An Modern Introduction to Dynamical Systems

Richard J. Brown

Contents

Preface

The following text comprises the content of a course I have designed and have been running at Johns Hopkins University since the Spring of 2007. It is a senior (read: 400-level) analysis course in the basic tools, techniques, theory and development of what is sometimes called the modern theory of dynamical systems. The modern theory, as best as I can define it, is a focus on the study and structure of dynamical systems as little more than the study of the properties of one-parameter groups of transformations on a topological space, and what these transformations say about the properties of either the space or the group that is acting. It is a pure mathematical endeavor in that we study the material simply for the structure inherent in the constructions, and not for any particular application or outside influence. It is understood that many of the topics comprising this theory have natural, beautiful and important applications, some of which actually dictate the need for the analysis. But the true motivation for the study is little more than the fact that it is beautiful, rich in nuance and relevance in many tangential areas of mathematics, and that it is there.

When I originally pitched this course to the faculty here at Hopkins, there was no course like it in our department. We have a well-developed engineering school, filled with exceptionally bright and ambitious students, which along with strong natural and social science programs provide a ready audience for a course on the pure mathematical study of the theory behind what makes a mathematical model and why do we study them. We have a sister department here at Homewood, the Applied Mathematics and Statistics Department, which also offers a course in dynamical systems. However, their course seemed to focus on the nature and study of particular models that arise often in other classes, and then to mine those models for relevant information to better understand them. But as a student of the field, I understood that a course on the very nature of using functions as models and then studying their properties in terms of the dynamical information inherent in them was currently missing from our collective curriculum. Hence the birth of this course.

In my personal and humble opinion, it continues to be difficult to find a good text that satisfies all of the properties I think would constitute the perfect text for a course such as this one: (1) a focus on the pure mathematical theory of the abstract dynamical system, (2) advanced enough that the course can utilize the relevant topological, analytical and algebraic nature of the topic without requiring so much prerequisite knowledge as to limit enrollment to just mathematicians, (3) rich enough to develop a good strong story to tell which provides a solid foundation for later individual study, and (4) basic enough so that students in the natural sciences and engineering can access the entirety of the content given only the basic foundational material of vector calculus, linear algebra and differential equations.

It is a tall order, this is understood. However, it can be accomplished, I believe, and this text is my attempt at accomplishment.

The original text I chose for the course is the text *A First Course in Dynamics*, by Boris Hasselblatt and Anatole Katok (Cambridge University Press: 2003). A wonderfully designed story-line from two transformational mathematicians in the field, I saw the development line they took, from the notion of simple dynamics to the more complicated, as proper and intuitive. I think their focus on using the properties of functions and that of the spaces they are acting upon to develop material is the correct one for this "modern" approach. And their reuse of particular examples over and over again as the story progresses is a strong one. However, in the years I have been teaching and revising the course, I have found myself, adding material, redesigning the focus and the schedule, and building in a slightly different storyline. All of this diverging from the text. Encouraged by my students and my general thrill at the field, I decided to create my version of a text. This manuscript is this version.

What the reader will find in this text is my view of the basic foundational ideas that comprise a first (and one semester) course in the modern theory of dynamical systems. It is geared toward the upper-level undergraduate student studying either mathematics, or engineering or the natural and social sciences with a strong emphasis in learning the theory the way a mathematician would want to teach the theory. It is a proof-based course. However, when I teach the course, I do understand that some of my students do not have experience in writing mathematics in general and proofs in particular. Hence I use the content of the course as a way to also introduce these students to the development of ideas instead of just calculation. It is my hope that these students, upon finishing this course, will begin to look at the models and analysis they see in their other applied classes with an eye to the nature of the model and not just to its mechanics. They are studying to be scholars in their chosen field. Their ability to really "see" the mathematical structure of their tools will be necessary for them to contribute to their field.

This course (this text) is designed to be accessible to a student who has had a good foundational course in the following:

- vector calculus, at least up to the topics of surface integration and the "big three" theorems of Green, Stokes and Gauss;
- linear algebra, through linear transformations, kernels and images, eigenspaces, orthonormal bases and symmetric matrices; and
- differential equations, with general first and second order equations, linear systems theory, nonlinear analysis, existence and uniqueness of first order solutions, and the like.

While I make it clear in my class that analysis and algebra are not necessary prerequisites, this course cannot run without a solid knowledge of the convergence of general sequences in a space, the properties of what makes a set a topological space, and the workings of a group. Hence in the text we introduce these ideas as needed, sometimes through development and sometimes simply through introduction and use. I have found that most of these advanced topics are readily used and workable for students even if they are not fully explored within the confines of a university course. Certainly, having sat through courses in advanced algebra and analysis will be beneficial, but I believe they are not necessary. The text to follow, like all proper endeavors in mathematics, should be seen as a work in progress. The storyline, similar to that of Hasselblatt and Katok, is to begin with basic definitions of just what is a dynamical system. Once the idea of the dynamical content of a function or differential equation is established, we take the reader a number of topics and examples, starting with the notion of simple dynamical systems to the more complicated, all the while, developing the language and tools to allow the study to continue. Where possible and illustrative, we bring in applications to base our mathematical study in a more general context, and to provide the reader with examples of the contributing influence the sciences has had on the general theory. We pepper the sections with exercises to broaden the scope of the topic in current discussion, and to extend the theory into areas thought to be of tangential interest to the reader. And we end the text at a place where the course I teach ends, on a notion of dynamical complexity, topological entropy, which is still a active area of research. It is my hope that this last topic can serve as a landing on which to begin a more individualized, higher-level study, allowing the reader to further their scholarly endeavor now that the basics have been established.

I am thankful to the mathematical community for facilitating this work, both here at Hopkins and beyond. And I hope that this text contributes to the learning of high-level mathematics by both students of mathematics as well as students whose study requires mathematical prowess.

CHAPTER 1

What is a Dynamical System?

1.1. Definitions

As a mathematical discipline, the study of dynamical systems most likely originated at the end of the 19th century through the work of Henri Poincare in his study of celestial mechanics (footnote this: See Scholarpedia[History of DS]). Once the equations describing the movement of the planets around the sun are formulated (that is, once the mathematical model is constructed), looking for solutions as a means to describe the planets' motion and make predictions of positions in time is the next step. But when finding solutions to sets of equations is seemingly too complicated or impossible, one is left with studying the mathematical structure of the model to somehow and creatively narrow down the possible solution functions. This view of studying the nature and structure of the equations in a mathematical model for clues as to the nature and structure of its solutions is the general idea behind the techniques and theory of what we now call dynamical systems. Being only a 100+ years old, the mathematical concept of a dynamical system is a relatively new idea. And since it really is a focused study of the nature of functions of a single (usually), real (usually) independent variable, it is a subdiscipline of what mathematicians call real analysis. However, one can say that dynamical systems draws its theory and techniques from many areas of mathematics, from analysis to geometry and topology, and into algebra. One might call mathematical areas like geometry, topology and dynamics second generation mathematics, since they tend to bridge other more pure areas in their theories. But as the study of what is actually means to model phenomena via functions and equations, dynamical systems is sometimes called the mathematical study of any mathematical concept that evolves over time. So as a means to define this concept more precisely, we begin with arguably a most general and yet least helpful statement:

Definition 1.1. A *dynamical system* is a mathematical formalization for any *fixed rule* which describes the dependence of the position of a point in some *ambient space* on a *parameter*.

- The *parameter* here, usually referred to as "time" due to its reference to application in the sciences, take values in the real numbers. Usually, these values come in two varieties:
  (1) discrete (think of the natural numbers $\mathbb{N}$ or the integers $\mathbb{Z}$), or
  (2) continuous (defined by some single interval in $\mathbb{R}$).

  The parameter can sometimes take values in much more general spaces, for instance, like subsets of $\mathbb{C}$, $\mathbb{R}^n$, the quaternions, or indeed any set with the structure of an algebraic group. However, classically speaking, a dynamical system really involves a parameter that takes values only in a subset of $\mathbb{R}$. We will hold to this convention.

- The *ambient space* has a "state" to it in the sense that all of its points have a marked position which can change as one varies the parameter. Roughly, every point has a position relative to the other points and a complete set of (generalized) coordinates on the space often provide this notion of position. Fixing the coordinates and allowing the parameter to vary, one can create a functional relationship between the points at one value of the parameter and those at another parameter value. In general, this notion of relative point positions in a space and functional relationships on that space involves the notion of a topology on a set. A topology gives a set the mathematical property of a space; It endows the elements of a set with a notion of nearness to each other and allows for functions on a set to have properties like continuity, differentiability, and such. We will expound more on this later. We call this ambient space the *state space*: it is the set of all possible states a dynamical system can be in at any parameter value (at any moment of time.)
- The *fixed rule* is usually a recipe for going from one state to the next in the ordering specified by the parameter. For discrete dynamical systems, it is often given as a function. The function, from the state space to itself, takes each point to its next state. The future states of a point are found by applying the same function to the state space over and over again. If possible, the past states of a point can be found by applying the inverse of the function. This defines the dynamical system recursively via the function. In continuous systems, where it is more involved to define what the successor to a parameter value may be, the continuous movement of points in a space may be defined by a differential equation, equal to that of the function in a discrete system in that it describes implicitly the method of going from one state to the next (defined only infinitesimally). The solution to an ODE (or system of them) would be a function whose domain contains the points of the state space and the parameter and taking values back in the state space (the codomain). Often, this latter function is called the *evolution* of the system, providing a way of going from any particular state to any other state reachable from that initial state via a value of the parameter. As we will see, such a function can be shown to exist, and its properties can often be studied, but in general, it will NOT be known *a priori*, or even knowable *a posteriori*.

Remark 1.2. It is common in this area of mathematics that the terms fixed-rule and evolution are used more or less interchangeably, and both referring to the same objects without distinction. In this book, we will differentiate the two as described above. Namely, the fixed rule will remain the recursively defined recipe for movement within a dynamical system, and the evolution will be reserved for the functional form of the movement of points. Thus the ODE is simply the fixed-rule, while the general solution, if it can be found, is the evolution, for example.

While this idea of a dynamical system is far too general to be very useful, it is instructive. Before creating a more constructive definition, let's look at some classical examples:

1.1.1. Ordinary Differential Equations (ODEs). Given the first-order, autonomous (vector)-ODE,

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}), \tag{1.1.1}$$

a solution, if it exists, is a vector of functions $\mathbf{x}(t) = [x_1(t)\ x_2(t)\ \cdots\ x_n(t)]^T$ parameterized by a real variable $t \in \mathbb{R}$ where the common domain of the coordinate functions is some subinterval of $\mathbb{R}$. [Develop this idea of autonomous verses non-autonomous ODEs as a dynamical system].

Recall that one fairly general form for first-order systems of ODEs is $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t)$, with the independent variable t explicitly represented on the right-hand side. An ODE (or system) is called *autonomous* if t is not explicitly represented, and *non-autonomous* when t is explicit in the ODE. For example $\dot{x} = \frac{tx^2}{4} = f(x,t)$ is non-autonomous, while $\dot{x} = \frac{x}{4} = f(x)$ is autonomous. We sometimes also call autonomous systems time-invariant when the independent variable actually does represent time. The important property of an autonomous ODE is that the laws of motion at any point in time are the same as at any other point in time; the laws of motion are invariant under translations in time. If a first-order system like Equation 1.1.1 is autonomous, then the right-hand side represents a vector field in the state space that doesn't change in time. This means that going from one state to the next will be the same no matter when the motion starts. With time explicit in the function $\mathbf{f}(\mathbf{x}, t)$, the vector field would be changing as time progresses. See Figure 1.

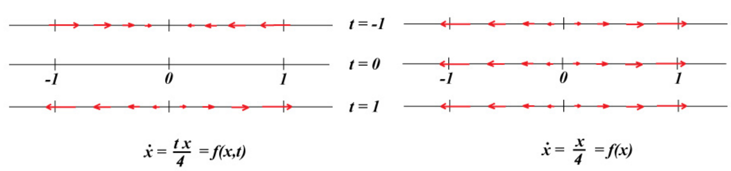


FIGURE 1. Vector fields on $\mathbb{R}$ under non-autonomous (left) and autonomous (right) ODEs.

Of course, it is always possible to render a non-autonomous, vector-ODE into an autonomous one simply by creating a new state variable equal to the independent time variable, rendering it a dependent variable, and creating a new independent variable as the new time. This increases the size of the state vectors by one, and introduces a new first-order ODE into the system. However, there are two points to consider: (1) the newest ODE is just the time derivative set to 1, and (2) there is really no advantage to this rendering in terms of solvability. Plus, this really is little different from the standard trick of turning a second-order ODE into a first-order system by creating a new dependent variable and setting it equal to the velocity of the "other" dependent variable. Although, in this latter case, there are distinct advantage to doing so. We will explore this later. Here, for our autonomous system in Equation 1.1.1, we have:

- The ODE itself is the fixed rule, describing the infinitesimal way to go from one state to the next by an infinitesimal change in the value of the parameter t. Solving the ODE means finding the unknown function

$\mathbf{x}(t)$, at least up to a set of constants determined by some initial state of the system. The inclusion of initial data provide this initial state of the variables of the system, making the system an Initial Value Problem (IVP). A solution to an IVP, $\mathbf{x}(t)$, for valid values of t, provides the various "other" states that the system can reach (either forward or backward in time) as compared to the initial state. Collecting up all the functions $\mathbf{x}(t)$ for all valid sets of initial data (basically, finding the expression that writes the constants of integration of the general solution to the ODE in terms of the initial data variables), into one big function IS the evolution.

- This type of a dynamical system is called *continuous*, since the parameter t will take values in some domain (an interval with non-empty interior) in $\mathbb{R}$. Dynamical systems like this arising from ODEs are also called *flows*, since the various IVP solutions in phase space look like the flow lines of a fluid in phase space flowing along the slope field (vector field defined by the ODE).
- In this particular example, the state space is the n-dimensional space parameterized by the n-dependent variables that comprise the vector $\mathbf{x}(t)$. Usually, these coordinate functions are simply (subsets of) $\mathbb{R}$, so that the state space is (a subset of) R^n. But there is no restriction that coordinates be rectilinear and no restriction that the state space be Euclidean. In fact, flows on spheres and other non-Euclidean spaces are very interesting to study. Regardless of the properties of the state space, solutions live in it as parameterized curves. These solution curves are often called *trajectories*. We also call this state space the *phase space*.

REMARK 1.3. One should be careful about not confusing a state space, the space of all possible states of a system, with the *configuration space* of, say, a physical system governed by Newton's Second Law of Motion. For example, the set of all positions of a pendulum at any moment of time is simply the circle. This would be the configuration space, the space of all possible configurations. But without knowing the velocity of the pendulum at any particular configuration, one cannot predict future configurations of the system (See Figure 2). The state space, in the case of the pendulum, involves both the position of the pendulum and its velocity (we will see why in a later chapter.) For a standard ODE system like the general one above, the state space, phase space and configuration space all coincide. We will elaborate more on this later.

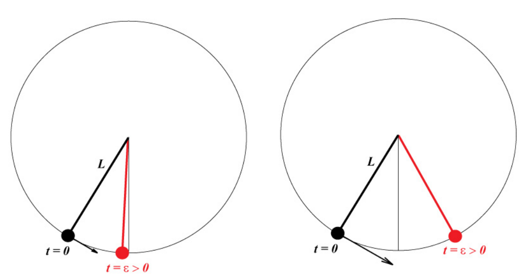


FIGURE 2. Different velocities render different future configurations.

So autonomous, first-order ODEs are examples of continuous dynamical systems (actually differentiable dynamical systems). Solving the ODE (finding the vector of functions $\mathbf{x}(t)$), means finding the rule which stipulates any state of a point in some other parameter value given the state of the point at a starting state. But as we will soon see, when thinking of ODEs as dynamical systems, we have a different perspective on what we are looking for in solving the ODE.

1.1.2. Maps. Given any set X and a function $f : X \to X$ from X to itself, one can form a dynamical system by simply applying the function over and over (iteratively) to X. When the set has a topology on it (a mathematically precise notion of an "open subset", allowing us to talk about the positions of points in relation to each other), we can then discuss whether the function f is continuous or not. When X has a topology, it is called a space, and a continuous function $f : X \to X$ is called a *map*.

- We will always assume that the sets we specify in our examples are spaces, but will detail the topology only as needed. Mostly our state spaces will exist as subsets of real space $\mathbb{R}^n$. Here, one such notion of the nearness of points will result from a precise definition of a distance between points given by a metric. In this context, there should be little confusion. Here the *state space* is X, with the positions of its points given by coordinates on X (defined by the topology.)
- the fixed rule is the map f, which is also sometimes called a *cascade*.
- In a purely formal way, f defines the evolution (recursively) by composition with itself. Indeed, $x \in X$, define $x_0 = x$, and $x_1 = f(x_0)$. Then
$$x_2 = f(x_1) = f(f(x_0)) = f^2(x_0),$$
and for all $n \in \mathbb{N}$, (the natural numbers)
$$x_n = f(x_{n-1}) = f(f(x_{n-2})) = \overbrace{f(f(\cdots f(f(x_0))\cdots))}^{n \text{ times}} = f^n(x_0).$$
- Maps are examples of *discrete dynamical systems*. Some examples of discrete dynamical systems you may have heard of include discretized ODEs, including difference equations and time-t maps. Also, fractal constructions like Julia sets and the associated Mandelbrot arising from maps of the complex plane to itself (although, precisely speaking, the Mandelbrot Set is actually a kind of parameter space of a dynamical system, recording particular information about an entire family of parameterized maps. Some objects that are not considered to be constructed by dynamical systems (at least not directly) include fractals like Sierpinski's carpet, Cantor sets, and Fibonacci's Rabbits (given by a second order recursion). Again, we will get to these.

Besides these classic ideas of a dynamical system, there are much more abstract notions of a dynamical system:

1.1.3. Symbolic Dynamics. Given a set of symbols $M = \{A, B, C, \ldots\}$, consider the "space" of all bi-infinite sequences of these symbols (infinite on both sides)
$$\Omega_M = \left\{(\ldots, x_{-2}, x_{-1}, x_0, x_1, x_2, \ldots)\middle|\, i \in \mathbb{Z}, x_i \in M\right\}.$$
One can consider Ω_M as the space of all functions from $\mathbb{Z}$ to M: each function is just an assignment of a letter in M to each integer in Z. Now let $f : \Omega_M \to \Omega_M$

be the *shift* map: on each sequence, it simply takes $i \mapsto i+1$; each sequence goes to another sequence which looks like a shift of the original one.

NOTE. *We can always consider this (very large) set of infinite sequences as a space once we give it a topology like I mentioned. This would involve defining open subsets for this set, and we can do this through ϵ-balls by defining a notion of distance between sequences (a metric). For those who know analysis, what would be a good metric for this set to make it a space using the metric topology? We will define one when we discuss metrics later on. For now, simply think of this example as something to think about. Later and in context, we will focus on this type of dynamical system and it will make more sense.*

This discrete dynamical system is sometimes used as a new dynamical system to study the properties of an old dynamical system whose properties were hard to study. Other times it is used as a model space for a whole class of dynamical systems that behave similarly. We will revisit this dynamical system and its uses later.

Sometimes, in a time-dependent system, the actual dynamical system will need to be constructed before it can be studied.

1.1.4. Billiards. Consider two mass-m point-beads moving at constant (possibly different) speeds along a finite length wire, with perfectly elastic collisions both with each other and with the walls. Recall that this means that the total kinetic energy and the total momentum of the constituent parts is each collision is preserved, so that while energy may be transferred between the point-beads in a collision, no energy is absorbed, either by a wall or another bead. Note that this also means that, while the velocity of a bead may reverse direction at times (after a collision with a wall) or switch with the velocity of the other bead (when the beads collide), there are only a few distinct velocities take by the beads in the system. In this manner, velocity is not really a variable in this system. So the state space is still only the set of all positions of the beads. The velocities do play a role in movement around the state space, however.

As an exercise (this is Exercise 1 below), parameterize the wire from wall to wall as a closed, bounded interval in $\mathbb{R}$. What does the state space look like in this case, then? Taking the position of each bead as a coordinate, the state space is just a triangle in the plane. Work this out. What are the vertices of this triangle? Does it accurately describe ALL of the states of the system? Are the edges of the triangle part of the state space? Are the vertices? And once you correctly describe the state space, what will motion look like in it as the beads move along the wire at their designated velocities? In other words, How does the dynamical system evolve?

We will revisit this model in detail later as an early example of a type of dynamical system called a billiard.

EXERCISE 1. One way to view the state space, the set of all states of the two point-beads, is to simply view each bead's position as a coordinate along the wire (in a closed subset of $\mathbb{R}$). Then the state of the system at a moment in time can be viewed as a point in the plane. Parameterize the wire by the interval $[0,1]$. Then construct the state space as a closed subset of $\mathbb{R}^2$. For given bead velocities v_1 and v_2, describe the motion in the state space by drawing in a representative trajectory. Do this for the following data:

a. $v_1 = 0$, $v_2 \neq 0$.
b. v_1/v_2 a rational number.
c. v_1/v_2 an irrational number.

By looking at trajectories that get close to a corner, can you describe what happens to a trajectory that intersects a corner directly?

Now consider a single point-ball moving at a constant velocity inside a closed, bounded region of $\mathbb{R}^2$, where the boundary is smooth and collisions with the boundary are specular (mirror-like, where again this means that the angle of incidence is equal to the angle of reflection). See Figure 3. Some questions to ponder:

(1) How does the shape of the region affect the types of paths the ball can traverse?
(2) Are there closed paths (periodic ones)?
(3) Can there be a dense path (one that eventually gets arbitrarily close to any particular point in the region?

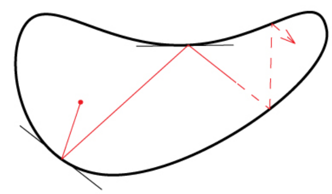
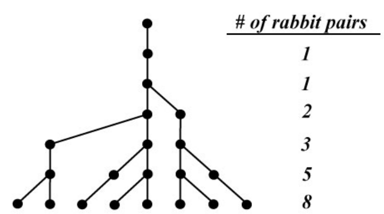


FIGURE 3. An example of a non-convex billiard.

There is a method to study the type of dynamical system called a billiard by creating a discrete dynamical system to record movement and collecting only essential information. In this discrete dynamical system, regardless of the shape of the region, the state space is a cylinder. Can you see it? If so, what would be the evolution?

1.1.5. Higher-order recursions. Maps as dynamical systems are examples of first-order recursions, since for $f : X \to X$, $x_n = f(x_{n-1})$ and each element of a sequence $\{x_n\}_{n\in\mathbb{N}}$ only depends on the previous element. The famous Rabbits of Leonardo of Pisa is a beautiful example of a type of growth that is not exponential, but something called *asymptotically exponential*. We will explore this more later. For now, though, we give a brief description: Place a newborn pair of breeding rabbits in a closed environment. Rabbits of this species produce another pair of rabbits each month after they become fertile (and they never die nor do they experience menopause). Each new pair of rabbits (again, neglect the incest, gender and DNA issues) becomes fertile after a month and starts producing each month starting in the second month. How many rabbits are there after 10 years?



Month	a_n	j_n	b_n	total pairs
1	0	0	1	1
2	0	1	0	1
3	1	0	1	2
4	1	1	1	3
5	2	1	2	5
6	3	2	3	8
7	5	3	5	13

Given the chart in months, we see a way to fashion an expression governing the number of pairs at the end of any given month: Start with r_n, the number of pairs of rabbits in the nth month. Rabbits here will come in three types:

Adults a_n, juveniles j_n, and newborns b_n, so that $r_n = a_n + j_n + b_n$. Looking at the chart, we can see that there are constraints on these numbers:

(1) the number of newborns at the $(n+1)$st stage equals the number of adults at the nth stage plus the number of juveniles at the nth stage, so that
$$b_{n+1} = a_n + j_n.$$
(2) This is also precisely equal to the number of adults at the $(n+1)$st stage, so that
$$a_{n+1} = a_n + j_n.$$
(3) and finally, the number of juveniles a the $(n+1)$st stage is just the number of newborns at the nth stage, so that
$$j_{n+1} = b_n.$$

Thus, we have
$$r_n = a_n + j_n + b_n = (a_{n-1} + j_{n-1}) + b_{n-1} + (a_{n-1} + j_{n-1}).$$
And since in the last set of parentheses, we have $a_{n-1} = a_{n-2} + j_{n-2}$ and $j_{n-1} = b_{n-2}$, we can substitute these in to get
$$\begin{aligned} r_n = a_n + j_n + b_n &= (a_{n-1} + j_{n-1}) + b_{n-1} + (a_{n-1} + j_{n-1}) \\ &= a_{n-1} + j_{n-1} + b_{n-1} + a_{n-2} + j_{n-2} + b_{n-2} = r_{n-1} + r_{n-2}. \end{aligned}$$

Hence the pattern is ruled by a second-order recursion $r_n = r_{n-1} + r_{n-2}$ with initial data $r_0 = r_1 = 1$. Being a second order recursion, we cannot go to the next state from a current state without also knowing the previous state. This is an example of a model which is not a dynamical system as stated. We can make it one (in the same fashion that one would use to turn a higher-order ODE into a first-order system, that is), but we will need a bit more structure, which we will introduce later.

Now, with this general idea of what a dynamical system actually is, along with numerous examples, we give a much more accurate and useful definition of a dynamical system:

Definition 1.4. A dynamical system is a triple $(\mathcal{S}, \mathcal{T}, \Phi)$, where $\mathcal{S}$ is the state space (or phase space), $\mathcal{T}$ is the parameter space, and
$$\Phi : (\mathcal{S} \times \mathcal{T}) \longrightarrow \mathcal{S}$$
is the evolution.

Some notes:

- In the previous discussion, the fixed rule was a map or an ODE which would only define recursively what the evolution would be. In this definition, Φ defines the entire system, mapping where each point $s \in \mathcal{S}$ goes for each parameter value $\tau \in \mathcal{T}$. It is the functional form of the fixed rule, unraveling the recursion and allowing one to go from a starting point to any point reachable by that point given a value of the parameter.

- In ODEs, Φ plays the role of the *general* solution, as a 1-parameter family of solutions (literally a 1-parameter family of transformations of phase space): In this general solution, one knows for ANY specified starting value where it will be for ANY valid parameter value, all in one function of two variables.

  EXAMPLE 1.5. In the Malthusian growth model, $\dot{x} = kx$, with $k \in \mathbb{R}$, and $x(t) \geq 0$ a population, the general solution is given by $x(t) = x_0 e^{kt}$, for $x_0 \in \mathbb{R}_0^+ = [0, \infty)$, the unspecified initial value at $t = 0$. (The notation $\mathbb{R}_0^+$ comes from the strictly positive real numbers $\mathbb{R}^+$ together with the value 0.) Really, the model works for $x_0 \in \mathbb{R}$, but if the model represents population growth, then initial populations can ONLY be nonnegative, right? Here, $\mathcal{S} = \mathbb{R}_0^+$, $\mathcal{T} = \mathbb{R}$ and $\Phi(s,t) = se^{kt}$.

  EXAMPLE 1.6. Let $\dot{x} = -x^2 t$, $x(0) = x_0 > 0$. Using the technique commonly referred to as separation of variables, we can integrate to find an expression for the general solution as $x(t) = \frac{1}{\frac{t^2}{2} + C}$. And since $x_0 = \frac{1}{C}$ (you should definitely do these calculations explicitly!), we get
  $$\Phi(x_0, t) = \frac{1}{\frac{t^2}{2} + \frac{1}{x_0}} = \frac{2x_0}{x_0 t^2 + 2}.$$
  Here, we are given $\mathcal{S} = \mathbb{R}^+$, and we can choose $\mathcal{T} = \mathbb{R}$. Question: Do you see any issues with allowing $x_0 < 0$? Let $x_0 = -2$, and describe the particular solution on the interval $t \in (0, 2)$.

  EXERCISE 2. Integrate to find the general solution above for the Initial Value Problem $\dot{x} = -x^2 t$, $x(0) = x_0 > 0$.

- In discrete dynamics, for a map $f : X \to X$, we would need a single expression to write $\Phi(x, n) = f^n(x)$. This is not always easy or doable, as it would involve finding a functional form for a recursive relation. Try doing this with f a general polynomial of degree more than 1.

  EXAMPLE 1.7. Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = rx$, for $r \in \mathbb{R}_+$. Then $\Phi(x, n) = r^n x$.

  EXAMPLE 1.8. For Leonardo of Pisa's (also known as Fibonacci, in case you recognized the pattern of the sequence) rabbits, we will have to use the recursion to calculate every month's population to get to the 10-year mark. However, if we could find a *functional* form for the recursion, giving population in terms of month, we could than simply plug in $12 \cdot 10 = 120$ months to calculate the population after 10 years. The latter functional form is the evolution Φ in the definition of a dynamical system above. How does one find this? We will see.

EXERCISE 3. Find a closed form expression for the evolution of $f(x) = rx + a$, in the case where $-1 < r < 1$ and a are constants. Also determine the unique point where $f(x) = x$ in this case.

EXERCISE 4. For $g(x) = x^2 + 1$, write out the first four iterates $g^i(x)$, $i = 1, 2, 3, 4$. Then look for a pattern with which to write out the nth iterate, $g^n(x)$. Do you see the difficulty? Now if your only interest was to know what the convergence properties of $\{g^n(x)\}_{n\in\mathbb{N}}$ was for arbitrary starting values x_0, what can you assert? And can you prove your assertions?.

In general, finding Φ (in essence, solving the dynamical system) is very difficult if not impossible, and certainly often impractical and/or tedious. However, it is often the case that the purpose of studying a dynamical system is not to actually solve it. Rather, it is to gain insight as to the structure of its solutions. Really, we are trying to make *qualitative* statements about the system rather than quantitative ones. Think about what you did when studying nonlinear systems of first order ODEs in any standard undergraduate course in differential equations. Think about what you did when studying autonomous first order ODEs.

Before embarking on a more systematic exploration of dynamical systems, here is another less rigorous definition of a dynamical system:

DEFINITION 1.9. Dynamical Systems as a field of study attempts to understand the structure of a changing mathematical system by identifying and analyzing the things that do not change.

There are many ways to identify and classify this notion of an unchanging quantity amidst a changing system. But the general idea is that if a quantity within a system does not change while the system as a whole is evolving, then that quantity holds a special status as a *symmetry*. Identifying symmetries can allow one to possibly locate and identify solutions to an ODE. Or one can use a symmetry to create a new system, simpler than the previous, where the symmetry has been factored out, either reducing the number of variables or the size of the system.

More specifically, here are some of the more common notions:

- Invariance: First integrals: Sometimes a quantity, defined as a function on all or part of the phase space, is constant along the solution curves of the system. If one could create a new coordinate system of phase space where one coordinate is the value of the first integral, then the solution curves correspond to constant values of this coordinate. The coordinate becomes useless to the system, and it can be discarded. The new system then has less degrees of freedom than the original. Phase space volume: In a conservative vector field, as we will see, if we take a small ball of points of a certain volume and then flow along the solution curves to the vector field, the ball of points will typically bend and stretch in very complicated ways. But it will remain an open set, and its total volume will remain the same. This is *phase volume preservation*, and it says a lot about the behavior and types of solution curves.
- Symmetry: Periodicity: Sometimes individual solution curves (or sets of them) are closed, and solutions retrace their steps over certain intervals of time. If the entire system behaves like this, the direction of the flow contains limited information about the solution curves of the system. One can in a sense factor out the periodicity, revealing more about the remaining directions of the state space. Or even near an isolated singular periodic solution, one can discretize the system at the period of the periodic orbit. This discretized system has a lower order, or number of variables, then the original.
- Asymptotics: In certain autonomous ODEs (systems where the time is not explicitly expressed in the system), one can start at any moment in

time and the evolution depends only on the starting time. In systems like these, the long-term behavior of solutions may be more important than where they are in any particular moment in time. In a sense, one studies the asymptotics of the system, instead of attempting to solve. Special solutions like equilibria and limit cycles are easy to find, and their properties become important elements of the analysis.

EXAMPLE 1.10. In an exact differential equation

$$M(x,y)\,dx + N(x,y)\,dy = M(x,y) + N(x,y)\frac{dy}{dx} = 0,$$

we have $M_y = \frac{\partial M}{\partial y} = \frac{\partial N}{\partial x} = N_x$. We know then that there exists a function $\phi(x,y)$, where $\frac{\partial \phi}{\partial x} = M$ and $\frac{\partial \phi}{\partial y} = N$. Indeed, given a twice differentiable function $\phi(x,y)$ defined on a domain in the plane, it's level sets are equations $\phi(x,y) = C$, for C a real constant. Each level set defines y implicitly as a function of x. Thinking of y as tied to x implicitly, differentiate $\phi(x,y) = C$ with respect to x and get

$$\frac{d\phi}{dx} = \frac{\partial \phi}{\partial x} + \frac{\partial \phi}{\partial y}\frac{dy}{dx} = 0.$$

This last equation will match the original ODE precisely if the two above properties hold. The interpretation then is: The solutions to the ODE correspond to the level sets of the function ϕ. We can say that that solutions to the ODE “are forced to live” on the level sets of ϕ. Thus, we can write the general solution set (at least implicitly) as $\phi(x,y) = C$, again a 1-parameter family of solutions. Here ϕ is a *first integral* of the flow given by the ODE, a concept we will define precisely in context later.

EXERCISE 5. Solve the differential equation $12-3x^2+(4-2y)\frac{dy}{dx} = 0$ and express the general solution in terms of the initial condition $y(x_0) = y_0$. This is your function $\phi(x,y)$.

EXAMPLE 1.11. Newton-Raphson: Finding a root of a (twice differentiable) function $f:\mathbb{R}\to\mathbb{R}$ leads to a discrete dynamical system $x_n = g(x_{n-1})$, where

$$g(x) = x - \frac{f(x)}{f'(x)}.$$

One here does not need to actually solve the dynamical system (find a form for the function Φ). Instead, all that is needed is to satisfy some basic properties of f to know that if you start sufficiently close to a root, the long-term (asymptotic) behavior of any starting point IS a root.

EXERCISE 6. One can use the Intermediate Value Theorem in single variable calculus to conclude that there is a root to the polynomial $f(x) = x^3 - 3x + 1$ in the unit interval $I = [0,1]$ (check this!). For starting values every tenth on I, iterate $g(x)$ to estimate this root to three decimal places (it converges quite quickly!). Now try to explain what is happening when you get to both $x_0 = .9$ and $x_0 = 1$.

EXAMPLE 1.12. Autonomous ODEs: One can integrate the autonomous first-order ODE

$$y' = f(y) = (y-2)(y+1), \quad y(0) = y_0,$$

since it is separable, and the integration will involve a bit of partial fraction decomposing. The solution is

$$y(t) = \frac{Ce^{3t} + 2}{1 - Ce^{3t}}. \tag{1.1.2}$$

EXERCISE 7. Calculate Equation 1.1.2 for the ODE in Example 1.12.

EXERCISE 8. Now find the evolution for the ODE in Example 1.12 (this means write the general solution in terms of y_0 instead of the constant of integration C.)

But really, is the explicit solution necessary? One can simply draw the phase line as in Figure 4,

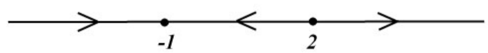


FIGURE 4. The phase line of $y' = (y-2)(y+1)$.

From this schematic view of the long-term tendencies of each solution, one can glean a lot of information about the solutions of the equation. For instance, the equilibrium solutions occur at $y(t) \equiv -1$ and $y(t) \equiv 2$, and that the equilibrium at -1 is asymptotically stable (the one at 2 is unstable). Thus, if long-term behavior is all that is necessary to understand the system, then we have:

$$\lim_{t\to\infty} y(t) = \begin{cases} -1 & \text{if} \quad y_0 < 2 \\ 2 & \text{if} \quad y_0 = 2 \\ \infty & \text{if} \quad y_0 > 2. \end{cases}$$

In both these last two examples, actually solving the dynamical system isn't necessary to gain important and possibly sufficient information about the system.

1.2. The viewpoint

Dynamical Systems, as a field of study, is a type of mathematical analysis; the study of the formal properties of sets of objects and the structures defined on them (think functions, spaces, etc.) You encountered analysis in your calculus classes, although there, one focuses on the more technical aspects of calculating and determining the properties of functions defined on a particular set, the real line $\mathbb{R}$. Indeed, the properties of functions and the spaces that serve as their domains (and codomains) are intimately intertwined in sometimes obvious and often subtle ways. For example, in a celebrated theorem by Luitzen E. J. Brouwer [1912], any continuous function from a compact, convex space to itself must contain at least one point where its image under the function is the same as the point itself (a *fixed point* of the function.) This fact has been celebrated in various ways over the years; often cited as the meaning of the phrase "one cannot comb the hair of a bowling ball". This property has enormous implications for not simply the function we apply to the space, but for the space itself. The consequences of a theorem like this are evident even on the beginning stages of math, like calculus and differential equations.

In general, studying how a map moves around the points of the space is to study the *dynamical content* of the map. Where the points go, upon repeated iteration of a map on a space, or how solutions of a differential equation behave once their parameter domain is known is to study the system *dynamically.* If most or all of the solutions tend to look alike, or if the diversity of the ways a collection of iterates of a point under a map is small, then we say that the dynamics are *simple.* In essence, they are easy to describe, or it does not take a lot of information to describe them. In contrast, if different solutions to the ODE can do many different things, or if it takes a lot of information to describe how many different ways a map can move distinct points around in a space, we say that the dynamics are *complex* or *complicated.* One may say that a dynamical system is more *interesting* if it is more complicated to describe, although that is certainly a subjective term.

Solving a dynamical system, or finding an explicit expression for the evolution, is typically not the general goal of an analysis of a dynamical system. Many non-linear systems of ODEs are difficult if not impossible to solve. Rather, the goal of an analysis of a dynamical system is the general description of the movement of points under the map or the ODE.

In the following chapters, we will develop a language and methods of analysis to study the dynamical content of various kinds of dynamical systems. We will survey both discrete and continuous dynamical systems that exhibit a host of phenomena, and mine these situations for ways to classify and characterize the behavior of the iterates of a map (or solutions of the ODE). We will show how the properties of the maps and the spaces they use as domains affect the dynamics of their interaction. We will start with situations that display relatively simple dynamics, and progress through situations and applications of increasing complexity (complicated behavior). In all of these situations, we will keep the maps and spaces as easy to define and work with as possible, to keep the focus directly on the dynamics.

Perhaps the best way to end this chapter is on a more philosophical note, and allow a possible raison d'etre for why dynamical systems even exists as a field of study enmeshed in the world of analysis, topology and geometry:

Definition 1.13. Dynamical systems is the study of the information contained in and the effects of groups of transformations of a space.

For a discrete dynamical system defined by a map on a space, the properties of the map as well as those of the space, will affect how points are moved around the space. As we will see, maps with certain properties can only do certain things, and if the space has a particular property, like the compact, convex space above, then certain things must be true (or may not), like a fixed-point free transformation. Dynamics is the exploration of these ideas, and we will take this view throughout this text.

CHAPTER 2

Simple Dynamics

2.1. Preliminaries

2.1.1. A simple system. To motivate our first discussion and set the playing field for an exploration of some simple dynamical systems, recall some general theory of first-order autonomous ODEs in one dimension: Let

$$\dot{x} = f(x), \quad x(0) = x_0$$

be an IVP (again, an ODE with an initial value) where the function $f(x)$ is a differentiable function on all of $\mathbb{R}$. From any standard course in differential equations, this means that solutions will exist and be uniquely defined for all values of $t \in \mathbb{R}$ near $t = 0$ and for all values of $x_0 \in \mathbb{R}$. Recall that the general solution of this ODE will be a 1-parameter family of functions $x(t)$ parameterized by x_0. In reality, one would first use some sort of integration technique (as best as one can; remember this ODE is always separable, although $\frac{1}{f(x)}$ may not be easy to integrate. As an example, consider $f(x) = e^{x^2}$) to find $x(t)$ parameterized by some constant of integration C. Then one would solve for the value of C given a value of x_0. Indeed, one could solve generally for C as a function of x_0, and then substitute this into the general solution, to get

$$x(t, x_0) : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$$

as the evolution. Then, for each choice of x_0, we would get a function $x_{x_0}(t) : \mathbb{R} \to \mathbb{R}$ as the particular solution to the IVP. We will use the notation with a subscript for x_0 to accentuate that the role of x_0 is that of a parameter. Specifying a value means solving the IVP for that value of x_0. Leaving x_0 unspecified means that we are looking for a particular solution at a fixed value of x_0. The resulting graph of $x_{x_0}(t)$ would "live" in the tx-plane as a curve (the trajectory) passing through the point $(0, x_0)$. Graphing a bunch of representative trajectories gives a good idea of what the evolution looks like. You did this in your differential equations course when you created phase portraits.

EXAMPLE 2.1. Let $\dot{x} = kx$, with $k \in \mathbb{R}$ a constant. Here, a general solution to the ODE is given by $x(t) = Ce^{kt}$. If, instead, we were given the IVP $\dot{x} = kx$, $x(0) = x_0$, the particular solution would be $x(t) = x_0 e^{kt}$. The trajectories would look like graphs of standard exponential functions (as long as $k \neq 0$) in the tx-plane. Below in Figure 1 are the three cases which look substantially different from each other: When $k > 0$, $k = 0$, and $k > 0$.

Recall in higher dimensions, $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, we typically do not graph solutions explicitly as functions of t. Rather, we use the t-parameterization of solutions $\mathbf{x}(t) = \begin{bmatrix} x_1(t) & x_2(t) & \cdots & x_n(t) \end{bmatrix}^T$ to trace out a curve directly in the $\mathbf{x}$-space. This space, whose coordinates are the set of dependent variables $x_1, x_2, \ldots, x_n$, is called

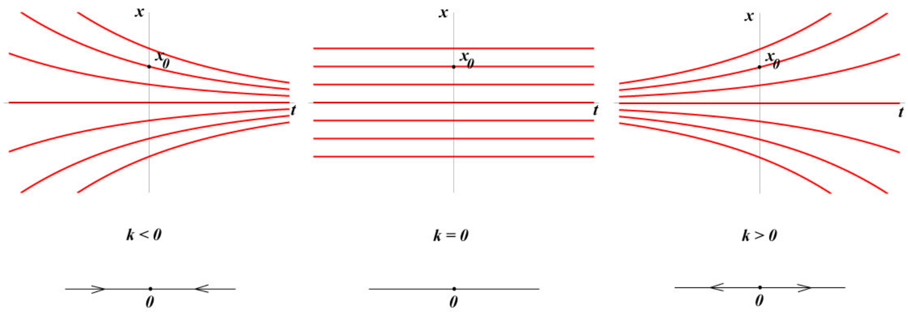


FIGURE 1. Sample solutions for $x_{x_0}(t) = x_0 e^{kt}$.

the phase-space (sometimes the tx-plane from above, or more generally the $t\mathbf{x}$-space is called the *trajectory space* to mark the distinction). The diagrams in the plane that correspond to linear systems with a saddle at the origin, or a spiral sink are examples of phase planes with representative trajectories. Often, particularly in phase space, trajectories are also called *orbits*.

EXAMPLE 2.2. The linear system IVP $\dot{x} = -y$, $\dot{y} = x$, $x(0) = 1$, $y(0) = 0$ has the particular solution $x(t) = \cos t$, $y(t) = \sin t$. Graphing the trajectory, according to the above, means graphing the curve in the txy-space, a copy of $\mathbb{R}^3$. While informative, it may be a little tricky to fully "see" what is going on. But the orbit, graphed in the xy-plane, which is the phase space, is the familiar unit circle (circle of radius 1 centered at the origin). Here $t \in \mathbb{R}$ is the coordinate directly on the circle, and even the fact that it overwrites itself infinitely often is not a serious sacrifice to understanding. See Figure 2.

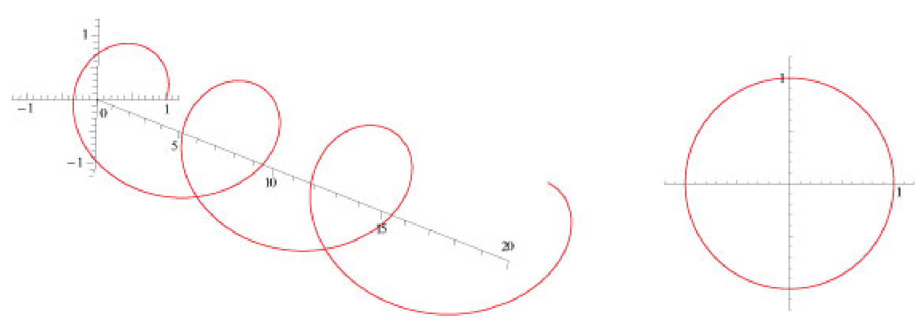


FIGURE 2. Solution curve $x(t) = \cos t$, $y(t) = \sin(t)$ in trajectory space and the phase plane.

Even for autonomous ODEs in one-dependent variable, we designed a schematic diagram called a phase-line to give a qualitative description of the "motion" of solutions to $\dot{x} = f(x)$.

EXAMPLE 2.3. The phase lines for $\dot{x} = kx$ for the three cases in Figure 1 are below the graphs. The proper way to think of these lines is as simply a copy of the vertical axis (the x-axis in this case of the tx-plane) in each of the graphs, marking the equilibrium solutions as special points, and indicating the direction of change of the x-variable as t increases. All relevant information about the long-term behavior is encoded in these phase lines. In fact, these lines ARE the 1-dimensional phase spaces of the ODE, and the arrows simply indicate the direction of the parameterized $x(t)$ inside the line. It is hard to actually see the parameterized curves, since they all run over the top of each other. This is why we graph solutions in 1-variable ODEs using t explicitly, while for ODEs in two or more dependent variables, we graph using t implicitly, as the coordinate directly ON the curve in the phase space.

2.1.2. The time-t map. Again, for $\dot{x} = f(x)$, $x(0) = x_0$, the general solution $x(t, x_0) : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is a 1-parameter family of solutions, written as $x_{x_0}(t)$, parameterized by x_0. However, we can also think of this family of curves in a much more powerful way: As a 1-parameter family of transformations of the phase space! To see this, rewrite the general solution as $\varphi(t, x_0) : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ instead of the possibly confusing notation $x(t, x_0)$. Now instead of thinking of x_0 as the parameter, fixing the second argument and varying the first as the independent variable, do it the other way: Fix a value of t, and allow the variable $x_0 = x$ (the starting point) to vary. Then we get for $t = t_0$:

$$\varphi(t_0, x) : \mathbb{R} \times \mathbb{R} \to \mathbb{R}, \quad \varphi_{t_0}(x) : \mathbb{R} \to \mathbb{R}, \quad x(0) \longrightarrow x(t_0).$$

As t varies, every point $x \in \mathbb{R}$ (thought of as the initial point $x(0)$, gets "mapped" to its new position at $x(t_0)$. Since all solutions are uniquely defined, this is a function for each value of t_0, and will have some very nice properties. But this alternate way of looking at the solutions of an ODE, as a family of transformations of its phase space, is the true *dynamical view*, and one we will explore frequently.

