f^n(x)$? Just based on heuristics, we should conclude that if $0 < |\alpha| < 1$, then it will converge. If $1 > \alpha > 0$, we cobweb directly to the fixed point. If $-1 < \alpha < 0$, we spiral to the fixed point, i.e. we cobweb but on alternating sides of the fixed point. This difference motivates the following definitions:

Definition: Let p be a fixed point of f. We say p is hyperbolic if $|f'(p)| \neq 1$. If p is hyperbolic, then we say it is attracting if $|f'(p)| < 1$, and repelling if $|f'(p)| > 1$.

Lemma: *Suppose p is an attracting fixed point, and $f \in C^1$. Then there exists an open interval $J \in P$ such that*

$$\lim_{n\to\infty} f^n(x) = p \ \ \forall x \in J$$

Proof: Choose J and $0 < \lambda < 1$ such that $|f'(x)| < \lambda \ \forall \ x \in J$ and $p \in J$. We claim $f(J) \subset J$. To see this, take $x, p \in J$. Then, by FTC,

$$\begin{aligned} |f(x) - f(p)| &= \int_x^p f'(y)dy \\ &\leq \lambda(x - p) \\ &\leq x - p \\ &\Rightarrow f(J) \subset J \end{aligned}$$

So attracting fixed points are contained in an interval which converges to p.
Definition: The stable set of an attracting fixed point p is

$$W^s(p) = \left\{ x | \lim_{n\to\infty} f^n(x) = p \right\}$$

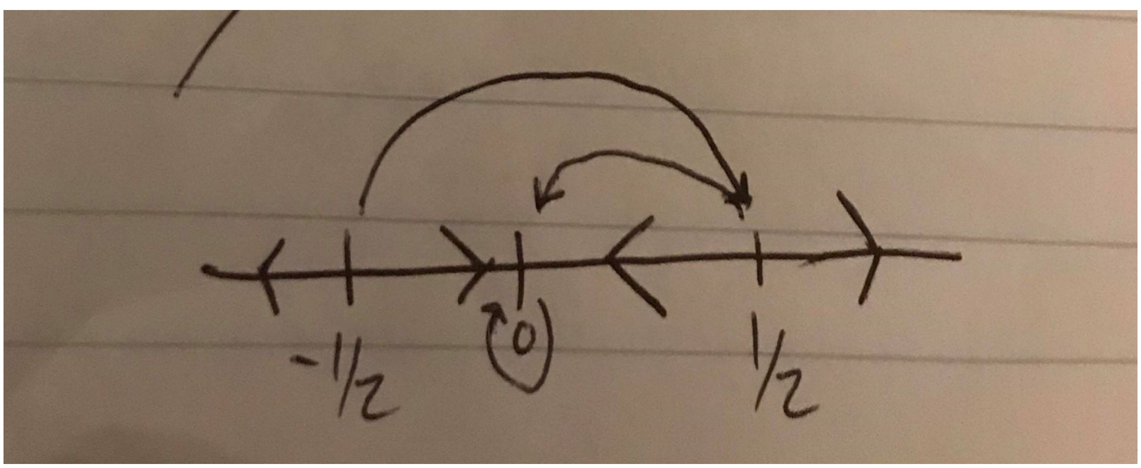
Similarly, the local stable set of p, W is the connected component of $W^s(p)$.
Example: $f(x) = x^2 + \frac{x}{2}$ has a fixed point at

$$\begin{aligned} x^2 + \frac{x}{2} &= x \\ x^2 - \frac{x}{2} &= 0 \\ \Rightarrow x(x - \frac{1}{2}) &= 0 \\ \Rightarrow x \in \{0, \frac{1}{2}\} \end{aligned}$$

0 is an attracting fixed point, but $\frac{1}{2}$ is repelling, because

$$f'(0) = \frac{1}{2}, \quad \text{and} \quad f'\left(\frac{1}{2}\right) = \frac{3}{2}$$

And $W^s_{loc}(p) = (-1, \frac{1}{2})$. The phase diagram is



We should note, in general, phase diagrams can omit information about a dynamical system. In particular, phase diagrams have a difficult time talking about *rate* of convergence. We will now see an important family of examples:
Example: For $\mu > 0$, define $F_\mu(x) = \mu x(1 - x)$. The dynamics of the system rely heavily on the value of μ, as we will see in the coming lecture.