Place a picture of how points move around phase space at time t_0. Viewing this as solely a transformation of phase space is the dynamical view.

Let X denote any particular topological space. For now, though, just think of X as some subset of the real space $\mathbb{R}^n$, something you are familiar with.

DEFINITION 2.4. For $f : X \to X$ a map, define the set

$$\mathcal{O}_x = \left\{ y \in X \;\middle|\; y = f^n(x), \;\; n \in \mathbb{N} \right\}$$

as the (forward) orbit of $x \in X$ under f.

Some notes:

- We define $\mathcal{O}_x$ as a set of points in X, but it is really more than just a set. It is a collection of points in X ordered, or parameterized, by the natural numbers: a sequence. Hence we often write $\mathcal{O}_x = \left\{x, f(x), f^2(x), \ldots\right\}$ to note the order, or for $x_{n+1} = f(x_n)$, $\mathcal{O}_x = \left\{x_0, x_1, x_2, \ldots\right\}$.
- If f is invertible, we can also then define the backward orbit

$$\mathcal{O}^-_x = \left\{ y \in X \;\middle|\; y = f^{-n}(x), \;\; n \in \mathbb{N} \right\},$$

or the full orbit

$$\mathcal{O}_x = \left\{ y \in X \,\middle|\, y = f^n(x), \;\; n \in \mathbb{Z} \right\},$$

rewriting the forward orbit then as $\mathcal{O}_x^+$. However, at times, even for invertible maps, we are only concerned with the forward orbit and simply write $\mathcal{O}_x$, using context for clarity.

Consider the discrete dynamical system $f : \mathbb{R} \to \mathbb{R}$, given by $f(x) = rx$, $r > 0$. What do the orbits look like? Basically, for $x \in \mathbb{R}$, we get

$$\mathcal{O}_x = \left\{ x, rx, r^2x, r^3x, \dots, r^nx, \dots \right\}.$$

In fact, we can "solve" this dynamical system by constructing the evolution

$$\Phi(x, n) = r^n x.$$

Do the orbits change in nature as one varies the value of r? How about when r is allowed to be negative? How does this relate to the ordinary differential equation $\dot{x} = kx$?

DEFINITION 2.5. For $t \geq 0$, the *time-t* map of a continuous dynamical system is the transformation of state space which takes $x(0)$ to $x(t)$.

EXAMPLE 2.6. Let $k < 0$ in $\dot{x} = kx$, with $x(0) = x_0$. Here, the state space is $\mathbb{R}$ (the phase space, as opposed to the trajectory space $\mathbb{R}^2$), and the general solution is $\Phi(x_0, t) = x_0 e^{kt}$ (the evolution of the dynamical system is $\Phi(x, t) = xe^{kt}$. Notice that

$$\Phi(x, 0) = x, \quad \text{while} \quad \Phi(x, 1) = e^k x.$$

Hence the time-1 map is simply multiplication by $r = e^k$. The time-1 map is the discrete dynamical system on $\mathbb{R}$ given by the function above $f(x) = rx$. In this case, $r = e^k$, where $k < 0$, so that $0 < r = e^k < 1$. See Figure 2.1.2. Now how do the orbits behave?

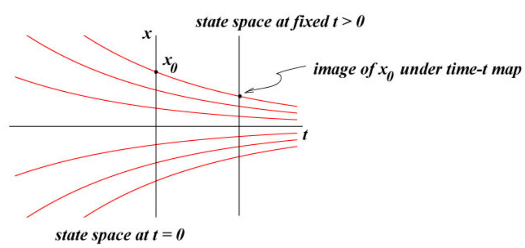


FIGURE 3. The time-t map for some positive time of $\dot{x} = kx$, $k < 0$.

EXERCISE 9. Given any dynamical system, describe the time-0 map.

DEFINITION 2.7. For a discrete dynamical system $f : X \to X$, a *fixed point* is a point $x_* \in X$, where $f(x_*) = x_*$, or where

$$\mathcal{O}_{x_*} = \left\{ x_*, x_*, x_*, \dots \right\}.$$

The orbit of a fixed point is also called a trivial orbit. All other orbits are called non-trivial.

In our example above, $f : \mathbb{R} \to \mathbb{R}$, $f(x) = e^k x$, $k < 0$, we have $x = 0$ as the ONLY fixed point. This corresponds nicely with the unique particular solution to the ODE $\dot{x} = kx$ corresponding to the equilibrium $x(t) \equiv 0$.

So what else can we say about the "structure" of the orbits? That is, what else can we say about the "dynamics" of this dynamical system? For starters, the forward orbit of a given x_0 will look like the graph of the discrete function $f_{x_0} : \mathbb{N} \to \mathbb{R}^2, f_{x_0}(n) = x_0 e^{kn}$. Notice how this orbit follows the trajectory of x_0 of the continuous dynamical system $\dot{x} = kx$. Here, f is the time-1 map of the ODE. Notice also that, as a transformation of phase space (the x-axis), f is not just a continuous function but a differentiable one, with $0 < f'(x) = e^k < 1$, $\forall x \in \mathbb{R}$. The orbit of the fixed point at $x = 0$, as a sequence, certainly converges to 0. But here ALL orbits have this property, and we can say

$$\forall x \in \mathbb{R}, \quad \lim_{n\to\infty} \mathcal{O}_x = 0, \text{ or } \mathcal{O}_x \longrightarrow 0.$$

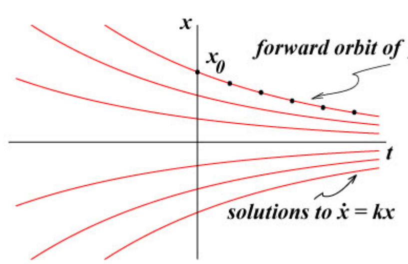


FIGURE 4. The forward orbit of $f(x) = xe^k$ lives on a solution to $\dot{x} = kx$, $k < 0$.

This gives a sense of what we will mean by a dynamical system exhibiting *simple dynamics*: If with very little effort or additional structure, one can completely describe the nature of all of the orbits of the system. Here, there is one fixed point, and all orbits converge to this fixed point.

DEFINITION 2.8. For a discrete dynamical system, a smooth curve (or set of curves) ℓ in state space is called an *orbit line* if $\forall x \in \ell$, $\mathcal{O}_x \subset \ell$.

EXAMPLE 2.9. The orbit lines for time-t maps of ODEs are the trajectories of the ODE.

EXERCISE 10. Go back to Figure 1. Describe completely the orbit structure of the discrete dynamical system $f(x) = rx$ for other two cases, when $r = 1$ and $r > 1$ (corresponding to $r = e^k$, for $k = 0$ and $k > 0$, respectively). That is, classify all possible different types of orbits, in terms of whether they are fixed or not, where they go as sequences, and such. You will find that even here, the dynamics are simple, but at least for the $k > 0$ case, one has to be a little more careful about accurately describing where orbits go.

EXERCISE 11. As in the previous exercise, describe the dynamics of the discrete dynamical system $f(x) = rx$, when $r < 0$ (again, there are cases here). In particular, what do the orbit lines look like in this case? You will find that this case does not, in general, correspond to a time-t map of the ODE $\dot{x} = kx$ for any value of k (why not?)

EXERCISE 12. Show that there does not exist a first-order, autonomous ODE where the map $f(x) = rx$ corresponds to the time-1 map, when $r < 0$.

EXERCISE 13. Construct a second-order ODE whose time-1 map is $f(x) = rx$, where $r < 0$ is any given constant.

EXERCISE 14. For the discrete dynamical system $f : \mathbb{R} \to \mathbb{R}$, $f(x) = rx + b$, calculate the evolution in closed form. Then completely describe the orbit structure when $b \neq 0$, noting in particular the different cases for different values of $r \in \mathbb{R}$.

REMARK 2.10. Is there really any difference in the dynamical content of the discrete dynamical systems given by the linear map $f : \mathbb{R} \to \mathbb{R}$, $f(x) = kx$, and the affine map $g(x) = kx + b$? By a linear change of variables (a translation here), one can change one to the other in a way that orbits go to orbits. And if this is so, is there any real need to study the more general affine map g once we know all of the characteristics of maps like f? We will explore this idea later in the concept of *topological conjugation*; the idea that two dynamical systems can be equivalent. For now, a simple exercise:

EXERCISE 15. For $k \neq 1$, and f and g as in Remark 2.10 find a linear change of variables $h : \mathbb{R} \to \mathbb{R}$ that satisfies the condition that $g \circ h(x) = h \circ f(x)$ (equivalently, that $f(x) = h^{-1} \circ g \circ h(x)$.

EXERCISE 16. Given $\dot{x} = f(x)$, $f \in C^1(\mathbb{R})$, recall that an *equilibrium solution* is defined as a constant function $x(t) \equiv c$ which solves the ODE. They can be found by solving $f(x) = 0$ (remember this?) Instead, define an equilibrium solution $x(t)$ as follows: A solution $x(t)$ to $\dot{x} = f(x)$ is called an *equilibrium solution* if there exists $t_1 \neq t_2$ in the domain of $x(t)$ where $x(t_1) = x(t_2)$. Show that this new definition is equivalent to the old one.

EXERCISE 17. For the first-order autonomous ODE $\frac{dp}{dt} = \frac{p}{2} - 450$, do the following:

- Solve the ODE by separating variables. Justify explicitly why the absolute value signs are not necessary when writing the general solution as a single expression.
- Calculate the time-1 map for this ODE flow.
- Discuss the simple dynamics of this discrete dynamical system given by the time-1 map.

2.1.3. Contractions. The above questions are all good to explore. For now, the above example $f(x) = e^k x$, where $k < 0$, is an excellent example of a particular class of dynamical systems which we will discuss presently.

DEFINITION 2.11. A *metric* on a subset of Euclidean space $X \subset \mathbb{R}^n$ is a function $d : X \times X \to \mathbb{R}$ where

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

One such choice of metric is the "standard Euclidean distance" metric

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2},$$

where $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$. Note that for $n = 1$, this metric reduces to $d(x, y) = \sqrt{(x - y)^2} = |x - y|$.

Exercise 18. Explicitly show that the standard Euclidean distance metric is indeed a metric by showing that it satisfies the three conditions.

Exercise 19. On $\mathbb{R}^n$, define a notion of distance by $d(\mathbf{x},\mathbf{y}) = \sum_{i=1}^n |x_i - y_i|$. Show this notion of distance is a metric. (For $n = 2$, this is sometimes called the taxicab or the Manhattan metric. Can you see why?)

Exercise 20. Again, on $\mathbb{R}^n$, show that

$$d(\mathbf{x},\mathbf{y}) = \max\{|x_1 - y_1|, |x_2 - y_2|, \dots, |x_n - y_n|\}$$

is a metric. (This norm is referred to as the maximum metric, or max metric.)

Remark 2.12. One can define notions of distance between points in vector spaces via vector norms. All of the previous three examples above are members of a family of distances defined by vector norms called the $\mathbf{L}^p$-norms or simply **p**-norms. Euclidean distance corresponds to $\mathbf{p} = 2$, the Manhattan distance is $\mathbf{p} = 1$, and the last, the maximum metric comes from the ∞-norm or maximum norm. All are defined via the **p**-norm

$$\|x\|_p = \left(\sum |x_i|^p\right)^{\frac{1}{p}}.$$

The ∞-norm is so-named since it is the norm formed by letting $\mathbf{p} \longrightarrow \infty$.

Exercise 21. On $\mathbb{R}^2$, consider a notion of distance defined by the following:

$$d(\mathbf{x},\mathbf{y}) = \begin{cases} |x_1 - y_1| & \text{if} \quad x_1 \neq y_1 \\ |x_2 - y_2| & \text{if} \quad x_1 = y_1. \end{cases}$$

This is similar to a lexicographical ordering of points in the plane. Show that this notion of distance is NOT a metric on $\mathbb{R}^2$.

Exercise 22. The original definition of a circle as a planar figure comes directly from Euclid himself: A *circle* is the set of points in the plane equidistant from a particular point. Naturally, using the Euclidean metric, a circle is what you know well as a circle. Show that circles in the taxicab metric on $\mathbb{R}^2$ are squares whose diagonals are parallel to the coordinate axes.

Exercise 23. Following on the previous exercise, construct a metric on $\mathbb{R}^2$ whose circles are squares whose sides are parallel to the coordinate axes. (Hint: Rotate the taxicab metric.)

Exercise 24. Let S be a circle of radius $r > 0$ centered at the origin of $\mathbb{R}^2$. It's circumference is $2\pi r$. Euclidean distance in the plane does restrict to a metric directly on the circle. Here instead, construct a metric on the circle using arc-length, and verify that it is a metric. (Be careful about measuring distance correctly.)

Remark 2.13. When discussing points in Euclidean space, it is conventional to denote scalars (elements of $\mathbb{R}$) with a variable in italics, and vectors (elements of $\mathbb{R}^n$, $n > 1$) as a variable in boldface. Thus $\mathbf{x} = (x_1, x_2, \dots, x_n)$. In the above definition of a metric, we didn't specify whether X was a subset of $\mathbb{R}$ or something larger. In the absence of more information regarding a space X, we will always use simple italics for its points, so that $x \in X$, even if it is possible that $X = \mathbb{R}^5$, for example. We will only resort to the vector notation when it is assured that we are specifically talking about vectors of a certain size. This is common in higher mathematics like topology.

DEFINITION 2.14. A map $f : X \subset \mathbb{R}^n \to \mathbb{R}^m$, is called *Lipschitz continuous* (with constant λ), or λ-Lipschitz, if

$$d(f(x), f(y)) \leq \lambda d(x, y), \quad \forall x, y \in X. \tag{2.1.1}$$

Some notes:

- The set X can always inherit the metric on $\mathbb{R}^n$ simply by declaring that the distance between two points in X is defined by the their distance in $\mathbb{R}^n$ (See Exercise 24). So subsets of $\mathbb{R}^n$ are always metric spaces. This includes the subset of $\mathbb{R}^m$ that is the range of f, often called the *image* of f, denoted $f(X) \in \mathbb{R}^m$. One can always define a different metric on X (or its image) if one wants. But the fact that X is a metric space comes for free, as they sometimes say.
- λ is a bound on the stretching ability (comparing the distances between the images of points in relation to the distance between their original positions) of f on X. This is actually a stronger form of continuity called *uniform continuity*: Lipschitz functions are always continuous, but there are continuous functions that are not Lipschitz. Basically, if the ratio of the distance between two images $f(x)$ and $f(y)$ to the distance between x and y is bounded by λ across all of X, then f is λ-Lipschitz.
- To get a better sense for Lipschitz continuity, consider the following: On a bounded interval in $\mathbb{R}$, polynomials are always Lipschitz continuous. But on $\mathbb{R}$ itself, only the constants and the linear polynomials are λ-Lipschitz. Rational functions, on the other hand, even though they are continuous and differentiable on their domains, are not Lipschitz continuous on any interval whose closure contains a vertical asymptote. $\sin x$ is 1-Lipschitz on $\mathbb{R}$, but $\tan x$ is not Lipshitz continuous on its domain. And a function like e^x?
- It should be obvious that $\lambda > 0$. Why?
- We can define
$$\mathrm{Lip}(f) = \sup_{x \neq y} \frac{d(f(x), f(y))}{d(x, y)},$$
which is the infimum of all λ's that satisfy Equation 2.1.1. When we speak of specific values of λ for a λ-Lipschitz function, we typically use $\lambda = \mathrm{Lip}(f)$, if known.

EXERCISE 25. Show for $f : \mathbb{R} \to \mathbb{R}$ that Lipschitz continuity implies continuity.

EXERCISE 26. Let $f(x) = \frac{1}{x}$. Show f is Lipschitz continuous on any domain (a, b), $a > 0$, $a < b \leq \infty$, and for any particular choice of a and b, produce the constant λ. Then show that f is not Lipschitz continuous on $(0, \infty)$.

EXERCISE 27. Show that $h(x) = |x|$ is Lipschitz continuous on R and produce $\mathrm{Lip}(h)$.

EXERCISE 28. Show that $g(x) = e^x$ is NOT Lipschitz continuous on $\mathbb{R}$.

EXERCISE 29. For a given non-negative λ, construct a function whose domain in all of $\mathbb{R}$, that is precisely λ-Lipschitz continuous on $I = (-\infty, 2) \cup (2, \infty)$ but not Lipschitz continuous.

EXERCISE 30. Produce a function that is continuous on $I = [-1, 1]$ but not Lipschitz continuous there.

It should be clear now that, for a real-valued function on $\mathbb{R}$, since we are measuring ratios of image distances to point distances, the derivative of f, if it exists, can say a lot about Lipschitz continuity. One must be careful about the domain, however:

PROPOSITION 2.15. *Let $f : I \subset \mathbb{R} \to \mathbb{R}$ be differentiable on an open interval I, where $\forall x \in I$, we have $|f'(x)| \leq \lambda$. Then f is λ-Lipschitz.*

PROOF. Really, this is simply an application of the Mean Value Theorem: For a function f differentiable on a bounded, open interval (a, b) and continuous on its closure, there is at least one point $c \in (a, b)$ where $f'(c) = \frac{f(b)-f(a)}{b-a}$, the average total change of the function over $[a, b]$. Here then, for any $x, y \in I$ (thus ALL of $[x, y] \in I$ even when I is neither closed nor bounded), there will be at least one $c \in I$ where

$$d\left(f(x), f(y)\right) = |f(x) - f(y)| = |f'(c)||x - y| \leq \lambda|x - y| = \lambda d(x, y).$$

□

DEFINITION 2.16. A λ-Lipschitz function $f : X \subset \mathbb{R}^n \to \mathbb{R}^m$ on a metric space X is called a *contraction* if $\lambda < 1$.

Note here that the definition here for a contraction as well as the general definition above of Lipschitz continuity both allow for the domain and range to be two different metric spaces. When using the function f as the fixed rule of a discrete dynamical system, however, we want the codomain and domain to be the same, and would define $f : X \to X$ to be a contraction if it is Lipschitz continuous with $\lambda < 1$.

EXAMPLE 2.17. Back to the previous example $f : \mathbb{R} \to \mathbb{R}$, $f(x) = e^k x$, the time-1 map of the ODE $\dot{x} = kx$. Given that $f'(x) = e^k$ everywhere, in the case that $k < 0$, the map f is a contraction on ALL of $\mathbb{R}$. Indeed, using the Euclidean metric in $\mathbb{R}$, we have

$$d(f(x), f(y)) = \left|e^k x - e^k y\right| = \left|e^k(x - y)\right| = \left|e^k\right||x - y| = e^k|x - y| = \lambda|x - y|$$

for all $x, y \in \mathbb{R}$, where $\lambda = e^k < 1$.

EXERCISE 31. Without using derivative information, show that $f(x) = ax + b$ is a-Lipschitz on $\mathbb{R}$.

EXERCISE 32. Again without using derivative information, show that the monomial x^n, $n \in \mathbb{N}$ is na^{n-1}-Lipschitz on the interval $[0, a]$

EXERCISE 33. Find a for the largest interval $[0, a]$ where $f(x) = 3x^2 - 2$ is a contraction.

Before we continue, we need to clarify some of the properties of the intervals we will be using in our dynamical systems. Here are a couple of definitions:

DEFINITION 2.18. A subset $U \in \mathbb{R}$ is called bounded if there exists a number $M > 0$ so that $\forall x \in U$, we have $|x| < M$.

DEFINITION 2.19. An interval I is called *closed* in $\mathbb{R}$ if it contains all of its limit points. If the interval is bounded (as a subset of $\mathbb{R}$), then this means that I includes its endpoints. But closed intervals need not be bounded. Hence closed intervals in $\mathbb{R}$ take one of the forms $[a, b]$, $(-\infty, b]$, $[a, \infty)$ or $(-\infty, \infty)$, for $-\infty < a \leq b < \infty$.

DEFINITION 2.20. Let $I = [a, b] \in \mathbb{R}$ be a closed, bounded interval, and $f : I \to I$ a map. We say f is *continuously differentiable*, or simply *differentiable* on I, if f is differentiable on (a, b) and differentiable from the right at a and from the left at b. In essence, this means that $f'(x)$ is continuous on I.

PROPOSITION 2.21. *Let $f : I \to I$ be continuously differentiable on I, a closed, bounded interval, with $|f'(x)| < 1$ $\forall x \in I$. Then f is a contraction.*

PROOF. $f'(x)$ is continuous on I so it will achieve its maximum there by the Extreme Value Theorem, and

$$\max_{x \in I} |f'(x)| = \lambda < 1.$$

Now apply Proposition 2.15. □

NOTE. *If I is not closed, or is not bounded, this may NOT be true. See Exercise 34.*

EXERCISE 34. Show $f(x) = 2\sqrt{x}$ is NOT a contraction on $(1, \infty)$.

DEFINITION 2.22. For $f : X \to X$ a map, a point $x \in X$ is called *periodic* (with period n) if $\exists n \in \mathbb{N}$ such that $f^n(x) = x$. The smallest such natural number is called the *prime period* of x.

Notes:

- If $n = 1$, then x is a fixed point.
- Define

$$\mathrm{Fix}(f) = \left\{ x \in X \,\middle|\, f(x) = x \right\}$$

$$\mathrm{Per}_n(f) = \left\{ x \in X \,\middle|\, f^n(x) = x \right\}$$

$$\mathrm{Per}(f) = \left\{ x \in X \,\middle|\, \exists n \in N \text{ such that } f^n(x) = x \right\}.$$

Keep in mind that these sets are definitely not mutually exclusive, and for $m, n \in \mathbb{N}$, $\mathrm{Per}_m(f) \subset \mathrm{Per}_n(f)$ precisely when $m \big| n$ (when n is an integer multiple of m.)

2.2. The Contraction Principle

Lipschitz continuous functions, in a rough sense, are fairly well-behaved when it comes to iteration. Basically, the distance between the images of two points in a domain of a Lipschitz continuous function are bounded relative to the original distance between the two points. Thus upon iteration, in the situation when the domain and the codomain are the same, two orbits that start close together cannot diverge faster than exponentially (recall that near a vertical asymptote of a rational function, this is not so.) When the Lipschitz constant is less than 1, then the images of points are actually closer together than their original distances. This has great consequences when trying to understand how orbits behave when viewing a map as a discrete dynamical system. Basically, in a Lipschitz contraction, all of the orbits can be described rather simply, as they all wind up going to the same place in the long term. We will say, in this case, that the map exhibits simple dynamics.

2.2.1. Contractions on intervals. For discrete dynamical systems defined on (subsets of) the real line, we can generalize that contractions have some special properties. One very special feature is that the fixed point set is always non-empty, but very small in size:

THEOREM 2.23 (the Contraction Principle). *Let $I \subset \mathbb{R}$ be a closed interval, and $f : I \to I$ a λ-contraction. Then f has a unique fixed point x_0, and*

$$|f^n(x) - x_0| \leq \lambda^n |x - x_0|.$$

Some notes before we prove this theorem:

- $\forall x \in I$, $\mathcal{O}_x \longrightarrow x_0$ as a sequence exponentially due to the factor λ^n, where $0 < \lambda < 1$.
- As stated, this is only a result valid on subsets of $\mathbb{R}$, for now.
- As a way to understand this theorem, first think about why a contraction cannot have more than on fixed point:

  EXERCISE 35. Show, without using the Contraction Principle, that a contraction cannot have two fixed points.

- For the proof of the theorem, we will need a convenient fact from Analysis: That the idea of a sequence converging in real space is the same as the fact that the sequence is *Cauchy*:

DEFINITION 2.24. For $N \in \mathbb{N}$, a sequence $\{x_1, x_2, \ldots\} \in \mathbb{R}^N$ is called *Cauchy* if $\forall \epsilon > 0$, $\exists A > 0$ such that $\forall m, n \geq A$, $d(x_m, x_n) < \epsilon$.

PROPOSITION 2.25. *A sequence in $\mathbb{R}^N$, $N \in \mathbb{N}$ converges iff it is Cauchy.*

PROOF OF CONTRACTION PRINCIPLE (IN $\mathbb{R}$). The proof of this theorem will consists of three parts: (1) That the orbit of an arbitrary point converges to something in the interval; (2) that any two orbits also converge (point-wise); and (3) that the thing all of these orbits converge to is actually an orbit (the orbit of a fixed point).

We start with the first of these. Choose $x \in I$. It should be obvious since f is a map on I that $\mathcal{O}_x \in I$. Now, for $m, n \in \mathbb{N}$, where we can assume without any loss of generality that $m \geq n$. Then

$$\begin{aligned}
|f^m(x) - f^n(x)| &= \big| f^m(x) - f^{m-1}(x) + f^{m-1}(x) - f^{m-2}(x) + f^{m-2}(x) - \ldots \\
&\qquad - f^{n+1}(x) + f^{n+1}(x) - f^n(x) \big| \\
&= \left| \sum_{r=n}^{m-1} \left(f^{r+1}(x) - f^r(x) \right) \right|.
\end{aligned}$$

This is a clever form of "addition by zero", and allows us to see additional structure not obvious in the calculation. Thus

$$\begin{aligned}
|f^m(x) - f^n(x)| &= \left| \sum_{r=n}^{m-1} \left(f^{r+1}(x) - f^r(x) \right) \right| \\
&\le \sum_{r=n}^{m-1} \left| f^{r+1}(x) - f^r(x) \right| \quad \text{(triangle inequality in the matric space)} \\
&\le \sum_{r=n}^{m-1} \lambda^r |f(x) - x| \quad \text{(r applications of the Lipschitz condition)} \\
&= |f(x) - x| \sum_{r=n}^{m-1} \lambda^r \\
&= |f(x) - x| \frac{\lambda^n - \lambda^m}{1-\lambda} \quad \text{(partial sum of a geometric series)} \\
&\le \frac{\lambda^n}{1-\lambda} |f(x) - x|.
\end{aligned}$$

The conclusion from all of this is that, as n gets large (under the assumption that $m \ge n$, this means that m gets large also), the right hand side of the last inequality gets small. And n can always be chosen large enough so that the distance between late terms in the sequence is less than some chosen ϵ. Hence, $\mathcal{O}_x$ is a Cauchy sequence. Hence it must converge to something. As I is closed, it must converge to something in I. Let's call this number x_0.

Now, applying the Lipschitz condition to the nth iterate of f leads directly to the condition that

$$|f^n(x) - f^n(y)| \le \lambda^n |x - y|,$$

so that every two orbits also converge to each other. Hence every orbit converges to this number x_0.

And finally, we can show that x_0 is actually a fixed point solution for f. To see this, again $\forall x \in I$, and $\forall n \in \mathbb{N}$,

$$\begin{aligned}
|x_0 - f(x_0)| &= \left| x_0 - f^n(x) + f^n(x) - f^{n+1}(x) + f^{n+1}(x) - f(x_0) \right| \\
&\le |x_0 - f^n(x)| + \left| f^n(x) - f^{n+1}(x) \right| + \left| f^{n+1}(x) - f(x_0) \right| \\
&\le |x_0 - f^n(x)| + \lambda^n |x - f(x)| + \lambda |f^n(x) - x_0| \\
&= (1+\lambda) |x_0 - f^n(x)| + \lambda^n |x - f(x)|.
\end{aligned}$$

Again, the steps are straightforward. The first step is another clever addition of zero, and the second is the triangle inequality. The third involves using the Lipschitz condition on two of the terms, and the last is a clean up of the leftovers. However, this last inequality is the most important. It must be valid for every choice of $n \in \mathbb{N}$. Hence choosing an increasing sequence of values for n, we see that as $n \to \infty$, both $|x_0 - f^n(x)| \longrightarrow 0$ and $\lambda^n \longrightarrow 0$. Hence, it must be the case that $|x_0 - f(x_0)| = 0$ or $x_0 = f(x_0)$. Thus, x_0 is a fixed point solution for f on I and the theorem is established. $\square$

Example 2.26. $f(x) = \sqrt{x}$ on the interval $[1, \infty)$ is a $\frac{1}{2}$-contraction. Why? $f \in C^1$ and $0 < f'(x) = \dfrac{1}{2\sqrt{x}} \le \dfrac{1}{2}$ on $[1, \infty)$, and strictly less than $\frac{1}{2}$ on $(1, \infty)$. Thus, by Proposition 2.15, f is $\frac{1}{2}$-Lipschitz. So by the Contraction Principle, then, there is a unique fixed point for this discrete dynamical system. Can you find it?

EXERCISE 36. Without using any derivative information (that is, without using Propositions 2.15 or 2.21 above), show that $f(x) = \sqrt{x}$ is a $\frac{1}{2}$-contraction on $[1, \infty)$.

EXERCISE 37. Find all periodic points (to an accuracy of $\frac{1}{1000}$) of the discrete dynamical system given by the map $f(x) = \ln(x-1) + 5$ on the interval $I = [2, 100]$.

2.2.2. Contractions in several variables. The Contraction Principle in several variables is basically the same as that of one variable. The only difference, really, is that the absolute signs are replaced by the more general metric in $\mathbb{R}^n$. To start, recall in any metric space X, we can define a small open set via the strict inequality:

$$B_\epsilon(x) = \left\{ y \in X \,\middle|\, d(x,y) < \epsilon \right\}.$$

DEFINITION 2.27. A subset $U \in X$ is called *open* if $\forall x \in U$, $\exists \epsilon > 0$ such that $B_\epsilon(x) \in U$. And $U \in X$ is called *closed* if its complement in X is open.

Note that for any $x \in U$, where U is open in X, we say U is a *neighborhood* of x in X, and write $U(x) \in X$.

DEFINITION 2.28. a point $x \in X$ is called a *boundary point* of a subset $U \in X$ if every neighborhood of x contains at least one point in U and one point not in U.

DEFINITION 2.29. a subset $U \in X$ is called *closed* in X if it contains all of its boundary points in X.

In a loose sense, one can think of a closed subset of real space as a set of solutions to either equations or inequalities of the form $\leq$ or $\geq$. In this fashion, curves in the plane and surfaces in $\mathbb{R}^3$ are closed sets, although ones without interior points (every point is a boundary point). Often, in vector calculus, though, the closed sets constructed as domains for functions are open sets together with their closure, formed by adding to the open set the set of all boundary points.

EXAMPLE 2.30. For $U(x) = B_\epsilon(x)$ the open ϵ-ball centered at $x \in \mathbb{R}^n$, its closure is

$$\overline{U}(x) = \overline{B}_\epsilon(x) = \left\{ y \in X \,\middle|\, d(x,y) \leq \epsilon \right\}.$$

The boundary of $\overline{U}(x)$, sometimes written $\partial\overline{U}$, is the set of all points $y \in \mathbb{R}^n$ where $d(x,y) = \epsilon$. From vector calculus, recall that this is the (n-1)-dimensional sphere in $\mathbb{R}^n$ of radius ϵ centered at x.

Keep this in mind as we generalize the Contraction Principle from $\mathbb{R}$ to $\mathbb{R}^n$, $n > 1$.

THEOREM 2.31 (The Contraction Principle). *Let $X \subset \mathbb{R}^n$ be closed and $f : X \to X$ a λ-contraction. Then f has a unique fixed point $x_0 \in X$ and $\forall x \in X$,*

$$d\left(f^n(x), x_0\right) \leq \lambda^n d\left(x, x_0\right).$$

Notes:

- Again, we say here that the "dynamics are simple". The orbit of every point in x does exactly the same thing: Converge exponentially to the fixed point solution x_0.
- This also means that every orbit converges to every other orbit also!

- What about periodic points? Can contractions have Periodic points other than fixed points. The answer is no:

  EXERCISE 38. Show that a contraction cannot have a non-trivial periodic point (periodic point of prime period greater than 1.)

- The Contraction Principle is also called the Contraction Mapping Theorem, or sometimes the Banach Fixed Point Theorem.

Recall in dimension-1, the derivative of f (if it exists) can help to define the Lipschitz constant (Recall Propositions 2.15 and 2.21 above.) How about in several dimensions? In this context, recall that for a C^1 function $f : X \subset \mathbb{R}^n \to \mathbb{R}^n$, the derivative at $x \in X$, $df_x : T_x\mathbb{R}^n \to T_{f(x)}\mathbb{R}^n$ is a linear map from $T_x\mathbb{R}^n$ to $T_{f(x)}\mathbb{R}^n$. The points in the domain X where the $n \times n$-matrix df_x is of maximal rank are called regular. At a domain point, we can use the Euclidean norm for vectors to define a matrix norm for df_x as the maximal stretching ability of the unit vectors in the tangent space:

$$||df_x|| = \max_{||\mathbf{v}||=1} ||df_x(\mathbf{v})|| .$$

This non-negative number is not difficult to find for an $n \times n$ matrix A: If A is symmetric, then $||A||$ is just the spectral radius $\rho(A)$, the absolute value of the largest (in magnitude) eigenvalue of A. For general $n \times n$ A, it is $\sqrt{\rho(A^T A)}$.

With this, there are two "derivative"-versions of the Contraction Principle of note. The first is a sort of "global version" since the result holds over the entire domain. We will require that the space be strictly convex, however (See Figure 5).

DEFINITION 2.32. A subset $X \in \mathbb{R}^n$ is *convex* if for any two points $x, y \in X$, the straight line segment joining x and y lies entirely in X. X is called *strictly convex*, if for any two boundary points $x, y \in X$, the line segment joining x to y intersects the boundary of X only at x and y.

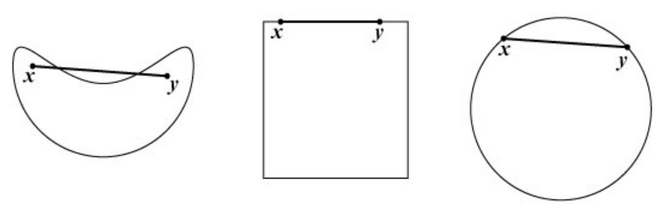


FIGURE 5. Convexity: Non-convex, convex, but not strictly, and strictly convex sets, respectively.

THEOREM 2.33 (Global version). *If X is the closure of a strictly convex set in $\mathbb{R}^n$, and $f : X \to \mathbb{R}^n$ a C^1-map with $||df_x|| \le \lambda < 1$, $\forall x \in X$, then f has a unique fixed point x_0, and $\forall x \in X$,*

$$d(f^n(x), x_0) \le \lambda^n d(x, x_0).$$

Note here that, technically, we really want that f is differentiable on the interior of X and continuous on the boundary.

EXERCISE 39. Let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be the affine map $f(x,y) = \left(\frac{1}{2}x + \frac{1}{3}y, -\frac{1}{2}x + \frac{1}{2}y + 1\right)$. Find the fixed point of f, and show that f is a contraction on $\mathbb{R}^2$.

This next version is a local version, and presents a very useful tool for the analysis of what happens near (in a neighborhood of) a fixed point:

THEOREM 2.34 (Local version). *Let f be differentiable with a fixed point x_0 such that all of the eigenvalues of the derivative matrix df_{x_0} have absolute values less than 1. Then there exists $\epsilon > 0$ and an open neighborhood $U(x_0) = B_\epsilon(x_0)$ such that on the closure $\overline{U}$ of U, $f(\overline{U}) \subset \overline{U}$ and f is a contraction on $\overline{U}$.*

Thus, on $\overline{U}$, which will be a strictly convex set, the global version above applies. To see these two versions in action, here are two applications:

2.2.3. Application: Newton-Raphson Method.

An iterative procedure for the location of a root of a C^2-function $f : \mathbb{R} \to \mathbb{R}$, called the Newton-Raphson Method, is an application of the local version of the contraction principle. This method is part of a series of approximation methods that utilize the Taylor Polynomials of a function to help identify important features of the function, and is typically found in most standard calculus texts as an application using the tangent line approximation to a function. Here, let f be C^1 near an unknown root x_*. Then, for a point x_0 near the root, the tangent line approximation to f at x_0 is

$$f(x) - f(x_0) = f'(x_0)\,(x - x_0)\,.$$

Under relatively mild conditions for f and for x_0 "close" enough to x_*, the tangent line function $f(x) = f(x_0) + f'(x_0)\,(x - x_0)$ will also have a root. Call this point x_1. Solving the tangent line function for x_1, we get

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

Again, under mild conditions on f, the number x_1 lies closer to x_* than x_0, and can serve as a new approximation to x_*.

Repeating this procedure yields a discrete dynamical system given by $x_{n+1} = g(x_n)$, where $g(x) = x - \frac{f(x)}{f'(x)}$. Some things to note:

- If f is C^2 near the root x_*, then as long as $f'(x)$ does not vanish in a neighborhood of x_*, then g is C^1 there. And
$$g'(x) = 1 - \frac{(f'(x))^2 - f(x)f''(x)}{(f'(x))^2} = \frac{f(x)f''(x)}{(f'(x))^2}.$$
- If there exists an open interval containing x_*, and if on this interval, there are positive constants δ, M, where $|f'(x)| > \delta$ (f' doesn't get too small), and $|f''(x)| < M$ (f'' doesn't get too large), then $g'(x)$ is bounded.
- Notice that if x_* is considered a root of f, then $g'(x_*) = \frac{f(x_*)f''(x_*)}{(f'(x_*))^2} = 0$ and $g(x_*) = x_* - \frac{f(x_*)}{f'(x_*)} = x_*$. Hence g fixes the root of f and, by continuity, "near x_*" $g'(x)$ will remain small in magnitude.

All this points to the contention that there will exist a small (closed) ϵ-neighborhood $\overline{B}_\epsilon(x_*)$ about x_*, where $|g'(x)| < 1$ for all $x \in \overline{B}_\epsilon(x_*)$. (Remember that the derivative of g is 0 at x_*, is continuous in a neighborhood of x_* and cannot grow too quickly around x_* due to the two constraints on the derivatives of f. Once we have this, then by Proposition 2.21, g will form a contraction on $\overline{B}$. Thus we have:

Proposition 2.35. *Let $f:\mathbb{R}\to\mathbb{R}$ be C^2 with a root x_*. If $\exists\delta, M>0$, such that $|f'(x)|>\delta$ and $|f''(x)|<M$ on an open interval containing x_*, then*

$$g(x) = x - \frac{f(x)}{f'(x)}.$$

is a contraction near x_* with fixed point x_*.

2.2.4. Application: Existence and Uniqueness of ODE solutions. A global version example of the contraction principle above involves the standard proof of the existence and uniqueness of solutions to the first-order IVP

$$\dot{y} = f(t,y), \quad y(t_0) = y_0 \tag{2.2.1}$$

in a neighborhood of (t_0, y_0) in the t,y-plane. The proof uses a dynamical approach, again with a first approximation and then successive iterations using a technique attributed to Charles Picard and known as Picard Iterations.

Theorem 2.36 (Picard-Lindelöf Theorem). *Suppose $f(t,y)$ is continuous in some rectangle*

$$R = \left\{(t,y)\in\mathbb{R}^2 \mid \alpha<t<\beta, \gamma<y<\delta\right\},$$

containing the initial point (t_0,y_0), and f is Lipschitz continuous in y on R. Then, in some interval $t_0-\epsilon<t<t_0+\epsilon$ contained in $\alpha<t<\beta$, there is a unique solution $y=\phi(t)$ of Equation 2.2.1.

To prove this theorem, we will need to understand a bit about how spaces of functions behave. To start, recall from linear algebra that an (real) *operator* is simply a function $f:U\to V$ whose domain and codomain are (real) vector spaces. An (real) operator is called *linear* if $\forall x,y\in U$ and $c_1,c_2\in\mathbb{R}$, we have

$$f(c_1x+c_2y) = c_1f(x)+c_2f(y).$$

Linear operators where both $\dim(U)=n$ and $\dim(V)=m$ are finite-dimensional can be represented by matrices, so that $f(x)=Ax$, A an $m\times n$ matrix. Real-valued continuous functions on $\mathbb{R}$ also form a vector space using addition of functions and scalar multiplication as the operations. One can form linear operators on spaces of functions like this one also, but the operator is not represented a as matrix. A good example is the derivative operator $\frac{d}{dx}$ which acts on the vector space of all differentiable, real-valued functions of one independent variable, and takes them to other (in this case, at least) continuous functions. Think

$$\frac{d}{dx}(x^2+\sin x) = 2x+\cos x.$$

This operator is linear due to the Sum and Constant Multiple Rules found in any standard single variable calculus text. There are numerous technical difficulties in discussing linear operators in general, but for now, simply accept this general description.

Back to the case at hand, any possible solution $y=\phi(t)$ (if it exists) to Equation 2.2.1 must be a differentiable function that satisfies

$$\phi(t) = y_0 + \int_{t_0}^{t} f(s,\phi(s))\,ds \tag{2.2.2}$$

for all t in some interval containing t_0.

EXERCISE 40. Show that this is true (Hint: Simply differentiate to recover the ODE.)

At this point, existence of a solution to the ODE is *assured* in the case that $f(t,y)$ is continuous on R, as the integral will then exist at least on some smaller interval $t_0 - \epsilon < t < t_0 + \epsilon$ contained inside $\alpha < t < \beta$. Note the following:

- One reason a solution may not exist all the way out to the edge of R? What if the edge of R is an asymptote in the t variable?
- A function does not have to be continuous to be integrable (step functions are one example of integrable functions that are not continuous. However, the integral of a step function IS continuous. And if $f(t,y)$ included a step-like function in Equation 2.2.1, solutions may still exist and be continuous.

As for uniqueness, suppose $f(t,y)$ is continuous as above, and consider the following operator T, whose domain is the space of all differentiable functions on R, which takes a function $\psi(t)$ to its image $T(\psi(t))$ (which we will denote $T\psi$ to help remove some of the parentheses) defined by

$$T\psi = y_0 + \int_{t_0}^{t} f(s,\psi(s))\,ds.$$

We can apply T to many functions $\psi(t)$ and the image will be a different function $T\psi$ (but still a function of t; see Example 2.37 below). However, looking back at Equation 2.2.2, if we apply T to an actual *solution* $\phi(t)$ to the IVP, the image $T\phi$ should be the same as ϕ. A solution will be a fixed point of the discrete dynamical system formed by T on the space of functions defined and continuous on R, since $T\phi = \phi$.

EXERCISE 41. Find all fixed points for the derivative operator $\frac{d}{dx}$ whose domain is all differentiable functions on $\mathbb{R}$.

Hence, instead of looking for solutions to the IVP, we can instead look for fixed points of the operator T. How do we do this? Fortunately, this operator T has the nice property that it is a contraction.

PROOF OF THEOREM. By assumption, $f(t,y)$ is Lipschitz continuous in y on R. Hence there is a constant $M > 0$ where

$$|f(t,y) - f(t,y_1)| \leq M|y - y_1|, \quad \forall y, y_1 \in R.$$

Choose a small number $\epsilon = \frac{C}{M}$, where $C < 1$. And define a distance within the set of continuous functions on the closed interval $I = [t_0 - \epsilon, t_0 + \epsilon]$ by

$$d(g,h) = \max_{t \in I} \left| g(t) - h(t) \right|.$$

Then we have

$$d(Tg,Th) = \max_{t\in I}\left|Tg(t) - Th(t)\right| \tag{2.2.3}$$

$$= \max_{t\in I}\left|y_0 + \int_{t_0}^{t} f(s,g(s))\,ds - y_0 - \int_{t_0}^{t} f(s,h(s))\,ds\right| \tag{2.2.4}$$

$$= \max_{t\in I}\left|\int_{t_0}^{t} f(s,g(s)) - f(s,h(s))\,ds\right| \tag{2.2.5}$$

$$\leq \max_{t\in I}\int_{t_0}^{t}\left|f(s,g(s)) - f(s,h(s))\right| ds \tag{2.2.6}$$

$$\leq \max_{t\in I}\int_{t_0}^{t} M\,|g(s) - h(s)|\; ds \tag{2.2.7}$$

$$\leq \max_{t\in I}\int_{t_0}^{t} M\cdot d(g,h)\,ds \tag{2.2.8}$$

$$\leq \max_{t\in I}\left\{M\cdot d(g,h)\cdot|t-t_0|\right\} \tag{2.2.9}$$

Exercise 42. The justifications of going from Step 2.2.5 to 2.2.6, Step 2.2.6 to 2.2.7, and Step 2.2.8 to 2.2.9 are adaptations of major concepts and/or theorems from Calculus I-II to functions of more than one independent variable. Find what theorems these are and show that these are valid justifications. Can you see now why the Lipschitz continuity of f is a necessary hypothesis to the theorem?

Exercise 43. Justify why the remaining steps are true.

Now notice in the last inequality that since $I = [t_0 - \epsilon, t_0 + \epsilon]$, we have that

$$|t - t_0| \leq \epsilon = \frac{C}{M}.$$

Hence

$$\begin{aligned} d(Tg,Th) &\leq \max_{t\in I}\left\{M\cdot d(g,h)\cdot|t-t_0|\right\} \\ &\leq M\cdot d(g,h)\cdot\frac{C}{M} = C\cdot d(g,h). \end{aligned}$$

Hence T is a C-contraction and there is a unique fixed point ϕ (which is a solution to the original IVP) on the interval I. Here

$$\phi(t) = T\phi(t) = y_0 + \int_{t_0}^{t} f(s,\phi(s))\,ds.$$

□

We can actually use this construction to construct a solution to an ODE:

Example 2.37. Solve the IVP $y' = 2t(1+y)$, $y(0) = 0$ using the above Picard iterations construction.

Here, $f(t,y) = 2t(1+y)$ is a polynomial in both t and y, so that f is obviously continuous in both variables, as well as Lipschitz continuous in y, on the whole plane $\mathbb{R}^2$. Hence unique solutions exist everywhere. To actually find a solution, start with an initial guess. An obvious one is

$$\phi_0(t) = 0.$$

Notice that this choice of $\phi_0(t)$ does not solve the ODE. But since the operator T is a contraction, iterating will lead us to a solution: Here $\phi_{n+1}(t) = T\phi_n(t)$. We get

$$\begin{aligned}
\phi_1(t) &= T\phi_0(t) = y_0 + \int_0^t 2s(1+\phi_n(s))\,ds = \int_0^t 2s(1+0)\,ds = t^2\\
\phi_2(t) &= T\phi_1(t) = y_0 + \int_0^t 2s(1+\phi_1(s))\,ds = \int_0^t 2s(1+s^2)\,ds = t^2 + \frac{1}{2}t^4,\\
\phi_3(t) &= T\phi_2(t) = y_0 + \int_0^t 2s(1+\phi_2(s))\,ds\\
&= \int_0^t 2s\left(1+s^2+\frac{1}{2}s^4\right)ds = t^2+\frac{1}{2}t^4+\frac{1}{6}t^6,\\
\phi_4(t) &= T\phi_3(t) = y_0 + \int_0^t 2s(1+\phi_3(s))\,ds\\
&= \int_0^t 2s\left(1+s^2+\frac{1}{2}s^4+\frac{1}{6}s^6\right)ds = t^2+\frac{1}{2}t^4+\frac{1}{6}t^6+\frac{1}{24}t^8.
\end{aligned}$$

Exercise 44. Find the pattern and write out a finite series expression for $\phi_n(t)$. Hint: Use induction.

Exercise 45. Find a closed form expression for $\lim\limits_{n\to\infty}\phi_n(t)$ and show that it is a solution of the IVP.

Exercise 46. Now rewrite the original ODE in a standard form as a first-order linear equation, and solve.

To understand why Lipschitz continuity in the dependent variable y is a necessary condition for uniqueness of solutions, consider the following example:

Example 2.38. Let $\dot{y} = y^{\frac{2}{3}}$, $y(0) = 0$ a first-order, autonomous IVP. It should be clear that $y(t) \equiv 0$ is a solution. But so is

$$y_c(t) = \begin{cases} \frac{1}{27}(t+c)^3 & t < -c \\ 0 & t \geq -c \end{cases} \qquad \forall c \geq 0.$$

There are lots of solutions passing through the origin in ty-trajectory space. Solutions exist but are definitely not unique here. What has failed in establishing uniqueness of solutions to this IVP in the Picard-Lindelöf Theorem? Here $f(y) = y^{\frac{2}{3}}$ is certainly continuous at $y = 0$, but it is NOT Lipshitz continuous there. In fact, $f'(y) = \frac{2}{3}y^{-\frac{1}{3}}$ is not differentiable at $y = 0$, and $\lim\limits_{y\to 0} f'(y) = \infty$.

Exercise 47. Verify that the family of curves $y_c(t)$ above solve the IVP, and derive this family by solving the IVP as a separable ODE.

Exercise 48. Verify that $f(y) = y^{\frac{2}{3}}$ is not Lipschitz continuous at $y = 0$.

2.2.5. Application: Heron of Alexandria. Start with the beautiful idea that the *arithmetic mean* of 2 positive real numbers $a, b > 0$, namely, $\frac{a+b}{2}$ is always greater than the *geometric mean*, defined as the square root of their product:

$$\frac{1}{2}(a+b) \geq \sqrt{ab}'$$

with equality only when $a = b$. Sometimes this is called the AMGM Inequality.

Geometrically, one can visualize the AMGM Inequality via the following (See Figure 6): Let B_a and B_b be 2 disks of respective diameters $a, b > 0$ resting on the real line and touching (assume $a \geq b$ just for the sake of argument.) Then the line connecting their center has length precisely $\frac{a+b}{2}$, the arithmetic mean. Using this line as the hypotenuse of a right triangle by dropping a vertical from the center of the larger ball (in Figure 6, it is B_a), the vertical side has length $\frac{a-b}{2}$ and the horizontal side has length $\sqrt{ab}$. Do you see the result now?

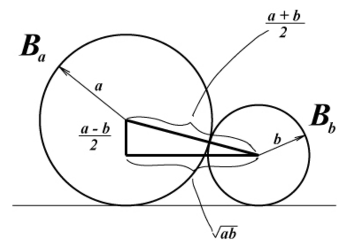


FIGURE 6. Visualization of the AMGM Inequality.

EXERCISE 49. Calculate the lengths of the two sides adjacent to the right angle here.

EXERCISE 50. Show algebraically that, for $a, b > 0$, $\frac{1}{2}(a+b) \geq \sqrt{ab}$, with equality iff $a = b$.

A Greek mathematician and engineer, Heron of Alexandria is credited with being the first to write down an iterated method of approximating the numeric value of the principle square root of a positive number in the first century of the Common Era in his work "Metrica". The method itself evidently goes back to the Babylonians. The method of Heron is simple enough: Let $N > 0$. We consider a way to approximate $\sqrt{N}$.

(1) Choose a whole number $a > 0$ near $\sqrt{N}$ as an approximation. A good choice would be the root of the closest perfect square. Then a^2 approximates N. But then $\frac{N}{a}$ also approximates $\sqrt{N}$, and. in fact, lies "on the other side" of $\sqrt{N}$, meaning that $\sqrt{N}$ lies between the two numbers a and $\frac{N}{a}$, no matter which is larger. See Figure 1.
(2) The algebraic mean $\frac{1}{2}\left(a + \frac{N}{a}\right)$ is a new approximation to $\sqrt{N}$ which is closer to $\sqrt{N}$ than at least the farther of a and $\frac{N}{a}$.
(3) If a better approximation to $\sqrt{N}$ is desired, repeat with $\frac{1}{2}\left(a + \frac{N}{a}\right)$ as the new a.

In our modern terminology, for $N > 0$, let $f : \mathbb{R}_+ \to \mathbb{R}_+$ be the function $f(x) = \frac{1}{2}\left(x + \frac{N}{x}\right)$. For $x_0 > 0$ any guess, let $x_{i+1} = f(x_i)$. Then f has a unique fixed point at $x = \sqrt{N}$.

Now there are many ways to show that this method converges $\forall x_0 > 0$ and $\forall N > 0$. For example, one could calculate the length of the intervals $\ell_n = \left\|\left(x_n, \frac{N}{x_n}\right)\right\|$ and show that $\lim_{n\to\infty} \ell_n = 0$, while $\sqrt{N} \in \left(x_n, \frac{N}{x_n}\right)$, $\forall n \in \mathbb{N}$.

EXERCISE 51. Show $\ell_{n+1} \leq \frac{1}{2}\ell_n$, $\forall n \in \mathbb{N}$. Thus the interval lengths decay exponentially.

Also, an interesting fact arises from the iterations of f: The new estimate at each stage is ALWAYS an overestimate. This is due directly to the AMGM Inequality. Thus the sequence $\{x_i\}_{i=1}^{\infty}$ is a decreasing sequence, bounded below by $\sqrt{N}$. Thus, it must converge.

EXERCISE 52. Show that this is enough to establish that $\lim_{n\to\infty} x_n = \sqrt{N}$.

EXERCISE 53. Show that Heron's Method is equivalent to the Newton-Raphson Method when locating the positive root of $g(x) = x^2 - N$.

EXERCISE 54. Show that the AMGM Inequality provides a constructive proof of the following optimization problems from calculus:

a. Among all positive numbers with a given product, the sum is minimal when the numbers coincide.
b. Among all positive numbers with a given sum, the product is maximal when the numbers coincide.
c. Among all quadrilaterals with the same perimeter, the square has the largest area.

EXERCISE 55. Show that, for $N > 0$, the family of functions

$$f_N : \mathbb{R}_+ \to \mathbb{R}_+, \quad f_N(x) = \frac{1}{2}\left(x + \frac{N}{x}\right)$$

are not contractions. Then find the largest interval $I_N \subset \mathbb{R}$, containing $\sqrt{N}$, where $f_N\big|_{I_N}$ is a $\frac{1}{2}$-contraction.

EXERCISE 56. Approximate $\sqrt{110}$using Herons method, to an accuracy of .001. Try this using a starting value of 10 and then again for a starting value of 1, noting the difference in convergence properties.

2.3. Interval Maps

The Contraction Principle above is a facet of some dynamical systems which display what is called "simple dynamics": With very little information about the system (map or ODE), one can say just about everything there is to say about the system. Another way to put this is to say that, in a contraction, all orbits do exactly the same thing. Which is, they all converge to the same fixed point (equilibrium solution in the case of a continuous dynamical system.) We can build on this idea by now beginning a study of a relatively simple family of discrete dynamical systems that display slightly more complicated behavior.

Let $f : I \to I$ be a continuous map, where $I = [0,1]$ (we will say f is a C^0-map on I, or $f \in C^0(I,I)$). The graph of f sits inside the unit square $[0,1]^2 = [0,1] \times [0,1] \subset \mathbb{R}^2$.

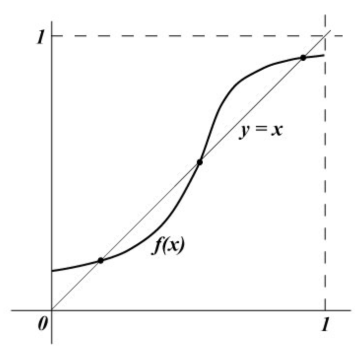


FIGURE 7. A typical C^0-map $f : [0,1] \to [0,1]$.

2.3.1. Cobwebbing.

This graph intersects the line $y = x$ at precisely the points where $y = f(x) = x$, or the fixed points of the dynamical system given by f on I. Recall that the dynamical system is formed by iterating f on I, and all of the forward iterates of x_0 under f comprise the orbit of x_0, $\mathcal{O}_{x_0}$, where

$$\mathcal{O}_{x_0} = \left\{x_0,\ x_1 = f(x_0),\ x_2 = f(x_1) = f^2(x_0),\ \dots\right\}$$

One can track $\mathcal{O}_{x_0}$ in I (and in $[0,1]^2$) visually via the notion of *cobwebbing*. We use the example of $f(x) = x^2$ on I to illustrate:

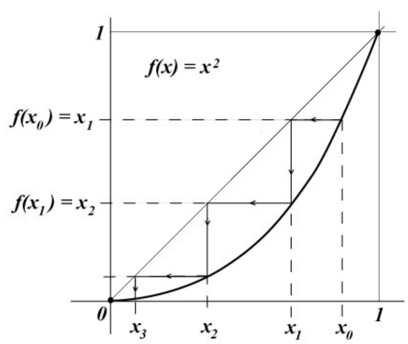


FIGURE 8. A cobweb of $f(x) = x^2$ on $[0,1]$.

Choose a starting value $x_0 \in I$. Under f, the next term in the orbit is $x_1 = f(x_0)$. Vertically, it is the height of the graph of f over x_0. Making it the new input value to f means finding its corresponding place on the horizontal axis. This is easy to see visually. The vertical line $x = x_0$ crosses the graph of f at the height $x_1 = f(x_0)$. The horizontal line $y = x_1 = f(x_0)$ crosses the diagonal $y = x$ precisely at the point (x_1, x_1). Taking the output x_1 and making it the new input to f (iterating the function) means finding where the vertical line through this point will again intersect the graph of f (at one point). That point will be at the height $x_2 = f(x_1)$ and constitutes the second value of the sequence $\mathcal{O}_{x_0}$. By only zigzagging this way – moving vertically from the $y = x$ line to the graph of f and then horizontally back to the $y = x$ line – we can document the orbit of the point x_0 without actually calculating the explicit function values. While this technique is only as accurate as the drawing of the graph, it is an excellent way to "see" an orbit without calculating it.

Specific to this example $f(x) = x^2$ on I, we have two fixed points: $x = 0$ and $x = 1$ (the graph crosses the $y = x$ line at these points). And if x_0 is chosen to be strictly less than 1, then we can easily conclude via the cobweb that $\mathcal{O}_{x_0} \longrightarrow 0$. Visually, it makes sense. Analytically, it is also intuitive; squaring a number between 0 and 1 always makes it smaller but keeping it positive. However, can you *prove* that every orbit goes to 0 except for the orbit $\mathcal{O}_1$? We will do something like this shortly.

EXERCISE 57. Cobweb the following functions and describe the dynamics as completely as you are able:

a. $f(x) = x^3$ on $[-1,1]$	b. $g(x) = \ln(x+1)$ on $[0,\infty)$	c. $h(x) = \frac{x^2-3}{x-2}$ on $\mathbb{R}$
d. $k(x) = -\frac{5}{3}(x^2 - x)$ on $[0,1]$	e. $\ell(x) = -\frac{10}{3}(x^2 - x)$ on $[0,1]$	.

2.3.2. Fixed point stability. What happens to orbits near a fixed point of a discrete dynamical system (equilibrium solutions to a continuous one) are of profound importance in an analysis of a mathematical model. Often, the fixed points are the only easily discoverable orbits of a hard-to-solve system. They play the role of a "steady-state" of the system. And like for functions in general, knowledge of a function's derivatives at a point say important things about how a function behaves near a point. To begin this analysis, we will need some definitions which will allow us to talk about the nature of fixed points in terms of what happens around them. This language is a lot like the way we classified equilibrium solutions in the ODEs class. For the moment, think of X as simply an interval in $\mathbb{R}$ with the metric just the absolute value of the difference between two points. These definitions are for all metric spaces X, though. The only caveat here is that in higher dimensions, there are some more subtle things that can happen near a fixed point, as we will see.

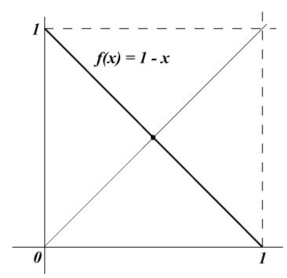
DEFINITION 2.39. Let x_0 be a fixed point of the C^0-map $f : X \to X$. Then x_0 is said to be

- *Poisson stable* if $\forall \epsilon > 0$, $\exists \delta > 0$ such that $\forall x \in X$, if $d(x, x_0) < \delta$, then $\forall n \in \mathbb{N}$ $d(f^n(x), x_0) < \epsilon$.
- *asymptotically stable*, an *attractor*, or a *sink* if $\exists \epsilon > 0$ such that $\forall x \in X$, if $d(x, x_0) < \epsilon$, then $\mathcal{O}_x \longrightarrow x_0$.
- a *repeller* or a *source* if $\exists \epsilon > 0$ such that $\forall x \in X$, if $0 < d(x, x_0) < \epsilon$, then $\exists N \in \mathbb{N}$ such that $\forall n > N$, $d(f^n(x), x_0) > \epsilon$.

Do we need semi-stable here for one-dimensional fixed points?

The basic idea behind this classification is the following: Asymptotically stable means that there is a neighborhood of the fixed point where f restricted to that neighborhood is a contraction with x_0 as the sole fixed point. Poisson stable means that given any small neighborhood of the fixed point, I can choose a smaller neighborhood where if I start in the smaller neighborhood, the forward orbit never leaves the larger neighborhood. Asymptotically stable points are always Poisson stable, but not necessarily vice versa. And a fixed point is a repeller if in a small neighborhood of the fixed point, all points that are not the fixed point itself have forward orbit that leave the neighborhood and never return.

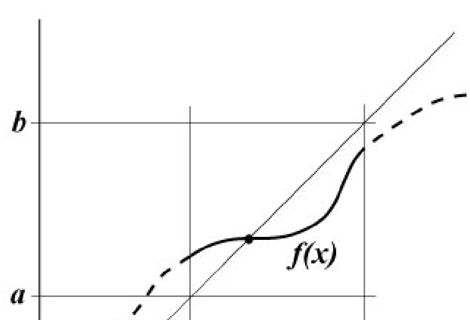
EXAMPLE 2.40. $f(x) = 1 - x$ on $[0, 1]$ has a unique fixed point $x_* = \frac{1}{2}$. This fixed point is Poisson stable, but not asymptotically stable. To see this, simply let $\delta = \epsilon$ and write out the definition.



REMARK 2.41. Recall from your Differential Equations class (or wait until Chapter 4), in the classification of 2×2, first-order, homogeneous, linear ODE systems, one can classify the type of the equilibrium solution at the origin via a knowledge of the eigenvalues of the coefficient matrix. In this classification, the sink (negative eigenvalues) was the asymptotically stable equilibrium, the source was the repeller, and the center (recall where the two eigenvalues were purely imaginary complex conjugates) was the Poisson stable equilibrium,

EXAMPLE 2.42. Back to Figure 8 the graph of $f(x) = x^2$ on $[0, 1]$. This dynamical system has two fixed points. One can see visually via the cobweb that $x = 0$ is asymptotically stable. Also, $x = 1$ is unstable and a repeller.

EXERCISE 58. Show analytically that $x = 0$ is an attractor while $x = 1$ is a repeller. That is, show that the fixed points satisfy the respective definitions.



In this class, we will spend a fair amount of time on the maps $f : [0, 1] \to [0, 1]$. There are basically two reasons for this: 1) They have applications beyond simple interval maps, and 2) maps of the unit interval are really all one need study when studying intervals. To see the second point, let $f : \mathbb{R} \to \mathbb{R}$, but suppose that there exists a closed interval $[a,b]$, $b > a$ (a single point is considered a closed interval, so the condition that $b > a$ means something with an interior), where

$$f\big|_{[a,b]} : [a,b] \to [a,b].$$

- Dynamically speaking, what happens to $f(x)$ under iteration on $[a,b]$ is no different from what happens to $g(y)$ on $[0,1]$ under the linear transformation of coordinates $y = \frac{x-a}{b-a}$, where one must transform both the input and output variables appropriately.)

EXERCISE 59. Find the map $g : [0,1] \to [0,1]$ on the unit interval in the plane that is equivalent dynamically to the map $f(x) = x^3$ on $[-1,1]$.

- Let f be continuous on an unbounded interval I, either on one side or both. Then in the case that f has bounded image (possibly f has horizontal asymptotes, but this is not necessary), then one can simply study the new dynamical system formed by f, where the domain is the interval $f(I)$, the set of first image points $f(x)$, for $x \in I$.
- Combining both of these items into one may also be useful: For example, a coordinate transformation can map an unbounded interval to a bounded one (e.g., $g(x) = \frac{1}{x}$, mapping $[1,\infty)$ onto $(0,1]$, or $h(x) = \frac{2}{\pi}\tan^{-1}(x)$, taking $\mathbb{R}$ to the interval $(-1,1)$). Under proper care with regard to the orbits, one can transform the dynamical system to one on the bounded interval. We shall elaborate on this later.

It is usually an $f(x)$ from above that appears in applications, and mathematically we usually only study maps like $g(y)$. For an example, let's go back to the Newton Raphson Method for root location. First, a quick definition:

DEFINITION 2.43. A fixed point x_0 for $f : I \to I$ where $f \in C^1$ is called *superattracting* if $f'(x_0) = 0$. (Why? See the picture.)

PROPOSITION 2.44. *Let $f : \mathbb{R} \to \mathbb{R}$ be C^2 with a root r. If $\exists \delta > 0$, $M > 0$ such that $|f'(x)| > \delta$ and $|f''(x)| < M$ on a neighborhood of r, then r is a superattracting fixed point of*

$$F(x) = x - \frac{f(x)}{f'(x)}.$$

PROOF. As we have already calculated, $F'(r) = \dfrac{f(r)f''(r)}{\left[f'(r)\right]^2} = 0$. $\square$

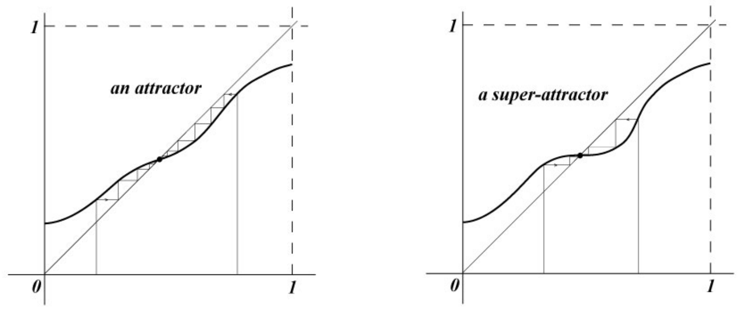


FIGURE 9. Follow the cobwebs to see how quickly nearby orbits converge to a superattractor vis a vis an attractor.

We can go further, and I will state this part without proof: Since f is C^2, then F is C^1. Calculating $F'(x)$ and knowing that it is both continuous and 0 at $x = r$, there will be a small, closed interval $[a, b]$, $b > a$ with r in the interior, where $|F'(x)| < 1$. One can show that, restricted to this interval,

$$F\big|_{[a,b]} : [a, b] \to [a, b]$$

is a λ-contraction, with a superattracting fixed point at r. In this case, all orbits of F converge exponentially to r by λ^2, even thought F is simply a λ-contraction.

As a final note before moving on to interval maps, a logical question to ask here is: Just how large is the interval $[a, b]$ on which F is a contraction? This is important for the Newton-Raphson Method. Convergence is guaranteed when one chooses an approximation to a root that is "close enough". But what does close enough actually mean? Of course, this depends severely on the properties of the original function f. But to gauge the edges of an interval like $[a, b]$, we offer:

DEFINITION 2.45. Let $f : X \to X$ be a discrete dynamical system with an attracting fixed point $x \in X$. Then the set

$$\mathcal{B}_x = \left\{ y \in x \;\middle|\; \mathcal{O}_y \longrightarrow x \right\}$$

is called the *basin of attraction* of x for f.

Essentially, $\mathcal{B}_x$ is the collection of all starting values of orbits that converge to x. Often, this set is quite easy to describe; For $f : X \to X$ a contraction with fixed point x_0, $\mathcal{B}_{x_0} = X$. Everything converges to the unique fixed point. However, as we will see, there are instances where this set is very complicated. As an indication of possible issues, consider the root search for a function $f : \mathbb{R} \to \mathbb{R}$ with at least two roots, both of which satisfy the conditions for convergence of the method "near" each the root. The individual basins of attraction for the two attractors should be open intervals, given the conditions for f. But must they come in one piece? The two basins cannot intersect, however (why not?) But can they butt up against each other? And if a starting point is in neither basin, there does it's orbit go. IN time, we will explore these issues and more. But for now, let's move on to maps that are slightly more complicated than contractions.

2.3.3. Monotonic maps. Interval maps are quite general, and display tons of diverse and interesting behavior. To begin exploring this behavior, we will need to restrict our choice of maps to those exhibiting somewhat mild behavior. The first type designation is as follows:

DEFINITION 2.46. A map $f : [a,b] \to [a,b]$ be C^0. We say f is

- *increasing* if for $x > y$, we have $f(x) > f(y)$,
- *nondecreasing* if for $x > y$, we have $f(x) \geq f(y)$,
- *non-increasing* if for $x > y$, we have $f(x) \leq f(y)$,
- *decreasing* if for $x > y$, we have $f(x) < f(y)$.

It is easy and intuitive to see how these definitions work. You should draw some examples to differentiate these types. You should also work to understand how these different types affect the dynamics a lot. For example, increasing and non-decreasing maps can have many fixed points (actually, the map $f(x) = x$ has ALL points fixed!). While all non-increasing maps (hence all decreasing maps also) can have only one fixed point each. Further, increasing maps cannot have points of period two (why not?), while there does exist a decreasing map with ALL points of period two (can you find it?). We will explore these in time. For now, we will start with a fact shared by ALL interval maps:

PROPOSITION 2.47. *For the C^0 map $f : [a,b] \to [a,b]$, where $a, b \in \mathbb{R}$, f must have a fixed point.*

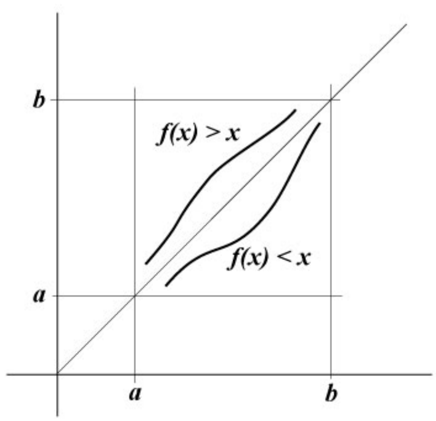


FIGURE 10. An interval map on $[a,b]$

Visually, this should make sense. Imagine trying to draw the graph of a continuous function in the unit square in a way that it does NOT intersect the diagonal, "fixed point" line. When you get tired of trying, read on.

PROOF. Suppose for now that f has no fixed point on (a,b) (this seems plausible, since our example above $f(x) = x^2$ on $[0,1]$ satisfies this criterion). Then, it must be the case that for all $x \in (a,b)$, either **(1)** $f(x) > x$, or **(2)** $f(x) < x$. This means that the entire graph of f lies above the diagonal, or below it, respectively (See Figure 10).

If we are in situation **(1)**, then for any choice of $x \in (a,b)$, $\mathcal{O}_x$ will be an increasing sequence in $[a,b]$. It is also bounded above by b since the entire sequence lives in $[a,b]$. Recall from Calculus the Monotone Sequence Theorem: Every bounded, monotone infinite sequence converges. By definition, monotone means either (strictly) increasing or decreasing. For case **(2)**, the orbit is strictly decreasing, and will be bounded below since the entire sequence live in the closed interval $[a,b]$. Thus we can say in each instance that $\mathcal{O}_x \longrightarrow x_0$, for some $x_0 \in [a,b]$ (we can also say $\lim_{n\to\infty} f^n(x) = x_0$).

What can this point x_0 look like? Well, for starters, it must be a fixed point! To see this,

$$f(x_0) = f\left(\lim_{n\to\infty} f^n(x)\right) = \lim_{n\to\infty} f^{n+1}(x) = \lim_{n\to\infty} f^n(x) = x_0.$$

Since there are no fixed points in (a,b) by assumption, it must be the case that either $x_0 = b$ (case **(1)**), or $x_0 = a$ (case **(2)**). $\square$

EXERCISE 60. Reprove Proposition 2.47 using the Intermediate Value Theorem on the function $g(x) = f(x) - x$.

This last proof immediately tells us the following:

PROPOSITION 2.48. *Let the C^0-map $f : [a,b] \to [a,b]$ be non-decreasing, and suppose there are no fixed points on (a,b). Then either*

- *exactly one end point is fixed and $\forall x \in [a,b]$, $\mathcal{O}_x$ converges to the fixed end point, or*
- *Both end points are fixed, one is an attractor and the other is a repeller.*

And if in the second case above, f is also increasing, then $\forall x \in (a,b)$, $\mathcal{O}_x$ is forward asymptotic to one end point, and backward asymptotic to the other.

EXAMPLE 2.49. Let $g(x) = \sqrt{x}$ on the closed interval $[1,b]$ for any $b > 1$. This is an example of the first situation in the proposition. Here, actually, we can let $g(x)$ have the closed interval $[1,\infty)$ as its domain, and we gain the same result. One does have to be careful here, though, as the related function $k(x) = \frac{\sqrt{x}}{2}$ on $[1,\infty)$ has no fixed points at all.

EXAMPLE 2.50. For the second case, think $f(x) = x^2$ on $[0,1]$, illustrated in Figure 8.

EXERCISE 61. For $h(x) = x^3$ on $I = [-1,1]$, both endpoints are fixed and repellers. But h does not satisfy the hypotheses of the proposition. (why not?) Hence the proposition does not apply to this function. However, h may be seen as two separate dynamical systems existing side-by-side. Justify this and show that the proposition holds for each of these "sub"-dynamical systems.

Place some more exercises here. Two important notes here:

- First, we can immediately see that the basin of attraction of the fixed point in the first case of Proposition2.48 is the entire interval $[a,b]$, while in the second case it is everything except for the other fixed point (the repeller.) One property of a basin of attraction is that it is an open set (every point in the set contains a small open interval which is also completely in the set.) This does not seem to be the case here. But it, in fact, is. The interval $[a,b]$, as a subset of $\mathbb{R}$ is closed (and bounded). But as a domain (not sitting inside $\mathbb{R}$ for the purpose of serving as the plug-in points for f), it is both open and closed as a topological space. In a sense, there is no outside to $[a,b]$ as a domain for f. It has open subsets, and these all look like one of $[a,c)$, (c,d), or $(d,b]$, for $a < c < d < b \in [a,b]$ and their various unions and finite intersections. But $[a,b]$ can be written as the union of two overlapping open sets. Hence it is open, as a subset of $[a,b]$. And so is $(a,b]$, for example. Hence the basins of attraction in both of these cases are in fact open subsets of $[a,b]$.

- Recall from calculus (or even pre-calculus) that an increasing function $f : I \to \mathbb{R}$ is called *injective*, or *one-to-one*, since for any two points $x, y \in I$, if $f(x) = f(y)$, then $x = y$. One would say its graph satisfies the Horizontal Line Test (remember this?) And one-to-one functions have inverses $f^{-1}(x)$, or we say that f is invertible. However, for $I = [a, b]$, a function $f : I \to I$, specifically from I back to itself, is only invertible if it satisfies certain precise criteria: It must be one-to-one and *onto*, or *surjective*, and either $f(a) = a$ and $f(b) = b$, or $f(a) = b$ and $f(b) = a$. Using the above examples, $f(x) = x^2$ on $I = [0, 1]$ and $h(x) = x^3$ on $I = [-1, 1]$ are invertible, but $g(x)$ is NOT invertible on $[1, b]$ for any $b > 1$. And, the function given in Figure 7 is also not invertible, even as it is strictly increasing on $I = [0, 1]$.

Exercise 62. Show that an invertible map $f : I \to I$ on $I = [a, b]$, for $b > a$, must satisfy all of the following (Hint: All of these can be shown by assuming the property does not hold and then finding a contradiction.)

a. f is injective (one-to-one),
b. f is surjective (onto; the range must be all of I),
c. f must satisfy either $f(a) = a$ and $f(b) = b$, or $f(a) = b$, and $f(b) = a$.

Hence we can revise the last statement as: *And if, in addition, f is invertible, then $\forall x \in (a, b)$, $\mathcal{O}_x$ is forward asymptotic to one end point, and backward asymptotic to the other.* Recall that if $\mathcal{O}_x$ is forward asymptotic to a point x_0, we write $\mathcal{O}_x^+ \to x_0$. If backward asymptotic, we write $\mathcal{O}_x^- \to x_0$. This notion of an orbit converging in both forward time and backward time is of enough importance in dynamics that we classify some differing ways in which this can happen.

2.3.4. Homoclinic/heteroclinic points. Using the notation for the forward and backward orbits introduced above, we can close in on a special property of nondecreasing interval maps. First, we can now say definitively that

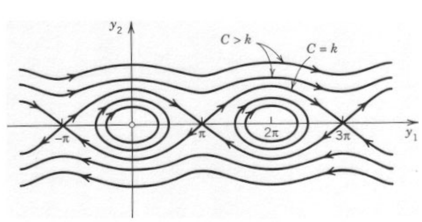
- In the first case of Proposition 2.48, $\forall x \in [a, b]$, either $\mathcal{O}_x^+ \longrightarrow a$ or $\mathcal{O}_x^+ \longrightarrow b$. Here, f will certainly look like a contraction. Must it be? Keep in mind that the definition of a contraction is very precise, and maps that behave like contractions may not actually be contractions. Think $f(x) = x^2$ on the closed interval $[0, .6]$. What is $\text{Lip}(f)$ here?
- For f in the second case of Proposition 2.48, and with f invertible (increasing), then $\forall x \in (a, b)$, either $\mathcal{O}_x^+ \longrightarrow a$ and $\mathcal{O}_x^- \longrightarrow b$, or $\mathcal{O}_x^+ \longrightarrow b$ and $\mathcal{O}_x^- \longrightarrow a$.

Let's elaborate on this idea of forward/backward orbits. Suppose now that $f : X \to X$ is a C^0-map on some subset of $\mathbb{R}^n$, and suppose $\exists x \in X$, where

$$\mathcal{O}_x^- \longrightarrow a \text{ and } \mathcal{O}_x^+ \longrightarrow b.$$

Definition 2.51. x is said to be *heteroclinic* to a and b if $a \neq b$, and *homoclinic* to a if $a = b$.

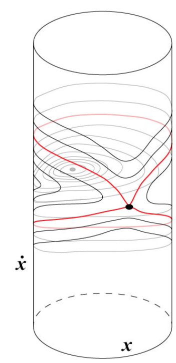
You have certainly seen heteroclinic and homoclinic orbits before. Think of the phase portrait of the undamped pendulum. It is the famous



picture at left. Here the separatrices are heteroclinic orbits from the unstable equilibrium solution at $(2n\pi, 0)$, $n \in \mathbb{Z}$, and $(2(n \pm 1)\pi, 0)$. Although, in reality, there is a much more accurate picture of the phase space of the undamped pendulum. The vertical variable (representing the instantaneous velocity of the pendulum ball, actually) takes values in $\mathbb{R}$, while the horizontal variable (representing the angular position of the pendulum ball with respect to downward vertical position) is in reality 2π-periodic. Truly, it takes values in the circle S^1:

$$S^1 = \text{ unit circle in } \mathbb{R}^2 = \left\{ e^{2\pi i\theta} \in \mathbb{C} \,\middle|\, \theta \in [0,1) \right\}. \tag{2.3.1}$$

Thus, the phase space is really a cylinder, and in actuality only has two equilibrium solutions; one at $(0,0)$, and the other at $(\pi, 0)$. In this view, which we will elaborate on later, there are only two separatrices (both in red at right), and both are homoclinic to the unstable equilibrium at $(\pi, 0)$. Also, it becomes clear once the picture is understood that ALL orbits of the undamped pendulum, except for the separatrices, are periodic. However, the period of these orbits is certainly not all the same. And there is NO bound to how long a period may actually be. See if you can fully grasp this.



Exercise 63. Show that there cannot exist homoclinic points for f a nondecreasing map on a closed, bounded interval.

Exercise 64. Construct an example (with an explicit expression) of a continuous C^0-map of S^1 that contains a homoclinic point. (Hint: In class we already have an example of an interval map that, when modified, will satisfy the C^0 construction.

Remark 2.52. Any continuous map on the unit interval with both endpoints fixed can be viewed as a map on the circle by thinking of the interval and identifying 0 and 1 (you can use the map $x \mapsto e^{2\pi i x}$ explicitly to see this). But can you construct one that does not fix the endpoints? Can you construct one that is also differentiable on all of the circle? And how does one graph (visualize) this function?

It turns out, forcing a map of an interval to be nondecreasing and forcing the interval to be closed really restricts the types of dynamics that can happen. We have:

Proposition 2.53. *Let $f : [a,b] \to [a,b]$ be C^0 and nondecreasing. Then $\forall x \in [a,b]$, either x is fixed, or asymptotic to a fixed point. And if f is increasing (and thus invertible), then $\forall x \in [a,b]$, either x is fixed or heteroclinic to adjacent fixed points.*

Clearly, the dynamics, although more complicated than for contraction maps, are nonetheless rather simple for nondecreasing interval maps (and even more so for invertible interval maps). Thus goes the second stop in our exploration of dynamical systems from simple to complex.

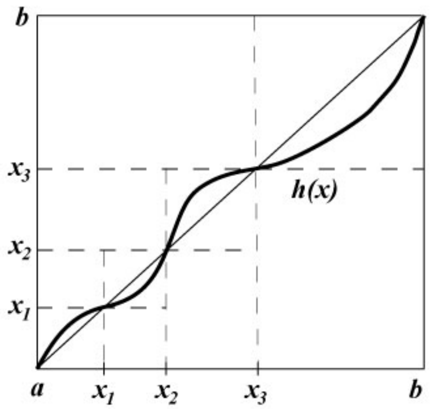


FIGURE 11. Increasing functions trap orbits between fixed points.

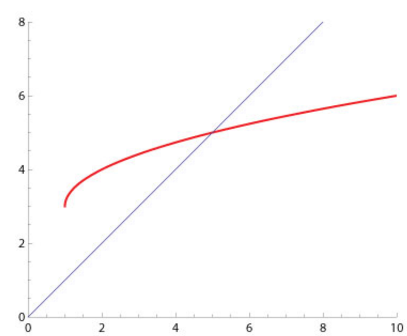
Before moving on, here are a few other things to think about: First, in the case of non-decreasing maps, the orbits of points in between fixed points are trapped in the interval bounded by the fixed points (See Figure **??**). This has enormous implications, and severely restricts what orbits can do (we say it restricts the "orbit structure" of the map). It makes Proposition 2.48 above much more general and consequential, since any nondecreasing interval map is now simply a collection of disjoint interval maps, each of one of the types in the proposition. And second, it helps to establish not only the types of fixed points one can have for an interval map, but the way fixed points relate to each other within an interval map.

EXERCISE 65. Let f be a nondecreasing map on a closed interval. Show if $x_0 \neq y_0$ are two fixed points of f, then $\forall x \in (x_0, y_0)$, $\mathcal{O}_x^+ \subset (x_0, y_0)$.

EXERCISE 66. Let f be a C^0 nondecreasing map on a closed interval. Show that for every $n \in \mathbb{N}$, $Per_n(f) = Fix(f)$. That is, there are no non nontrivial periodic points of nondecreasing interval maps.

EXERCISE 67. Let f be a C^0 nondecreasing map on $[a, b]$, and $x_0 < y_0$ be two adjacent fixed points (i.e., there are no fixed points in (x_0, y_0)). Show that x_0 and y_0 cannot both be attractors or repellers.

EXAMPLE 2.54. For $f(x) = \sqrt{x-1} + 3$ on $I = [1, \infty)$, determine the set $Per(f)$.



We can address this question in a number of ways. First, notice that f is a strictly increasing function since $f'(x) = \frac{1}{2\sqrt{x-1}} > 0$ on the interior of I (it is not defined at $x = 1$.) Hence, every orbit is monotonic (not necessarily strictly, though. Why is this true?) Hence, as a generalization of the conditions of Proposition 2.53, every point is either fixed or has an orbit asymptotic to a fixed point, OR is unbounded (the consequence of an unbounded domain). At this point, you can conclude that there are no non-trivial periodic points, like in the above exercise. Now since the derivative is a strictly decreasing function, if there is a fixed point at all, it will be unique (think about this also). There is a fixed point x_0 of this map (the technique of Exercise 60 will work here.)

And, since $f(x) > x$ on $[1, x_0)$, and $f(x) < x$ on (x_0, ∞), the fixed point x_0 is an attractor. Hence $Per(f) = \{x_0\} = \left\{\frac{1}{2} + \frac{\sqrt{33}}{2}\right\}$, solving $f(x) = x$ algebraically.

Perhaps there is an easier way: $f(x)$ is certainly NOT a contraction, as for all pairs of points $x \neq y \in \left[1, \frac{5}{4}\right]$, the distance between images is actually greater than the original distances (check this!). In fact, on $[1, \infty)$, f is not even Lipschitz, although you should show this fact.

EXERCISE 68. Show that $f(x) = \sqrt{x-1} + 3$ on $I = [1, \infty)$ is not a contraction.

However, consider that the image of I, $f(I) = [3, \infty)$, and restricted to $f(I)$, f is actually a λ-contraction, with $\lambda = \frac{1}{2\sqrt{3}}$. Hence, one could simply start iterating after the first iterate, knowing that the long-term behavior of orbits, fixed and periodic points, convergent orbits and stability, will all be the same. Thus, the map $f : f(I) \to f(I)$ is a contraction, and hence will have a unique fixed point and no other periodic points. Thus the fixed point x_0 found above is precisely all of $Per(f)$.

REMARK 2.55. It would be tempting to call the map f above *eventually contracting*, since it is a contraction on a forward iterate of the domain. However, this is not the case here. As we will see in soon enough, there is a technical condition that makes a map an eventual contraction, and there is a pathology at $x = 1$ here (pay attention to the derivative as one approaches 1 from numbers larger than 1). Suffice it to say that the language is not completely settled here.

Now take an increasing interval map f and vary it slightly. Usually, the dynamical behavior of the "perturbed" map stays the same (the number and type of fixed points does not change, even though their position may vary a bit). This may not be the case for a non-decreasing map: A slightly perturbed increasing map will remain increasing, while a slightly perturbed map with a flat interval may not remain nondecreasing. Think about that.

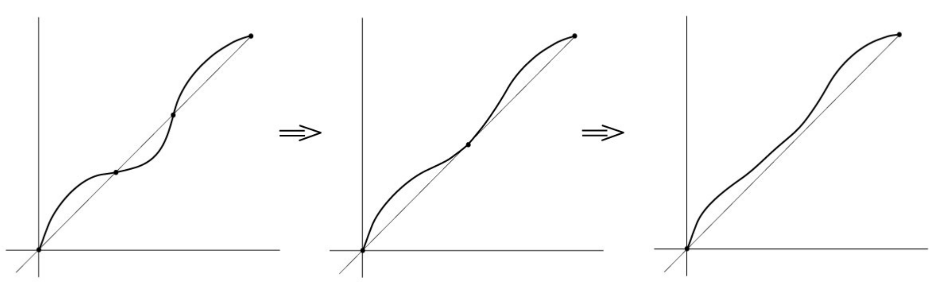
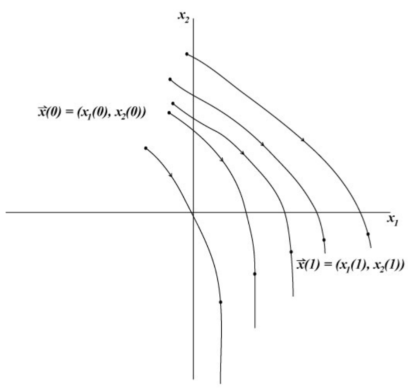


FIGURE 12. A bifurcation in an interval map.

Sometimes, a small change in an increasing map may lead to a big change is the number and type of fixed points (i.e., a big dynamical change!). Consider the three graphs above. Do you recognize this behavior? Have you ever seen a bifurcation in a mechanical system with a parameter?

Maybe place here a discussion of the notion of structural stability for maps, where increasing maps are structurally stable and nondecreasing maps are not in general.

2.4. First Return Maps

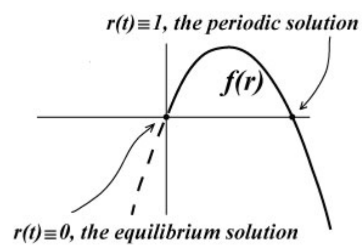


Recall that the time-1 map of an ordinary differential equation defines a discrete dynamical system on the phase space. Indeed, for $\mathbf{x} \in \mathbb{R}^n$, the system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ defines the map $\phi_1 : \mathbb{R}^n \to \mathbb{R}^n$, where $\phi_1 : \mathbf{x}(0) \mapsto \mathbf{x}(1)$ is a transformation of $\mathbb{R}^n$. Really, $t = 1$ is only one such example, and any t will work, so long as the system solutions are defined (and unique, for the most part).

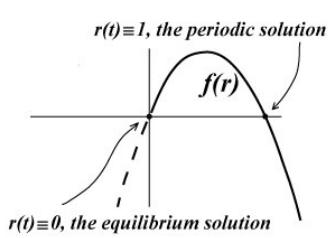
There is another kind of discrete dynamical system that comes from a continuous one: the First Return Map. One can view the first return map as a local version (only defined near interesting orbits) of the more globally defined time-t map (defined over all of phase space). Let's start with a 2-dimensional version. Consider the first-order system of ODEs in the plane in polar coordinates:

$$\begin{aligned} \dot{r} &= r(1-r) \\ \dot{\theta} &= 1 \end{aligned} . \tag{2.4.1}$$

Without solving this system (although this is not difficult as the equations are uncoupled; see Exercise 69), we can say a lot about how solutions behave:



- The system is autonomous, so when you start does not matter, and the vector field is constant over time,
- The only equilibrium solution is at the origin. The second equation in the system really states that no point is fixed when θ is uniquely defined (on $[0, 2\pi)$, that is) for a choice of point in the plane. But the origin is special in polar coordinates.
- Considering the first equation in the system $\dot{r} = r(1-r) = f(r)$, $r(t) \equiv 1$ is another solution that corresponds to $f(r) = 0$. However, this solution is only fixed in r. It is a periodic solution called a *cycle*. What is the period?
- $r(t) \equiv 1$ is asymptotically stable as a cycle, and is called a *limit cycle*. Can you see why?



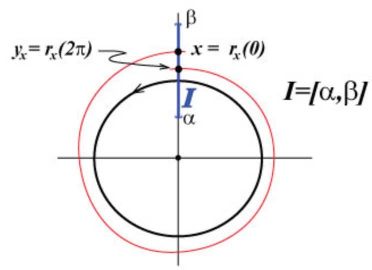
Now define

$$I = \left\{ [\alpha, \beta] \subset \text{ vertical axis } \middle| 0 < \alpha < 1, \beta > 1 \right\}.$$

For each $x \in I$, Call $r_x(t)$ the solution of Equation 2.4.1 passing through x at $t = 0$, so that

$x = r_x(0)$. Let y_x be the point in I which corresponds to the earliest positive time that the resulting $r_x(t)$ again crosses I. (1) It must cross again (why?), and (2), really $y_x = r_x(2\pi)$. Then the map $\phi : x \mapsto y_x$ defines a discrete dynamical system on I.

Some properties of this discrete dynamical system should be clear:



- The dynamics are simple on I: There is a unique fixed point at $x = 1$ corresponding to the limit cycle crossing. This fixed point is asymptotically stable so that $\forall x \in I$, $\mathcal{O}_x \longrightarrow 1$. Thus this discrete dynamical system behaves like a contraction on I.
- The same can be said for the system

$$\begin{array}{rcl} \dot{r} & = & g(r) = r\left(\frac{1}{2} - r\right)(r-1)\left(\frac{3}{2} - r\right) \\ \dot{\theta} & = & -1 \end{array}, \tag{2.4.2}$$

but only if I is chosen more carefully: Here choose

$$I = \left\{ [\alpha, \beta] \subset \text{ vertical axis } \middle| \frac{1}{2} < \alpha < 1 < \beta < \frac{3}{2} \right\}.$$

- You should draw pictures to verify this. In this last system, what happens near the cycles $r(t) \equiv \frac{1}{2}$ and $r(t) \equiv \frac{3}{2}$? Is there some kind of discrete dynamical system in the form of a first return map near there also?

EXERCISE 69. Solve the system in Equation 2.4.1. (Hint: It is uncoupled, so you can solve each equation separately.)

This part may be either discarded, or placed somewhere else (Chapter 4, where we do this stuff in detail), or modified to allude to the later work.

Recall also that for any "nice" ODE in $\mathbb{R}^n$ (the definition of nice here is mathematical and subtle, though for now just think of one where the vector field is differentiable), in a neighborhood of an equilibrium solution, one can "linearize" the system. This means that, when possible, one can associate to this system a linear system whose equilibrium solution at the origin has the same properties as that of the original system, at least near the equilibrium in study. Think of the tangent line approximation of a function at a point and you get the idea of linearization even for ODEs. Indeed, for the C^1-system

$$\begin{array}{rcl} \dot{x} & = & f(x,y) \\ \dot{y} & = & g(x,y), \end{array}$$

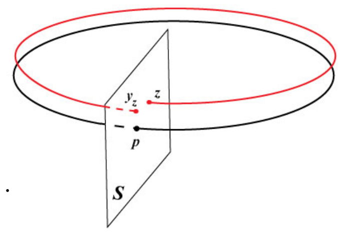
if (x_0, y_0) is an equilibrium solution, then the associated system

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x}(x_0,y_0) & \frac{\partial f}{\partial y}(x_0,y_0) \\ \frac{\partial g}{\partial x}(x_0,y_0) & \frac{\partial g}{\partial y}(x_0,y_0) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

is the linearized ODE system in a neighborhood of (x_0, y_0).