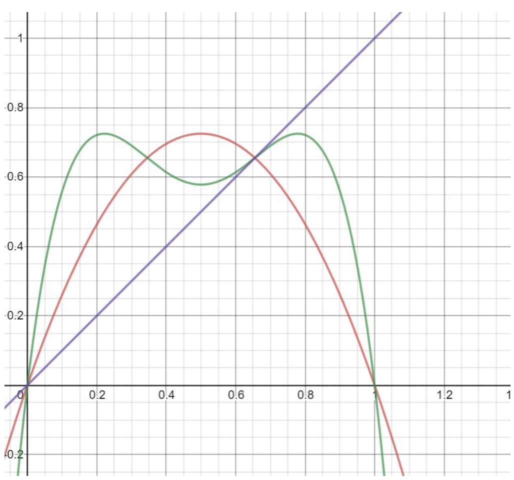
CHAPTER 4

F_μ and bifurcations

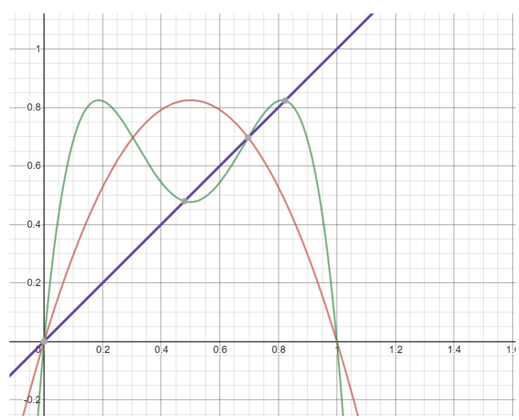
Lecture 4, Jan 30. Considering the family of functions $F_\mu(x) = \mu x(1-x)$, if we let $p = \frac{\mu-1}{\mu}$, it is a fixed point. 0 is always a fixed point, and $F'_\mu(0) = \mu$, so if $\mu > 1$, then 0 is a repelling fixed point, and if $|\mu| < 1$, it is an attracting fixed point. We can also calculate

$$F'_\mu(p) = 2 - \mu$$

so again, the nature of the fixed point depends on μ. If $1 < \mu < 3$, p is attracting, and all points are attracted to p.
A bifurcation is when the dynamics of an indexed family of maps changes significantly. For example, there is a bifurcation of F_μ at $\mu = 3$. If $\mu < 3$, then there are 2 period 2 orbits:



$\mu = 2.9$



$\mu = 3.1$

In the above, the green plot is F^2, the red is F, and the blue is $y = x$. As you can observe, there are no period 2 points when $\mu < 3$, and two when $\mu > 3$. In both cases, we have only two fixed points, p and 0. However, the dynamics of p also depend on μ. If $\mu < 3$, p is an attracting fixed point. If $\mu > 3$, it is a repelling fixed point. For the reasons discussed above, would say F_μ has a bifurcation at $\mu = 3$. Bifurcations in this context are not a rigorously defined term, although it can be made rigorous in some settings.

If $\mu > 4$, then for $x > 0$, $f(x) < 1$ and $f^2(x) < 0$. 0 is a repelling fixed point. If we define

$$\Lambda = \{x \in [0,1] \mid F^n(x) \in [0,1] \; \forall \, n \in \mathbb{N}\}$$

then $\Lambda \cong C$, where C is the cantor set.

Definition: A perfect space is one with no isolated points.

Theorem: *X compact, totally disconnected*[1]*, and perfect* $\Rightarrow X \cong C$.