It is the main point of the celebrated Grobman-Hartman Theorem and the idea behind what is called *local linearization* that under certain conditions of the ODE system (that the vector field is C^1; this is the "nice" I mentioned before) and for certain classes of values of the eigenvalues of the linearized system, the origin of the linearized system, as an equilibrium solution, with be of the same type and have the same stability as that of (x_0, y_0). Keep this in mind.

Now, let $\mathbf{x} \in \mathbb{R}^n$ and $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ be an ODE system. Define $\phi_t : \mathbb{R}^n \to \mathbb{R}^n$ the transformation of phase space given by the time-t map. Note that this is a "slice", for fixed t, of the flow in trajectory space given by $\phi(\mathbf{x}, t) : \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$, so that



$$\phi_{t_0}(\mathbf{x}) = \phi(\mathbf{x}, t)\Big|_{t=t_0}.$$

Let $\mathbf{p}$ be a T-periodic point which is NOT an equilibrium solution (think of the point $\left(1, \frac{\pi}{2}\right)$ in the $r\theta$-plane of the system in Equation 2.4.1 above. There, $T = 2\pi$.

Then $\mathbf{p} \in Fix(\phi^T)$, but $\mathbf{f}(\mathbf{p}) \neq \mathbf{0}$.

LEMMA 2.56. *1 is an eigenvalue of the matrix $D\phi^T_{\mathbf{p}}$.*

The proof is quite straightforward and really a vector calculus calculation. However, the implications are what is interesting here.

- For the time-T map which matches the period of the cycle perfectly, the point $\mathbf{p}$ appears as a fixed point (every point on the cycle will share this property.)
- The time-T map $\phi^T : \mathbb{R}^n \to \mathbb{R}^n$ is a transformation which takes the entire phase space to itself, and is in general non-linear.
- Since $\mathbf{p}$ is fixed by ϕ^T, the derivative map $D\phi^T_{\mathbf{p}} : T_{\mathbf{p}}\mathbb{R}^n \to T_{\mathbf{p}}\mathbb{R}^n$ is simply a linear transformation of the tangent space to $\mathbb{R}^n$ at the point $\mathbf{p}$.

PROOF. The directional derivative of ϕ^T in the direction of the cycle (the curve parameterized by t, really) is the vector field $\mathbf{f}(\mathbf{p})$, and

$$\mathbf{f}(\mathbf{p}) = \mathbf{f}\left(\phi^T(\mathbf{p})\right) = \frac{d}{ds}\left(\phi^s(\mathbf{p})\right)\Big|_{s=T} = \frac{d}{ds}\left[\phi^T \circ \phi^s(\mathbf{p})\right]\Big|_{s=0} = D\phi^T_{\mathbf{p}}\left(\mathbf{f}(\mathbf{p})\right).$$

The first equality is because $\mathbf{p}$ is T-periodic, the second is due to the definition of a vector field given by an ODE, the third is because of the autonomous nature of the ODE and the last..., well, work it out. The end effect is that we have constructed the standard eigenvalue/eigenvector equation $\lambda\mathbf{v} = A\mathbf{v}$, where here $\lambda = 1$, $\mathbf{v} = \mathbf{f}(\mathbf{p})$ and the derivative matrix is A. □

DEFINITION 2.57. For $\mathbf{p}$ a T-periodic point, call the other eigenvalues of $D\phi^T_{\mathbf{p}}$ the eigenvalues of $\mathbf{p}$.

REMARK 2.58. The cycle (the periodic solution to the ODE system) becomes sort-of-like an equilibrium solution in many ways. It is another example of a closed, bounded solution that limits to itself. Solutions that start nearby may or may not stay nearby and may even converge to it. This give cycles the property of stability, much like equilibria. Many mechanical systems do exhibit asymptotically stable equilibrium states that are not characterized by the entire system staying still (think of the undamped pendulum, or more precisely a damped pendulum with just the right forcing function). How to analyze the neighboring solutions to see if a cycle is stable or not requires watching the evolution of these nearby solutions. The time-t map, and its cousin the First Return map, are ways to do this.

Maybe find an example of a damped pendulum with a forcing function where there is a asymptotically stable limit cycle (like a clock pendulum).

We end this section with a result that goes back to the notion of a contraction map:

PROPOSITION 2.59. *If* $\mathbf{p}$ *is a periodic point with all of its eigenvalues of absolute value strictly less than* 1*, then* $\mathcal{O}_{\mathbf{p}}$ *is an asymptotically stable limit cycle.*

Here the tangent linear map at $\mathbf{p}$ carries the infinitesimal variance in the vector field, which in turn betrays what the neighboring solutions do over time. The non-infinitesimal version theoretically plays the same role. Construct X, the closure of a small open subset of $\mathbb{R}^{n-1}$ centered at $\mathbf{p}$ and normal to the vector field $\mathbf{f}(\mathbf{p})$ at $\mathbf{p}$ (See the picture). Due to the continuity of the vector field given by the ODE system, all solutions of the ODE system that start in X sufficiently close to $\mathbf{p}$ will leave X, circle around, and again cross X. In the case where all solutions cross again at time-T (the period of $\mathbf{p}$), then the time-T map defines a discrete dynamical system on X. In the case where this is not the case (usual in nonlinear systems), then we neglect where the nearby solutions are at time T and simply look for where they again cross X. This latter case is the difference between a time-T map and the first return map. However, both of these constructions coalesce nicely into the infinitesimal version. We will revisit this point maybe later.

EXERCISE 70. Solve the first order, autonomous non-linear ODE system in cylindrical coordinates $\dot{r} = \frac{1}{2}r(1-r)$, $\dot{\theta} = 1$, $\dot{z} = -z$ and show that there exists an asymptotically stable limit cycle (Hint: Since the system is incoupled, you can solve each ODE separately.) What are the eigenvalues of the 2π-periodic point at $p = (1,0,0)$?

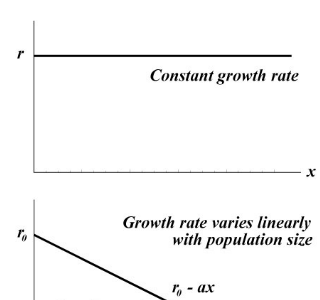
We definitely need the picture here alluded to in the previous paragraph.

2.5. A Quadratic Interval Map: The Logistic Map

Like linear functions defined on the unit interval, discrete dynamical systems constructed via maps whose expression is a quadratic polynomial have many interesting properties. The ideal model for a study of quadratic maps of the interval is the Logistic Map. Before defining it, however, let's motivate its prominence.

Consider the standard linear map on the real line, $f : \mathbb{R} \to \mathbb{R}$, $f(x) = rx$. As a model for population growth (or decay), we restrict the domain to be non-negative (for a realistic population size) and the values for the parameter r to be positive, so that $f_r : [0,\infty) \to [0,\infty)$, where $r \geq 0$. Hence the recursive model is $x_{n+1} = f(x_n) = rx_n$, and again $\mathcal{O}_x = \{y \in [0,\infty) \,|\, y = f^n(x) = r^n x, n \in \mathbb{N}\}$. It is a good model for population growth when the population size is not affected by any environmental conditions or resource access, and is considered "ideal" growth. One way to view this is to say that in this case, "the growth factor r is constant and independent of the size of the population (see figure).

However, realistically speaking, unlimited population growth is unsustainable in any limited environment, and hence the actual growth



factor winds up being dependent on the actual size of the population. Things like crowding and the finite allocation of resources typically mean that larger population sizes usually experience a dampened growth factor over time vis a vis small populations (think of a small number of fish in a large pond as opposed to a very large number of fish in the same pond). Hence a better model to simulate populations over time is to allow the growth factor to vary with the population size. The easiest way to do this is to replace the constant growth factor r, with one that varies linearly with population size. Here then replace r with the expression $r_0 - ax$, where r_0 is an ideal growth factor (for very small populations near 0), and a is a positive constant (see figure). The model becomes

$$f : \mathbb{R} \to \mathbb{R}, \quad f(x) = (r_0 - ax)x,$$

or with a change in variables

$$f : \mathbb{R} \to \mathbb{R}, \quad f(y) = \lambda y(1 - y).$$

Keep in mind the limitations of the model as a guide to studying populations, however. For λ a positive constant, f is positive only on the interval $[0,1]$. And really only some values of λ make this a good model for populations. To understand the last statement, you will need to actually see how λ relates to the constants r_0 and a, and to study the graph of $r_0 - ax$ above as it relates to a population x.

Exercise 71. Do the change of variables that takes $f(x) = (r_0 - ax)x$ to $f(y) = \lambda y(1 - y)$, writing λ as a function of a and r_0.

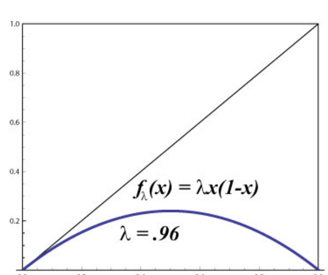
Hence we will begin to study the dynamics of the map $f : [0,1] \to [0,1]$, $f(x) = \lambda x(1 - x)$, called the logistic map. We will eventually see just how rich and complex the dynamics can actually be. For now, however, we will only spend time on the values of λ where the dynamics are simple to describe. First some general properties:

- f is only a map on the unit interval when $\lambda \in [0,4]$. Why does it fail for other values of λ?

  Exercise 72. Show that the logistic map $f(x) = \lambda x(1 - x)$ does not produce a dynamical system on the interval $[0,1]$, for $\lambda \notin [0,4]$.

- λ is sometimes called the *fertility constant* in population dynamics.
- We will use the notation f_λ to emphasize the dependence of f on the parameter.

Proposition 2.60. *For $\lambda \in [0,1]$, $\forall x \in [0,1)$, we have $\mathcal{O}_x \longrightarrow 0$.*



Visually, the graph of f_λ is a parabola opening down with horizontal intercepts at $x = 0, 1$. The vertex is at $\left(\frac{1}{2}, \frac{\lambda}{4}\right)$. And for $\lambda \in [0,1]$, the entire graph of f_λ lies below the diagonal $y = x$ (see figure at left). Cobweb to see where the orbits go.

PROOF. The fixed points of f_λ satisfy $f_\lambda(x) = x$, or $\lambda x(1-x) = x$. This is solved by either $x = 0$ or $x = \frac{\lambda-1}{\lambda} = 1 - \frac{1}{\lambda}$. Hence for $0 \leq \lambda \leq 1$, the only fixed point on the interval $[0,1]$ is $x = 0$. Also, $\forall x \in [0,1]$, $f_\lambda(x) < x$. This implies that $\mathcal{O}_x$ is a decreasing sequence. As it is obviously bounded below, it must converge.

Now choose a particular $x \in [0,1]$ and notice that $f_\lambda(x) < \frac{1}{2}$. Thus, after one iteration of the map, every orbit lies inside of the subinterval $\left[0, \frac{1}{2}\right]$. So after one iteration, $f_\lambda|_{[0,1]}$ is a discrete dynamical system on a closed, bounded interval which is nondecreasing and has no fixed points on the interior $\left(0, \frac{1}{2}\right)$. Then by Proposition 2.48, the only fixed point is $x = 0$ and all orbits converge to it. □

Some notes:

- Both conditions, that the interval be closed, and that the map be nondecreasing, are necessary to apply Proposition 2.48. Since the original map f_λ was not nondecreasing, and the interval was open at 1, we needed to modify the situation a bit to fit the lemma. The nice structure of the graph of f_λ allowed for this by looking for a future iterate where the map would be nondecreasing. This is a common idea, and the basis for the notion of a map being *eventually nondecreasing*. Look for this in other maps in this class and beyond.
- The orbit $\mathcal{O}_1$ is special:
$$\mathcal{O}_1 = \{1, 0, 0, 0, \ldots\}.$$
The point $x = 1$ is called a pre-image of the fixed point $x = 0$. This is often seen in maps which are not one-to-one. The orbit $\mathcal{O}_1$ is called *eventually fixed*. There also exist *eventually periodic* points also. Both of these can not exist in invertible maps (why?), but it is easy to see that the quadratic map f_λ is NOT invertible on $[0,1]$. But for now, realize that Proposition 2.60 is actually valid for $x \in [0,1]$, including $x = 1$. I left it out originally due to its special nature.
- Were this logistic map with this range of λ to be used to model populations, one can conclude immediately the following:

  All starting populations are doomed!

  Think about that.

Now, let's change our parameter range a bit, and consider some higher parameter values:

PROPOSITION 2.61. *For $\lambda \in [1,3]$, $\forall x \in (0,1)$, $\mathcal{O}_x \longrightarrow 1 - \frac{1}{\lambda}$.*

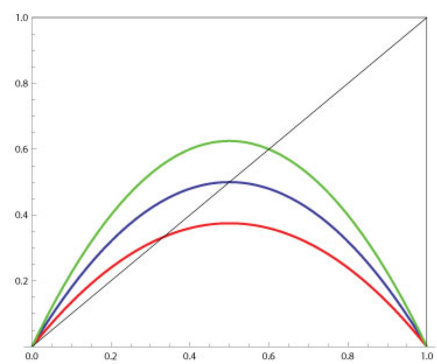
REMARK 2.62. if this is true, then $\lambda = 1$ is a bifurcation value for the family of maps f_λ, since

- for $\lambda \in [0,1]$, the fixed point $x = 0$ is an attractor, and
- for $\lambda \in [1,3]$, the fixed point $x = 0$ is a repeller (do you see this?).

The idea of the proof is that on this range of values for λ, the graph of f_λ intersects the line $y = x$ at two places, and these places are the two roots of $x = \lambda x(1-x)$ (see proof of Proposition 2.60 above.) At right are the graphs for three typical logistic maps, for $\lambda = 1.5$, $\lambda = 2$, and $\lambda = 2.5$. It turns out that showing the fixed point $x_\lambda = 1 - \frac{1}{\lambda}$ is attractive is straightforward and left as an exercise.

EXERCISE 73. Show that for $\lambda \in (1,3)$, there is an attracting fixed point of the logistic map $f_\lambda(x)$ at $x = 1 - \frac{1}{\lambda}$.

However, showing that almost every orbit converges to x_λ is somewhat more involved. I won't do the proof in class, as it is in the book. You should work through it to understand it well, since it raises some interesting questions. Like:



(1) What generates the need for the two cases they describe in the book?
(2) For what value(s) of λ is the attracting fixed point super-attracting?
(3) The endpoints of the interval $\lambda \in [1,3]$ are special and related in a very precise and interesting way. The property they share indicates a central property of attractive fixed an periodic points of C^1-maps of the interval. Can you see this?

This proof needs to be here in detail and we need to remove the references to the text.

Once we surpass the value $\lambda = 3$ for $\lambda \in [0,4]$, things get trickier. We will suspend our discussion of interval maps here for a bit and develop some more machinery first.

2.6. More general metric spaces

There are easy-to-describe-and-visualize dynamical systems that occur on subsets of Euclidean space which are not Euclidean. As long as we have a metric on the space, it remains easy to discuss how points move around by their relative distances from each other. So let's generalize a bit and talk about metric spaces without regard to how they sit in a Euclidean space. To this end, let X be a metric space.

DEFINITION 2.63. An ϵ-ball about a point $x \in X$ is the set

- (open) $B_\epsilon(x) = \{y \in X \mid d(x,y) < \epsilon\}$, and
- (closed) $\overline{B}_\epsilon(x) = \{y \in X \mid d(x,y) \leq \epsilon\}$.

DEFINITION 2.64. A sequence $\{x_i\}_{i=1}^\infty \subset X$ is said to converge to $x_0 \in X$, if $\forall \epsilon > 0$, $\exists N \in \mathbb{N}$ such that $\forall i \geq N$, $d(x_i, x_0) < \epsilon$.

DEFINITION 2.65. A sequence is Cauchy if $\forall \epsilon > 0$, $\exists N \in \mathbb{N}$ such that $\forall i, j \geq N$, $d(x_i, x_j) < \epsilon$.

REMARK 2.66. A metric space X is called *complete* if every Cauchy sequence converges.

DEFINITION 2.67. A map $f : X \to X$ on a metric space X is called an *isometry* if

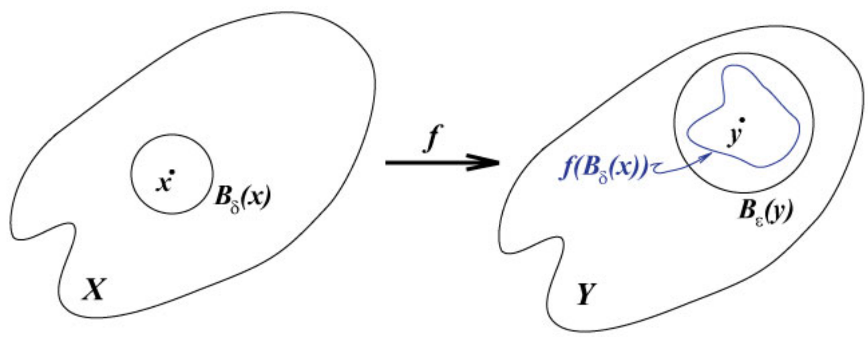
$$\forall x, y \in X, \quad d(f(x), f(y)) = d(x, y).$$

We can generalize this last definition to maps where the domain and the range are two different spaces: $f : X \to Y$, where both X and Y are metric spaces:

DEFINITION 2.68. A map $f : X \to Y$ between two metric spaces X, with metric d_X, and Y, with metric d_Y, is called an *isometry* if

$$\forall x_1, x_2 \in X, \quad d_Y(f(x_1), f(x_2)) = d_X(x_1, x_2).$$

DEFINITION 2.69. A map $f : X \to Y$ between two metric spaces is called *continuous* at $x \in X$ if $\forall \epsilon > 0$, $\exists \delta > 0$, such that if $\forall y \in X$ where $d_X(x, y) < \delta$, then $d_Y(f(x), f(y)) < \epsilon$.



This gives us an easy way to define what makes a function continuous at a point when the spaces are not Euclidean but are metric spaces. We will need this as we talk about common spaces in dynamical systems that are not like $\mathbb{R}^n$ but still allow metrics on them. That a map on a space is continuous is a vital property to possess if we are to be able to really talk at all about how orbits behave under iteration of the map. For now, though, more definitions:

DEFINITION 2.70. A continuous bijection (remember that a bijection is a continuous map which is also an injection, or one-to-one map, as well as a surjection, or onto map) $f : X \to Y$ with a continuous inverse is called a *homeomorphism*.

EXAMPLE 2.71. For any metric space (or any topological space in general!) X, the *identity map* on X ($f : X \to X$, $f(x) = x$) is a homeomorphism. It is obviously continuous (for any $\epsilon > 0$, choose $\delta = \epsilon$), one-to-one and onto, and it is its own inverse.

EXAMPLE 2.72. The map $h : [0, 1) \to S^1$ given by $h(x) = e^{2\pi i x}$ is continuous, one-to-one and onto. It also has an inverse, but the inverse is NOT continuous.

EXERCISE 74. Show that h in Example 2.72 is continuous, one-to-one and onto. Construct h^{-1} and show that it is not continuous.

EXAMPLE 2.73. Recall that linear maps $f:\mathbb{R}\to\mathbb{R}$, $f(x)=ax+b$ are, of course, invertible, as long as $a\neq 0$. However, be careful of the domain: Let $f(x)=\frac{1}{2}x+\frac{1}{2}$ on $I=[0,1]$. Here, f is certainly injective (it is non-decreasing, and a contraction!). But it cannot have an inverse on I, since it is NOT onto I. In fact, if the domain is $I=[0,1]$, then the range of f is $\left[\frac{1}{2},1\right]$. So think about the following: For $f:I\to I$ to be a homeomorphism on a bounded $I=[a,b]$, it must be both one-to-one and onto I. What does that imply about the images of the endpoints?

EXERCISE 75. Show that for a homeomorphism $f:[a,b]\to[a,b]$, it must be the case that either $f(a)=a$ and $f(b)=b$, or $f(a)=b$ and $f(b)=a$.

REMARK 2.74. When a homeomorphism exists between two spaces, the two spaces are called *homeomorphic* and mathematically they are considered equivalent, or the same space. Anything defined on a space or with it can be defined or used on any other space that is homeomorphic to it. It is the chief way for mathematicians to classify spaces according to their properties.

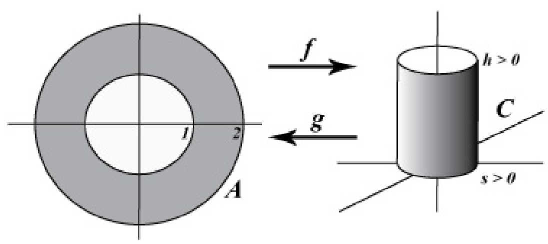


FIGURE 13. The annulus $A\in\mathbb{R}^2$ and the cylinder $C\in\mathbb{R}^3$ are homeomorphic spaces.

EXERCISE 76. Given the annulus A and the cylinder C in Figure 13 (both include their boundary circles, although the dimensions of C are given only as radius $s>0$ and height $h>0$), show that they are homeomorphic by explicitly constructing the maps $f:\mathbb{R}^2\to\mathbb{R}^3$ which takes A to C, and $g:\mathbb{R}^3\to\mathbb{R}^2$ which takes C to A. (Hint: Use the obvious coordinate systems for the Euclidean spaces.)

Typically, on a metric space X, there are many metrics that one can define. However, like the above notion of homeomorphism, many of them are basically the same, and can be treated a equivalent. Others, maybe not. To understand this better,

DEFINITION 2.75. Let d_1 and d_2 be two metrics on a metric space X. Then we say d_1 and d_2 are *isometric* if $\forall x,y\in X$, $d_1(x,y)=d_2(x,y)$.

DEFINITION 2.76. Two metrics d_1 and d_2 on a metric space X are called *(uniformly) equivalent* if the identity map and its inverse are both Lipschitz continuous.

To elaborate on this last definition, we consider $f:X\to X$, $f(x)=x$, to be the map that takes points in X using the metric d_1 to points in x using the other metric d_2. This is like considering X as two different metric spaces, one with d_1

and the other with d_2. In essence, then the definition says that $\exists C, K \geq 0$ such that $\forall x, y \in X$, both 1) $d_2(f(x), f(y)) \leq C d_1(x,y)$ and 2) $d_1\left(f^{-1}(x), f^{-1}(y)\right) \leq K d_2(x,y)$ hold. This simplifies using the identity map to 1) $d_2(x,y) \leq C d_1(x,y)$ and 2) $d_1(x,y) \leq K d_2(x,y)$ everywhere. Of course, what this really only means is that there are global bounds (over the space X, that is) on how the two metrics differ.

EXAMPLE 2.77. On $\mathbb{R}^2$, the Euclidean metric and the Manhattan metric are uniformly equivalent. Recall from Remark 2.12 that these are the metrics that come, respectively, from the $\mathbf{p} = 2$-norm and the $\mathbf{p} = 1$-norm. Call d_2 the Euclidean metric and d_1 the Manhattan metric. Then we can show uniform equivalence via the following:

$$(d_1(\mathbf{x}, \mathbf{y}))^2 = (|x_1 - y_1| + |x_2 - y_2|)^2 \geq |x_1 - x_2|^2 + |y_1 - y_2|^2 = (d_2(\mathbf{x}, \mathbf{y}))^2.$$

Hence $(d_2(\mathbf{x}, \mathbf{y}))^2 \leq (d_1(\mathbf{x}, \mathbf{y}))^2$ so that $d_2(\mathbf{x}, \mathbf{y}) \leq d_1(\mathbf{x}, \mathbf{y})$.

Going the other way is a little trickier. Given the Cauchy-Schwartz Inequality from Linear Algebra, $(x \cdot y)^2 \leq \|x\| \cdot \|y\|$, we can say

$$\begin{aligned} d_1(\mathbf{x}, \mathbf{y}) &= |x_1 - y_1| + |x_2 - y_2| \\ &= |x_1 - y_1| \cdot 1 + |x_2 - y_2| \cdot 1 \\ &\leq \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}\sqrt{1^2 + 1^2} = \sqrt{2}\, d_2(\mathbf{x}, \mathbf{y}). \end{aligned}$$

Hence

$$d_1(\mathbf{x}, \mathbf{y}) \leq \sqrt{2}\, d_2(\mathbf{x}, \mathbf{y}) \text{ and } d_2(\mathbf{x}, \mathbf{y}) \leq d_1(\mathbf{x}, \mathbf{y}).$$

EXERCISE 77. Going back to Remark 2.12, show that the Euclidean and maximum metrics are also uniformly equivalent on $\mathbb{R}^2$.

One last definition:

DEFINITION 2.78. A map $f : X \to Y$ is called *eventually contracting* if $\exists C > 0$, such that if $\forall x, y \in X$ and $\forall n \in \mathbb{N}$,

$$d(f^n(x), f^n(y)) \leq C\lambda^n d(x,y)$$

for some $0 < \lambda < 1$.

There are many maps that are definitely not contractions, yet ultimately behave like one. Here is one:

EXAMPLE 2.79. Let $f_2(x) = 2x(1-x)$ be the $\lambda = 2$-logistic map, restricted to the open interval $(0,1)$ (this cuts out the repelling fixed point at 0). This is the one with the super-attracting fixed point at $x = \frac{1}{2}$. Here f_2 is definitely NOT a contraction. You can see this visually by inspecting the graph: Should the graph of a function have a piece which is sloping up or down at a grade more than perfectly diagonal, the function will stretch intervals there. See the graph. To see this analytically, let $x = \frac{1}{8}$ and $y = \frac{1}{4}$. Then $f_2(x) = \frac{7}{32}$ and $f_2(y) = \frac{3}{8} = \frac{12}{32}$. Then

$$d(f_2(x), f_2(y)) = \left|\frac{12}{32} - \frac{7}{32}\right| = \frac{5}{32} \leq C\frac{4}{32} = C\frac{1}{8} = C\left|\frac{1}{4} - \frac{1}{8}\right| = Cd(x,y)$$

only when C is some number greater than 1. However, eventually, every orbit gets close to the only fixed point at $x = \frac{1}{2}$ where the derivative is very flat. The function f_2, restricted to the interval $\left[\frac{3}{8}, \frac{5}{8}\right]$ is a $\frac{1}{2}$-contraction (Can you show this? Use the derivative!). And one can also show that f_2 is 2-Lipschitz on all of $(0,1)$. Thus, one can conclude here that f_2 is eventually contracting on $(0,1)$, and $\forall x, y \in (0,1)$,

$$d\left(f(x), f(y)\right) \le 4\left(\frac{1}{2}\right)^n d(x,y).$$

EXERCISE 78. Go back to the Example 2.54 $f(x) = \sqrt{x-1}+3$ on the interval $I = [1,\infty)$. Show that f is NOT an eventual contraction (Hint: Try to find a value for C in Definition 2.78 that works in a neighborhood of $x = 1$.) Now show that f IS an eventual contraction on any closed interval $[b,\infty)$, for $1 < b \le 5$.

Here are some of the more common non-Euclidean metric spaces encountered in dynamical systems:

2.6.1. The n-sphere. The n-dimensional sphere

$$S^n = \left\{\mathbf{x} \in \mathbb{R}^{n+1} \;\middle|\; \|\mathbf{x}\| = 1\right\}.$$

Here, the $n \in \mathbb{N}$ denotes the "size" of the space and not the space in which it exists in this definition. In fact, the space called the n-sphere doesn't actually exist in any space unless we define it that way. We typically place the n-sphere in $\mathbb{R}^{n+1}$ so that we can place coordinates on it easily. For example, we can parameterize the 2-sphere via the function $\Psi_2 : D \to \mathbb{R}^3$, where $D = [0,2\pi] \times [0,\pi]$, by

$$\Psi_2(\theta_1,\theta_2) = \left(\cos\theta_1 \sin\theta_2, \sin\theta_1 \sin\theta_2, \cos\theta_2\right), \tag{2.6.1}$$

as in Figure 14. And in general, for $n \in \mathbb{N}$, the spherical coordinate system in $\mathbb{R}^n$ provides a ready parameterization for S^n: In spherical coordinates (ρ,θ), the unit sphere is simply the set of all points in $\mathbb{R}^{n+1}$ with first coordinate 1. Hence all of the other (angular) coordinates $\theta = (\theta_1,\ldots,\theta_n)$ parameterize the n-sphere.

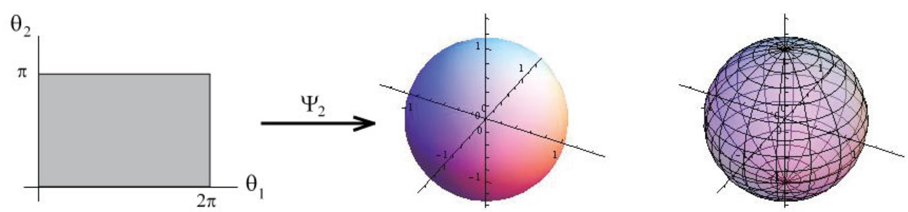


FIGURE 14. A common parameterization of the 2-sphere S^2.

EXERCISE 79. Find the pattern in parameterizing S^1 by $\Psi_1 : [0,2\pi] \to \mathbb{R}^2$, where $\Psi_1(\theta) = (\cos\theta, \sin\theta)$ and S^2 by Equation 2.6.1, to construct an explicit parameterization Ψ_n of S^n using the n-angles of the spherical coordinate system in $\mathbb{R}^{n+1}$. What are the ranges of each of your angular coordinates θ_i, $i = 1,\ldots,n$?

EXERCISE 80. For the parameterization in Equation 2.6.1, one can view this as a wraparound of the box at left of the figure onto the 2-sphere at right. Identify where the four edges of the box go under the map Ψ_2, and draw the images of these four edges (the "seam" of the parameterization) onto the 2-sphere.

2.6.2. The unit circle. Really this is the 1-dimensional sphere

$$S^1 = \left\{ \mathbf{x} \in \mathbb{R}^2 \,\middle|\, ||\mathbf{x}|| = 1 \right\}.$$

However, we also can interpret the circle as the unit-modulus complex numbers

$$\begin{aligned} S^1 &= \left\{ z \in \mathbb{C} \,\middle|\, |z| = 1 \right\}, \\ &= \left\{ e^{i\theta} \in \mathbb{C} \,\middle|\, \theta \in [0, 2\pi) \right\}, \end{aligned}$$

(compare this to Equation 2.3.1) and also in a more abstract sense as

$$S^1 = \left\{ x \in \mathbb{R} \,\middle|\, x \in [0,1] \text{where } 0 = 1 \right\}.$$

This last definition requires a bit of explanation. From Set Theory, we have the following:

Definition 2.80. Given a set X, a *partition* $\mathcal{P}_X$ on X is a set of disjoint, exhaustive subsets of X.

Remark 2.81. You are already familiar with the term partition from Calculus. For example, in the development of the notion of a definite integral of a function defined on a closed, bounded interval $I = [a,b] \in \mathbb{R}$, $a < b$, one defines a *partition* of I into smaller intervals via a finite set of points in I, $a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b$, so that

$$[a,b] = [a = x_0, x_1] \cup [x_1, x_2] \cup \ldots \cup [x_{n-1}, x_n = b].$$

This is slightly different from our current definition in that the intervals overlap on their edges. For our current purpose in more generally, we require that the partition elements be mutually exclusive and exhaustive.

Definition 2.82. An *equivalence relation* R on a set X is a collection of elements of $X \times X$, denoted (x,y) or $x \sim_R y$, (or simply $x \sim y$, when R is understood,) such that

(1) $x \sim_R x$, $\forall x \in X$,
(2) $x \sim_R y$ iff $y \sim_R x$, $\forall x, y \in X$, and
(3) if $x \sim_R y$ and $y \sim_R z$, then $x \sim_R z$, $\forall x, y, z \in X$.

Each $x \in X$ is an element of a unique *equivalence class* $[x]$, a subset of X where

$$[x] = \left\{ y \in X \,\middle|\, y \sim_R x \right\}.$$

And the set of all equivalence classes of X under R form a partition of X.

Furthermore, the set of all partition elements form a new set, called the *quotient set* of the equivalence relation. Given X a topological space, one can always make the quotient set into a space using the topology of X (it is called the quotient topology). But it is a much deeper question exactly when the quotient set, made into a space using the topology of X, has the same properties as that of X. But for now, we say that for X a set with an equivalence relation R, the quotient set is denoted $Y = X/R$, or sometimes $Y = X/$.

Place some relatively simple equivalence relation examples here.

Example 2.83. Any function $f : X \to \mathbb{R}$ defines an equivalence relation on X. Each element of the partition is simply the collection of all point that map to the same point in the range of X:

$$[x] = \left\{ y \in X \,\middle|\, f(y) = f(x) \right\}.$$

Recall in Calculus, we defined the inverse image of a point in the range of a function as

$$f^{-1}(c) = \left\{ x \in X \,\middle|\, f(x) = c \right\}.$$

Hence we can say that here that the equivalence class of a point $x \in X$ given by the function $f : X \to \mathbb{R}$ is simply the inverse image of the image of x, or $[x] = f^{-1}(f(x))$. This is well-defined regardless of whether f even has an inverse, since the inverse image of a function is only defined as a set. Think about this.

Using this last example, we have one more definition of S^1. namely, let $r : \mathbb{R} \to S^1$ be a function $r(x) = e^{2\pi i x}$. Then $r(x) = r(y)$ iff $x - y \in \mathbb{Z}$.

EXERCISE 81. Show that for r defined here, that $r(x) = r(y)$ iff $x - y \in \mathbb{Z}$.

In this case, each point on the circle has as its inverse image under r all of the points in the real line that are the same distance from the next highest integer (see Figure 15. Thus the map r looks like the real line $\mathbb{R}$ infinitely coiled around the circle. In this way, we commonly say that

$$S^1 = \mathbb{R}/\mathbb{Z}.$$

REMARK 2.84. Now the more abstract definition

$$S^1 = \left\{ x \in \mathbb{R} \,\middle|\, x \in [0,1] \text{ where } 0 = 1 \right\}$$

should make more sense. In a way, one can take the unit interval in $\mathbb{R}$, pluck it out of $\mathbb{R}$, curve it around to the point where one can "join" its two endpoints together to make the circle. We say the two end points are now *identified*, and the space S^1, while still having a well-defined parameter on it, isn't sitting in an ambient space anymore. Note also that the new space is still closed and bounded, but it has NO boundary. Intervals in $\mathbb{R}$ cannot have these properties simultaneously. But the circle, still one dimensional, is not a subset of $\mathbb{R}$.

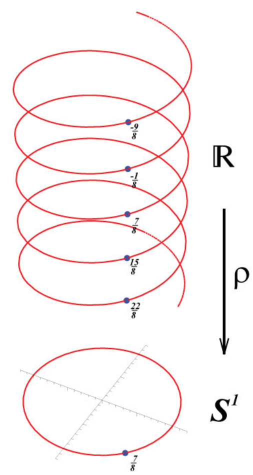


FIGURE 15. The map $\rho : \mathbb{R} \to S^1$.

Another interesting consequence of this idea involves how to view functions on S^1 via a vis those on $\mathbb{R}$. Let $f : \mathbb{R} \to \mathbb{R}$ be any function which is T-periodic (thus it satisfies $f(x+T) = f(x)$, $\forall x \in \mathbb{R}$). Then f induces a function $g : S^1 \to \mathbb{R}$ on S^1, given by $g(t) = f(tT)$. Conversely, any function defined on S^1 may be viewed as a periodic function on $\mathbb{R}$, a tool that will prove very useful later on.

EXERCISE 82. Show that a T-periodic, continuous map $f : \mathbb{R} \to \mathbb{R}$ will induce a continuous map $g : S^1 \to \mathbb{R}$. Then show that a T-periodic, differentiable map $f : \mathbb{R} \to \mathbb{R}$ will induce a differentiable map $g : S^1 \to \mathbb{R}$. Hint: Parameterize the circle correctly.

EXERCISE 83. Show that you cannot have a continuous surjective contraction on S^1. However, construct a continuous, non-trivial contraction on the circle S^1. (Hint: A continuous map cannot break the circle in its image or it would not be continuous at the break. But it can fold the circle.)

2.6.3. The cylinder. Relatively simple to describe *product* spaces show up often as the state spaces of dynamical systems. Define the cylinder as $C = S^1 \times I$, where $I \subset \mathbb{R}$ is some interval. Here I can be closed, open or half-closed, and can be bounded or all of $\mathbb{R}$. In fact, by the above discussion, any function $f : \mathbb{R}^2 \to \mathbb{R}^2$ which is T period in one of its variables, may be viewed as a function on a cylinder (think of the phase space of the undamped pendulum). Sometimes we call a cylinder whose linear variable is all of $\mathbb{R}$ the infinite cylinder.

2.6.4. The 2-torus. The 2-dimensional torus $\mathbb{T}^2$ (or just the torus T when there is no confusion) $\mathbb{T} = S^1 \times S^1$ is another surface. Like before, any function $f : \mathbb{R}^2 \to \mathbb{R}^2$ which is periodic in each of its variables, may be viewed as a function on a torus. Conversely, a function on the torus may be studied instead as a doubly periodic function on $\mathbb{R}^2$. We will have occasion to use this fact later. For now, consider the parameterization of this surface as a subset of $\mathbb{R}^3$ by $\Phi_2 : [0, 2\pi] \times [0, 2\pi] \to \mathbb{R}^3$, where

$$\Phi_2(\theta_1, \theta_2) = \left((2 + \cos\theta_1)\cos\theta_2, (2 + \cos\theta_1)\sin\theta_2, \sin\theta_1\right).$$

Figure 16 shows the parameterization, along with the images of the bottom and left edges of the parameter space.

Exercise 84. Show that the function $g(x, y) = \left(2\cos 2x, 4y^2 - y^4\right)$, from the plane to itself, can be made into a function on the standard infinite cylinder $C = S^1 \times \mathbb{R}$. Show also that by limiting the domain appropriately, one can use g to construct a continuous function on the torus $T = S^1 \times S^1$.

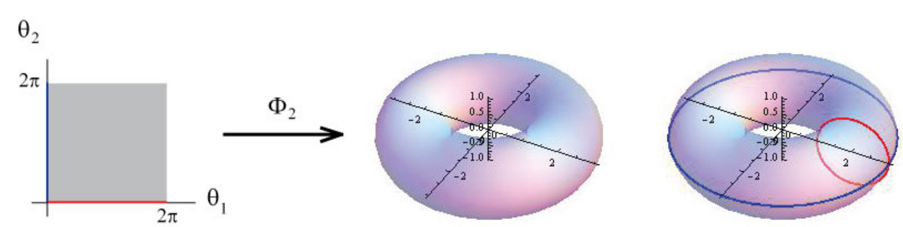


Figure 16. One parameterization of the 2-torus $\mathbb{T}^2$.

One thing to note: Where are the images of the top and right edges? Since both parameters are 2π-periodic, they are the same, respectively, as the bottom and left edges. Envision taking the square piece of paper at left, bending it into a cylinder to identify the left and right edges and then bending the cylinder to identify the top and bottom circles. That is a torus. And if you were an ant walking on the torus and you approached the red line and crossed it, how would your shadow path look back in the parameter square? Understanding this will be very important at some point soon. Think about it.

We end this chapter with one last example of a particular set that has an interesting property constructed via a contraction map. For now, we will only define the set and identify the property. In time, we shall return to this set, as it is quite ubiquitous in dynamical systems theory.

2.7. A Cantor Set

Simple dynamics also allows us to define some of the physical properties of the spaces that a map can "act" on. For example, consider the following subset of the unit interval $[0,1]$, defined by what is called a *finite subdivision rule*; A method of recursively dividing a polygon or other geometric shape into smaller and smaller pieces.

Define $C_0 = [0,1]$. On C_0, remove the open middle third interval (this is the finite subdivision rule) $\left(\frac{1}{3},\frac{2}{3}\right)$ and call the remainder C_1. Hence $C_1 = \left[0,\frac{1}{3}\right] \cup \left[\frac{2}{3},1\right]$. Continue removing the middle third from each remaining closed interval from the previous C_n to construct C_{n+1} (See Figure 17. Then define

$$C = \bigcap_{n=0}^{\infty} C_n.$$

C is called the Ternary Cantor Set.

FIGURE 17. The first few steps in the construction of the Ternary Cantor Set

It has the following rather remarkable properties:

- There are no positive-length intervals in C. Indeed, at each step, you are removing open intervals whose total length is exactly $\frac{1}{3}\left(\frac{2}{3}\right)^{n-1}$. Sum these lengths over the natural numbers (this forms a geometric series) and you get $\frac{1}{3}\left(\frac{1}{1-\frac{2}{3}}\right) = 1$.
- It is easy to see that, although many intervals of points of $[0,1]$ are not in C, there are many that remain. Any point in $[0,1]$ that is a multiple of a power of $\frac{1}{3}$ (the end points of the intervals at each stage) is in C. What is more surprising is that there are many others. One way to see this is to use an alternate description of C as the set of points in the unit interval whose ternary expansion base 3 (like our decimal expansion but using only 0's 1's, and 2's) has no 1's. Note that here, the number $\frac{1}{3} = (.100000\ldots)_3 = (.022222222\ldots)_3$, so $\frac{1}{3} \in C$.

  EXERCISE 85. In the normal decimal expansion, show $1.0 = .999\overline{999}$.

  Why no 1's are allowed? Because we take them out. Think about this: In a ternary expansion of points in $[0,1]$, the middle third (every point in the open interval $\left(\frac{1}{3},\frac{2}{3}\right)$) will have an expansion $(.1****...)_3$ where not all of the *'s are 0. But these points are not in C. Make sense?

EXERCISE 86. Show that the number $x = \frac{1}{4} \in C$ via its ternary expansion (it is not a boundary point of any open subinterval removed at any intermediate stage C_n.)

- It is also easy to see that there are infinitely many points in C (all of the multiples of powers of $\frac{1}{3}$ that are the boundary points of removed intervals.) What is more interesting is that the number of points in C is uncountable (so not like the natural number kind of infinity. This is more like the number of points in $[0,1]$ kind of infinity.) In fact, there are AS MANY points left in C as there were in the original $[0,1]$! (Wrap your head around that mathematical fact! We will return to this in a minute.)
- One can give C a topology so that it is a space, like $[0,1]$ is a space (the subspace topology it inherits from $[0,1]$ will do.) Then one can say that anything homeomorphic to a Cantor Set is a Cantor Set. Hence this one example will share its properties with all other Cantor sets defined similarly.
- Continuing on the ternary expansion theme, we see

$$C = \left\{ x \in \mathbb{R} \;\middle|\; x = (0.x_1x_2x_3\ldots)_3 = \sum_{n\in\mathbb{N}} x_n \cdot 3^{-n}, \quad x_n = 0 \text{ or } 2 \right\}.$$

There is a wonderful, real-valued function defined on $[0,1]$ which exposes some of its properties of C. This function, the Cantor-Lebesgue function, relies on this series description of Cantor numbers: Define $F : [0,1] \to [0,1]$, by

$$F(y) = \sum_{n\in\mathbb{N}} x_n 2^{-(n+1)}, \quad \text{where} \quad x = (.x_1x_2x_3\ldots)_3 = \min_{x\in C} x \geq y.$$

In essence, all points in $[0,1]$ situated in a positive-length gap between points in C are mapped to a constant; the constant being a multiple of a power of $\frac{1}{2}$ corresponding uniquely to that gap and the function value of the Cantor point at the upper end of the gap. The resulting graph of F is an example of what is called a Devil's Staircase (see Figure 18). It is a continuous function whose derivative is almost everywhere 0, which we will not show here. It is also a surjective function on $[0,1]$ To see this, let $z \in [0,1]$. Then its binary expansion is $z = (0.a_1a_2a_3a_4\ldots)_2$, where $z = \sum_{n\in\mathbb{N}} a_n 2^{-n}$. But then $y = \sum_{n\in\mathbb{N}} (2a_n)3^{-n} \in C$ and $F(y) = z$. Conclusion? There are at least as many points in C as there are in $[0,1]$. There cannot be more, so the cardinality of C and $[0,1]$ are the same! Strange, eh? Cantor sets are quite popular in analysis due to there ability to provide counterexamples to seemingly intuitive, but untrue beliefs.
- The map $f : C \to C$, $f(x) = \frac{x}{3}$ is a contraction whose sole fixed point is at 0.

EXERCISE 87. Show that for $f(x) = \frac{x}{3}$, $f(C) \subset C$.

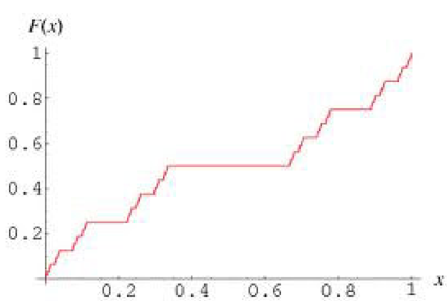


FIGURE 18. A Devil's Staircase: The graph of

This last point is special:

DEFINITION 2.85. A set on which there exists a contraction map which is a homeomorphism onto its image has the property of *self-similarity*.

The set is also called *re-scalable*. This is easy to see for a contraction on $\mathbb{R}$.

EXERCISE 88. Find an expression for a contraction on C whose fixed point is $x = 1$.

EXERCISE 89. Show that the map $f(x) = 1 - \frac{x}{3}$ is a contraction on C, and verify that its fixed point is in C.

In fact, all of these contractions are maps that take the interval C_0 onto one of the subintervals of C_1 (the last exercise reversing the order of points), indicating that the subset of C inside each of the subintervals in C_1 looks exactly like the parent set C in C_0. In this way, one can build a (linear) map taking C_0 onto any subinterval in any C_n. In this way, the part of C in any subinterval of C_n looks exactly the same as C. That such a complicated set can be defined by such a simple single finite subdivision rule is quite remarkable.

EXAMPLE 2.86. Let $f : \mathbb{R} \to \mathbb{R}$ be defined by the function $f(x) = \frac{2}{\pi} \arctan x$. Here, f is a homeomorphism onto its image which is $(0, 1)$. Viewed as a dynamical system, it is easy to see from the picture that iterating the map quickly leads to the conclusion that f is a contraction. What is the Lipschitz constant in this case? (hint: use the derivative!).