The totally disconnected requirement, ironically enough, gives the best intuition for understanding this theorem. If you have a totally disconnected space, then to specify a point in the space, you break the space up into two open subspaces. Break each subspace up into two further subspaces, and continue this process until each point is identified with an infinite sequence of 1's and 0's, indicating which subspace contains the point. The space of such sequences is homeomorphic to the Cantor set[2].

[1] Anyone know where this property comes up more? I've only seen it twice in my "mathematical career" now.

[2] This intuition is still a little shaky in my mind. I wouldn't trust it.

CHAPTER 5

F_μ part II

Lecture 5, Feb 4. A further way to realize $\Lambda \cong C$ is to make the following definitions:

$$O_1 = F^{-1}(1, \infty), \quad O_n = F^{-n+1}(O_1)$$

In this case,

$$\Lambda = [0,1] / \bigcup_{i=1}^{\infty} O_i$$

Λ is closed in the real line becuase it is a closed set minus many open sets. To show it's totally disconnected, we assume $\mu > 2 + \sqrt{5}$. This implies

$$|F'(x)| > 1 \; \forall \; x \neq 0$$

Now suppose $J = [a, b] \subset \Lambda$. Then

$$\text{length}(F(J)) = \int_a^b F'(x)dx = F(b) - F(a)$$

by FTC. But J is contained entirely on either the left or right of O_1, since it is a connected interval. Then there exists $\lambda > 1$ such that $|F'(x)| \geq \lambda$, so that

$$\text{length}(F(J)) \geq \lambda(b - a)$$

However, $F(J)$ is also an interval, since F is continuous. By induction, we may show that

$$\text{length}(F^n(J)) \geq \lambda^n(\text{length}(J)) \mapsto \infty$$

since $\lambda > 1$. But by definition, $F^n(J) \subset [0,1]$, so it cannot have infinite length. Therefore Λ is totally disconnected. The spirit of this is that F stretches intervals too much, and therefore Λ cannot contain any intervals. To show Λ is perfect, suppose x_0 is an isolated point, and that it is bordered by O_m and O_n, i.e. $\sup(O_m) = \inf(O_n) = x_0$. Each O_n is a disjoint union of intervals, and

$$F^n(\text{intervals in } O_m)$$
$$= \left(1, \frac{1}{4}\mu\right]$$
$$\Rightarrow F^n(x_0) = 1, \quad \text{since it is an endpoint. } F \text{ sends all endpoints to 1.}$$

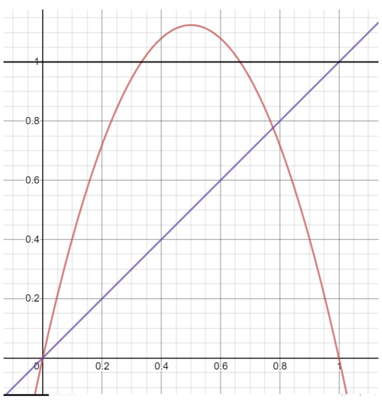
Similarly, $F^m(x_0) = 1$, and $F(1) = 0$. $\Rightarrow n = m$. To see this implication, if, WLOG, $n > m$ then $F^n(x_0) = F^j(F^m(x_0))$, where $j + m = n$. But by assumption,

$$F^m(x_0) = 1$$
$$\Rightarrow F^j(F^m(x_0)) = F^j(1) = 0 \equiv F^n(x_0)$$

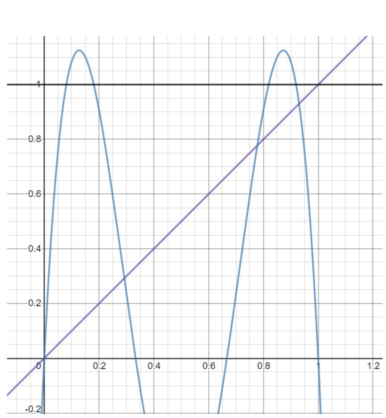
a contradiction, so $n = m$. Further, this implies[1] that F^n has a local min at x_0, in particular $(F^n)'(x_0) = 0$. But by iterated chain rule, this implies $F'(x_0) = 0$, but F has only one critical point, at $x_0 = \frac{1}{2}$. But $\frac{1}{2} \notin \Lambda$ because $F\left(\frac{1}{2}\right) = \mu/4$, which if μ is large, is not in $[0, 1]$, so Λ can not have any isolated points. This proves the theorem.