EXERCISE 90. In fact, every open interval $I = (a, b) \in \mathbb{R}$, for $a < b$, is homeomorphic to every other open interval in $\mathbb{R}$: Show that the map $f : I \to \mathbb{R}$, $f(x) = \frac{x - \frac{a+b}{2}}{(x-a)(b-x)}$ is a homeomorphism. Then show that, for $a = -1$, and $b = 1$, f is a contraction with unique fixed point $x = 0$.

Hence, $\mathbb{R}$ is a self-similar set. What is more interesting, though, are self-similar sets defined via some finite subdivision rules, like the Cantor set. Repeating the rule on smaller and smaller regions inside some original set creates the self-similarity. Then the contraction used to verify self-similarity is simply the map taking one stage in the recursive construction to a future stage. Iterating the map uncovers the recursive structure on finer and finer scales, and ultimately all orbits converge to a single point in the set. Here are a few more examples:

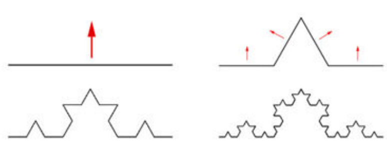


FIGURE 19. The first three iterates of the Koch Curve construction.

2.7.1. The Koch Curve: Swedish mathematician Helge von Koch described a special planar curve in a 1904 paper "On a continuous curve without tangents, constructible from elementary geometry" (original French title: Sur une courbe continue sans tangente, obtenue par une construction gomtrique lmentaire). The curve, which is continuous everywhere, but differentiable nowhere, was an oddity at the time and now we see it as the first fractal description in mathematics. To construct it, let $K_0 = [0, 1]$ be the unit interval, and again think of K_0 being divided into three equal parts. Then construct K_1 by removing the middle third of K_0, and adding in an equilateral triangle minus the base.

Then K_1 looks like four connected equal line segments with a peaked-point in the middle, as in Figure 19. Then, on each of these four line segments of K_1, remove the middle third and, again, add in the upper two sides of an equilateral triangle, to create K_2. Repeat this procedure for all $n \in \mathbb{N}$ to create $K = \lim_{n\to\infty} K_n$. The result is in Figure 20. Can you see the contraction which makes the curve self similar?

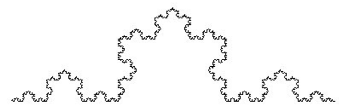


FIGURE 20. The Koch Curve K.

2.7.2. Sierpinski Gasket: and Carpet, and the famous Mandelbrot Set. We will look at these in turn.

CHAPTER 3

Linear Maps of $\mathbb{R}^2$

3.1. Linear, First-order ODE Systems in the Plane

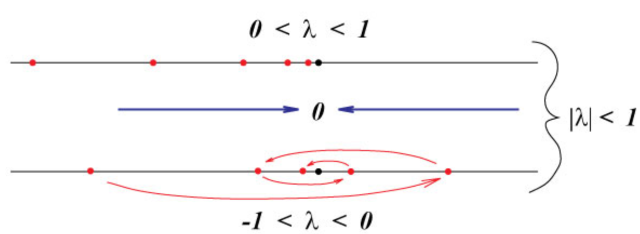
In this section, review all of the linear theory and do a full classification of linear systems, including a good drawing of the parameter space involving the trace and determinant of the coefficient matrix A of the system x' = Ax. All of this is stated in dynamics terms, with a focus on the qualitative nature of solution behavior, stability and classification of the equilibrium at the origin. When not isolated, when 0 is an eigenvalue, one can solve for the equilibria set via the linear space which contains the eigenvectors of the null eigenvalue. But also, one can solve for the zeros of the vector field directly. We culminate this part with expressions of the fundamental matrix, both as a generator of all solutions and as a single solution to the matrix form of the ODE. With some linear algebra, we discuss the relationship between the eigenvectors as generators of a coordinate system in the phase plane vis a vis the standard rectilinear coordinate system which governs the choice of variables x and y. Rectifying these two yields a new fundamental matrix, which is what we call the exponential matrix. We establish what this matrix is, how it functions, how it is constructed and what are its properties. The example already given will do well here.

3.2. Local Linearization and Equilibria Stability

Here, do the full local-linearization treatment leading up to the Poincare-Lyapunov Theorem and the Hartman-Grobman Theorem. DO many examples, and also use the Competing Species bifurcation analysis example from ODE class.

3.3. Linear Planar Maps

Recall for the moment the linear map of $\mathbb{R}$ defined by $f(x) = \lambda x$ (this can also be written $x \overset{f}{\longmapsto} \lambda x$). One can classify dynamical behavior of this map by the magnitude of λ, neglecting the reflection of the real line given by $f(x) = -x$. This is because it is the magnitude only that determines the dynamical structure of the linear map: Hence we classify by whether $|\lambda| < 1$, $|\lambda| = 1$, or $|\lambda| > 1$. Respectively, the origin is a sink and asymptotically stable, all points are either fixed ($\lambda = 1$) or at most of order 2 ($\lambda = -1$) and all are



stable but not asymptotically stable, or the origin is a source. However, the actual orbit structure is also affected by the sign of λ. Indeed, when $\lambda < 0$, the sequence becomes an alternating sequence and successive terms of the orbit will flip in sign (See the Figure). However, dynamically, the orbits will converge or diverge regardless of the sign of λ. To see this, create a new dynamical system using the square of the map $f^2(x) = \lambda^2 x$. Then the new factor $\lambda^2 > 0$ always. The orbits of f^2 will stay on the side of the origin they started on. For f, however, the orbits will flip back and forth from one side of the origin to the other (see the figure). Hold this thought as we move on to linear maps of $\mathbb{R}^2$.

Now let's move up to linear maps of the plane and see if we can classify dynamical behavior in a similar fashion, by creating a small set of types. So let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear map. Then for $\mathbf{v} = \begin{bmatrix} x \\ y \end{bmatrix}$, we have $\mathbf{f}(\mathbf{v}) = A\mathbf{v}$, or

$$\begin{bmatrix} x \\ y \end{bmatrix} \overset{f}{\longmapsto} A \begin{bmatrix} x \\ y \end{bmatrix}, \qquad \text{where } A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

- Here, $\mathbf{v}$ is an eigenvector of A, with eigenvalue λ if $\mathbf{v}$ satisfies the vector equation $A\mathbf{v} = \lambda\mathbf{v}$, or equivalently $(A - \lambda I)\,\mathbf{v} = \mathbf{0}$, where I is the 2-dimensional identity matrix.
- Recall the characteristic equation of A: $\det(A - \lambda I) = 0$, which can also be written

$$\lambda^2 - (\operatorname{tr} A)\,\lambda + \det A = 0.$$

  The solution to this equation (the roots to the polynomial $\det(A - \lambda I)$) are the eigenvalues of A.

A good question to ask is: What information is conveyed by $\mathbf{v}$ and λ about the discrete dynamical system formed by iterating $\mathbf{f}$ on $\mathbb{R}^2$?

There is an easy classification of matrix types for A, and the classes are determined by the data of A:

I. Two real distinct eigenvalues $\lambda \neq \mu$.

II. One real eigenvalue λ. In this case, there are two possibilities:
   - $A = \lambda I$. Here A is called a *homothety* or a scaling.
   - A is conjugate to $\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$. When $\lambda = 1$, this map is often called a *shear*.

III. Two complex conjugate eigenvalues $\lambda = a + ib = \rho e^{i\theta}$, and $\mu = a - ib = \rho e^{-i\theta}$, where $\rho^2 = a^2 + b^2$, and $\tan\theta = \frac{b}{a}$. Here, the effect of A is by rotation by θ and a scaling by ρ. When $\rho = 1$, the effect is a pure rotation (see below).

Note here that in the case of a shear above, the fact that the upper right-hand entry is non-zero is vital. But the value of 1 is not:

EXERCISE 91. Show if a 2×2 matrix A is conjugate to $\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$, it is conjugate to $\begin{bmatrix} \lambda & s \\ 0 & \lambda \end{bmatrix}$ for any non-zero $s \in \mathbb{R}$.

In actuality, there is a subclassification of these matrices which is of special interest to dynamicists. First, we identify two factors:

- Every 2×2 matrix A is a scaled version of a determinant-(± 1) matrix, since $\frac{1}{\sqrt{|\det A|}} A = C$, where $|\det C| = 1$. Here, then, we can view a transformation of the plane corresponding to the matrix A as a composition of such a matrix C with a pure scaling matrix $S = \begin{bmatrix} s & 0 \\ 0 & s \end{bmatrix}$ so that $A = \begin{bmatrix} \sqrt{|\det A|} & 0 \\ 0 & \sqrt{|\det A|} \end{bmatrix} C$.
- Every determinant-(-1) matrix C is a composition of a determinant 1 matrix B with the reflection $R = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, so that $B = RC$ and $\det B = 1$.

It turns out that 2×2 matrices of determinant 1 play a special role in dynamical systems. Hence their types have specific names. A 2×2, determinant 1 matrix A can be either:

I. **Hyperbolic:** A has two real distinct eigenvalues λ and μ where $\lambda = \frac{1}{\mu}$ (necessarily $|\lambda| > 1 > |\mu|$). Thus A is diagonalizable over the real numbers and $A \cong \begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix}$.

II. **Parabolic:** A has one real eigenvalue $\lambda = 1$ but is not diagonalizable (the 1-eigenspace is only 1-dimensional.) In this case, A is conjugate to $\begin{bmatrix} 1 & s \\ 0 & 1 \end{bmatrix}$ for some non-zero $s \in \mathbb{R}$.

III. **Elliptic:** A has two, unit-modulus complex conjugate eigenvalues $\lambda = e^{i\theta}$, and $\mu = e^{-i\theta}$. Here A is again not diagonalizable (over the real numbers) and A is only conjugate to $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$, for $a, b \in \mathbb{R}$, $a^2 + b^2 = 1$.

Some notes here:

- Both the Identity Matrix $I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and its negative $\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$ show up in this classification. Do you see where?
- Compose a hyperbolic matrix with the reflection R from above and you still get something that looks hyperbolic, but the two eigenvalues only satisfy $\lambda = -\frac{1}{\mu}$. But compose a parabolic matrix with R and you get a hyperbolic one. And when you compose an elliptic matrix with R, you get another elliptic matrix. But the angle of rotation has changed. Think about this.

Exercise 92. Show that for a parabolic matric P, the composition PR is hyperbolic.

Exercise 93. Show that for an elliptic matrix E, the composition ER is elliptic also. Compute the rotation angle of ER in terms of the rotation angle of E.

Exercise 94. Let A be a hyperbolic 2×2 matrix. Show $|\operatorname{tr} A| > 2$.

Exercise 95. Find the elliptic matrix which rotates the plane through an angle of $\frac{\pi}{6}$ radians.

Geometrically, choose a representative from each case above of determinant-1. Then it will be easier to see the effect on points in the plane. Figure 1 is useful here, noting the following: In the left-hand case, $A = \begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix}$, for $\lambda > 1 > \mu > 0$; in the middle case, $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$; and is the right, $A = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$, for eigenvalues $\lambda = a \pm ib$, and $a^2 + b^2 = 1$.

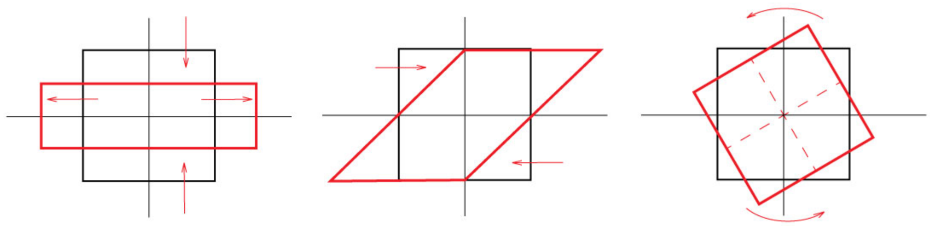


FIGURE 1. $f(\mathbf{x}) = A\mathbf{x}$ for A hyperbolic, parabolic and elliptic, respectively.

We will presently embark on a detailed study of linear maps of the plane corresponding to these basic types. However, let's begin by returning to a previous notion to dispel an incorrect but seemingly intuitive belief about linear planar maps.

DEFINITION 3.1. The *spectral radius* of a matrix A is the quantity ρ_A, where

$$\rho_A = \left\{ \max_i |\lambda| \;\middle|\; \lambda \text{ is an eigenvalue of } A \right\}.$$

Here, ρ_A is related to the matrix norm $||A|| = \max_{||\mathbf{v}||=1} ||A\mathbf{v}||$, but they are not equal in general (they are equal when the matrix A is symmetric, though.). The fact that the matrix norm is not equal to the modulus of its largest eigenvector is a clue to the following:

PROPOSITION 3.2. *A linear transformation A of $\mathbb{R}^n$ is eventually contracting if $\rho_A < 1$.*

EXERCISE 96. Construct an explicit example of a linear map on $\mathbb{R}^2$ which is eventually contracting but NOT a contraction. (Hint: Pay attention to the types in the classification of matrices above.)

EXERCISE 97. Show that no determinant-1 linear map on $\mathbb{R}^2$ can be eventually contracting. (Note: This has enormous implications in the mathematical models of physics and engineering, as it restricts the stability classification of both fixed points of maps and the equilibrium and periodic solutions of ODE systems.)

Some things to consider:

- Any positive determinant linear map on $\mathbb{R}^2$ can be written as a linear combination of the three types in the figure (sum of scalar multiples of the three, that is). And any negative determinant linear map will have a square which is positive, like in linear maps of $\mathbb{R}$, the dynamics (up to flipping across some line) will be similar in nature. Hence a detailed study of these three determinant-1 types and their scaled versions is necessary to explore the dynamical structure of linear maps of the plane.

- Diagonalizing a matrix (conjugating it, as best as one can, to one where the eigenvalues are prominent) is really just a linear coordinate change. Hence this process can be viewed simply as a change of the metric on $\mathbb{R}^2$. However, the new metric is always uniformly equivalent to the old one (you should show this!) It turns out that the process of diagonalization **does not** change the dynamical structure of the system!

EXERCISE 98. Show that, for a non-degenerate 2×2-matrix A, the metric $d_A(\mathbf{x}, \mathbf{y}) = d(A\mathbf{x}, A\mathbf{y})$ is uniformly equivalent to the Euclidean metric d on $\mathbb{R}^2$.

Hence, like maps on closed intervals, where we only needed to study maps on the unit interval, we can always limit our analysis to certain types of linear maps to capture all of the possible dynamical behavior of the entire family of linear planar maps. With this in mind, we begin our survey.

Let $\mathbf{v} \in \mathbb{R}^2$ and consider $\mathcal{O}_{\mathbf{v}}$ under the linear map $f(\mathbf{v}) = A\mathbf{v}$. On iteration,

$$\mathbf{v} \longmapsto A\mathbf{v} \longmapsto A(A\mathbf{v}) = A^2\mathbf{v} \longmapsto \cdots \longmapsto A^n\mathbf{v} \longmapsto \cdots.$$

Hence, the orbit of $\mathbf{v}$ will depend critically on the data associated to A.

3.3.1. Sinks & sources. Suppose that the two eigenvalues of A are real and distinct, so that $\lambda \neq \mu$. Then there exists a matrix B, where $A \overset{\text{conj}}{\cong} B = \begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix}$. Suppose further that $0 < |\lambda| < |\mu| < 1$ (A is nondegenerate. We will treat the 0-eigenvalue case separately). Then, by the previous proposition, the origin is a sink and all orbits tend to $\mathbf{0}$. That is, $\forall \mathbf{v} \in \mathbb{R}^2$, $\mathcal{O}_{\mathbf{v}} \longrightarrow \mathbf{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. A deeper question is, however, how the orbit evolves as it tends to the origin.

For this, let's restrict the case further to the case where both eigenvalues are positive, so $0 < \lambda < \mu < 1$. Then, for $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$, the nth term in the orbit sequence is $B^n\mathbf{v} = \begin{bmatrix} \lambda^n & 0 \\ 0 & \mu^n \end{bmatrix}\mathbf{v} = \begin{bmatrix} \lambda^n v_1 \\ \mu^n v_2 \end{bmatrix}$. A typical orbit would live entirely within one quadrant of the plane, like the black dots in Figure 2. Changing the sign of λ and/or μ without changing the magnitude would create orbits like the red dots in Figure 2. Can you work out the signs of the two eigenvalues to create the orbit in the figure? Some observations:

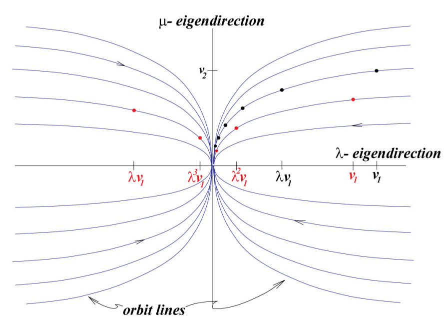


FIGURE 2. Phase Portrait for $f(\mathbf{v}) = B\mathbf{v}$, where $0 < \lambda < \mu < 1$.

- Since we have diagonalized the matrix A to get B, the nth term is easy to calculate as we have uncoupled the coordinates.
- The smaller eigenvalue λ means the first coordinate sequence in the orbit will have a faster decay toward 0 than that of the second coordinate. In the plane, this will imply a curved path toward the origin, with the

orbit line (recall the definition of the orbit line?) bending toward the μ-eigendirection (why is this the case? Think about which coordinate is decaying faster.)

- Given the orbit lines in the figure, do you recognize the phase portrait? From the classification of 2-dimensional first-order, homogeneous linear systems (with constant coefficents), systems with this phase portrait have a node at the origin which is asymptotically stable. This is a sink.
- In the figure, the first few elements of $\mathcal{O}_{\mathbf{v}}$ are plotted in black.
- How does the phase portrait change if one or both of the eigenvalues of B are negative? As a hint, the orbit lines do not change at all. But the orbits, themselves? In red in the figure are the first few elements of $\mathcal{O}_{\mathbf{v}}$ in the case that $0 < -\lambda < \mu < 1$. Do you see the effect?

We can actually calculate the equations for the orbit lines: Let $x = v_1$ and $y = v_2$. Then the orbit lines satisfy the equation

$$|y| = C|x|^{\alpha}, \quad \text{where} \quad \alpha = \frac{\log|\mu|}{\log|\lambda|}.$$

EXERCISE 99. Derive this last equation for the orbit lines.

EXERCISE 100. The map above $f(\mathbf{v}) = B\mathbf{v}$, (the original one above, where $0 < \lambda < \mu < 1$, that is) is the time-1 map of a first-order, linear homogeneous 2×2 system of ODEs. Find such a system and compare the matrix in the ODE system to B.

EXERCISE 101. Sketch the phase portrait in the case that $|\lambda| > |\mu| > 1$.

3.3.2. Star nodes. Suppose now that the linear map has a matrix with only 1 eigenvalue but 2 independent eigenvectors. That is, $A \overset{\text{conj}}{\cong} B = \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \lambda I_2$, a homothety. Here, it should be clear that for any starting vector $\mathbf{v} \in \mathbb{R}^2$, the nth orbit element is $B^n\mathbf{v} = \lambda^n\mathbf{v}$. In the case that $0 < |\lambda| < 1$, we have $\mathcal{O}_{\mathbf{v}} \longrightarrow \mathbf{0}$, for every $\mathbf{v} \in \mathbb{R}^2$. What does the motion look like in this case?

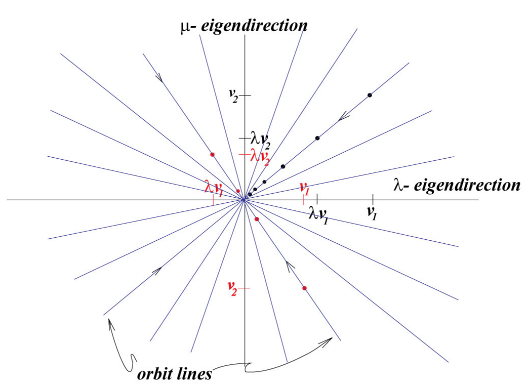


FIGURE 3. Star node phase portrait for $f(\mathbf{v}) = B\mathbf{v}$, where $0 < \lambda < 1$.

Since we are simply rescaling the initial vector $\mathbf{v}$, motion will be along the line through the origin given by $\mathbf{v}$. And the orbits will decay exponentially along these orbit lines. The phase portrait is that of a *star node* in this case, which is a sink, or implosion, when $|\lambda| < 1$, and a source, or explosion, when $|\lambda| > 1$. Again, think about what the orbits look like when $\lambda < 0$. Does anything change if the eigenvectors were $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, and $\mathbf{v}_2 = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$?

3.3.3. Degenerate Nodes. Suppose now that the linear map has a matrix with only 1 linearly independent eigenvector. In this case, $A \overset{\text{conj}}{\cong} B = \begin{bmatrix} \lambda & \lambda \\ 0 & \lambda \end{bmatrix}$ (remember that A is conjugate to a matrix with λ's on the main diagonal and ANY non-zero real number in the upper right-hand corner. We choose λ again here so that the calculations simplify a bit without losing detail.) Then $B^n = \begin{bmatrix} \lambda^n & n\lambda^n \\ 0 & \lambda^n \end{bmatrix} = \lambda^n \begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix}$, and

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \overset{B^n}{\longmapsto} \lambda^n \begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \lambda^n \begin{bmatrix} v_1 + nv_2 \\ v_2 \end{bmatrix}.$$

Here, the presence of the summand nv_2 has a twisting effect on $\mathcal{O}_{\mathbf{v}}$ even though the exponential factor λ^{n-1} still dominates the long-term orbit behavior.

One can see this twisting effect in the orbit lines. Indeed, for ease of argument, assume $\lambda > 0$ and combine the two equations defining points in the plane along the curve containing $x = v_1$ and $y = v_2$ at $n = 0$ through n: $x = \lambda^n (v_1 + nv_2)$, and $y = \lambda^n v_2$. Solving the second for n, we get $n = \frac{\ln y - \ln v_2}{\ln \lambda}$. With substitution into the first, we have

$$\begin{aligned} x &= \lambda^n (v_1 + nv_2) \\ &= \frac{y}{v_2}\left(v_1 + \left(\frac{\ln y - \ln v_2}{\ln \lambda}\right) v_2\right) \\ &= y\frac{v_1}{v_2} + \frac{y \ln y}{\ln \lambda} - \frac{y \ln v_2}{\ln \lambda} \\ &= y\left(\frac{v_1}{v_2} - \frac{\ln v_2}{\ln \lambda} + \frac{\ln y}{\ln \lambda}\right) = y\left(C + \frac{\ln y}{\ln \lambda}\right). \end{aligned}$$

Thus, if $0 < \lambda < 1$, then $\forall \mathbf{v}$, $\mathcal{O}_{\mathbf{v}} \longrightarrow \mathbf{0}$, but the orbit lines twist toward (but do NOT rotate around) the origin!

This phase portrait exhibits a *degenerate node*, picture in Figure 5. Notice a few things here: First, these equations for the invariant lines cannot express the solutions for any points whose starting coordinates include $v_2 = 0$. But the original equations would allow such a solution. Here $x = \lambda^n v_1$ and $y = 0$, for all $n \in \mathbb{N}$. These straight line solutions along the λ-eigendirection, are extraneous, in that they are hidden by an assumption in the method of solution (namely dividing by v_2 in solving for n.) Be careful not to neglect such solutions. Second, the single dimension eigenspace in this case is the only subspace comprising "linear motion" (compare the nodes above.) In solving for the general solution to the corresponding linear 2×2 system

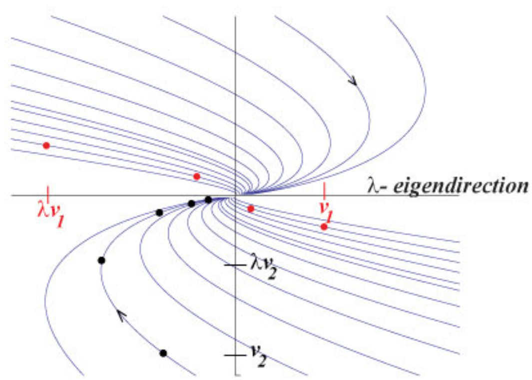


FIGURE 4. Degenerate node phase portrait for $f(\mathbf{v}) = B\mathbf{v}$, where $0 < \lambda < 1$.

of first order ODEs, one would employ a calculation involving a *generalized eigenvector* to construct the solution. And lastly, while it is easy to also understand the case where $\lambda > 1$, what changes in the calculation of the orbit curves for $\lambda < 0$?

EXERCISE 102. Recalculate the equations for the invariant curves of a degenerate node in the case that $\lambda < 0$.

3.3.4. Spirals and centers. Suppose that the linear map has two complex conjugate eigenvalues $\lambda = \rho e^{i\theta} = a + ib$, and $\mu = \rho e^{-i\theta} = a - ib$. Here in general, $\rho^2 = a^2 + b^2$. Then $A \cong B \cong \rho \begin{bmatrix} \frac{a}{\rho} & \frac{b}{\rho} \\ -\frac{b}{\rho} & \frac{a}{\rho} \end{bmatrix}$, where B is a constant ρ times a pure rotation. This scaling affects the rotational effect of the map. The orbit lines are

- spirals toward the origin if $0 < \rho < 1$,
- spirals away from the origin if $\rho > 1$, and
- Concentric circles if $\rho = 1$ (the eigenvalues then are purely imaginary).

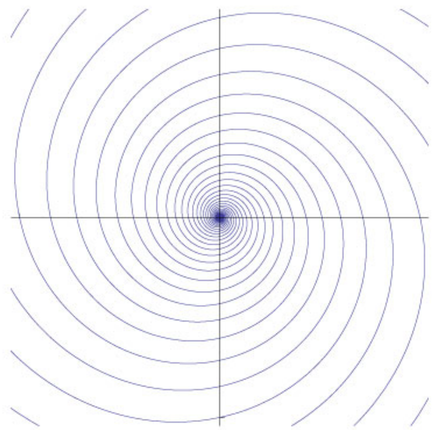


FIGURE 5. Spiral sink phase portrait for $f(\mathbf{v}) = B\mathbf{v}$, where $\lambda = a + ib$, and $|\lambda| < 1$.

EXERCISE 103. Write down an explicit expression for the orbit lines in this case.

3.3.5. Saddles. Now suppose A is a 2×2 matrix with eigenvalues $0 < |\mu| < 1 < |\lambda|$. Then $A \overset{\text{conj}}{\cong} B = \begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix}$ like the other examples in Case I, but the orbit lines are different. In fact, writing out the nth term in $\mathcal{O}_{\mathbf{v}}$ for a choice of $\mathbf{v} \in \mathbb{R}^2$, we see that there are four types:

(1) $\mathcal{O}^+_{\mathbf{v}} \longrightarrow \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ and $\mathcal{O}^-_{\mathbf{v}} \longrightarrow \infty$,
(2) $\mathcal{O}^+_{\mathbf{v}} \longrightarrow \infty$ and $\mathcal{O}^-_{\mathbf{v}} \longrightarrow \infty$,
(3) $\mathcal{O}^+_{\mathbf{v}} \longrightarrow \infty$ and $\mathcal{O}^-_{\mathbf{v}} \longrightarrow \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, and
(4) $\mathcal{O}^+_{\mathbf{v}} \longrightarrow \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ and $\mathcal{O}^-_{\mathbf{v}} \longrightarrow \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

With B as our matrix, the eigenvectors $\mathbf{v}_\lambda = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{v}_\mu = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ lie on the coordinate axes, and for a choice of $\mathbf{v} \in \mathbb{R}^2$, the nth term is again $B^n\mathbf{v} = \begin{bmatrix} \lambda^n v_1 \\ \mu^n v_2 \end{bmatrix}$. Can you envision the orbit lines and motion along them? Do you recognize the phase portrait? Can you classify the type and stability of the origin?

Again, place a picture here, some representative orbits, and discuss the equations of the orbit lines.

Consider now the hyperbolic matrix $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$. Here the characteristic equation is $r^2 - r - 1 = 0$, which is solved by $r = \frac{1 \pm \sqrt{5}}{2}$, giving us the eigenvalues

$$\lambda = \frac{1+\sqrt{5}}{2} > 1, \quad \text{and} \quad \mu = \frac{1-\sqrt{5}}{2} \in (-1, 0).$$

The eigenspace of λ is the line $y = \frac{1+\sqrt{5}}{2} x = \lambda x$, and for an eigenvector, we choose $\mathbf{v}_\lambda = \begin{bmatrix} 1 \\ \lambda \end{bmatrix}$.

Now let $f : \mathbb{R}^2 \to \mathbb{R}^2$, be the linear map $f(\mathbf{v}) = A\mathbf{v} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \mathbf{v}$. Then, for $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, we get

$$\mathcal{O} = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ 5 \end{bmatrix}, \begin{bmatrix} 5 \\ 8 \end{bmatrix}, \begin{bmatrix} 8 \\ 13 \end{bmatrix}, \ldots \right\}.$$

Do you see the patterns? Call $\mathbf{v}_n = \begin{bmatrix} x_n \\ y_n \end{bmatrix} = \begin{bmatrix} x_n \\ x_{n+1} \end{bmatrix}$ and the sequences $\{x_n\}_{n \in \mathbb{N}}$ and $\{y_n\}_{n \in \mathbb{N}}$ are Fibonacci with $y_n = x_{n+1}$. Notice that the sequence of ratios

$$\left\{ \frac{y_n}{x_n} \right\}_{n \in \mathbb{N}} = \left\{ \frac{x_{n+1}}{x_n} \right\}_{n \in \mathbb{N}}$$

has a limit, and $\lim_{n \to \infty} \frac{y_n}{x_n} = \frac{1+\sqrt{5}}{2} = \lambda$.

Recall how to find this limit: Use the second-order recursion inherent in the Fibonacci sequence, namely $a_{n+1} = a_n + a_{n-1}$, and the ratio to calculate a first-order recursion. This first-order recursion will correspond to a map, which one can study dynamically. Indeed, Let $r_{n+1} = \frac{x_{n+1}}{x_n}$, Then

$$r_{n+1} = \frac{x_{n+1}}{x_n} = \frac{x_n + x_{n-1}}{x_n} = 1 + \frac{1}{\frac{x_n}{x_{n-1}}} = 1 + \frac{1}{r_n}.$$

So $r_{n+1} = f(r_n)$, where $f(x) = 1 - \frac{1}{x}$. The only non-negative fixed point of this map is the sole solution to $x = f(x) = 1 - \frac{1}{x}$, or $x^2 - x - 1 = 0$, which is $x = \frac{1+\sqrt{5}}{2}$. Note that really there are two solutions and the other one is indeed μ. However, since we are talking about populations, the negative root doesn't really apply to the problem.

Example 3.3. Recall the Lemmings problem, with its second-order recursion $a_{n+1} = 2a_n + 2a_{n-1}$. Here the sequence of ratios of successive terms $\left\{ \frac{a_{n+1}}{a_n} \right\}_{n \in \mathbb{N}}$ has the limit $1 + \sqrt{3}$.

Here are two rhetorical questions:

(1) What is the meaning of these limits?
(2) How does the hyperbolic matrix in the above Fibonacci sequence example help in determining the limit?

To answer these, let's start with the sequence

$$\{b_n\} = \{1, 1, 2, 3, 5, 8, 13, 21, \ldots\}.$$

As before, we see

$$\mathbf{v}_{n+1} = \begin{bmatrix} b_{n+1} \\ b_{n+2} \end{bmatrix} = \begin{bmatrix} b_{n+1} \\ b_{n+1} + b_n \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} b_n \\ b_{n+1} \end{bmatrix} = A \begin{bmatrix} b_n \\ b_{n+1} \end{bmatrix} = A\mathbf{v}_n,$$

where $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$. This is precisely the matrix that (1) moves the second entry into the first entry slot, and (2) creates a new entry two by summing the two entries.

Here, we have associated to the second-order recursion $b_{n+2} = b_{n+1} + b_n$ the matrix $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$ and the first-order *vector* recursion $\mathbf{v}_{n+1} = A\mathbf{v}_n$.

Remark 3.4. This is basically a *reduction of order* technique, much like the manner with which one would reduce a second-order ODE into a system of 2 first-order ODES, written as a single vector ODE.

Remark 3.5. Note here that the second-order recursion is NOT a dynamical system, since one needs not only the previous state to determine the next state, but the previous two states. However, transformed into a first-order vector recursion, the new linear system is now a dynamical system.

This is actually used to construct a function which gives the nth term of a Fibonacci sequence in terms of n (rather than only in terms of the $(n-1)$st term):

Proposition 3.6. *Given the second order recursion $b_{n+2} = b_{n+1} + b_n$ with the initial data $b_0 = b_1 = 1$, we have*

$$b_n = \frac{\lambda^{n+1} - \mu^{n+1}}{\lambda - \mu},$$

where $\lambda = \frac{1+\sqrt{5}}{2}$ and $\mu = \frac{1-\sqrt{5}}{2}$.

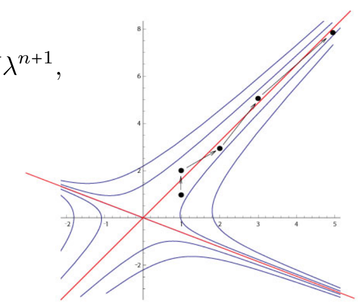
Prove this here.

We showed that λ and μ were the eigenvalues of a matrix $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$, and that the linear map on $\mathbb{R}^2$ given by A, $\mathbf{v}_{n+1} = A\mathbf{v}_n$, is in fact the first-order vector recursion for the second-order recursion in the proposition under the assignment $\mathbf{v}_n = \begin{bmatrix} b_n \\ b_{n+1} \end{bmatrix}$. This reduction-of-order technique for the study of recursions is quite similar to (and is the discrete version of) the technique of studying the solutions of a single, second-order, homogeneous, ODE with constant coefficients by instead studying the system of two first-order, linear, constant-coefficient, homogeneous ODEs. In fact, this analogy is much more robust, which we will see in a minute.

First, a couple of notes:

- For very large n,

$$b_n = \frac{\lambda^{n+1} - \mu^{n+1}}{\lambda - \mu} \sim K\lambda^{n+1},$$



since $0 < |\mu| < 1 < |\lambda|$. Thus the growth rate of terms in the Fibonacci sequence is not exponential. It does, however, tend to look more and more exponential as n gets large. In fact, we can say the Fibonacci sequence displays *asymptotic exponential* growth, or that the sequence grows *asymptotically exponentially.*

- Start with the initial data $\mathbf{v}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, and plot $\mathcal{O}_{\mathbf{v}_0}$ in the plane. What you will find is that the iterates of $\mathcal{O}^+_{\mathbf{v}}$ will live on two curves of motion (there will be flipping here across the $y = \lambda x$ eigenline. See the figure above.) Why does this happen? And $\mathcal{O}^+_{\mathbf{v}}$ will tend toward the λ-eigenline as they grow off of the page (see the figure below). Getting closer to the λ-eigenline means that the growth rate is getting closer to the growth rate ON the λ-eigenline. But on this line, growth is purely exponential!. With growth factor $\lambda > 1$.

  EXERCISE 104. If we neglect the application of a rabbit population, the discrete dynamical system we constructed above is invertible. Calculate the first few pre-images of the vector $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, and plot them on the figure below. Then calculate the orbit line equations for the orbit line on which the sequence lives. Hint: you may need to solve the original second-order ODE to do this.

- Every other point $\mathbf{v}_0 = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$ is really just another set of initial data for the second-order recursion (or the first-order vector version). Start taking iterates and plot and you will see that these orbits will also live on either one or will flip between two curves of motion and the phase diagram in the figure will tell you the ultimate fate of the orbits.

EXERCISE 105. For the Fibonacci vector recursion, find non-zero, explicit starting data for which the limit is NOT infinity (i.e., which lead to a sequence which does NOT run off of the page as n goes to infinity.)

EXERCISE 106. Show that for any non-zero, integer-valued initial values $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$, $\mathcal{O}_{\mathbf{v}} \longrightarrow \infty$ for the Fibonacci vector recursion.

Thus we can say the following:

PROPOSITION 3.7. *All populations governed by the second-order Fibonacci recursion experience asymptotic exponential growth limiting to a growth factor of* $\frac{1+\sqrt{5}}{2}$.

In general, let $a_{n+2} = pa_n + qa_{n+1}$ (careful of the order of the terms in this expression). Then we can construct a first-order vector recursion

$$\mathbf{v}_{n+1} = \begin{bmatrix} a_{n+1} \\ a_{n+2} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ p & q \end{bmatrix} \begin{bmatrix} a_n \\ a_{n+1} \end{bmatrix} = A\mathbf{v}_n, \text{ for } A = \begin{bmatrix} 0 & 1 \\ p & q \end{bmatrix}.$$

The characteristic equation of A is $r^2 - qr - p = 0$, with solutions $r = \frac{q \pm \sqrt{q^2+4p}}{2}$.

PROPOSITION 3.8. *If* $\begin{bmatrix} 0 & 1 \\ p & q \end{bmatrix}$ *has two distinct eigenvalues* $\lambda \neq \mu$, *then every solution to the second-order recursion* $a_{n+2} = pa_n + qa_{n+1}$ *is of the form*

$$a_n = x\lambda^n + y\mu^n$$

where $x = \alpha v_1$ *and* $y = \beta w_1$, $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ *and* $\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$ *are respective eigenvectors of* λ *and* μ, *and* α *and* β *satisfy the vector equation*

$$\begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \alpha\mathbf{v} + \beta\mathbf{w}.$$

REMARK 3.9. Hence the general second-order recursion and the first-order vector recursion carry the same information, and the latter provides all of the information necessary to completely understand the former. The method of solution is quickly discernable: Given a second-order recursion, calculate the data from the matrix A in the corresponding first-order vector recursion, including the eigenvalues and a pair of respective eigenvectors. Use this matrix data along with the initial data given with the original recursion to calculate the parameters in the functional expression for a_n.

Here is an example going back to our Fibonacci Rabbits Problem. Is essence, we use Proposition 3.8 to essentially prove Proposition 3.6.

EXAMPLE 3.10. Go back to the original Fibonacci recursion $a_{n+2} = a_{n+1} + a_n$, with initial data $a_0 = a_1 = 1$. The matrix $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$ has $\lambda = \frac{1+\sqrt{5}}{2}$ and $\mu = \frac{1-\sqrt{5}}{2}$ (as before) and using the notation of Proposition 3.1.13, one can calculate representative eigenvectors as $\mathbf{v} = \begin{bmatrix} 1 \\ \lambda \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} 1 \\ \mu \end{bmatrix}$. Thus $v_1 = w_1 = 1$. To calculate α and β, we have to solve the vector equation

$$\begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \alpha\mathbf{v} + \beta\mathbf{w}, \quad \text{or} \quad \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \alpha \begin{bmatrix} 1 \\ \lambda \end{bmatrix} + \beta \begin{bmatrix} 1 \\ \mu \end{bmatrix}.$$

This is solved by $\alpha = \frac{1-\mu}{\lambda-\mu}$ and $\beta = \frac{\lambda-1}{\lambda-\mu}$.

EXERCISE 107. Verify this calculation.

Hence we have $x = \alpha v_1 = \frac{1-\mu}{\lambda-\mu}$ and $y = \beta w_1 = \frac{\lambda-1}{\lambda-\mu}$, and our formula for the nth term of the sequence is

$$a_n = \frac{(1-\mu)\lambda^n + (\lambda-1)\mu^n}{\lambda-\mu}.$$

This does not look like the form in Proposition 3.6, however. But consider that the term

$$(1-\mu) = \frac{2}{2} - \frac{1-\sqrt{5}}{2} = \frac{1+\sqrt{5}}{2} = \lambda,$$

and similarly $(\lambda - 1) = -\mu$, we wind up with

$$a_n = \frac{(1-\mu)\lambda^n + (\lambda-1)\mu^n}{\lambda-\mu} = \frac{\lambda\cdot\lambda^n + (-\mu)\cdot\mu^n}{\lambda-\mu} = \frac{\lambda^{n+1} - \mu^{n+1}}{\lambda-\mu},$$

and we recover Proposition 3.6 precisely.

EXERCISE 108. Perform this calculation for the second-order recursion in the Lemmings Problem, and use it to calculate the population of lemmings today, given that the initial population was given in 1980.

And lastly, at least for this section, we have a way to ensure when a second order recursion with non-zero initial data exhibits asymptotically exponential growth; namely, when the phase portrait of the associated first-order vector recursion has a saddle point at the origin.

PROPOSITION 3.11. *Let $a_{n+2} = pa_n + qa_{n+1}$ be a second-order recursion, where $p, q \in \mathbb{N}$ and $0 < p \leq q$. Then all populations with non-zero initial conditions exhibit asymptotically exponential growth with asymptotic growth factor given by the spectral radius of* $A = \begin{bmatrix} 0 & 1 \\ p & q \end{bmatrix}$.

PROOF. By the discussion above, the only thing to ensure is that the two eigenvalues of A are $\lambda = \frac{q+\sqrt{q^2+4p}}{2} > 1$ and $-1 < \mu = \frac{q-\sqrt{q^2+4p}}{2} < 0$ and $\lambda \notin \mathbb{Q}$. It is obvious that $\lambda > 1$ since $q \geq 1$. And since $\det A < 0$, $\mu < 0$. So we claim $\mu \geq -1$. Since $q \geq p > 0$, we have $q + 1 > p$ implies $4q + 4 > 4p$ which implies $q^2 + 4q + 4 = (q + 2)^2 > q^2 + 4p$ which implies $q + 2 > \sqrt{q^2 + 4p}$ which implies $-2 < q - \sqrt{q^2 + 4p}$ which implies $-1 < \mu$. □

3.4. The Matrix Exponential

These calculations lead to a very important discussion on the relationship between the matrices found in first-order, 2-dimensional homogeneous linear systems (with constant coefficients) of ODEs and the corresponding matrices of the discrete, time-1 maps of those systems. The central questions is: Why is it that for a ODE system with coefficient matrix A, the **sign** of the eigenvalues determines the stability of the equilibrium solution at the origin. But for a linear map of $\mathbb{R}^n$, it is the size of the **absolute values** of the eigenvalues that determine the stability of the fixed point at the origin. The matrix of the time-1, ODE system is NOT the same matrix as the coefficient matrix of the system. The two matrices are certainly related, but they are not identical. Furthermore, ANY ODE system has a time-1 map. But only certain types of linear maps correspond to the time-1 maps of ODE systems. To understand better why, let's start with an example:

EXAMPLE 3.12. Calculate the time-1 map of the ODE system

$$\dot{\mathbf{x}} = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix} \mathbf{x}, \quad \mathbf{x}(0) = \mathbf{x}^0 = \begin{bmatrix} x_1^0 \\ x_2^0 \end{bmatrix}.$$

This system is uncoupled and straightforward to solve. Using linear system theory, the eigenvalues of the matrix $A = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}$ are $\lambda = 2$ and $\mu = -1$, and, since A is diagonal, we can choose the vectors $\mathbf{v}_\lambda = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{v}_\mu = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Hence the general solution is

$$\mathbf{x}(t) = c_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} e^{2t} + c_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} e^{-t},$$

or $x_1(t) = c_1 e^{2t}$ and $x_2(t) = c_2 e^{-t}$. For the choice of any initial data, the particular solution is $x_1(t) = x_1^0 e^{2t}$ and $x_2(t) = x_2^0 e^{-t}$, and the evolution of this continuous dynamical system is

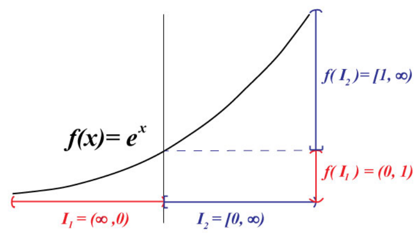
$$\varphi(\mathbf{x}, t) = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} e^{2t} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} e^{-t} = \begin{bmatrix} e^{2t} & 0 \\ 0 & e^{-t} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} e^{2t} & 0 \\ 0 & e^{-t} \end{bmatrix} \mathbf{x}.$$

The time-1 map is then $\varphi(\mathbf{x},1) = \varphi_1(\mathbf{x}) : \mathbf{x}(0) \longmapsto \mathbf{x}(1)$, or the linear map

$$\varphi_1 : \mathbb{R}^2 \to \mathbb{R}^2, \quad \varphi_1(\mathbf{x}) = B\mathbf{x},$$

where $B = \begin{bmatrix} e^2 & 0 \\ 0 & e^{-1} \end{bmatrix}$ is the matrix associated to the linear map.

Do you see the relationship between the ODE matrix A and the time-1 linear map matrix B? The type and stability of the equilibrium solution at the origin of this linear system given by A is that of a saddle, and unstable. The time-1 map must also be a saddle, as the orbit lines of the time-1 map coincide precisely to the solution curves of the ODE system. It is the sign of the eigenvalues (non-zero entries of A in this case) that determine the type and stability of the origin of the ODE system. However, it is the "size" (modulus) of the eigenvalues of B which determine the type and stability of the fixed point at the origin in the linear map given by B. Some notes:



- Notice that the *exponential map*, $\exp : x \mapsto e^x$ takes $\mathbb{R}$ to $\mathbb{R}^+$ (see the figure above) and maps all non-negative numbers, $\mathbb{R}_0^+ = \{0\} \cup \mathbb{R}^+$, to the interval $[1,\infty)$ and all negative numbers to $(0,1)$. This is no accident, and exposes a much deeper meaning of the exponential map.
- One might conclude that there could not be a time-1 map of a linear, constant coefficient, homogeneous ODE system with negative eigenvalues. And you would be correct in this hyperbolic case. In general?
- One might also conclude that for any 2×2-matrix A, the associated time-1 map B would simply be the exponentials of each of the entries of A. Here, you must definitely be much more careful, as we shall see.

EXERCISE 109. Let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear map $f(\mathbf{x}) = B\mathbf{x}$, where $B = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$ and both $a > 0$ and $b > 0$. Determine a linear, 2-dimensional ODE system that has f as its time-1 map. For $a > 0$, show that B cannot correspond to a time-1 map of an ODE system if $b \le 0$. Can B correspond to a time-1 map of an ODE system if both $a < 0$ and $b < 0$? Hint: The answer is yes.

For a moment, recall the 1-dimensional linear, homogeneous, constant coefficient ODE $\dot{x} = ax$, for $a \in \mathbb{R}$ a constant. The evolution is $x(t) = x_0 e^{at}$, the ODE is solved by an exponential function involving a. For the nth order linear, homogeneous, constant coefficient case, one creates an equivalent system $\dot{\mathbf{x}} = A\mathbf{x}$, a single vector ODE whose solution also seems exponential in nature (exponentials have the appeal that the derivative is proportional to the original function). That is, it is tempting to write the evolution as $\mathbf{x}(t) = \mathbf{x}^0 e^{At}$, since if it is the case that $\frac{d}{dt}\left[\mathbf{x}^0 e^{At}\right] = A\mathbf{x}^0 e^{At}$, then this expression solves the ODE. However, it is not yet clear what it means to take the exponential of a matrix.

DEFINITION 3.13. For an $n \times n$ matrix A, define $e^A = \displaystyle\sum_{n=0}^{\infty} \frac{A^n}{n!}$.

This definition obviously comes directly from the standard definition of the exponential e^x via its Maclauren Series. Also, it seems to make sense in that one can certainly sum matrices, take them to positive integer powers, and divide them by scalars. The question of whether this converges or not is unclear, though. It really is a question of whether each entry, written as a series will converge. While this is basically a calculus question, we will not elaborate here but will state without proof the following: The definition above is well-defined and the series converges absolutely for all $n \times n$ matrices A.

PROPOSITION 3.14. $\dfrac{d}{dt}\left[\mathbf{x}^0 e^{At}\right] = A\mathbf{x}^0 e^{At}$.

PROOF. Really, this is just the definition of a derivative:

$$\begin{aligned}
\frac{d}{dt}e^{At} &= \lim_{h\to 0}\frac{e^{A(t+h)} - e^{At}}{h} == \lim_{h\to 0}\frac{e^{At}e^{Ah} - e^{At}}{h} = e^{At}\lim_{h\to 0}\frac{e^{Ah} - 1}{h} \\
&= e^{At}\lim_{h\to 0}\frac{1}{h}\left(\sum_{n=0}^{\infty}\frac{(Ah)^n}{n!} - I\right) = e^{At}\lim_{h\to 0}\frac{1}{h}\sum_{n=1}^{\infty}\frac{(Ah)^n}{n!} \\
&= e^{At}\lim_{h\to 0}\sum_{n=1}^{\infty}\frac{A^n h^{n-1}}{n!} = e^{At}\lim_{h\to 0}A\sum_{n=1}^{\infty}\frac{A^{n-1}h^{n-1}}{n!} \\
&= Ae^{At}\lim_{h\to 0}\left(\frac{I}{1!} + \frac{Ah}{2!} + \frac{A^2h^2}{3!} + \frac{A^3h^3}{4!} + \dots\right).
\end{aligned}$$

At this point, every term in the remaining series has an h in it except for the $n = 1$ term, which is I. So

$$\frac{d}{dt}e^{At} = e^{At}\lim_{h\to 0}A\sum_{n=1}^{\infty}\frac{A^{n-1}h^{n-1}}{n!} = Ae^{At}.$$

□

Hence the expression e^{At} behaves a lot like the exponential of a scalar and in fact does solve the vector ODE $\dot{\mathbf{x}} = A\mathbf{x}$, with initial condition $\mathbf{x}(0) = \mathbf{x}^0$. However, contrary to Example 3.12, it is not in general true that the exponential of a matrix is simply the matrix of exponentials of the entries.

EXAMPLE 3.15. Find the evolution for $\dot{\mathbf{x}} = \begin{bmatrix} 4 & -2 \\ 3 & -3 \end{bmatrix}\mathbf{x}$.

Here, the characteristic equation is $r^2 - r - 6 = 0$, with solutions giving eigenvalues of $\lambda = 3$ and $\mu = -2$. Calculating eigenvectors, we choose $\mathbf{v}_\lambda = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $\mathbf{v}_\mu = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$. Thus the general solution is

$$\mathbf{x}(t) = c_1\begin{bmatrix} 2 \\ 1 \end{bmatrix}e^{3t} + c_2\begin{bmatrix} 1 \\ 3 \end{bmatrix}e^{-2t} = \begin{bmatrix} 2e^{3t} & e^{-2t} \\ e^{3t} & 3e^{-2t} \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}. \tag{3.4.1}$$

Writing this in terms of $\mathbf{x}^0$ (in essence, finding the evolution), we get the linear system

$$\begin{bmatrix} x_1^0 \\ x_2^0 \end{bmatrix} = \begin{bmatrix} 2c_1 + c_2 \\ c_1 + 3c_2 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}.$$

Solving for c_1 and c_2 in terms of the initial conditions involves inverting the matrix, and $\begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}^{-1} = \frac{1}{5}\begin{bmatrix} 3 & -1 \\ -1 & 2 \end{bmatrix}$. Hence the evolution is

$$\begin{aligned}
\mathbf{x}(t) &= \begin{bmatrix} 2e^{3t} & e^{-2t} \\ e^{3t} & 3e^{-2t} \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \\
&= \begin{bmatrix} 2e^{3t} & e^{-2t} \\ e^{3t} & 3e^{-2t} \end{bmatrix}\begin{bmatrix} \frac{3}{5} & -\frac{1}{5} \\ -\frac{1}{5} & \frac{2}{5} \end{bmatrix}\begin{bmatrix} x_1^0 \\ x_2^0 \end{bmatrix} \\
&= \begin{bmatrix} \frac{6}{5}e^{3t} - \frac{1}{5}e^{-2t} & -\frac{2}{5}e^{3t} + \frac{2}{5}e^{-2t} \\ \frac{3}{5}e^{3t} - \frac{3}{5}e^{-2t} & -\frac{1}{5}e^{3t} + \frac{6}{5}e^{-2t} \end{bmatrix}\begin{bmatrix} x_1^0 \\ x_2^0 \end{bmatrix}.
\end{aligned}$$

Hence we can also say now that

$$e^{At} = \begin{bmatrix} \frac{6}{5}e^{3t} - \frac{1}{5}e^{-2t} & -\frac{2}{5}e^{3t} + \frac{2}{5}e^{-2t} \\ \frac{3}{5}e^{3t} - \frac{3}{5}e^{-2t} & -\frac{1}{5}e^{3t} + \frac{6}{5}e^{-2t} \end{bmatrix}, \quad \text{for} \quad A = \begin{bmatrix} 4 & -2 \\ 3 & -3 \end{bmatrix}$$

and that the time-1 map of this ODE is the linear map given by

$$e^{A} = \begin{bmatrix} \frac{6}{5}e^{3} - \frac{1}{5}e^{-2} & -\frac{2}{5}e^{3} + \frac{2}{5}e^{-2} \\ \frac{3}{5}e^{3} - \frac{3}{5}e^{-2} & -\frac{1}{5}e^{3} + \frac{6}{5}e^{-2} \end{bmatrix}.$$

So how does one square these calculations into a general understanding of e^A? Via the properties of of a matrix exponential and a bit of standard linear algebra:

PROPOSITION 3.16. *Let $A_{n\times n}$ be diagonalizeable. Then $A = SBS^{-1}$, where*

- *$B_{n\times n}$ is diagonal, and*
- *the columns of $S_{n\times n}$ form an eigenbasis of A.*

PROPOSITION 3.17. *If $A_{n\times n}$ is diagonalizeable, then $e^A = Se^BS^{-1}$, where both B and e^B are diagonal.*

PROOF. Note that since

$$e^A = \sum_{n=1}^{\infty} \frac{A^n}{n!} \quad \text{and} \quad \left(SAS^{-1}\right)^n = SA^nS^{-1},$$

we have

$$Se^BS^{-1} = S\left(\sum_{n=1}^{\infty}\frac{B^n}{n!}\right)S^{-1} = \sum_{n=1}^{\infty}\frac{SB^nS^{-1}}{n!} = \sum_{n=1}^{\infty}\frac{\left(SBS^{-1}\right)^n}{n!} = \sum_{n=1}^{\infty}\frac{A^n}{n!} = e^A.$$

□

EXAMPLE 3.18. Back to the previous system, with $\dot{\mathbf{x}} = A\mathbf{x}$, and $A = \begin{bmatrix} 4 & -2 \\ 3 & -3 \end{bmatrix}$. The general solution, written in Equation 3.4.1 was

$$\begin{aligned}
\mathbf{x}(t) &= \begin{bmatrix} 2e^{3t} & e^{-2t} \\ e^{3t} & 3e^{-2t} \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \\
&= \begin{bmatrix} 2e^{3t} & e^{-2t} \\ e^{3t} & 3e^{-2t} \end{bmatrix}\begin{bmatrix} \frac{3}{5} & -\frac{1}{5} \\ -\frac{1}{5} & \frac{2}{5} \end{bmatrix}\begin{bmatrix} x_1^0 \\ x_2^0 \end{bmatrix} \\
&= \begin{bmatrix} \frac{6}{5}e^{3t} - \frac{1}{5}e^{-2t} & -\frac{2}{5}e^{3t} + \frac{2}{5}e^{-2t} \\ \frac{3}{5}e^{3t} - \frac{3}{5}e^{-2t} & -\frac{1}{5}e^{3t} + \frac{6}{5}e^{-2t} \end{bmatrix}\begin{bmatrix} x_1^0 \\ x_2^0 \end{bmatrix} = e^{At}\mathbf{x}^0.
\end{aligned}$$

But the middle equal sign in the last grouping can easily be written

$$\begin{aligned}
\mathbf{x}(t) &= \begin{bmatrix} 2e^{3t} & e^{-2t} \\ e^{3t} & 3e^{-2t} \end{bmatrix} \begin{bmatrix} \frac{3}{5} & -\frac{1}{5} \\ -\frac{1}{5} & \frac{2}{5} \end{bmatrix} \begin{bmatrix} x_1^0 \\ x_2^0 \end{bmatrix} \\
&= \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} e^{3t} & 0 \\ 0 & e^{-2t} \end{bmatrix} \begin{bmatrix} \frac{3}{5} & -\frac{1}{5} \\ -\frac{1}{5} & \frac{2}{5} \end{bmatrix} \begin{bmatrix} x_1^0 \\ x_2^0 \end{bmatrix} = Se^{Bt}S^{-1}\mathbf{x}^0,
\end{aligned}$$

where S is the matrix whose columns form an eigenbasis of A, and e^{Bt} is the exponential of the diagonal matrix B. Hence, as in the proposition, $e^{At} = Se^{Bt}S^{-1}$.

EXERCISE 110. Show that the time-1 map of the ODE system $\dot{\mathbf{x}} = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}\mathbf{x}$ is given by the linear map $f(\mathbf{x}) = B_1\mathbf{x}$, where $B_1 = \begin{bmatrix} e^{\lambda} & e^{\lambda} \\ 0 & e^{\lambda} \end{bmatrix}$, but the time-$t$ map in general is NOT given by the linear map $B_t = \begin{bmatrix} e^{\lambda t} & e^{\lambda t} \\ 0 & e^{\lambda t} \end{bmatrix}$.

EXERCISE 111. Find the time-1 map of the IVP $\dot{\mathbf{x}} = \begin{bmatrix} 0 & \alpha \\ -\alpha & 0 \end{bmatrix}\mathbf{x}$, and use it to construct a form for the exponential of a matrix with purely imaginary eigenvalues.

3.4.1. Application: Competing Species. A common biological model for understanding the possible interaction between 2 species in a closed environment that interact only in their competition for food. Not that one tends to eat the other. More like two herbivores both competing for limited food supplies. If two species did not interact at all, their respective population equations would fit the Logistic Model and be uncoupled:

$$\begin{aligned}
\dot{x} &= x(\alpha_1 - \beta_1 x) \\
\dot{y} &= y(\alpha_2 - \beta_2 y)
\end{aligned}$$

where $\alpha_1, \alpha_2, \beta_1, \beta_2 > 0$ are positive constants. We can model a simple interaction between these two species by adding in a cross-term, negative in sign (why?) and scaled by yet another parameter. We get:

$$\begin{aligned}
\dot{x} &= x(\alpha_1 - \beta_1 x - \gamma_1 y) \\
\dot{y} &= y(\alpha_2 - \beta_2 y - \gamma_2 x)
\end{aligned}$$

where now all $\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma_1, \gamma_2 > 0$ are again positive constants. What are the effects of these added terms? How may steady-state solutions (long-term behaviors where the populations of the species do not change over time)? Let's look at these models for a few parameter assignments to see. Do not worry so much about just how modelers came up with this idea of simply adding a term. Let's focus on the solutions for now.

Here are two sets of parameter assignments from Section 9.4:

(1) Let $\alpha_1 = \beta_1 = \gamma_1 = 1$, $\alpha_2 = .75$, $\beta_2 = 1$ and $\gamma_2 = .5$. The system is then

$$\begin{aligned}
\dot{x} &= x(1 - x - y) \\
\dot{y} &= y(.75 - y - .5x).
\end{aligned}$$

(2) Let $\alpha_1 = \beta_1 = \gamma_1 = 1$, $\alpha_2 = .5$, $\beta_2 = .25$ and $\gamma_2 = .75$. The system is now

$$\begin{aligned}\dot{x} &= x(1 - x - y)\\ \dot{y} &= y(.5 - .25y - .75x).\end{aligned}$$

Look at the slope fields for these two systems on the next two pages and try to identify the differences between these two.

Some questions:

Question 1. Where are the critical points in these systems?

(1) Here, we solve the system

$$\begin{aligned}0 &= x(1 - x - y)\\ 0 &= y(.75 - y - .5x).\end{aligned}$$

Of course the origin $\mathbf{a} = (0,0)$ is one solution. But so are $\mathbf{b} = (0, .75)$, $\mathbf{c} = (1, 0)$, and $\mathbf{d} = (.5, .5)$ (Verify that you know how to find these!).

(2) In this case, the system is

$$\begin{aligned}0 &= x(1 - x - y)\\ 0 &= y(.5 - .25y - .75x).\end{aligned}$$

Again, we have $\mathbf{e} = (0, 0)$. The others are $\mathbf{f} = (0, 2)$, $\mathbf{g} = (1, 0)$, and $\mathbf{h} = (.5, .5)$. See these?

Question 2. What are the type and stability of each of these equilibria?

- First note that, for any values of the parameters, the functions $F(x,y) = x(\alpha_1 - \beta_1 x - \gamma_1 y)$ and $G(x,y) = y(\alpha_2 - \beta_2 y - \gamma_2 x)$ are simply polynomials in x and y, and hence by the proposition we did in class, as long as the determinant of the matrix

$$A = \begin{bmatrix} \frac{\partial F}{\partial x}(x_0,y_0) & \frac{\partial F}{\partial y}(x_0,y_0) \\ \frac{\partial G}{\partial x}(x_0,y_0) & \frac{\partial G}{\partial y}(x_0,y_0) \end{bmatrix} = \begin{bmatrix} \alpha_1 - 2\beta_1 x_0 - \gamma_1 y_0 & -\gamma_1 x_0 \\ -\gamma_2 y_0 & \alpha_2 - 2\beta_2 y_0 - \gamma_2 x_0 \end{bmatrix},$$

where (x_0, y_0) is a fixed point, is not 0, the system is almost linear at (x_0, y_0).

(1) In our first case, we have

$$A = \begin{bmatrix} 1 - 2x_0 - y_0 & -x_0 \\ -.5y_0 & .75 - 2y_0 - .5x_0 \end{bmatrix},$$

and at the four critical points, we have

$$A_{\mathbf{a}} = \begin{bmatrix} 1 & 0 \\ 0 & .75 \end{bmatrix}, \quad A_{\mathbf{b}} = \begin{bmatrix} .25 & 0 \\ -.375 & -.75 \end{bmatrix}$$

$$A_{\mathbf{c}} = \begin{bmatrix} -1 & -1 \\ 0 & .25 \end{bmatrix}, \quad A_{\mathbf{d}} = \begin{bmatrix} -.5 & -.5 \\ -.25 & 0 \end{bmatrix}.$$

None of these have determinant 0, so the system is almost linear at all of these equilibria. The eigenvalues tell us that the corresponding linear systems have a source at $\mathbf{a}$, saddles at both $\mathbf{b}$ and $\mathbf{c}$, and a sink at $\mathbf{d}$. (Verify this!) By The Stability Theorem we did in class, the nonlinear equilibria will also have these types and their corresponding stability. The ONLY one of these that is stable is the asymptotically stable sink at $\mathbf{a}$.

(2) Contrast these fixed poitns with those at **e**, **f**, **g** and **h**. We play the same game, and we get the matrix

$$A = \begin{bmatrix} 1 - 2x_0 - y_0 & -x_0 \\ -.75y_0 & .5 - .5y_0 - .75x_0 \end{bmatrix}.$$

Thus we have the four linear systems given by

$$A_{\mathbf{e}} = \begin{bmatrix} 1 & 0 \\ 0 & .5 \end{bmatrix}, \quad A_{\mathbf{f}} = \begin{bmatrix} -1 & 0 \\ -1.5 & -.5 \end{bmatrix}$$

$$A_{\mathbf{g}} = \begin{bmatrix} -1 & -1 \\ 0 & -.25 \end{bmatrix}, \quad A_{\mathbf{d}} = \begin{bmatrix} -.5 & -.5 \\ -.325 & -.125 \end{bmatrix}.$$

Again, calculate the eigenvalues, and you should find that the linear systems again have an unstable node (a source) at the origin (**e**), sinks at **f** and **g**, and a saddle at **h**. All of these are such that the original nolinear equilibria share these characteristics.

Given all of this data, let's place the equilibria and think about how the stability of each would affect the nearby solutions. At the saddles, we would have to find the approximate directions of linear motion. The non-linear saddles will not have linear motion, but they will have something similar; a curve with very specific properties, namely that along one curve, all solution are asymptotic to the equilibrium in forward time. And there will be another curve where all solutions will be asymptotic to the equilibrium in backward time. All other nearby solutions eventually veer away from the equilibrium. The curves of forward and backward asymptotic solutions wind up being tangent at the equilibrium to the directions of linear travel form the linear saddle at that equilibrium. This give an idea of how the non-linear saddle is oriented. We have the two hand drawings below. Your homework now is to go onto JODE, or a similar graphing device, and actually compute slope fields and some numerical solutions to verify that this is more or less correct.

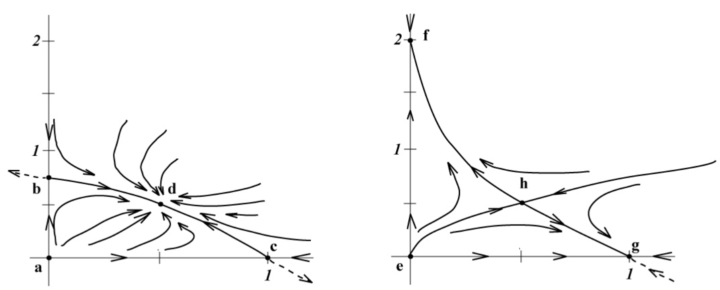


FIGURE 6. Competing Species phase portraits for $\delta = 0$ and $\delta = 1$.

Last question: Suppose we built a system that had sliders for each of the parameters so that we could watch how the phase portrait changed as we alter the parameters continuously. Notice in the two examples above that in both we had $\alpha_1 = \beta_1 = \beta_2 = \gamma_1 = 1$. But in the first example, we had $\alpha_2 = \frac{3}{4}$, $\beta_2 = 1$ and $\gamma_2 = \frac{1}{2}$, and in the second, $\alpha_2 = \frac{1}{2}$, $\beta_2 = \frac{1}{4}$ and $\gamma_2 = \frac{3}{4}$. Imagine if we smoothly slid the values of these parameters from their initial values to their final values, from one diagram to the second. One way to do this is with a single slider: δ, going from 0 to 1. We can use a linear parameterization from any vector $\mathbf{a} \in \mathbb{R}^3$ to any other vector $\mathbf{b} \in \mathbb{R}^3$ to "slide" these parameter values, via $\mathbf{x} = \mathbf{a} + \delta(\mathbf{b} - \mathbf{a})$. Then when $\delta = 0$, $\mathbf{x} = \mathbf{a}$ and when $\delta = 1$, $\mathbf{x} = \mathbf{b}$. Here, we do this with the three parameters above simultaneously:

$$\alpha_2 = \frac{3}{4} + \delta\left(\frac{1}{2} - \frac{3}{4}\right) = \frac{3}{4} - \frac{1}{4}\delta = \frac{3-\delta}{4}$$
$$\beta_2 = 1 + \delta\left(\frac{1}{4} - 1\right) = 1 - \frac{3}{4}\delta = \frac{4-3\delta}{4}$$
$$\gamma_2 = \frac{1}{2} + \delta(\frac{3}{4} - \frac{1}{2}) = \frac{1}{2} + \frac{1}{4}\delta = \frac{2+\delta}{4}.$$

Given this, we can continuously change the phase portrait and look for places where the number, type and/or stability of any of the equilibria change. Given the profound differences between the two phase portraits, we will find something for some intermediate value of rδ.

The Fixed Points, First in this analysis, understand that since we are changing the values of the parameters continuously, the fixed points (equilibria) will either stay where they are or move continuously also. So we can track them. We will do this by rewriting the vector field functions $F(x,y)$ and $G(x,y)$ in terms of δ instead of the parameters: Here, equilibria are the solutions to the equations

$$F(x,y) = 0 = x(1-x-y)$$
$$G(x,y) = y(\alpha_2 - \beta_2 y - \gamma_2 x)$$

become

$$F(x,y) = 0 = x(1-x-y)$$
$$G(x,y) = y\left(\frac{3-\delta}{4} - \frac{4-3\delta}{4}y - \frac{2+\delta}{4}x\right).$$

By inspection, we find:

(1) Whenever $y = 0$, $G(x,y) = 0$. Hence any equilibria along the x-axis will not depend on δ for position at all. Hence the equilibria at $(0,0)$ and $(1,0)$ do not move for $\delta \in [0,1]$.

(2) For the non-trivial equilibrium along the y-axis, where $x = 0$ but $y \neq 0$, $F(x,y) = 0$ but $G(x,y) = 0$ only when $\alpha_2 - \beta_2 y = 0$, so $y = \frac{\alpha_2}{\beta_2}$. IN terms of δ, there will be a critical point for the system when $x = 0$, and

$$y = \frac{\frac{3-\delta}{4}}{\frac{4-3\delta}{4}} = \frac{3-\delta}{4-3\delta}.$$

(3) Lastly, there seems to persist a critical point in the open first quadrant $x, y > 0$. This equilibrium will satisfy both

$$\left.\begin{array}{rcl} 1-x-y &=& 0 \\ \alpha_2 - \beta_2 y - \gamma_2 x &=& 0 \end{array}\right\} \Rightarrow \left\{\begin{array}{rcl} x+y &=& 1 \\ \gamma_2 x + \beta_2 y &=& \alpha_2 \end{array}\right. .$$

Combining these via $y = 1 - x$, we get

$$\gamma_2 x + \beta_2(1-x) = \alpha_2, \quad \text{or} \quad x = \frac{\alpha_2 - \beta_2}{\gamma_2 - \beta_2}.$$

In terms of δ, we get

$$x = \frac{\frac{3-\delta}{4} - \frac{4-3\delta}{4}}{\frac{2+\delta}{4} - \frac{4-3\delta}{4}} = \frac{\frac{-1+2\delta}{4}}{\frac{-2+4\delta}{4}} = \frac{1}{2}.$$

Thus $y = 1 - x = \frac{1}{2}$ and we see that the equilibrium strictly in the first quadrant does not move for $\delta \in [0,1]$ and is at $\left(\frac{1}{2}, \frac{1}{2}\right)$.

Thus the four critical points for this model, in terms of δ are

$$(0,0), \quad (1,0), \quad \left(0, \frac{3-\delta}{4-3\delta}\right), \quad \text{and} \quad \left(\frac{1}{2}, \frac{1}{2}\right).$$

Type and Stability. Now the analysis moves toward a classification of these equilibria for the various values of $\delta \in [0,1]$. Recall that at any critical point $\mathbf{x}^0 = (x_0, y_0)$ of an almost linear system, we can form the matrix A of an associated linear system, where

$$A = \begin{bmatrix} F_x\big|_{\mathbf{x}^0} & F_y\big|_{\mathbf{x}^0} \\ G_x\big|_{\mathbf{x}^0} & G_y\big|_{\mathbf{x}^0} \end{bmatrix} = \begin{bmatrix} 1 - 2x_0 - y_0 & -x_0 \\ -\gamma_2 y_0 & \alpha_2 - 2\beta_2 y_0 - \gamma_2 x_0 \end{bmatrix}.$$

Here, then

$$A_{(0,0)}(\delta) = \begin{bmatrix} 1 & 0 \\ 0 & \alpha_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \frac{3-\delta}{4} \end{bmatrix}.$$

Thus, in this case, $r_1 = 1 > 0$ and $r_2 = r_2(\delta) = \frac{3-\delta}{4} > 0$, $\forall \delta \in [0,1]$. By the Hartman-Grobman Theorem, the equilibrium at the origin is a source for all $\delta \in [0,1]$.

At the static fixed point at $(1,0)$, we have

$$A_{(1,0)}(\delta) = \begin{bmatrix} -1 & -1 \\ 0 & \alpha_2 - \gamma_2 \end{bmatrix} = \begin{bmatrix} -1 & -1 \\ 0 & \frac{1-2\delta}{4} \end{bmatrix}.$$

Eigenvalues of $A_{(1,0)}(\delta)$ are immediately available to us since the matrix is upper triangular, so the eigenvalues are the entries on the main diagonal:

$$r_1 = 1, \quad \text{and} \quad r_2 = \frac{1-2\delta}{4}.$$

One can readily show that, via the eigenvector equation, an eigenvector for $r_1 = -1$ is $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Along the "other" direction, we have an eigenvalue/eigenvector pair

$$r_2 = \frac{1-2\delta}{4}, \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ \frac{2\delta-5}{4} \end{bmatrix}.$$

The interesting effect is at $\delta = \frac{1}{2}$, where the non-horizontal eigendirection is seen to slow to a stop, creating a curve of equilibria emanating from $(1,0)$. As δ passes through $\frac{1}{2}$, the eigenvalue r_2 goes from positive to negative, and the saddle bifurcates to a sink, passing through the value where the node is not isolated. This is a planar bifurcation where an unstable node can become stable.

Now, for the critical point $\left(0, \frac{3-\delta}{4-3\delta}\right)$, we get

$$A_{\left(0,\frac{3-\delta}{4-3\delta}\right)}(\delta) = \begin{bmatrix} 1 - y_0 & 0 \\ \gamma_2 y_0 & \alpha_2 - \beta_2 y_0 \end{bmatrix} = \begin{bmatrix} 1 - \frac{3-\delta}{4-3\delta} & 0 \\ -\left(\frac{2+\delta}{4}\right)\left(\frac{3-\delta}{4-3\delta}\right) & \frac{3-\delta}{4} - 2\left(\frac{4-3\delta}{4}\right)\left(\frac{3-\delta}{4-3\delta}\right) \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1-2\delta}{4-3\delta} & 0 \\ \left(\frac{2+\delta}{4-3\delta}\right)\left(\frac{\delta-3}{4}\right) & \frac{\delta-3}{4} \end{bmatrix}.$$

Here, the eigenvalues are $r_1 = \frac{1-2\delta}{4-3\delta}$, and $r_2 = \frac{\delta-3}{4}$. For the eigenvector $\mathbf{v}_2 = [v_1, v_2]^T$ corresponding to r_2, we have the eigenvector system

$$\left(\frac{1-2\delta}{4-3\delta}\right)v_1 = \left(\frac{\delta-3}{4}\right)v_1$$
$$\left(\frac{2+\delta}{4}\right)\left(\frac{\delta-3}{4}\right)v_1 + \left(\frac{\delta-3}{4}\right)v_2 = \left(\frac{\delta-3}{4}\right)v_2.$$

This system is solve by $v_1 = 0$ and v_2 is anything non-trivial, so that the vector $\mathbf{v}_2$ is along the y-axis $\forall \delta \in [0,1]$.

For r_1, eigenvector system is

$$\frac{1-2\delta}{4-3\delta}v_1 = \frac{1-2\delta}{4-3\delta}v_1$$
$$\left(\frac{2+\delta}{4}\right)\left(\frac{\delta-3}{4}\right)v_1 + \left(\frac{\delta-3}{4}\right)v_2 = \left(\frac{1-2\delta}{4-3\delta}\right)v_2.$$

While this is a fairly messy calculation, we can boil it down to

$$v_1 = \frac{3\delta^2 - 21\delta + 16}{(2+\delta)(\delta-3)}v_2.$$

Upon inspection, one can readily see that both components will be non-zero for every $\delta \in [0,1]$, except for at one value: $\delta \sim .87$. At this point, one can show, $r_1 = r_2$, and there is only one eigendirection.

Exercise 112. Establish what is happening for this value of δ at the non-trivial critical point along the vertical axis.

Lastly, for this case, notice again, that one of the eigenvalues $r_1 = 0$, when $\delta = \frac{1}{2}$. This is precisely another instance of a bifurcation from a saddle to a sink, where one of the eigendirections slows down its repellent motion, stops and then reverses direction. Interesting....

And lastly, Let's analyze the stability, type and structure of the phase space at the point $\left(\frac{1}{2}, \frac{1}{2}\right)$. We have

$$A_{\left(\frac{1}{2},\frac{1}{2}\right)}(\delta) = \begin{bmatrix} 1 - 2x_0 - y_0 & -x_0 \\ -\gamma_2 y_0 & \alpha_2 - 2\beta_2 y_0 - \gamma_2 x_0 \end{bmatrix}$$
$$= \begin{bmatrix} -\frac{1}{2} & -\frac{1}{2} \\ -\frac{2+\delta}{8} & \frac{3-\delta}{4} - \frac{4-3\delta}{4} - \frac{2+\delta}{8} \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} & -\frac{1}{2} \\ -\frac{2+\delta}{8} & \frac{-4+3\delta}{8} \end{bmatrix},$$

with eigenvalues

$$r = -\frac{(-8+3\delta) \pm \sqrt{(-8+3\delta)^2 - (32-64\delta)}}{16} = \frac{(8-3\delta) \pm \sqrt{9\delta^2 + 16\delta + 32}}{16}.$$

One can easily see by inspection here that both of the eigenvalues here are real, with one of them remaining negative for all $\delta \in [0,1]$. The other one, however, is negative on $\delta \in \left[0, \frac{1}{2}\right)$, and positive on $\delta \in \left[\frac{1}{2}, 1\right)$, and 0, when $\delta = \frac{1}{2}$. This, again, denotes a bifurcation value for δ, with the equilibrium going from a sink to a saddle.

IN fact, at $\delta = \frac{1}{2}$, we have that strange situation where the three non-trivial critical points all have 0 as an eigenvalue of their linearization. This suggests a curve of critical points in the phase plane. We can actually of directly to the original differential equation to find these:

Let $\delta = \frac{1}{2}$. Then the critical points are all at

$$\begin{aligned}
F(x,y) &= x(1-x-y) = 0 \\
G(x,y) &= y\left(\frac{3-\frac{1}{2}}{4} - \left(\frac{4-\frac{3}{2}}{4}\right)y - \left(\frac{2+\frac{1}{2}}{4}\right)x\right) = 0 \\
&= y\left(\frac{5}{8} - \frac{5}{8}y - \frac{5}{8}x\right) = 0 \\
&= \frac{5}{8}y(1-x-y) = 0.
\end{aligned}$$

At $\delta = \frac{1}{2}$, there will be a line of critical points, ranging within the first quadrant along the line $y = 1-x$, from the equilibrium at $(1,0)$, through the equilibrium at $\left(\frac{1}{2},\frac{1}{2}\right)$ to the the fixed point at $(1,0)$. Can you envision this?

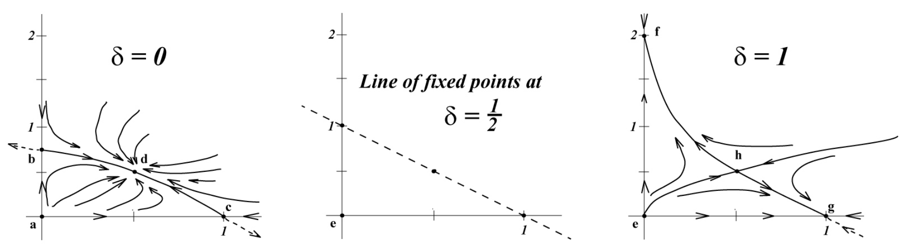


FIGURE 7. Can you draw the phase portrait for $\delta = \frac{1}{2}$?

EXERCISE 113. Figure **??** shows the equilibria of the non-linear ODE Competing Species model for $\delta = \frac{1}{2}$. The equilibria are at the origin and on the line $y = 1-x$. Draw in the rest of the phase portrait as accurately as possible by linearizing the four persistent equilibria and extrapolating what happens off of the line.

CHAPTER 4

Recurrence

So far, we have explored many systems and contexts where dynamical systems have exhibited simple behavior, or fairly simple behavior. We will now begin to explore more complicated behavior. However, to start, we will stay with maps of the type we have already introduced. But we will change the place on which they are acting. This, in and of itself, changes the nature of the orbits. It turns out that when the space is Euclidean, orbits can converge to something or wander away toward the edge of the space. However when a space is compact, roughly that its edges are not infinitely far away, if the edges in fact exist at all, then an orbit that does not converge to any particular thing must go somewhere within the space. How to describe where it goes will take us to behavior which is more complicated than what we have already seen. To begin, consider the definition:

Definition 4.1. For $f : X \to X$ a continuous map on the metric space X, a point $x \in X$ is called (*positively*) *recurrent* with respect to f if there exist a sequence of natural numbers $n_k \longrightarrow \infty$ where $f^{n_k}(x) \longrightarrow x$.

In the simple dynamical systems we studied so far, the only recurrent points were fixed and periodic points (this makes sense, right?). However, in the right context, non-periodic points can also be recurrent. This chapter begins a study of relatively simple maps that exhibit this more complicated behavior. And this behavior is captured in this notion of recurrence.

4.1. Rotations of the circle

Again, think of S^1 either as the set of unit modulus numbers of the complex plane

$$S^1 = \left\{ z \in \mathbb{C} \,\middle|\, z = e^{2\pi i\theta}, \theta \in \mathbb{R} \right\},$$

or as the quotient space of the real line modulo the integers, $S^1 = \mathbb{R}/\mathbb{Z}$. Recall, for $x, y \in \mathbb{R}$, denote $\overline{x}$, $\overline{y}$ their respective points in S^1 under the exponential map $\rho : \mathbb{R} \to S^1$, $\rho(\theta) = e^{2\pi i\theta}$.

- Here $\overline{x} = \overline{y}$ iff $x - y \in \mathbb{Z}$, or $x \equiv y \pmod 1$.
- $\overline{x}$, $\overline{y}$ are the equivalence classes of points in $\mathbb{R}$ under the equivalence relations imposed on $\mathbb{R}$ by the map ρ.

In this last interpretation, one can imagine S^1 to be the unit interval $[0, 1]$ in $\mathbb{R}$ where one agrees to identify the endpoints (hence the notation sometimes used when we say $0 = 1$).

One can define a metric on S^1 by simply inheriting the one it has as it sits in $\mathbb{C}$ (or if you will, $\mathbb{R}^2$). This is essentially the Euclidean metric and measures the straight line distance in the plane between two points. Really, this is the length of the chord, or secant line, joining the points. See Figure 1.

But also, we can define a distance between points by the arc length between them. In some ways, this is preferable, since in the abstract, S^1 doesn't really sit anywhere. There is no interior and exterior of S^1, unless you call the actual points in the curved line making the circle the interior points. The problem with using arc length to determine the distance between points is that there are two distinct paths going from one point to another. There must be a determination as to which one to choose. Choosing the minimal path is a nice choice, but how does one do this mathematically. The answer lies within the view that S^1 really is the real line $\mathbb{R}$ infinitely coiled up like a slinky by the exponential map ρ above, and length in $\mathbb{R}$ is easy to describe, and passes through this map, at least locally:

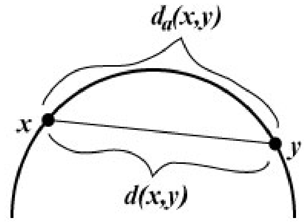


FIGURE 1. Equivalent metrics on S^1.

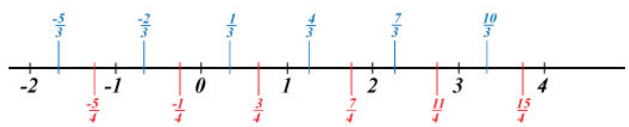


FIGURE 2. The equivalence classes in $\mathbb{R}$ of $\overline{x} = \frac{1}{3}$ and $\overline{y} = \frac{3}{4}$ in S^1.

Define

$$d(\overline{x}, \overline{y}) = \min\left\{ |x - y| \;\middle|\; x, y \in \mathbb{R}, x \in \overline{x}, y \in \overline{y} \right\}. \tag{4.1.1}$$

Figure 2 shows the equivalence classes of the points $\overline{x} = \frac{1}{3}$ and $\overline{y} = \frac{3}{4}$. Choosing arbitrary representatives x and y and calculating their distance in $\mathbb{R}$ will lead to many different results. However, the minimum distance between representatives of these two classes is well-defined and in this case, $d(\overline{x}, \overline{y}) = \frac{5}{12}$. Notice that really, the closest two distinct distances between two equivalence classes in $\mathbb{R}$ correspond precisely to the arc lengths in S^1 along the two distinct paths joining $\overline{x}$ and $\overline{y}$.

LEMMA 4.2. *These two metrics are equivalent.*

PROOF. We leave the proof as an exercise. □

EXERCISE 114. Show that the inherited Euclidean metric on $S^1 \subset \mathbb{R}^2$ is equivalent to the arc length metric in Equation 4.1.1.

Denote by R_α the translation of the points in S^1 by the real number α. We could say R_α denotes the rigid rotation of S^1 by the angle α, but, as we will see, the parameterization of S^1 becomes vitally important here. We have currently parameterized S^1 as the unit interval in $\mathbb{R}$, with $0 = 1$. So even though α technically can be any real number, rotating by α and rotating by $\alpha + n$, where $n \in \mathbb{Z}$ amounts to the same thing. (Note that this would definitely not be the case for a continuous dynamical system given by $\dot{x} = \alpha$, $x \in S^1$. Can you see why?) Here $R_\alpha(\overline{x}) = \overline{x} + \alpha \pmod 1 = \overline{x + \alpha}$. In complex notation, we view rotations as linear maps, with multiplication by the factor $z_\alpha = e^{2\pi i\alpha}$, so that $R_\alpha(z) = z_\alpha z$. In each case, then $R_\alpha : S^1 \to S^1$, with either

$$R_\alpha^n(\overline{x}) = R_{n\alpha}(\overline{x}) = \overline{x} + n\alpha \pmod 1 = \overline{x + n\alpha},$$

or $R_\alpha^n(z) = z_\alpha^n z$.

Q. What can we say about the dynamics of a circle rotation?
Q. What if $\alpha \in \mathbb{Q}$?
Q. What if $\alpha \notin \mathbb{Q}$?

The quick answers are that, when α is rational, all orbits are periodic, and all of the same period. When α is not rational, then there are no periodic orbits at all. The trick really is to understand well what R_α^n looks like for each n, and what it means for a point to be periodic in the circle.

EXERCISE 115. Let $R_\alpha : S^1 \to S^1$, be the rotation $r_\alpha(\overline{x}) = \overline{x} + \alpha = \overline{x + \alpha}$. Show that every orbit is periodic when $\alpha \in \mathbb{Q}$, and no orbit is periodic when $\alpha \notin \mathbb{Q}$.

The latter exercise creates a deeper concern: Without fixed or periodic points in S^1 for what I will call an *irrational rotation*, the question is, where do the orbits go? They cannot converge to a point in the circle, since in many cases (and really in general), if they converged to a point in S^1, then that point would have to be a fixed point (if orbits converge, they must converge to another orbit). The answer is that they go everywhere. And that tells one a lot about the dynamics.

REMARK 4.3. The above notion of an irrational rotation was based on the parameterization of S^1 given by the interval $[0,1)$. There, the rotation R_α was irrational as a rotation when α isn't rational as a number. However, the parameterization is critical here, and the rationality IS of the rotation really with respect to the integer 1, the maximum value of the parameter. To see this, suppose instead we parameterized S^1 via the interval $[0,2\pi)$, another rather common parameterization given by the map $\rho : \mathbb{R} \to S^1$, where $\rho(x) = e^{ix}$. Here, a rotation half way around the circle is given by R_π, where $\alpha = \pi$ is irrational (as a number!) Thus the rotation R_π is not irrational at all, as every point is 2-periodic. However, the rotation by 1, R_1 would have NO periodic orbits.

EXERCISE 116. Show there are no periodic orbits for the rotation R_1 on S^1 parameterized via the map $\rho : \mathbb{R} \to S^1$, where $\rho(x) = e^{ix}$.

The correct conclusion to draw here is that the rationality of the rotation R_α depends on the parameterization. We offer a definition to be clear.

DEFINITION 4.4. A rotation $R_\alpha : S^1 \to S^1$, where S^1 is parameterized by the interval $[0,T)$ for $T > 0$, is called *irrational* if $\frac{\alpha}{T} \notin \mathbb{Q}$. Otherwise, the rotation is called *rational*.

PROPOSITION 4.5. *For R_α an irrational rotation of S^1, all orbits are dense in S^1.*

(IDEA OF PROOF). Really, the idea is the following:

- Show the forward orbit of any $\overline{x}$ is not periodic (you will do this in the exercises).
- Show that $\forall \epsilon > 0$, $\exists N \in \mathbb{N}$, such that $d\left(R_\alpha^N(\overline{x}), \overline{x}\right) < \epsilon$.
- Show that this is true for all $\overline{x}$.

□

Note: All rotations are invertible, right? Really, they are all homeomorphisms. So define $R_\alpha^{-1}(\overline{x}) = R_{-\alpha}(\overline{x})$. To show density, we have to show that the orbit of $\overline{x}$ will visit any size open neighborhood of $\overline{x}$. Here is a nice technique for showing this:

4.1.1. Continued Fraction Representation. The continued fraction representation (CFR) of a real number is a representation of real numbers as a sequence of integers in a way which essentially determines the rationality of the number. This is very much like the standard decimal representations of real numbers, in that it also (our usual base-10 version is a good example) provides a ready way to represent all real numbers. However, the sequence of integers which represent a real number in a base-10 decimal expansion represent some rational numbers as finite-length sequences (think $\frac{11}{8} = 1.375$), and others as infinite length sequences (think $\frac{4}{9} = 0.44444\cdots$). The CFR instead is a base-free representation in which all and only rational number representations are the finite length sequences. Plus, the CFR is another nice way to approximate a real number by either truncating its sequence or simply not calculating the entire sequence.

Indeed, in the CFR, Any real number in $(0,1)$ can be written as $\frac{1}{s}$, where $s \in (1,\infty)$. More generally, then, any real number r can be written as an integer and a real number in $(0,1)$; as

$$r = n + \frac{1}{s}, \text{ where } n \in \mathbb{Z}, \text{ and } s \in (1,\infty).$$

If $s \in \mathbb{N}$, then this expression is considered the CFR of r (it is sometimes written then $r = [m;s]$; For example, $\frac{5}{2} = [2:2]$.

Now suppose $s \notin \mathbb{N}$. Then since $s \in (1,\infty)$, $s = m + \frac{1}{t}$, for $m \in \mathbb{N}$, and $t \in (1,\infty)$. Thus,

$$r = n + \frac{1}{\frac{1}{m+\frac{1}{t}}}, \text{ where } n \in \mathbb{Z}, \ m \in \mathbb{N}, \text{ and } t \in (1,\infty).$$

Again, if $t \in \mathbb{N}$, then we stop and $r = [n;m,t]$ is the CFR of r. If it is not, we again let $t = p + \frac{1}{u}$, for $p \in \mathbb{N}$ and $u \in (1,\infty)$ so

$$r = n + \frac{1}{\frac{1}{m+\frac{1}{p+\frac{1}{u}}}}, \text{ where } n \in \mathbb{Z}, \ m,p \in \mathbb{N}, \text{ and } u \in (1,\infty).$$

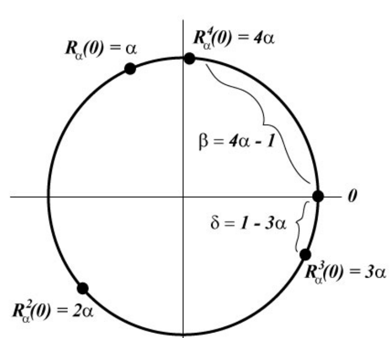
Again, if $u \in \mathbb{N}$, we stop and the CFR of r is $[n : m,p,u]$. If not, then we continue indefinitely. The CFR is a finite sequence iff $r \in \mathbb{Q}$.

Exercise 117. Compute the CFR of $-\frac{33}{13}$.

Exercise 118. Calculate the fraction whose CFR is $[0:3,5,7]$.

Example 4.6. So let R_α be a rotation of S^1 for $\alpha = \frac{1}{3+\frac{1}{5+\frac{1}{c}}}$, where $c > 1$, and $c \notin \mathbb{Q}$. Then it turns out that $\alpha \notin \mathbb{Q}$.

To see this, let's start the construction which would establish the middle bullet point in the above proof idea. To start, it should be obvious that $\frac{1}{4} < \alpha <$



$\frac{1}{3}$ (why?). In the figure, we can graph $R_\alpha(0)$, $R_\alpha^2(0)$, $R_\alpha^3(0)$, and $R_\alpha^4(0)$. One of the latter two winds up being the early, closest approach to 0 of the orbit $\mathcal{O}_0^+$. But which is smaller, $\delta = 1 - 3\alpha$, or $\beta = 4\alpha - 1$?
Visually, the closest approach to 0 is $R_\alpha^3(0) = 3\alpha$, but without the benefit of knowing what the choice of c is in general, it is not clear *a priori* whether $\delta = 1 - 3\alpha$ is actually smaller than $\beta = 4\alpha - 1$. Even without knowing c, we can still perform the comparison via the CFR:

$$\delta = 1 - 3\alpha = 1 - \frac{3}{3 + \frac{1}{5+\frac{1}{c}}} = \frac{3 + \frac{1}{5+\frac{1}{c}} - 3}{3 + \frac{1}{5+\frac{1}{c}}} = \frac{\frac{1}{5+\frac{1}{c}}}{3 + \frac{1}{5+\frac{1}{c}}} = \frac{1}{16 + \frac{3}{c}}.$$

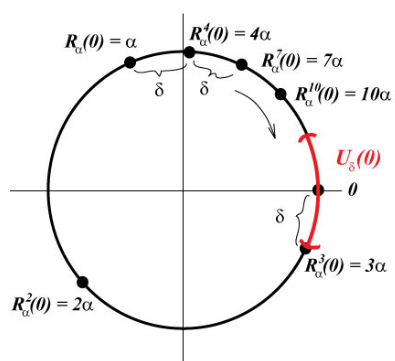
Exercise 119. Calculate $\beta = 4\alpha - 1$ in the same way as above, and show that it is larger than δ for any choice of $c > 1$.