$\square$

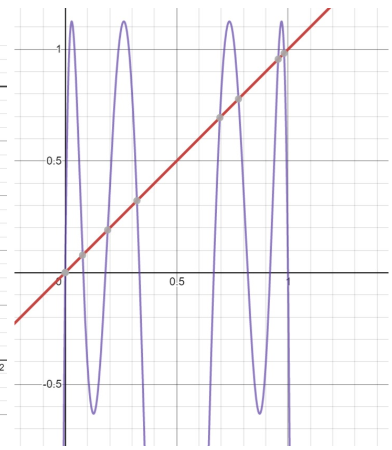
We may ask about the dynamics of F_μ in more generality. If we draw a couple of plots of iterations of F_μ, we have



F



F^2



F^3

We can now infer a number of facets about the dynamics of the system.

$$|\text{Per}_n(F_\mu)| = 2^n$$
$$|\text{Crit. Points}| = 2^n - 1$$
$$|\text{Local Min}(F^n)| = |\text{Local min}(F^{n-1})| + |\text{Local Max}(F^{n-1})|$$

Definition: $f : I \to I$ is topologically transitive[2] if for all pairs of non-empty open sets, $U, V \subset I, \exists\, n \in \mathbb{N} \mid f^n(U) \cap V \neq \emptyset$.
We should think of this intuitively as, rather than a transitive action on a set, which says nothing about topology, it is an action on the set of open sets which moves the open sets around enough, if you iterate enough times.
Theorem: f is topologically transitive iff it has a dense orbit.

[1] I actually can't see this implication at the moment, but it is like 3 am, if someone could email/dm me an explanation that would be appreciated.
[2] We are borrowing language from group theory again. Aren't group actions great?

Proof: Assume f has a dense orbit, O_x. Then there exists $n < m$ such that $f^n(x) \in U$, $f^m(x) \in V$, so that

$$\Rightarrow f^m(x) \in f^{n-m}(U)$$
$$\Rightarrow f^{n-m}(U) \cap V \neq \emptyset$$

For the opposite direction, we need to know (a simplified version of) Baire Category Theorem:

Theorem: If X is a complete metric space, and $\phi_1, \phi_2, \ldots$ is a sequence of dense open sets, then $\cap_{i=1}^{\infty} \phi_i$ is dense.

Now for the proof: We pick open sets $V_i \subset I$ such that for any open set $O \subset I$, $\exists\, i \mid V_i \subset O$[3]. Then we can consider

$$\bigcup_{n \geq 0} f^{-n}(V_i)$$

For all i, this is dense in I. So by Baire Category Theorem,

$$\bigcap_{i=1}^{\infty} \bigcup_{n=0}^{\infty} f^{-n}(V_i)$$

is dense. Let x be a point in the above dense space. Then we claim O_x is dense. We can see this because $\forall\, i,\ \exists\, n \mid x \in f^{-n}(V_i)$. But $f^n(x) \in V_i \Rightarrow$ the orbit is dense.

$\square$

For example, we saw already that all the orbits of the doubling map were dense. To see that this is equivalent to topological transitivity, if you take any two arcs on the circle, and suppose they have empty intersection. Then you can keep rotating one arc around the circle until it intersects the other arc, and hence the action of f is topologically transitive.