Hence, the third iterate is the first closest return of $\mathcal{O}_0^+$ to 0.

Q. Will the orbit ever get closer to 0?
Q. If it will, then which iterate?

These questions will help us to show the orbit will eventually get arbitrarily close to 0.

We could simply hunt for the next return. Or we can be clever and calculate it. Here is the idea: it took three steps to get within δ of the initial point 0. (We could say it took three steps to get δ-close to 0). If we now create an open δ-neighborhood of 0, $N_\delta(0)$, when will the first iterate occur when we will enter this neighborhood and thus get closer than δ to 0?



One way to ensure this is to look at the first step after our previous close approach. This is the fourth element of $\mathcal{O}_0$ and is $R_\alpha^4(0) = 4\alpha$. Here $4\alpha = \alpha + 3\alpha = \alpha + (1 - \delta)$, so that $4\alpha - 1 = \beta = \alpha - \delta$. One conclusion to draw from this is that $R_{3\alpha}$ takes α to 4α which is $\alpha - \delta$ (see figure). So $R_{3\alpha}(\alpha) = \alpha - \delta$, $R_{3\alpha}^2(\alpha) = \alpha - 2\delta$, and $R_{3\alpha}^n(\alpha) = \alpha - n\delta$. So for which n would we satisfy

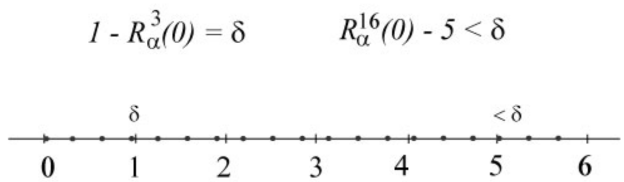
$$0 < \alpha - n\delta < \delta\ ?$$

Note that for some choice of n, the iterate will have to lie on the positive side of 0 in $N_\delta(0)$ (why?). Of course, this simplifies to $n\delta < \alpha < (n+1)\delta$, which is solved by simply taking the integer part of the fraction $\frac{\alpha}{\delta}$. Denote the greatest integer function by $\lfloor\cdot\rfloor$, so that, for example, $\lfloor\pi\rfloor = 3$. Then, the iterate n we are looking for is

$$n = \left\lfloor \frac{\alpha}{\delta} \right\rfloor = \left\lfloor \frac{\frac{1}{3+\frac{1}{5+\frac{1}{c}}}}{\frac{1}{16+\frac{3}{c}}} \right\rfloor = \left\lfloor 5 + \frac{1}{c} \right\rfloor = 5.$$

Hence we can say that $R^5_{3\alpha}(\alpha) = R^{15}_\alpha(\alpha) = R^{16}_\alpha(0)$ is within δ of 0 (See figure at right below). We could then use the actual distance between 0 and $R^{16}_\alpha(0)$ as our new δ, and look for iterates of R^{16}_α to find our next closest approach. Continuing this way, we create a subsequence of $\mathcal{O}_0$ which consists of exponentially increasing powers of the original R_α and this subsequence converges to 0. This is the basic approach to proving the second bullet point in the above proof idea.

On the real line, we see that our rotations by α is simply a translation by α. Approaching and getting closer to 0, means that our orbit will at some point come close to an integer value (ANY integer will do, as they all represent 0 in the circle!). See the figure here.



There is really a better way to understand this notion of visiting neighborhoods of points in S^1 under irrational rotations. This other way is by understanding the frequency with which an orbit visits a small open set under a rotation. This is called the *dynamical frequency*, and is a measure of how often an orbit visits a small open interval in S^1 relative to how much time it is outside of the interval.

Fix $\Delta \subset S^1$ an arc. Then for $x \in S^1$ and $n \in \mathbb{N}$, define

$$F_\Delta(x,n) = \#\left\{ k \in \mathbb{Z} | \, 0 \le k < n, R^k_\alpha(x) \in \Delta \right\}.$$

Here, the number sign # denotes the cardinality of the set. For example, in the above figure with our choice of α, and $\Delta = N_\delta(0)$, we have

$$F_\Delta(0,18) = F_{N_\delta(0)}(0,18) = \#\left\{0, 16\right\} = 2.$$

Note that for Δ small, then for any $x \in S^1$, F_Δ will be small. And for Δ large, F_Δ will be bigger, but always less than n. So we can say that $0 \le F_\Delta(x,n) \le n$, for every x and Δ. And for any choice of x and Δ, as n grows, F_Δ is monotonically increasing.

However, it is also true that for $\alpha \notin \mathbb{Q}$, $\lim\limits_{n\to\infty} F_\Delta(x,n) = \infty$. (Can you show this?) Hence instead of studying the frequency with which the orbit of a point visits an arc, we study the *relative frequency* of visits as n gets large, or the quantity

$$\frac{F_\Delta(x,n)}{n}.$$

Suppose on the orbit segment of a point x under the irrational rotation by α given by $\left\{R^i_\alpha(x)\right\}^m_{i=0}$, we found that given the arc Δ, that $R^{k_1}_\alpha(x), R^{k_2}_\alpha(x), R^{k_3}_\alpha(x) \subset \Delta$ and these were the only three. Then we know that the frequency $F_\Delta(x,m) = 3$, and the relative frequency $\frac{F_\Delta(x,m)}{m} = \frac{3}{m}$. In our example from above in the figures, the relative frequency of hits on the interval $N_\delta(0)$ on the orbit segment $\left\{R^i_\alpha(x)\right\}_{i=0}^{18}$ is $\frac{F_{N_\delta(0)}(0,18)}{18} = \frac{2}{18} = \frac{1}{9}$. The goal is to study the relative frequency of a rotation on any arc of any length and be able to say something meaningful about how often, on average, the *entire orbit* visits the arc.

Some notes:

- Define $\ell(\Delta)$ = length of Δ (under some metric).
- The relative frequency really does not depend on whether Δ is open, closed or neither (why not?). Is this true also for rational rotations?
- The convention is to take representatives for arcs to be of the "half-closed" form $[\cdot,\cdot)$. Then it is easy to see whether unions of arcs are connected or not.
- We study the overall relative frequency of entire orbits: This translates to a study of
$$\lim_{n\to\infty} \frac{F_\Delta(x,n)}{n}.$$
However, It is yet not entirely clear that this limit actually exists. We first address this point.

Definition 4.7. Given a sequence $s = \{a_n\}_{n\in\mathbb{N}}$, the extended number $\ell \in \mathbb{R}^* = \mathbb{R}\cup\{-\infty,\infty\}$ is a *limit point* of s if there exists a subsequence $\{a_{n_k}\}_{k\in\mathbb{N}}$ of s where
$$\lim_{k\to\infty} a_{n_k} = \ell.$$

It is a classical result in analysis that every bounded sequence in $\mathbb{R}^n$ has limit points (this is the celebrated Bolzano-Weierstrauss Theorem.) Using this extended notion of $\mathbb{R}$, we can say that every sequence in $\mathbb{R}$ has limit points in $\mathbb{R}^*$. Categorizing the limit points of a sequence provides important information about the extent of a sequence. One way to categorize them is to find their bounds:

Definition 4.8. For $\{a_n\}_{n\in\mathbb{N}}$, the *limit inferior* is
$$\liminf_{n\to\infty} a_n = \underline{lim}_{n\to\infty} a_n = \lim_{n\to\infty}\left(\inf_{m\geq n} x_m\right),$$
where $\inf\limits_{m\geq n} a_m$ is the infimum of the sequence $\{a_m\}_{m=n}^{\infty}$.

Really, it is the largest number smaller or equal to all of the remaining elements of the sequence. It is the largest number smaller than all but a finite number of the elements of the sequence. One can define the *limit superior* similarly as the smallest of the numbers larger than all but a finite number of elements of the sequence. We use the notation $\limsup\limits_{n\to\infty}$ or $\overline{\lim}_{n\to\infty}$ for the limit superior.

It should be intuitively obvious that

(1) the limits inferior and superior always exist in $\mathbb{R}^*$.
(2) For any sequence s, $\liminf s \leq \limsup s$.
(3) For any sequence s, $\liminf s = \limsup s$ iff $\lim s$ exists.

In essence, if $\liminf s = a$ and $\limsup s = b$, then the interval $[a,b]$ will contain all possible limit points of s (although in general not every point in $[a,b]$ may be a limit point of the sequence s.)

Consider the function $f : \mathbb{N} \to \mathbb{R}$ defined by $f(n) = \left(1+\frac{1}{n}\right)\sin n$. A priori, we do not know whether $\lim_{n\to\infty} f(n)$ exists or not (Really, though, think of the continuous version of this function in calculus. There isn't a horizontal asymptote for f). So we first define the limit inferior (respectively superior) for f. This type of limit either always exists or is $-\infty$ (resp. ∞). It is the largest (resp. smallest) number where no more than a finite number of terms in the sequence are smaller (resp. larger) than it on the entire sequence. maybe separate this out into an actual definition? Think of the envelope of a sequence being defined to allow some terms to be outside the envelope, but only a finite number of them. In the case of $f(n) = \sin n$, the $\liminf_{n\to\infty} f(n) = -1$. This makes sense, since if we try to "cut" the function at anything above -1, that small interval of values (think $[-1, -1+\epsilon)$) will be visited an infinite number of times eventually by f. Also, $\limsup_{n\to\infty} f(n) = 1$. It should be obvious that while these quantities may not be easy to calculate, not only should they exist (for the minute, think of an infinite limit as existing in the sense that the sequence is going somewhere), but it must be the case that for any sequence $\{x_n\}_{n\in\mathbb{N}}$,

$$\liminf_{n\to\infty} x_n \le \limsup_{n\to\infty} x_n.$$

And should they be equal, then $\lim_{n\to\infty} x_n$ in fact exists and is equal to the two limit bounds.

In our case, let A be a disjoint union of arcs. Then define

$$\overline{f}_x(A) = \limsup_{n\to\infty} \frac{F_A(x,n)}{n}, \quad \underline{f}_x(A) = \liminf_{n\to\infty} \frac{F_A(x,n)}{n}.$$

It turns out that these two quantities not only exist. They also are equal:

Proposition 4.9. *For any arc $\Delta \subset S^1$, and every $\overline{x} \in S^1$, and any irrational rotation R_α, $\alpha \notin \mathbb{Q}$ on S^1, we have*

$$f(\Delta) := \lim_{n\to\infty} \frac{F_\Delta(x,n)}{n} = \ell(\Delta).$$

idea. The proof relies on finding bounds for the quantities $\overline{f}_x(\Delta)$ and $\underline{f}_x(\Delta)$, and showing that it is always the case that $\overline{f}_x(\Delta) \le \ell(\Delta)$ and $\underline{f}_x(\Delta) \ge \ell(\Delta)$. This can only be the case if the limits superior and inferior are in fact equal, and equal to $\ell(\Delta)$. □

Complete this proof.

Notes: Let $R_\alpha : S^1 \to S^1$ be an irrational rotation. Then for $\overline{x} \in S^1$,

- the orbit $\mathcal{O}_{\overline{x}}$, as a sequence $\{R_\alpha^n(\overline{x})\}_{n\in\mathbb{N}}$, is called a *uniform distribution* or an *equidistribution* on S^1.
- the orbit $\mathcal{O}_{\overline{x}}$ in a sense "fills" every arc in S^1.

Hence, we say that any orbit of an irrational rotation of S^1 is uniformly distributed on S^1. This is our notion of a set being dense in another set, and for these orbits, one can actually "see" the notion of recurrence. To further understand this new type of dynamical behavior, we do an application. But first, let's continue with a little more nomenclature.

Definition 4.10. A set $Y \subset X$ is invariant under a map $f : X \to X$, if

$$f\big|_Y : Y \to Y.$$

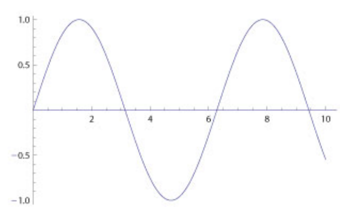
Definition 4.11. A continuous map $f : X \to X$ is called *topologically transitive* if $\exists x \in X$ such that $\mathcal{O}_x$ is dense in X. A non-invertible map is called topologically transitive if $\exists x \in X$ such that $\mathcal{O}_x^+$ is dense in X.

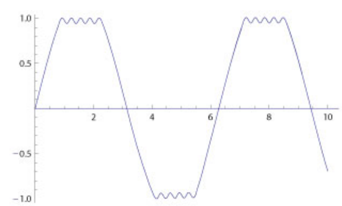
Definition 4.12. A continuous map $f : X \to X$ is *minimal* if $\forall x \in X$, $\mathcal{O}_x$ is dense in X (the forward orbit is dense for a noninvertible map).

Definition 4.13. A closed, invariant set is *minimal* is there does not exist a proper, closed invariant subset.

More notes:

- Like in the case of open and closed domains in vector calculus, a set is closed if it contains all of its limit points. And for any set X, the closure of X, denoted $\overline{X}$ is defined to be the closed set obtained by adding to X all of its limit points (think of adding the sphere which is the boundary of an open ball in $\mathbb{R}^3$). In the case of a minimal map $f : X \to X$, for any $x \in X$, we have $\overline{\mathcal{O}}_x = X$.
- Same is true for a topologically transitive map f, if one takes any point on the dense orbit.
- Irrational rotations of the circle are minimal!.





4.1.2. Application: Periodic Function Reconstruction via Sampling. Consider the two functions in the picture.

- Each is periodic and of the same period as the other.
- Each can be viewed as a real-valued smooth function on S^1. And each takes values in the interval $I = [-1, 1]$.
- Question: Are the values of these two functions equally distributed equally (or even evenly) on I?
- Question: If we knew the period and range of some unknown function, and needed to sample the function (create a sequence of function values) to see which of the above two function was the one we are seeking, how can we design our sampling to ensure we can differentiate between these two?

Dynamics attempts to answer this question. Let $\{x_n\}$ be a sequence (think of this sequence as a sampling of the function), and $a < b$ two real numbers. Define

$$F_{a,b}(n) = \#\left\{ k \in \mathbb{Z} \,\middle|\, 1 \leq k \leq n, a < x_k \leq b \right\}$$

as the number of times the sequence up to element n visits the interval $(a, b) \subset \mathbb{R}$. Really, this is the same definition of F as before on the arc $\Delta \subset S^1$. The only change in this case is that we are defining F in this context as an interval in $\mathbb{R}$.

Then define the *relative frequency* in the same way as before. In the figure, the relative frequency of $\{x_n\}$ on the interval $(a,b]$ shown is

$$\left.\frac{F_{a,b}(n)}{n}\right|_{n=6} = \frac{2}{6} = \frac{1}{3}.$$

We say that $\{x_n\}$ has an *asymptotic distribution* if $\forall a, b$, where $-\infty \leq a < b \leq \infty$, the quantity

$$\lim_{n\to\infty} \frac{F_{a,b}(n)}{n}$$

exists. In a sense, we are defining the percentage of the time that a sequence visits a particular interval.

In the case where the sequence has an asymptotic distribution, the function

$$\Phi_{\{x_n\}}(t) = \lim_{n\to\infty} \frac{F_{-\infty,t}(n)}{n}$$

is called the distribution function of the sequence $\{x_n\}$. Here Φ is monotonic, and measures how often the values of a sequence visit regions of the real line as one varies the height of an interval $(-\infty, t]$.

Definition 4.14. A real-valued function φ on a closed, bounded interval is called *piecewise monotonic* if the domain can be partitioned into finite many subintervals on which φ is monotonic. A real-valued function on $\mathbb{R}$ is *piecewise monotonic* if it is piecewise monotonic on every closed, bounded subinterval of $\mathbb{R}$.

Remark 4.15. Monotonic means strictly monotonic here. Really, this means that there are no flat (purely horizontal on an open interval) regions of the graph of φ. Think of functions like $f(x) = \sin x$, and polynomials of degree larger than 1, which are piecewise monotonic, and functions like

$$g(x) = \begin{cases} -(x+2)^2 & -4 \leq x < -2 \\ 0 & -2 \leq x \leq 0 \\ x^2 & 0 < x \leq 2 \end{cases},$$

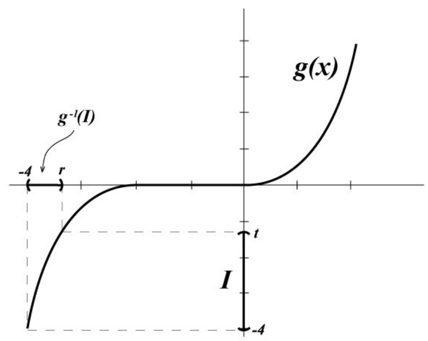
which is not piecewise monotonic (See the graph of $g(x)$ below).

When φ is piecewise monotonic, the pre-image of any interval I is a finite union of intervals in the domain (see the figure).

Definition 4.16. The φ-length of an interval I is

$$\ell_\varphi(I) := \ell\left(\varphi^{-1}(I)\right).$$

- This is the total length of all pieces of the domain that map onto I. In the figure, $\ell_\varphi(I) = \ell(A) + \ell(B)$.
- For piecewise monotonic functions φ, the φ-length is a continuous function of the end points of I (vary one end point of I continuously, and the φ-length of I also varies continuously. This doesn't work with flat regions since the φ-length ell_φ would then jump as one hits the value of the flat region.



Indeed, let's look at the $g(x)$ in the figure more closely. Here, one can calculate the φ-length. Indeed, choose the interval $I = [-4, t]$. Here, t is the function value, and there is only a single interval mapped onto i for any value of t.

For $t < 0$, this interval is given in the figure as the interval of the domain $g^{-1}(I) = [-4, r]$, where $g(r) = t$. Solving the equation $g(r) = t$ for r yields

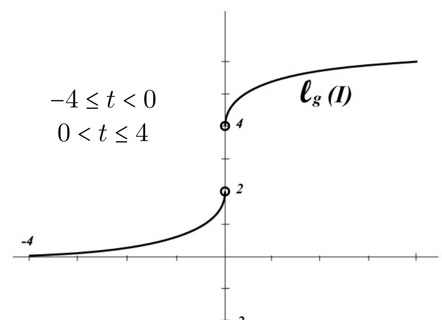
$$-(r+2)^2 = t \iff r = -\sqrt{-t} - 2$$

where we chose the negative branch of the square root function in the middle step to account for the domain restrictions. Here, the g-length of I,

$$\begin{aligned} \ell_g(I) &= \ell\left(g^{-1}([-4,t])\right) \\ &= -2 - \sqrt{-t} - (-4) = 2 - \sqrt{-t}. \end{aligned}$$

Now for $t > 0$, the same calculation yields $\ell_g(I) = 4 + \sqrt{t}$ for $I = [-4, t]$. Putting these two pieces of the g-length function together yields the graph of

$$\ell_g(I) = \begin{cases} 2 - \sqrt{-t} & -4 \le t < 0 \\ 4 + \sqrt{t} & 0 < t \le 4 \end{cases}$$



which has a jump discontinuity at $t = 0$. In fact, the only way to change $g(x)$ to make the g-length function continuous is to remove the middle piece of the $g(x)$ function and translate one or the other pieces right or left to again make $g(x)$ continuous. But that would have the effect of moving the two pieces of the graph of $\ell_g(I)$ together. The jump discontinuity becomes a hole in the graph, easily filled. But in this case, the changed $g(x)$ has been made piecewise monotone!

One can show that for a piecewise monotonic function φ, a distribution function for φ is

$$\Psi : \mathbb{R} \to \mathbb{R}, \quad \Psi_\varphi(t) = \ell_\varphi\left((-\infty, t)\right).$$

We can use this for:

THEOREM 4.17. *Let φ be a T-periodic function of $\mathbb{R}$ such that $\varphi_T = \varphi|_{[0,T]}$ is piecewise monotone. If $\alpha \notin \mathbb{Q}$ and $t_0 \in \mathbb{R}$, then the sequence $x_n = \varphi(t_0 + n\alpha T)$ has an asymptotic distribution with distribution function*

$$\Phi_{\{x_n\}}(t) = \frac{1}{T}\Psi_\varphi(t) = \frac{\ell\left(\varphi^{-1}((-\infty,t))\right)}{T}.$$

We won't prove this or study it in any more detail. But there is an interesting conclusion to draw from this. In the theorem, the sequence of samples of the T-periodic function φ has the same distribution function as the actual function φ, (defined over the period, that is) precisely when the sampling is taken at a rate which is an irrational multiple of the period T. In this way, the sequence, over the long term, will fill out the values of φ over the period in a dense way. In a way, one can recover the function φ from a sequence of regular samples of it only if the sampling is done in a way which ultimately allows for all regions of the period to be visited evenly. This is a very interesting result.

Here, do the example $f(x) = 2 + \cos(2n)$.

Exercise 120. Calculate the distribution function for the sequence $\cos n$.

4.2. Linear Flows on the Torus

Here is another type of dynamical system where circle rotations and their implications play a vital role. This one involves generalizing circle maps via a corresponding circle flow into more than one dimension.

To start, recall what a flow is: Let $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, $\mathbf{x}(0) = \mathbf{x}_0 \in \mathbb{R}^n$ be an IVP, where the vector field $\mathbf{f}(\mathbf{x})$ is C^1. This IVP defines a flow on $\mathbb{R}^n$. For $I \subset \mathbb{R}$ an interval containing 0, define a continuous map $\varphi : I \times \mathbb{R}^n \to \mathbb{R}^n$ that satisfies the following:

- $\forall T \in I$, $\varphi^t = \varphi(t, \cdot) : \mathbb{R}^n \to \mathbb{R}^n$ is a homeomorphism (for a given choice of t, this is is simply the time-t map of the IVP).
- $\forall s, t \in I$, where $s + t \in I$, one has
$$\varphi^s \circ \varphi^t(\mathbf{x}) = \varphi^{s+t}(\mathbf{x}).$$

Now suppose that $S^1 = \left\{e^{2\pi i x} \in \mathbb{C}\right\}$, and $\frac{dx}{dt} = \alpha$, $x(0) = x_0$ is an IVP defined on S^1 Be very careful here. Better to define S^1 in a different way to avoid the constants. This is solved by $x(t) = \alpha t + x_0$, which can also be written in flow form $\varphi^t_\alpha(x) = \alpha t + x$. Notice in this last expression, we have included the subscript α to denote the dependence of the flow on the value of the parameter α. here the time-1 map is just
$$\varphi^1_\alpha(x) = \alpha + x = R_\alpha(x), \quad x \in S^1.$$
The time-1 map is just a rotation map of the circle by α. Keep in mind, however, that the IVP will share the same time-1 map as the new IVP given by $\frac{dx}{dt} = \alpha + n$, $x(0) = x_0$ for n any positive integer. However, the flows will all be very different! (do you see this?) Linear flows on S^1 are not very interesting (are you starting to get used to this term in mathematics yet?). They differ only by speed (and possibly direction), and ultimately, all look like continuous rotations of the circle, whether α is rational or not. However, we can generalize this flow to a situation which does produce somewhat interesting dynamics.

Consider now a flow given by the pair of uncoupled circle ODEs:
$$\frac{dx_1}{dt} = \omega_1, \quad \frac{dx_2}{dt} = \omega_2.$$

This system, which can be written as the uncoupled vector ODE $\dot{\mathbf{x}} = \boldsymbol{\omega}$, or $\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix}$, can be viewed as defining a flow on the two-torus $\mathbb{T} = S^1 \times S^1$, and has the solution $\mathbf{x} = \begin{bmatrix} x_1 + \omega_1 t \\ x_2 + \omega_2 t \end{bmatrix}$.

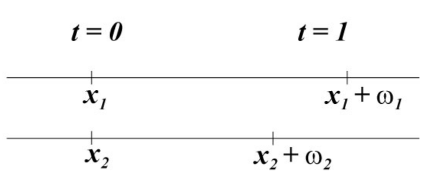


Figure 3. A pair of translation flows on $\mathbb{R}$.

In flow notation, we can write either

$$T^t_{\boldsymbol{\omega}}(x_1,x_2) = (x_1+\omega_1 t, x_2+\omega_2 t)\,, \text{ or } \varphi^t_{\boldsymbol{\omega}}(\mathbf{x}) = \mathbf{x}+\boldsymbol{\omega}t.$$

Graphically, solutions are simply translations along $\mathbb{R}$ (Figure 3) or as straight line motion in $\mathbb{R}^2$ (Figure 4). Note that in this last interpretation, the slope of the solution line is $\gamma = \frac{\omega_2}{\omega_1}$.

However, each of these uncoupled ODEs also can be considered as a flow on S^1, and hence the system can be considered a flow on $S^1 \times S^1 = \mathbb{T}$. Suppose, for example, that $1 < \omega_1 < 2$, while $0 < \omega_2 < 1$. The flow from time $t = 0$ to time $t = 1$ would take the origin on one circle to the point $1 - \omega_1$, and the flow line would start at $x_1 = 0$ and travel once around the circle before stopping to ω_1. The flow on the other circle would take $x_2 = 0$ partway around the circle to ω_2. Viewed via the two periodic coordinates of $\mathbb{T}$, we have the flow line in Figure 5

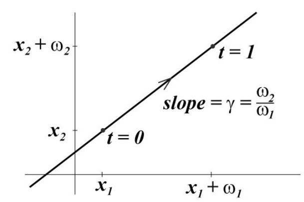


FIGURE 4. A translational flow in the plane.

Another way to see this is to go back to the plane and consider the equivalence relation given by the exponential map on each coordinate. The set of equivalence classes are given by the unit square in the plane, under the idea that the left side of the square (the side on the $x_1 = 0$ line) and the right side (the $x_1 = 1$ side) are considered the same points (this is the $0 = 1$ idea of the circle identification). Similarly, the top and bottom of the square are to be identified. Then the flow line at the origin under the ODE system is a straight line of slope γ emanating from the origin and meeting the right edge of the unit square at the point $(1,\gamma)$. But by the identification, we can restart the graph of the line at the same height on the left side of the square (at the point $(0,\gamma)$. Continuing to do this, we will eventually reach the top of the square. But by the identification again, we will drop to the bottom point and continue the line as before. See Figure 6 In essence, we are graphing the flow line as it would appear on the unit square. When we pull this square out of the plane and bend it to create our torus $\mathbb{T}$, the flow line will come with it. Suppose $\gamma \notin \mathbb{Q}$. What can we say about the positive flow line?

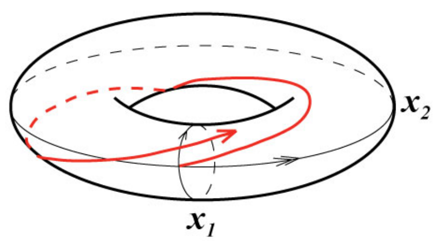


FIGURE 5. The flow line of the origin in $\mathbb{T}$.

PROPOSITION 4.18. *if* $\gamma = \frac{\omega_2}{\omega_1}$ *is irrational, then the flow is minimal. If* $\gamma \in \mathbb{Q}$*, then every orbit is closed.*

PROOF. Choose a point $\overline{\mathbf{x}} = (\overline{x}_1, \overline{x}_2)$ and let $S_{\overline{\mathbf{x}}}$ be a Poincare section (a first-return map for the flow.) Then, along $S_{\overline{\mathbf{x}}}$, the Poincare map is precisely R_γ (see the left side of Figure 7). Since $\gamma \notin \mathbb{Q}$, $\mathcal{O}_{\overline{\mathbf{x}}} \cap S_{\overline{\mathbf{x}}}$ is dense in $S_{\overline{\mathbf{x}}}$.

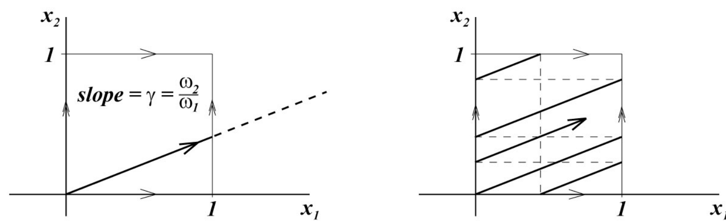


FIGURE 6. Wrapping the flow onto $\mathbb{T}$, the unit square with opposite sides identified.

So let $\overline{\mathbf{y}}$ be some arbitrary point in $\mathbb{T}^2$. We claim that arbitrarily close to $\overline{\mathbf{y}}$ is a point in $\mathcal{O}_{\overline{\mathbf{x}}}$. Indeed, choose $\epsilon > 0$ and construct an open ϵ-neighborhood of $\overline{\mathbf{y}}$ in $\mathbb{T}^2$, $N_\epsilon(\overline{\mathbf{y}})$. Then $\exists \overline{\mathbf{z}} \in S_{\overline{\mathbf{x}}}$ and a $t > 0$, such that $\overline{\mathbf{y}} = \boldsymbol{\omega} t + \overline{\mathbf{z}}$ But then $\exists \overline{\mathbf{u}} \in \mathcal{O}_{\overline{\mathbf{x}}} \cap S_{\overline{\mathbf{x}}}$, where $d(\overline{\mathbf{u}}, \overline{\mathbf{z}}) < \epsilon$. And then $\overline{\mathbf{v}} = \boldsymbol{\omega} t + \overline{\mathbf{u}} \in N_\epsilon(\overline{\mathbf{y}})$ and $\overline{\mathbf{v}} \in \mathcal{O}_{\overline{\mathbf{x}}}$. See the right side of Figure 7

REMARK 4.19. For a Poincare first return map, if every orbit intersects the Poincare section, we call the section a *global section*. Otherwise, it is a local section.

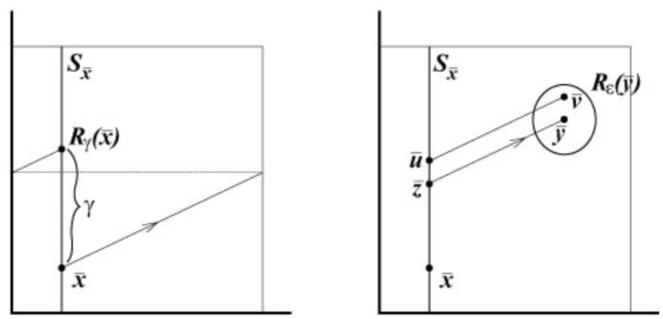


FIGURE 7. The orbit of $\overline{\mathbf{x}}$ is dense in $\mathbb{T}^2$.

Now let $\gamma = \frac{\omega_2}{\omega_1} \in \mathbb{Q}$. Then R_γ is a rational rotation on $S_{\overline{\mathbf{x}}}$. Then $\exists\ t_0 > 0$ such that $\overline{\mathbf{x}} = \boldsymbol{\omega} t_0 + \overline{\mathbf{x}} \mod 1$. Thus $\boldsymbol{\omega} t_0 = 0 \mod 1$ and $\mathcal{O}_{\overline{\mathbf{x}}}$ is closed. But then for any other point $\overline{\mathbf{y}} \in \mathbb{T}^2$, we have $\overline{\mathbf{y}} - \overline{\mathbf{x}} = \overline{\mathbf{z}} \mod 1$. Thus $\overline{\mathbf{x}} = \overline{\mathbf{y}} - \overline{\mathbf{z}} \mod 1$ and $\overline{\mathbf{y}} - \overline{\mathbf{z}} = \boldsymbol{\omega} t_0 + \overline{\mathbf{y}} - \overline{\mathbf{z}}$ so that $\overline{\mathbf{y}} = \boldsymbol{\omega} t_0 + \overline{\mathbf{y}} \mod 1$. Hence $\mathcal{O}_{\overline{\mathbf{y}}}$ is closed. $\square$

EXERCISE 121. Draw enough of the flow lines passing through the origin to indicate general behavior for the following values of γ:

$$\textbf{(a)}\ \gamma = 1, \quad \textbf{(b)}\ \gamma = 2, \quad \textbf{(c)}\ \gamma = \frac{1}{2}, \quad \textbf{(d)}\ \gamma = \frac{1+\sqrt{5}}{2}, \quad \textbf{(e)}\ \gamma = \frac{1-\sqrt{5}}{2}.$$

EXERCISE 122. Show that a linear map on the real line $f(x) = kx$ corresponds to a continuous map on S^1 iff k is an integer. Graph the circle map on the unit interval (with the endpoints identified) that corresponds to $f(x) = 3x$. Identify all fixed points. Can you find a period-3 orbit? Hint: Remember that two points in the real line correspond to the same point on the circle is their distance apart is an integer. This is the parameterization that we will always refer to by default.

EXERCISE 123. Let $h : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear map so that $h(\mathbf{x}) = A\mathbf{x}$, where $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $a, b, c, d \in \mathbb{Z}$. Do the following:

(a) Show h induces a map on the standard two torus $\mathbb{T} = S^1 \times S^1$. Hint: Two vectors in the plane are in the same equivalence class on the torus (correspond to the same point on $\mathbb{T}$ if they differ by a vector with integer entries.)

(b) Describe, as best as you can, the dynamics of the linear maps on $\mathbb{T}$ corresponding to

$$\textbf{i}.\ A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad \textbf{ii}.\ A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad \textbf{iii}.\ A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}.$$

4.2.1. Application: Lissajous Figures. We can look at toral flows in another way: Via an embedding of a torus into $\mathbb{R}^4$ with simultaneous rotations in the two orthogonal coordinate planes. To see this, think of $S^1 \in \mathbb{R}^2$ as a circle of radius r centered at the origin. Then we can represent $\mathbb{T}$ as the set

$$\mathbb{T} = \left\{ (x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \middle| x_1^2 + x_2^2 = r_1^2,\ \ x_3^2 + x_4^2 = r_2^2 \right\}.$$

Now recall a continuous rotation in $\mathbb{R}^2$ is given by the linear ODE system $\dot{\mathbf{x}} = B_\alpha \mathbf{x}$, where B is the matrix $\begin{bmatrix} 0 & \alpha \\ -\alpha & 0 \end{bmatrix}$ whose eigenvalues $\pm\alpha i$ are purely imaginary. Do this for each pair of coordinates (each of two copies of $\mathbb{R}^2$) to get the partially uncoupled system of ODEs on $\mathbb{R}^4$,

$$\dot{\mathbf{x}} = A\mathbf{x}, \quad \begin{bmatrix} \dot{x_1} \\ \dot{x_2} \\ \dot{x_3} \\ \dot{x_4} \end{bmatrix} = \begin{bmatrix} 0 & \alpha_1 & 0 & 0 \\ -\alpha_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \alpha_2 \\ 0 & 0 & -\alpha_2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}.$$

We will eventually see that this is the model for the spherical pendulum.For now, Figure 8 shows two planes, which under uncoupled rotations would leave concentric circles invariant. Joining these two planes only at the origin comprises $\mathbb{R}^4$ (hard to visualize, no?) The 2-torus $\mathbb{T}^2$ is a subspace of $\mathbb{R}^2$ invariant under this flow.

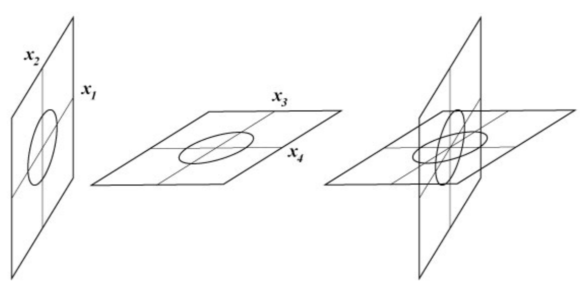


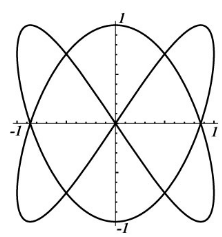
FIGURE 8. The x_1x_2 and x_3x_4-planes as subspaces of $\mathbb{R}^4$,

Some notes:

- The two circles $x_1^2 + x_2^2 = r_1^2$ and $x_3^2 + x_4^2 = r_2^2$ are invariant under this flow.
- We can define angular coordinates on $\mathbb{T}$ via the equations
$$x_1 = r_1 \cos 2\pi\varphi_1 \quad x_2 = r_1 \sin 2\pi\varphi_1$$
$$x_3 = r_2 \cos 2\pi\varphi_2 \quad x_4 = r_2 \sin 2\pi\varphi_2.$$
Then, restricted to these angular coordinates and with $\omega_i = \frac{\alpha_i}{2\pi}$, $i = 1, 2$, we recover
$$\dot{\varphi}_1 = -\omega_1, \quad \dot{\varphi}_2 = -\omega_2.$$
Motion is independent along each circle, and the solutions are $\varphi_i(t) = \omega_i(t - t_0)$.
- If $\frac{\alpha_2}{\alpha_1} = \frac{\omega_2}{\omega_1} \notin \mathbb{Q}$, then the flow is minimal.

Exercise 124. Do the change of coordinates explicitly to show that these two interpretations of linear toral flows are the same.

Now, for a choice of $\boldsymbol{\omega}$ and $r_1 = r_2 = 1$, project a solution onto either the (x_1, x_3) or the (x_2, x_4)-planes. The resulting figure is a plot of a parameterized curve whose two coordinate functions are cosine (resp. sine) functions of periods which are rationally dependent iff $\boldsymbol{\omega}$ is rational. In this case, the figure is closed, and is called a Lissajous figure. See the figure below for the case of two sine functions (projection onto the (x_2, x_4)-plane, in this case), where $\omega_1 = 2$ and $\omega_2 = 3$.



$\omega_1 = 2, \quad \omega_2 = 3$

Q. What would the figure look like if ω_1 and ω_2 were not rational multiples of each other?

Exercise 125. Draw the Lissajous figure corresponding to the x_2x_4-planar projection of the toral flow when $\omega_1 = 2$ and $\omega_2 = 5$. For these same values, draw the orbit also on the torus using the unit square in $\mathbb{R}^2$ representation (with sides identified appropriately), and as well on the "surface of a doughnut" representation in $\mathbb{R}^3$.

A nice physical interpretation of this curve is as the trajectory of a pair of uncoupled harmonic oscillators, given by
$$\ddot{x_1} = -\omega_1 x_1$$
$$\ddot{x_2} = -\omega_2 x_2.$$

Toral flows also show up as a means to study a completely different class of dynamical system; a billiard. We will eventually focus on some more general features of this class of dynamical systems. For now, we will introduce a particularly interesting example, where a linear flow on a torus allows us to answer a rather deep question.

4.2.2. Application: A polygonal Billiard. Consider the unit interval $I = [0,1]$ with two point masses x_1 and x_2, with respective masses m_1 and m_2 respectively, free to move along I but confined to stay on I. Without outside influence, these point masses will move at a constant, initial velocity. Eventually, they will collide with each other and with the walls (see left side of Figure 9). Assume also that theses collisions are elastic, with no energy absorption or loss due to friction. Here, elastic means that, upon a wall collision, a point mass' velocity will only switch sign. And upon a point mass collision, the two point masses will exchange velocities. For now, assume that $m_1 = m_2 = 1$.

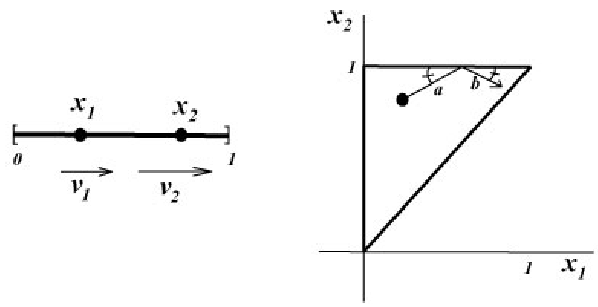


FIGURE 9. The state space (right) of two point masses in the unit interval (left).

The state space in $\mathbb{R}^2$ is

$$T = \left\{ (x_1, x_2) \in \mathbb{R}^2 \middle| 0 \leq x_1 \leq x_2 \leq 1 \right\}.$$

Here, T is the region in the unit square above the diagonal line which is the graph of the identity map on I (right side of Figure 9). The edges of the region T are included; since the point masses have no size, they can occupy the same position at the point of contact. An interesting question to ask yourself is: How does the state space change if the point masses had size to them?

Now, given an initial set of data, with initial positions and velocities v_1 and v_2, respectively, what is the evolution of the system? The answer lies in the study of these types of dynamical systems called billiards. Evolution will look like movement in T. A point in T comprises the simultaneous positions of the two particles, and movement in T will consist of a curve parameterized by time t. The idea is that this curve will be a line since the two velocities are constants. The slope of this line (in Figure 9, line a is the trajectory before any collisions have happened), will be $\frac{v_2}{v_1}$. (why?) Once a collisions happens, though, this changes. There are two types of collisions: Assuming that $\frac{v_2}{v_1}$ is the ratio of the velocities of the two point masses before a collision, we have

- When a point mass hits a wall, it "bounces off", traveling back into I with equal velocity and of opposite sign. Thus the new velocity is $-\frac{v_2}{v_1}$ (This is the slope of line b in Figure 9 above).
- When the two point masses collide, they exchange their velocities (really, think of billiard balls here). Thus the new velocity is $\frac{v_1}{v_2}$. Caution: This reciprocal velocity is NOT the slope of a perpendicular line, which would be the negative reciprocal.

Envision these collisions in the diagram and the resulting trajectory curves before and after each type of collision, as in the figure. What you see are perfect rebounds off of each of the three walls, where the angle of reflection equals the angle of incidence. An ideal billiard table, although one with no pockets. Which leads to the obvious question: What happens if a trajectory heads straight into a corner? For now, we will accept the stipulation that

- When the two point masses collide with a wall simultaneously, either at separate ends of I or at the same end, both velocities switch sign. While this will not change the slope of the trajectory, it will change the direction of travel along that piece of trajectory line.

Some questions to ask:

Q. Can there exist closed trajectories?
Q. Can there exist a dense orbit?
Q. The orbits of points in T will very much intersect each other and many trajectories will intersect themselves also. The phase space will get quite messy. Is there a way to better "see" the orbits of points more clearly?

The answer to the last question is yes, although this table is fairly special. Here, one can "unfold" the table:

- Think of the walls of T as mirrors. When a trajectory hits a wall, it rebounds off in a different direction. However, its reflection in the mirror simply continues its straight line motion. Think of a reflected region T across this wall. The trajectory looks to simply pass through the wall and continue on, as in Figure 10 below.
- Envision each collision that follows also via its reflection. Motion continues in a straight line fashion through each mirrored wall. By continuing this procedure, the motion will look linear for all forward time, no?
- This idea works because this particular triangle, under reflections, will eventually cover the plane in a way that only its edges overlap and all points in $\mathbb{R}^2$ are covered by at least one triangle. This is called a (regular) *tessellation* of the plane by T, and works only because T has some very special properties. See below.
- The unfolded trajectory is called a linear flow on the billiard table $\mathbb{R}^2$.

So what does a billiard flow in $\mathbb{R}^2$ look like? Obviously, it is just straight line motion at a slope $\frac{v_2}{v_1}$ forever since there are no collisions. The better question to ask is: What does this tell us about the original flow on the triangle T?

By continually unfolding (reflecting) the table T, on starts to notice that there

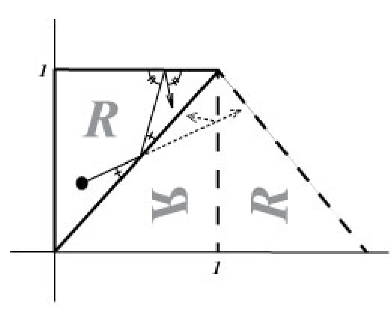


Figure 10. Starting to unfold the triangular billiard table.

are only 8 different configurations: the four orientations of T given by rotations by multiples of $\frac{\pi}{2}$ radians, and the reflection of each. If you collect up a representative of each of these configurations into a connected region, you wind up with enough information to characterize the entire flow in $\mathbb{R}^2$: Each time your $\mathbb{R}^2$ linear flow re-enters a region of a particular configuration of T, you can simply note the trajectory in your representative of that region. This region of representative configurations is called a *fundamental domain* for the flow. One such fundamental domain of this flow is the square of side length 2 in Figure 11. Noting the configurations, as the trajectory leaves the square, it enters a configuration exactly like that at the other side of the square. One can see the trajectory then re-enter the square from the other side. Similarly, when one leaves the square at the top, it enters a configuration represented at the bottom of the square. Thus one can continue the trajectory as if it had re-entered the square at the bottom.

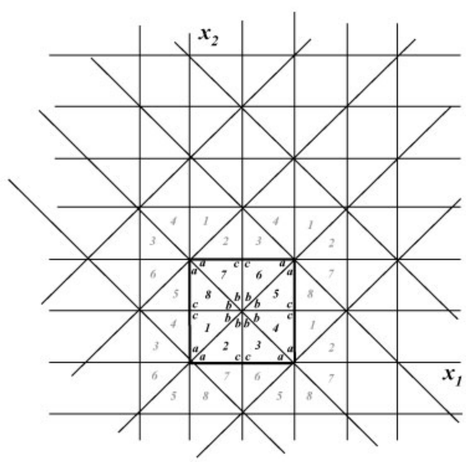


FIGURE 11. A fundemental domain in the fully unfolded table.

Note: There was a famous arcade video game called Asteroids, an Atari, Inc., game released in 1979, where a space ship was planted in the middle of a square screen. It could turn but not move. Various boulders (asteroids) would float in and out of the screen. Should an asteroid hit the ship, the game is over. The ship can fire a weapon at an asteroid, and if hit, would break into two smaller ones, which would go off in different directions. The asteroids (or pieces of asteroids) always traveled in a straight line. And as an asteroid left the screen, it would always reappear on the opposite side and travel in the same direction. Really, the asteroids were only exhibiting a linear toral flow. Who would have though that in playing this game, one was actually playing in a universe which was not the plane at all but rather the torus $\mathbb{T}$? See Figure 12.

Hence linear flows on $\mathbb{R}^2$ again look like toral flows on this fundamental domain, which comprises the space of configurations of T as one uses T to tile $\mathbb{R}^2$. So what do linear toral flows say about the trajectories on T?