We may also want to know if the shift map is topologically transitive. If (x_i) is a sequence of 0's and 1's, then

Definition: (x_i) is <u>normal</u> if, $\forall\, y \in \{0,1\}^n$,

$$2^{-m} = \lim_{n \to \infty} \frac{\left\{ \left| 1 \leq i \leq n \right| (x_i, x_{i+1}, \ldots, x_{im-1}) = y \right| \right\}}{n}$$

Then $\exists x \in \Sigma_2$ such that $|\forall y \in \{0,1\}^m$, $\exists i \mid (x_{i+1}, \ldots, x_{i+m}) = (y_1, \ldots, y_m) \Rightarrow \sigma^i(x) = (x_{i+1}, \ldots x_{i+m+1}, \ldots)$. Then any such x has a dense orbit.

[3]This statement confuses me. Can't we just take $V_i = \theta(I)$, i.e. the set of all open sets of I? Perhaps there is an issue of countability, but I don't think we rely on the cardinality being countable at all. Without further restriction on the V_i, I'm not sure what this statement is telling us, nor am I even sure I believe it is true/can be done.

CHAPTER 6

Lecture 6

Lecture 6, Feb 6. **Definition:** Let $f : \Omega \to \Omega$, $g : \Omega' \to \Omega'$ be cts. Then f and g are topologically conjugate, $f \sim g$, if there exists $h : \Omega \to \Omega'$ such that

$$\begin{array}{ccc} \Omega & \xrightarrow{f} & \Omega \\ {\scriptstyle h}\downarrow & & \downarrow{\scriptstyle h} \\ \Omega' & \xrightarrow[g]{} & \Omega' \end{array}$$

commutes.

We will sometimes say f and g are topologically conjugate via h. Topological conjugacy is the natural equivalence relation[1] on the set of maps in dynamical systems, because they preserve dynamical properties: For example, periodic points of f are mapped to periodic points of g: If p is a periodic point of f with prime period n, then

$$\begin{aligned} f^n(p) &= p \\ g^n h f^n(p) &= h f^n(p) \\ g^n(h(p)) &= h(p) \end{aligned}$$

i.e. $h(p)$ is also periodic with prime period n. In particular, this restricts to fixed points as well. We can also guess that the basins of attraction, $W_s(p)$, should be mapped to each other. This is the case, and is part of the reason we don't simply require a commutative diagram, but the stronger topological conjugacy: We would like to guess

$$h(W_s(p)) = W_s(h(p))$$

In fact, this is true because

$$\begin{aligned} \lim_{n\to\infty} f^n(p) &= p \\ \Rightarrow h\left(\lim_{n\to\infty} f^n(p)\right) &= h(p) \\ \Rightarrow \lim_{n\to\infty} h\left(f^n(p)\right) &= h(p) \\ \Rightarrow \lim_{n\to\infty} g^n h(p) &= h(p) \end{aligned}$$

[1] It is easy to verify that this is an equivalence relation.

But here, the second $\Rightarrow$ implication only holds when h is continuous. Limits are a topological property, so we must require our maps, h, to be homeomorphisms, instead of just invertible maps. This is a long winded way of saying that dynamical systems takes place in the topological or smooth category, not in an algebraic category.

Lemma: *Topological transitivity is also preserved by topological conjugation.*

Proof: Let U', V' be open in Ω'. Then $h^{-1}(U'), h^{-1}(V')$ are open, so that there exists n such that

$$f^n\left(h^{-1}(U')\right) \cap h^{-1}(V') \neq \emptyset$$
$$h^{-1}g^n(U') \cap h^{-1}(V') \neq \emptyset$$
$$\Rightarrow g^n(U') \cap V' \neq \emptyset$$

□

As one may note, this proof does not require h to be a homeomorphism, only to be continuous and surjective. In such a case, we call h a semi-conjugacy between f and g, or that f factors through h.

Though we preserve Top notions such as limits, and Set notions such as fixed points, we do not Diff/Man (analytic) notions, such as hyperbolicity[2].

[2] Still thinking of an example of this.