PROPOSITION 4.20. *If the ratio of initial velocities $\frac{v_2}{v_1} \in \mathbb{Q}$, then the orbit is closed (on $\mathbb{T}$ and thus also on T). If $\frac{v_2}{v_1} \notin \mathbb{Q}$, then the orbit is dense in T.*

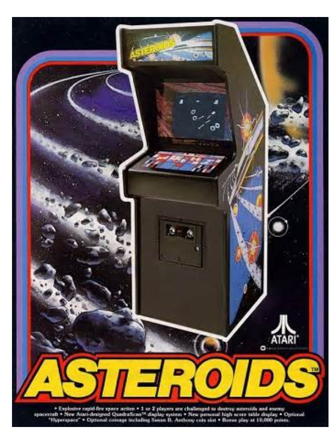


FIGURE 12. Space

Note: Now, it is easier to see what a collision in a corner will look like. Like I said, this table T is quite special in many ways. These ways do not generalize well. However, with T we can say much more:

PROPOSITION 4.21. *For any starting set of data (point mass positions and velocities), the trajectory will assume at most* 8 *different velocity ratios.*

Count them: There are two possible ratio magnitudes, each with two signs. That makes 4. But travel along the lines of each of these slopes can be in each of the two directions due to the reflected configurations in the fundamental domain.

How can one generalize these results to other tables:

- **Unequal masses.**
  - An elastic collision between unequal masses will not result in what would look like a reflection off of the diagonal wall in T. One could certainly accurately chart the collision as a change in direction off of the wall. However, when unfolding the table, the resulting flow in $\mathbb{R}^2$ will not be linear (each reflected trajectory through the diagonal wall will be a change in direction in the planar flow. You will see a piecewise linear flow in $\mathbb{R}^2$ and hence also on the fundamental domain $\mathbb{T}$. While this is workable, it is not such as easy leap to a conclusion.
  - One can also actually change the table. Use momenta to define the collision between the point masses, and alter the diagonal wall to be a perfect reflective wall. The resulting will not be linear. The new table will not tile the plane anymore, but in many cases the unfolded table will cover the plane with many holes (the reflecting curve will be concave, so will fit into the original T. The unfolded flow will look liner until it hits a hole, where it will reflect through he hole perpendicularly through its center axis and appear on the other side to continue at the same slope. I haven't worked out the details here (and a hat tip to Jonathan Ling who started to work on this idea), but there should be results here that are similar to the original table T, as long as one is careful with the analysis.This needs to be worked out in detail.
- **Other tessellations of $\mathbb{R}^2$.** It is easy to see that some shapes tessellate the plane while others do not. For regular polygons, only triangles, squares or hexagons tessellate the plane. Rectangles, and a few other non-regular triangles also work fine. Work out some good examples here. But examples are fairly rare. And in each case, one would need to find a fundamental domain and then interpret the resulting flow on that domain in terms of the original flow as well as that on the place. All good stuff, and are the initial ways one may study polygonal billiards. However, later, we will generalize our analysis of billiards in a completely different direction.

4.2.3. Application: The Kepler Problem. Scholarpedia[Celestial Mechanics] One more application of linear toral flows: The Kepler Problem: Consider two point masses moving in an inverse square gravitational field. Assume that they do not interact or influence each other. Then their equations of motion are second order homogeneous (and separable). Total energy is conserved (this is actually the solution function when using the method one is taught in ODEs to solve separable equations) as is total angular momentum. Hence flow is planar, and confined to an ellipse (hence each point mass has periodic flow of some period independent of the other). For two particles, the flow would be confined to a torus again (two separate periodic flows, although in this case, the flow would look a bit more like that of the flow in $\mathbb{R}^4$ above. The flow would be linear in a space where momenta is used instead of velocity. Then, one can easily say whether the two point mass system will ever reach its starting positions simultaneously again based on whether the ratio of the momenta is rational or not. Pretty easy result for such a complicated system. Two other thoughts:

- **Q.** Can you now see the similarity between the Kepler Problem and the HW assignment I gave you concerning the rotation of the earth and the lunar rotation?
- **Q.** What would the Kepler Problem look like for three point masses in the same field? Where would the resulting flow reside? Can we make a concluding statement about the flow in such an easy way as for only two point masses?

We will return to the last questions in short order. But first, let's move from continuous dynamical systems on a torus to discrete dynamical systems. Some surprising relationships occur between flows on a torus and the corresponding time-t maps.

4.3. Toral translations

Like Euclidean space $\mathbb{R}^n$, one can generalize the construction of the 2-torus $\mathbb{T} = S^1 \times S^1$ by considering a system of equations involving more than two angular coordinates; The n-dimensional torus, or the n-torus, denoted $\mathbb{T}^n$ is simply the n-fold product of n circles

$$\mathbb{T}^n = \overbrace{S^1 \times \cdots \times S^1}^{n \text{ times}}.$$

Then

$$\mathbb{T}^n = \mathbb{R}^n / \mathbb{Z}^n = \mathbb{R}/\mathbb{Z} \times \cdots \times \mathbb{R}/\mathbb{Z}.$$

Recall the Kepler Problem. With n point masses, the resulting flow may be seen as linear motion on $\mathbb{T}^n$.

Another way to view the n-torus is via an identification within $\mathbb{R}^n$. Remember the unit square with it opposite sides identified plays a good model for the 2-torus, $\mathbb{T} = \mathbb{T}^2$. The generalization works well here for all the natural numbers. Take the unit cube in $\mathbb{R}^3$. Picture here? Identify each of the opposite pairs of sides, squares in this case (think of a die, and identify two sides if their numbers add up to 7). The resulting model is precisely the $\mathbb{T}^3$. This works well if one wants to watch a linear flow on $\mathbb{T}^3$. Simply allow the flow to progress in the unit cube, and whenever one hits a wall, simply vanish and reappear on the opposite wall, entering back into the cube. See Figure 13. Note that here the origin is at the lower left corner, intersections with $\mathcal{O}_{(0,0,0)}$ with the top and bottom are in red, while intersections at the back and front walls are in blue. This represents a periodic orbit. Can you see this?

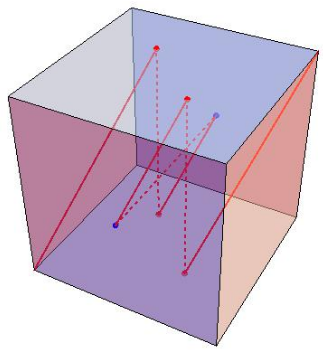


FIGURE 13. A periodic orbit of the translation in $\mathbb{T}^3$.

Note this also works well for $n = 1$: Take the unit interval and identify its two sides (the numbers $x = 0$ and $x = 1$). This is what I mean by the phrase $0 = 1$ on S^1, where the circle is the 1-torus.

Now, the vector exponential map $(\theta_1, \ldots, \theta_n) \overset{\exp}{\longmapsto} \left(e^{2\pi i\theta_1}, \ldots, e^{2\pi i\theta_n}\right)$ maps $\mathbb{R}^n$ onto $\mathbb{T}^n$. We can define a (vector) rotation on $\mathbb{T}^n$ by the vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)$, where $R_{\boldsymbol{\alpha}}(\mathbf{x}) = (x_1 + \alpha_1, \ldots, x_n + \alpha_n) = \mathbf{x} + \boldsymbol{\alpha}$. Normally, this is called a *translation* (by $\boldsymbol{\alpha}$) on the torus. In Figure 13, $\boldsymbol{\alpha} = (1, 2, 3)$, and only the orbit of the origin is shown. Note that it should be obvious that if all of the α_i's are rational, then the resulting map on $\mathbb{T}^n$ will have closed orbits. Questions to ask are: Are theses the only maps of this type that are linear? If one or more α_i's are not rational, can there still be periodic orbits? And if there cannot, are the non-periodic orbits dense in the torus?

Consider the linear flow in $\mathbb{T}^n$ whose time-1 map is $R_{\boldsymbol{\alpha}}$. This would be the flow whose ith-coordinate solution i $x_i(t) = x_i + \alpha_i t$. Again, with ALL of the α_i's rational, the flow would have all closed orbits. Now allow one of the coordinate rotation numbers to be irrational. We saw how it was the ratio of the two flow rates that determined whether the flow had closed orbits on $\mathbb{T}^2$. Does this hold in higher dimensions? Do the properties of the time-1 map still reflect accurately the properties of the flow? Does the irrationality of some or all of the coordinate rotations imply minimality of the map? of the flow? Really, all of these questions will rely on a good notion of measuring the relative ratios of the individual pairs of map rotations and flow rates. And how do we define these ratios in higher dimensions? By a notion of the rational independence of sets of numbers:

DEFINITION 4.22. A set of n real numbers $\{\alpha_i\}_{i=1}^n$ is said to be *rationally independent* if, given $k_1, \ldots, k_n \in \mathbb{Z}$, the only solution to

$$k_1\alpha_1 + \ldots + k_n\alpha_n = 0$$

is for $k_1 = \cdots = k_n = 0$.

Another way to say this is the following: For all nontrivial integer vectors $\mathbf{k} = (k_1, \ldots, k_n) \in \mathbb{Z}^n - \{\mathbf{0}\}$,

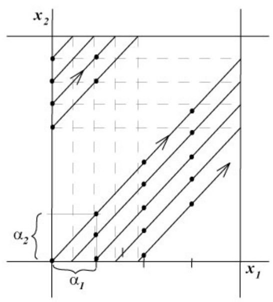
$$\sum_{n=1}^{n} k_i\alpha_i = \mathbf{k} \cdot \boldsymbol{\alpha} \neq 0.$$

We have the following:

PROPOSITION 4.23. *A toral translation on $\mathbb{T}^n$, given by $R_{\boldsymbol{\alpha}}$ is minimal iff the numbers α_1, …, α_n, 1 are rationally independent.*

PROPOSITION 4.24. *The flow on $\mathbb{T}^n$ whose time-1 map is the translation $R_{\boldsymbol{\alpha}}$ is minimal iff the numbers $\alpha_1, \ldots, \alpha_n$ are rationally independent.*

Do you see the difference? One way to view this is to restrict to the case of a 2-torus. Here, the second proposition says that the flow will be minimal if, in essence, $k_1\alpha_1 + k_2\alpha_2 = 0$ is only satisfied when $k_1 = k_2 = 0$. Really, if there were another solution, then it would be the case that $\frac{\alpha_2}{\alpha_1} = \frac{k_1}{k_2} \in \mathbb{Q}$.



On the other hand, the first proposition indicates that both α_1 and α_2 need to be rationally independent and also both rationally independent from 1! That means that not only do the two α's need to be rationally independent from each other, but neither α_1 nor α_2 can be rational (then it would be a rational multiple of 1). Hence a flow can be minimal on a torus, while the time-1 map isn't. Why is this so? Let's study the situation via an example.

EXAMPLE 4.25. On the two torus, let $\alpha_1 = \frac{1}{4}$ and $\alpha_2 = \frac{\pi}{16}$. The flow will be minimal here since $\frac{\alpha_2}{\alpha_1} = \frac{\pi}{4} \notin \mathbb{Q}$ (α_1 and α_2 are rationally independent). However, the time-1 map of this flow is $R_{\vec{\alpha}}$, and since

$$k_1 \cdot \frac{1}{4} + k_2 \cdot \frac{\pi}{16} + k_3 \cdot 1 = 0 \text{ is solved by } \begin{bmatrix} k_1 & k_2 & k_3 \end{bmatrix} = \begin{bmatrix} 4 & 0 & -1 \end{bmatrix},$$

the translation will not be minimal (the orbits are not dense in the torus). The fact that α_1 is already rational is the problem. The figure will tell the story. Essentially, the orbit coordinates of the translation in the x_1 direction will only take the values $0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}$, while the x_2-coordinates will "fill out" the vertical direction. The result is that the orbit of the translation will only be dense on the vertical lines corresponding to the x_1-coordinates of the orbit. This sits in contrast to the flow, in which every orbit will "fill" the torus.

EXERCISE 126. Given the flow in Example 4.25, show that on the global Poincare section corresponding to the set $S_{\frac{1}{2}} = \left\{\left(\frac{1}{2}, y\right) \in \mathbb{T}^2 \,\middle|\, y \in S^1\right\}$, the first return map is minimal.

Place some exercises here.

4.4. Invertible S^1-maps.

Let's return to maps on the circle, and try to gain more general information than by using simply rigid rotations. Again, think of S^1 as the identification space $S^1 = \mathbb{R}/\mathbb{Z}$, given by the level sets of the map

$$\pi : \mathbb{R} \to S^1, \quad \pi(x) = [x].$$

One easy way to think about $[x]$ is to simply take any real number and disregard the integer part. Thus $[2.13] = .13$, and $[e] = e - 2$. We note here that π is an example of a projection of $\mathbb{R}$ onto S^1:

DEFINITION 4.26. A map $f : X \to Y$ is called a *projection* if $\forall x \in X$, $f(x) = f^2(x)$. That is, if f equals its square (composition with itself.)

Note that a map or an operation that doesn't change the effect on inputs upon repeated application after the initial application is called *idempotence*. Think absolute value, for another example.

Proposition 4.27. *For any continuous map $f : S^1 \to S^1$, there exist an associated map $F : \mathbb{R} \to \mathbb{R}$, called a* lift *of f to $\mathbb{R}$, where*

$$f \circ \pi = \pi \circ F, \qquad \textit{equivalently } f([x]) = [F(x)].$$

Some Notes:

- One way to see this is via the commutative diagram

$$\begin{tikzcd} \mathbb{R} \arrow[r, "F"] \arrow[d, "\pi"'] & \mathbb{R} \arrow[d, "\pi"] \\ S^1 \arrow[r, "f"] & S^1 \end{tikzcd}$$

- The lift F is unique up to an additive constant (sort of like how the antiderivative of a function is unique only up to an additive constant, right?)
- The quantity
$$\deg(f) = F(x+1) - F(x)$$
is well-defined for all $x \in \mathbb{R}$ and is called the *degree* of f.
- If f is a homeomorphism, then $|\deg(f)| = 1$.
- The structure of F is quite special. It looks like the sum of a periodic function with the line $y = (\deg(f))\, x$. This is due to the structure of the projection π.

So just how much information about f can we learn by the study of the lifts of f? Certainly, maps on $\mathbb{R}$ are fairly easy to study (this is what the calculus is really all about, right? Although, in mathematics calculus is just what we call analysis.) And maps with the structure of the lifts F may be easier still. If we can use these lifts to say fairly general things about how an f may behave, this would be quite important. For example, this quantity $\deg(f)$ is defined solely by a choice of lift f. We will see just what information $\deg(f)$ conveys. For a moment, let's first take a look at why some of the assertions we just made are true.

- *Lifts always exist.* Simply construct one using the definition. This will be an exercise.
- *F is unique up to a constant.*

  Proof. Suppose $\overline{F}$ is another lift. Then
  $$\left[\overline{F}(x)\right] = f([x]) = [F(x)], \quad \forall x \in \mathbb{R}.$$
  This is just another way of saying that $\pi \circ \overline{F} = f \circ \pi = \pi \circ F$, $\forall x \in \mathbb{R}$. But then $\overline{F} - F$ ia <u>always</u> an integer! (why?) But $\overline{F} - F$ is the difference between two continuous functions, and hence is continuous. But a continuous function on $\mathbb{R}$ that take values in the integers is necessarily constant. □

- $\deg(f)$ *is well defined.*

  Proof. Here $\deg(f) = F(x+1) - F(x)$ is a continuous function on $\mathbb{R}$ that takes values in the integers (it must, due to the projection π and the commutativity of the diagram defining F). Thus it also is a constant for all $x \in \mathbb{R}$. □

- *If f is a homeomorphism, then* $|\deg(f)| = 1$.

Proof. Suppose that $|\deg(f)| > 1$. Then $|F(x+1) - F(x)| > 1$. And since $F(x+1) - F(x)$ is continuous, by the Intermediate Value Theorem, $\exists y \in (x, x+1)$ where $|F(y) - F(x)| = 1$. But then $f([y]) = f([x])$ for some $y \neq x$. Thus f cannot be one-to-one and hence cannot be a homeomorphism.

Now suppose that $|\deg(f)| = 0$. Then $F(x+1) = F(x)$, $\forall x$, and hence F is not one-to-one on the interval $(x, x+1)$. But then neither is f, and again, f cannot be a homeomorphism. □

- $F(x) - x\deg(f)$ *is periodic.*

Proof. It is certainly continuous (why?) To see that it is periodic (of period-1), simply evaluate this function at $x+1$:

$$\begin{aligned} F(x+1) - (x+1)\deg(f) &= (F(x) + \deg(f)) - (x+1)\deg(f) \\ &= F(x) - x\deg(f). \end{aligned}$$

□

Example 4.28. Let $f(x) = x$. Picture? This is the "identity" map on S^1, since all points are taken to themselves. A suitable lift for f is the map $F(x) = x$ on $\mathbb{R}$. To see this, make sure the definition works. Question: Are there any other lifts for f? What about the map $\overline{F}(x) = x + a$ for a a constant? Are there any restrictions on the constant a? The answer is yes. For a to be an acceptable constant, we would need the definition of a lift of be satisfied. Thus

$$\left[\overline{F}(x)\right] = [x + a] = f([x]) = [x].$$

So the condition that a must satisfy is $[x + a] = [x]$ on S^1. Hence, $a \in \mathbb{Z}$.

Exercise 127. For a real number $a \notin \mathbb{Z}$, can $\overline{F}(x) = x + a$ serve as a lift of a circle map? What sort of circle map?

Exercise 128. Find a suitable lift $F : \mathbb{R} \to \mathbb{R}$ for the rotation map $R_\alpha : S^1 \to S^1$ where $\alpha = 2\pi$ and verify that is works. Graph both F and R_α. Keep in mind that we are using $S^1 = \mathbb{R}/\mathbb{Z}$ as our model of the circle.

Example 4.29. Let $f(x) = x^n$. Thinking of x as the complex number $x = e^{2\pi i\theta}$, for $\theta \in \mathbb{R}$, then

$$f(x) = f\left(e^{2\pi i\theta}\right) = \left(e^{2\pi i\theta}\right)^n = e^{2\pi i(n\theta)}.$$

Hence a suitable lift is obviously $F(x) = nx$ (I say obviously, since the variable in the exponent <u>is</u> the lifted variable!) Question: This is a degree n map. For which values of n does the map f have an inverse" And what does the map f actually do for different values of n?

Example 4.30. Let f be a general degree-r map. Then $F(1) - F(0) = r = \deg(f)$. Suppose that $F(0) = 0$. Then $F(1) = r$ and if, for example, $r > 1$, it is now easy to see that there will certainly be a $y \in (0,1)$, where $F(y) = 1$. This was a fact that we used in the proof above to show that f cannot be a homeomorphism Draw a picture. In this case, where $r > 1$, at every point in $y \in (0,1)$ where $F(y) \in \mathbb{Z}$, we will have $\pi \circ F(y) = [F(y)] = 0$ on S^1. This won't happen when $r = 1$. When $r = 0$, the map F will be periodic, which is definitely not one-to-one. Question: What happens when $r < 0$? Draw some representative examples to see.

DEFINITION 4.31. Suppose that $f : S^1 \to S^1$ is invertible. Then

(1) if $\deg(f) = 1$, f is orientation preserving.
(2) if $\deg(f) = -1$, f is orientation reversing.

Recall from your study of vector calculus that orientation is a choice of direction in the parameterization of a space (really, it exists outside of any choice of coordinates on a space, but once you put coordinates on a space, you have essentially chosen an orientation for that space. This is true at least for those spaces that actually are orientable, that is (Moebius Band?) On $\mathbb{R}$, it is the choice of direction for the symbol ">". On a surface, it is a choice of side. In $\mathbb{R}^3$, one can use the Right Hand Rule. Etcetera. On S^1, orientation preserving really means that after one applies the map, points to the right of a designated point remain on that side. Orientation reversing will flip a very small neighborhood of a point.

Circle maps may or may not have periodic points. And given an arbitrary homeomorphism, without regard to any other specific properties of the map, one would expect that we can construct maps with lots of periodic points of any period. However, it turns out that circle homeomorphisms are quite restricted. because they must be one-to-one and onto, only certain things can happen. To explain, we will need another property of circle homeomorphisms to help us.

PROPOSITION 4.32. *Let $f : S^1 \to S^1$ be an orientation preserving homeomorphism, with $F : \mathbb{R} \to \mathbb{R}$ a lift. Then the quantity*

$$\rho(F) := \lim_{|n| \to \infty} \frac{F^n(x) - x}{n}$$

(1) exists $\forall x \in \mathbb{R}$,
(2) is independent of the choice of x and is defined up to an additive integer, and
(3) is rational iff f has a periodic point.

Given these qualities, the additional quantity $\rho(f) = [\rho(F)]$ is well-defined and is called the *rotation number* of f.

Some notes:

- This quantity and this proposition were proposed and proved by Henri Poincare back in the 1880's.
- $\rho(f)$ is also sometimes called the *map winding number*, although it is different from the winding number used in algebraic topology or complex analysis. Be careful here.
- $\rho(R_\alpha) = [\alpha]$.

  EXERCISE 129. Use this definition to calculate $\rho(R_\alpha)$ for the circle rotation R_α, $\alpha \in (0, 1)$.

- ρ represents in a way the average rotation of points in a circle homeomorphism.

It turns out that the rotation number is a very telling property of a circle homeomorphism. And like interval maps, there circle homeomorphisms can only do certain things:

PROPOSITION 4.33. *If $\rho(f) = 0$, then f has a fixed point.*

Another way of saying that if there is no average rotation of the circle map, then somewhere a point doesn't move under f. This is like the Intermediate value Theorem on a closed, bounded interval of $\mathbb{R}$ where a map is positive at one end point and negative at the other.

PROOF. We prove the contrapositive. Assume a circle homeomorphism $f : S^1 \to S^1$ has no fixed points. Then, given a lift $F : \mathbb{R} \to \mathbb{R}$, where $F(0) \in [0,1)$, if $F(x) - x \in \mathbb{Z}$ for any $x \in \mathbb{R}$, then f must have a fixed point (do you see this?) Hence $\forall x \in \mathbb{R}$, $F(x) - x \notin \mathbb{Z}$. But as a function on $\mathbb{R}$, the function $F(x) - x$ is continuous, and hence by the Intermediate Function Theorem,

$$0 < F(x) - x < 1, \quad \forall x \in \mathbb{R},$$

and on the interval $[0,1]$, $F(x) - x$ must achieve its maximum and minimum by the Extreme Value Theorem. SO there must exist a constant m, where

$$0 < m \le F(x) - x \le 1 - m < 1.$$

But by above, $F(x) - x$ is periodic, and hence this last inequality holds on all of $\mathbb{R}$. Choosing $x = 0$, we get that $m \le F(0) \le 1 - m$ implies $nm \le F^n(0) \le n(1-m)$ by additivity, so that

$$m \le \frac{F^n(0)}{n} \le 1 - m.$$

Since this hold for all $n \in \mathbb{N}$, it holds in the limit. But this limit is the rotation number $\rho(f)$ and is independent of the choice of initial point. Hence $rho(f) \ne 0$. $\square$

One can generalize quite readily to q-periodic points by looking at the fixed points of f^q, so that we get the following:

PROPOSITION 4.34. *If f is an orientation-preserving homeomorphism of S^1 then $\rho(f)$ is rational if and only if f has a periodic point.*

EXERCISE 130. Prove this.

But it gets even more restrictive. If f has a q-periodic point, then for a lift F, we have $F^q(x) = x + p$ for some $p \in \mathbb{Z}$. For example, let $f = R_{\frac{6}{7}}$. Then a suitable lift for f would necessarily satisfy $F^7(x) = x + 6$, ∀x∈ $\mathbb{R}$. Notice that there would be mo room for any other periodic points in this case. But this is true in general.

PROPOSITION 4.35. *Let $f : S^1 \to S^1$ be an orientation preserving homeomorphism. Then all periodic points must have the same period.*

This last point is quite restrictive. Essentially, if an orientation preserving homeomorphism has a fixed point, it cannot have periodic points of any other period, say. Note that this is not true of a orientation reversing map. For example, the map which flips the unit circle in $\mathbb{R}^2$ across the y-axis, will fix the two points $(0,1)$ and $(0,-1)$, while every other point is of order two. Perhaps the best way to look at this is the following. For a circle homeomorphism (orientation-preserving), the lift is an increasing map. And increasing interval maps can have many fixed points, but no n-periodic points for $n > 1$. And any point that is not fixed is forward asymptotic to a fixed point and backward asymptotic to another one. Here, it is quite typical that degree-1 circle homeomorphisms, when they have n-periodic points, have two distinct n-orbits, with one attractive and one repulsive (backwards attractive). We will be more precise about this later. But one of these periodic orbits becomes the ω-limit set for everything outside of the repulsive n-orbit, and the α-limit set of anything not on the stable periodic orbit is the unstable one.

EXAMPLE 4.36. Let $f : \S^1 \to S^1$ be the orientation-preserving homeomorphism given by the expression $f(x) = x + \frac{1}{2} + \frac{1}{4\pi}\sin(4\pi x)$. It is not easy to verify that f is indeed a homeomorphism analytically. But from the plot of f in Figure **??** on the left, one can verify its properties readily. Notice that f has no fixed points, but does have two easily to verify period-2 orbits (one can "see" them in the argument to sine), namely at $\left\{0, \frac{1}{2}\right\}$ and at $\left\{\frac{1}{4}, \frac{3}{4}\right\}$. And the stability of these period-2 orbits? Take a look at the graph of $f^2 = f \circ f$ in Figure **??** on the right. The period-2 orbits appear as fixed point here, and their stability is readily readable here from the derivative information. Just cobweb a bit to verify.

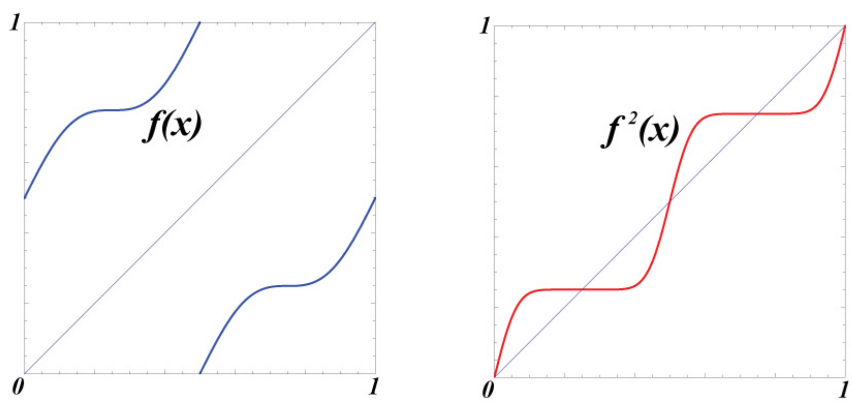


FIGURE 14. The circle homeomorphism f and its square $f^2 = f \circ f$.

EXERCISE 131. Show that any lift of the rotation $R_{\frac{6}{7}}$ must satisfy $F^7(x) = x+6$, $\forall x \in \mathbb{R}$, and explicitly construct two such lifts.

EXERCISE 132. Find the rotation number for the following invertible circle map:

$$f(x) = \begin{cases} \frac{x^2+1}{3} & 0 \le x < \frac{1}{3} \\ 6x - \frac{44}{27} & \frac{1}{3} \le x < \frac{10}{27} \\ 2x - \frac{4}{27} & \frac{10}{27} \le x < \frac{1}{2} \\ \frac{8}{5}\left(x - \frac{1}{2}\right) + \frac{23}{27} & \frac{1}{2} \le x < \frac{21}{27} \\ \frac{1}{6}\left(x - \frac{21}{27}\right) + \frac{35}{27} & \frac{21}{27} \le x < 1 \end{cases}$$

Perhaps we should find a way to end this nicely. This is enough for circle homeomorphisms for now. And ends our work in Chapter 4. There is a great section on frequency locking on page 141. Look it over at your leisure. We won't work through it in the course, but it is very interesting. Dynamically, it represents a situation where a linear flow on the torus (with its uncoupled ODEs) becomes the limiting system to a system of coupled ODEs, representing a nonlinear flow. Question: For this to be the case, must the resulting linear flow on the torus be a rational flow?

CHAPTER 5

Conservative Systems

In Chapter 4, we first looked at what was considered "recurrent" behavior at a point in a dynamical system, which roughly means that the orbit of a point passes arbitrarily close to the point. This worked well in the classification of circle rotations, since either the orbit of a point was closed (the orbit was periodic; the rotation was rational) or the orbit was dense (for an irrational rotation). In either case, every point was recurrent. The same was true for the linear toral flows and their time-t maps.

Contrast this with the dynamical systems that we studied in Chapters 2 and 3. Here, with examples like contracting maps and interval maps with sinks and sources, the only recurrent points were the fixed and periodic points, and there were very few of those in each system. More generally, maps can exhibit much more complicated behavior. To understand this behavior, we will have to broaden our idea of how to study such systems. This chapter begins this study.

To start, let's change our perspective. Given a dynamical system, let's not worry about how an individual orbit behaves so much as how whole families of nearby orbits evolve. This would be more like following all of the orbits that start in a small open subset of the state space over the evolution of the map. For a contraction, this would be easy and not very insightful. (Why is that again?) But for a general map, this idea can be quite interesting.

Somewhere here, place Poincare-Bendixson as a means to discuss recurrence in the plane verses on a cylinder or torus.

5.1. Incompressibility

The notion of incompressibility in a dynamical system means that positive volume domains in the state space do not change their volume as the orbits of their points evolve. This notion is also called *phase volume preservation*. Suppose we have a dynamical system where this property holds; as one evolves via a flow, or iterates via a map, the volume of a small domain does not change. Then the volume is said to be preserved by the flow (respectively, map), or the volume is invariant under the flow (respectively, map). Obvious examples include linear flows in $\mathbb{R}^n$, rotation maps on S^1 (remember that volume in a space like $\mathbb{R}$ or S^1 is just length, and in dimension 2 is just area), and linear toral flows. Examples which do not preserve volume include contraction maps, and flows (defined by ODEs) that include sinks and sources (saddles and centers, maybe, though).

In fact, if the map is an isometry, or the flow has all of its time-t map given by isometries, then the volume will be preserved. This should be obvious, as if all of the distances between the points of a small domain are preserved, the volume cannot change. The converse is not true however. Lots of maps and flows preserve volume but are not isometries.

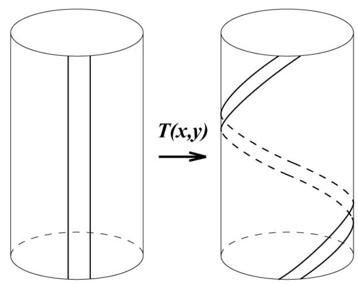
EXAMPLE 5.1. The linear twist map on the cylinder

$$T : S^1 \times [0,1] \longrightarrow S^1 \times [0,1], \quad T(x,y) = (x+y,y)$$

is an area-preserving map which is not an isometry. See the figure below.

EXERCISE 133. Show that (1) T is not an isometry, but (2) T preserves area on the cylinder.

EXERCISE 134. Show that the linear flow on the 2-torus is an isometry (Hint: Build the proper metric on the torus).



Now, let's consider a linear map on $\mathbb{R}^n$:

$$f : \mathbb{R}^n \to \mathbb{R}^n, \quad f(\mathbf{x}) = A\mathbf{x},$$

where A is a $n \times n$ matrix. Choose an orthonormal basis for $\mathbb{R}^n$. Then the standard cube C whose sides are the basis vectors (this is the "unit cube" relative to the basis) will be mapped by f to a parallelepiped. What would be the volume of this image? Well, here

$$\operatorname{vol}(f(C)) = |\det A|. \tag{5.1.1}$$

This is standard Linear Algebra, where according to f, each standard basis vector $\mathbf{e}_i$ is mapped to the ith column of A. Hence the image of the unit cube is the parallelepiped with edges the columns of A. Hence the determinant of A (in absolute value) is the volume of the image of the unit cube under f. Hence volume is preserved by f if $|\det A)| = 1$. What would be the conclusion one can draw from this? Read f as the linear model for the infinitesimal version of any smooth map on $\mathbb{R}^n$. Then we have:

PROPOSITION 5.2. *let $U \in \mathbb{R}^n$ be an open domain. A differentiable map $f : U \to \mathbb{R}^n$ preserves volume iff $|\det(Df_x)| = 1$, $\forall x \in U$.*

The *Jacobian matrix* of a function $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$ is the $n \times m$ matrix of partial derivatives of f, sometimes denoted $\frac{\partial(f_1,\dots,f_n)}{\partial(x_1,\dots,x_m)}$, or simply Df. The Jacobian is a matrix of functions whose ijth element is $\frac{\partial f_i}{\partial x_j}$. Evaluated at a point $x \in U$, $Df(x) = Df_x$ is a matrix of real numbers called the derivative of f at x. When $n = m$, the derivative matrix is square, and its determinant becomes an important property. In the square matrix case, it is common to refer to the determinant of this derivative matrix the *Jacobian* of f, $Jac(f)$.

DEFINITION 5.3. A map $f : U \to \mathbb{R}^n$, where $U \subset \mathbb{R}^n$ is a domain, preserves orientation if $\forall x \in U$, $Jac(f) > 0$.

"Nice" ODEs (where solutions exist and are unique everywhere, for example), are always orientation preserving. Recall the relationship between the time-1 map of any linear ODE system on $\mathbb{R}^2$ and its corresponding flow. The time-1 map is a linear transformation on the plane, and its matrix always has eigenvalues which were related to those of the original flow by the exponential map. Under the exponential map, the time-1 map will always have a positive Jacobian (why?).

More generally, let $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ be an ODE on $\mathbb{R}^n$ so that $\mathbf{f}$ is a C^1-vector field on $\mathbb{R}^n$. Let $\varphi : \Omega \times I \to \mathbb{R}^n$ be its corresponding flow, where $0 \in I \subset \mathbb{R}$ is open and $\Omega \subset \mathbb{R}^n$ is a domain (recall we often write $\varphi(\mathbf{x}, t) = \varphi^t(\mathbf{x})$ to accentuate the one-parameter group of transformations of phase space as well as the exponent notation of function iteration compatible with the additive group.) Here, for a fixed $t \in I$m $\varphi^t : \Omega \to \mathbb{R}^n$ is a diffeomorphism (a homeomorphism where both the map and its inverse are differentiable) onto its image, and $\varphi^0(\mathbf{x}) = \varphi(\mathbf{x}, 0) = \mathbf{x}$.

Denote by $Jac(\varphi^t)$ the Jacobian matrix of the flow map φ^t. Here $Jac(\varphi^t)$ is just the derivative of the time-t map of the flow and $Jac(\varphi^0) = I_n$. This way, $\det\left(Jac(\varphi^t)\right)$ measures the relative volume of bounded domains at time t in Ω relative to their original volumes at $t = 0$. Then $\frac{d}{dt}\big|_{t=0} \det\left(Jac(\varphi^t)\right)$ measures the instantaneous rate of change of volume along the flow. Recall that this is precisely one of the geometric interpretations of the divergence of a vector field:

$$\operatorname{div}(\mathbf{f}) = \frac{d}{dt}\bigg|_{t=0} \det\left(Jac(\varphi^t)\right),$$

when φ^t is the flow corresponding to the vector field $\mathbf{f}$.

Indeed, we know the following:

(1) Flows solve their corresponding ODE systems:

$$\frac{d}{dt}\varphi^t = \mathbf{f}(\varphi^t), \text{ and } \frac{d}{dt}\bigg|_{t=0} \varphi^t = \mathbf{f}(\mathbf{x}).$$

(2) $\frac{d}{dt}\big|_{t=0} Jac(\varphi^t) = Jac\left(\frac{d}{dt}\big|_{t=0}\varphi^t\right)$, since time and all phase space coordinates are independent.

Thus we have

$$\frac{d}{dt}\bigg|_{t=0} Jac(\varphi^t) = Jac\left(\frac{d}{dt}\bigg|_{t=0} \varphi^t\right) = Jac(\mathbf{f}).$$

We now have a beautiful result from (multi-) linear algebra (actually matrix calculus), whose proof we leave as an exercise, as it is purely constructive.

Proposition 5.4. *Let $A(t)$ be a C^1 family of $n \times n$-matrices (whose entries are C^1-functions of t) defined on some $I \subset \mathbb{R}$ containing 0, where $A(0) = I_n$. Then*

$$\frac{d}{dt}\bigg|_{t=0} \det A(t) = trace\left(\frac{d}{dt}\bigg|_{t=0} A(t)\right).$$

Exercise 135. Prove this, noting the special form for the derivative of a determinant of a matric of functions:

$$\frac{d}{dt}\det A(t) = \sum_i Row_i'\left(A(t)\right),$$

where $Row_i'\left(A(t)\right)$ is the matrix $A(t)$ with the ith row's entries replaced with their corresponding derivatives.

Then since $Jac(\varphi^t)$ is such a 1-parameter family of $n \times n$-matrices, we have

$$\frac{d}{dt}\bigg|_{t=0} \det\left(Jac(\varphi^t)\right) = \operatorname{trace}\left(\frac{d}{dt}\bigg|_{t=0} Jac(\varphi^t)\right) = \operatorname{trace}\left(Jac(\mathbf{f})\right) = \operatorname{div}(\mathbf{f}).$$

This leads immediately to the following conclusion:

PROPOSITION 5.5. *If the divergence of the vector field* $\mathbf{f}$ *vanishes (that is, if* $div(\mathbf{f}) = 0$*), then* $\mathbf{f}$ *preserves volume.*

THEOREM 5.6. *let* X *be a finite volume domain in* $\mathbb{R}^n$ *or* $\mathbb{T}^n$*, and* $f : X \to X$ *be an invertible, volume preserving* C^1*-map. Then* $\forall x \in X$ *and* $\forall \epsilon > 0$*,* $\exists n \in \mathbb{N}$ *such that*

$$f^n\left(B_\epsilon(x)\right) \cap B_\epsilon(x) \neq \varnothing.$$

PROOF. This can be easily seen as follows: Suppose $\exists x \in X$, and $\exists \epsilon > 0$ such that $\forall n \in \mathbb{N}$

$$f^n\left(B_\epsilon(x)\right) \cap B_\epsilon(x) = \varnothing.$$

Since f is volume preserving, we must have at the nth iterate:

$$\infty > \text{ vol}(X) > \sum_{i=1}^{n} \text{ vol}\left(f^i\left(B_\epsilon(x)\right)\right) = n \cdot \text{ vol}\left(B_\epsilon(x)\right).$$

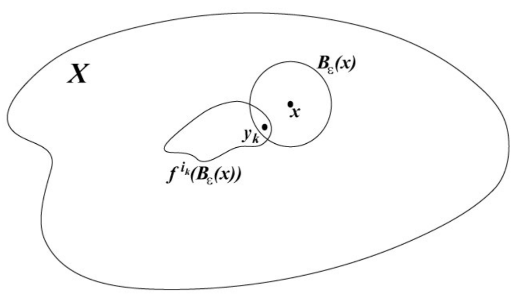
But for all choice of $\epsilon > 0$,

$$\lim_{n\to\infty} n \cdot \text{ vol}\left(B_\epsilon(x)\right) = \left(\text{vol}\left(B_\epsilon(x)\right)\right) \lim_{n\to\infty} n = \infty$$

since $\text{vol}\left(B_\epsilon(x)\right) > 0$. This contradiction establishes the proof. □

This gives us an immediate consequence:

COROLLARY 5.7. *For* $f : X \to X$ *as above,* $\forall x \in X$*, there exists a sequence* $\{y_k\} \longrightarrow x$ *and a sequence* $\{i_k\} \longrightarrow \infty$ *where* $\left\{f^{i_k}(y_k)\right\} \longrightarrow x$*.*



See the figure to get an idea of what is going on. Given any small neighborhood of $B_\epsilon(x) \subset X$, There will be a iterate (here, the i_k-th iterate) of f in the forward orbit of $B_\epsilon(x)$ which will intersect $B_\epsilon(x)$. Choose any point y_k in the intersection. Now choose a new $\epsilon > 0$ where $\epsilon < d(x, y_k)$, and repeat the procedure. Play this game for a decreasing sequence of ϵ's going to 0. At each stage, you produce a y_k close to x that has a forward iterate that is even closer. In the limit, you show that arbitrarily close to x is a recurrent point. As the choice of x does not matter, you have that recurrent points are almost everywhere.

EXERCISE 136. Produce this sequence.

Recall Definition 4.1 on a point being recurrent. We can extend that notion now to

DEFINITION 5.8. For $f : X \to X$ a continuous map of a metric space, a point $x \in X$ is called

- *positively recurrent* with respect to f if $\exists$ a sequence $\{n_k\} \longrightarrow \infty$ such that $\{f^{n_k}(x)\} \longrightarrow x$,
- if f is invertible, *negatively recurrent* if $\exists$ a sequence $\{n_k\} \longrightarrow -\infty$ such that $\{f^{n_k}(x)\} \longrightarrow x$,
- *recurrent* if it is both positively and negatively recurrent.

Definition 5.9. For $f : X \to X$ a continuous map of a metric space, the set

$$\omega(x) = \overline{\bigcap_{n\in\mathbb{N}} \left\{ f^i(x) \middle| i \geq n \right\}}$$

is the set of all accumulation points of the orbit of x. It is called the *ω-limit set* of $x \in X$ with respect to f. For f an invertible map on X, the set

$$\alpha(x) = \overline{\bigcap_{n\in\mathbb{N}} \left\{ f^{-i}(x) \middle| i \leq n \right\}}$$

is called the *α-limit set* of x with respect to f.

Note: $x \in X$ is positively recurrent if $x \in \omega(x)$ (if x is in its own ω-limit set). In fact, we can use this as a definition of the set of all recurrent points:

Definition 5.10. For $f : X \to X$ a continuous map of a metric space, call $\mathcal{R}_f(X)$ the set of all recurrent points of f on X, where

$$\mathcal{R}_f(X) = \left\{ x \in X \middle| x \in \omega(x) \right\}.$$

Definition 5.11. For $f : X \to X$ a continuous map of a metric space, a point $x \in X$ is called *wandering* if there exists $\epsilon > 0$ and $\exists N \in \mathbb{N}$ such that $\forall n > N$, $f^n(B_\epsilon(x)) \cap B_\epsilon(x) = \varnothing$. Points that are not wandering are called *non-wandering*.

Exercise 137. Develop a precise definition for a non-wandering point.

One can collect up all of the wandering and non-wandering points of a map $f : X \to X$, respectively, as $\mathcal{W}_f(X)$ and $\mathcal{NW}_f(X)$, and note immediately that $\mathcal{R}_f(X) \subset \mathcal{NW}_f(X)$ (recurrent points are non-wandering.) However, these sets are not the same: Non-wandering is a property of a point that is based on what happens to orbits near x. But non-recurrence is a property only of the point x. Indeed, there exist non-wandering points which are not recurrent:

Example 5.12. Let $\ddot{x} = x^2 - x$ be a second-order, autonomous ODE (Is the vector field conservative)? Or, if you prefer, the system $\dot{x} = y$, $\dot{y} = x^2 - x$.

Exercise 138. For this system, (1) Calculate the time-1 map as a transformation of the plane, and (2) solve the system by constructing an equivalent exact ODE.

Here, as in the Figure **??**, the shaded region consists of the non-wandering points. But this set is a closed region, and $\mathcal{NW}$ also contains the separatrix forming the orbit line containing $\mathcal{O}_{(-1,0)}$. The points on this orbit line are all homoclinic to the unstable equilibrium at $(1,0)$. They are not recurrent. But they are also non-wandering since any neighborhood of an initial point on this orbit line will contain pieces of periodic orbits.

Exercise 139. Let $T : [0,1] \to [0,1]$, $T(x) = 1 - 2|x - 1|$ (T is an example of what we call a tent map. See Equation 6.5.1.) Show that any point x which is a dyadic rational (a rational with a denominator which is a power of 2) is non-recurrent and non-wandering.

Exercise 140. Show, by construction, the rotation map $R_\alpha : S^1 \to S^1$ has the property that $\mathcal{R}_{R_\alpha}(S^1) = \mathcal{NW}_{R_\alpha}(S^1) = S^1$, $\forall \alpha \in \mathbb{R}$.

EXERCISE 141. Show the same by construction for a translation on $\mathbb{T}^2$.

THEOREM 5.13. *Let X be a closed finite-volume domain in $\mathbb{R}^n$ or $\mathbb{T}^n$ and $f : X \to X$ an invertible volume preserving map. Then the set of recurrent points for f is dense in X.*

Note; This does not mean that all points are recurrent, not that there may be tons of points whose ω-limit sets do not include the original point. It does mean that every point either is recurrent, or has a recurrent point arbitrarily close to it. We won't prove this here. The proof is in the book on page 160. Instead, let's skip ahead to Section 6.2.

This is the notion of Poincare Recurrence, see W[Poincare Rec]. Give a treatment here.

5.1.1. Lagrange. One way to understand Newtonian (read: Classical) mechanics is via a formulation developed by Joseph Louis Lagrange in the late 1700's. This approach is essentially a variational approach that says, roughly, that the path of a particle though a force field can be described not only via the equations of motion in the standard cartesian coordinates by forces determining the various constraints of the motion, but also via a set of independent generalized coordinates that completely characterize the motion of the particle; the choices of these generalized variables eliminate the need for the constraints. This approach usually reduces the number of coordinates needed to completely describe the motion (by parameterizing a subspace of Euclidean space on which motion is constrained) and in some cases greatly simplifies the process of solving the equations of motion. And this formulation works for conservative and nonconservative systems.

Indeed, if the motion in $\mathbb{R}^n$ is constrained to a subspace in a way that can be described via a set of transformational equations

$$\mathbf{x} = \mathbf{x}(q_1, \ldots, q_m, t),$$

then we can use this parameterization as a way to rewrite the system in terms of the generalized coordinates q_i, $i = 1, \ldots, m$. Here the $m \in \mathbb{N}$ is called the *number of degrees of freedom* of the system.

EXAMPLE 5.14. For the Pendulum in $\mathbb{R}^2$, we have $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$. Motion is constrained to the circle of radius ℓ about the origin. A choice of generalized coordinate is the angular coordinate θ along this circle, where $\theta \in S^1$. We place the origin on the circle at the bottom pt $\begin{bmatrix} 0 \\ -\ell \end{bmatrix}$ for physical reasons (why?) Then

$$\mathbf{x}(\theta(t)) = \begin{bmatrix} \ell \sin\theta \\ -\ell\cos\theta \end{bmatrix} \quad \text{and} \quad \dot{\mathbf{x}}(\theta, \dot\theta) = \begin{bmatrix} \ell\dot\theta\cos\theta \\ \ell\dot\theta\sin\theta \end{bmatrix}.$$

With this choice of coordinate, the pendulum has only one degree of freedom.

For a system of n-particles with constraints, we can list the transformational equations $\mathbf{x}_i = \mathbf{x}_i(q_1, \ldots, q_m, t)$, $i = 1, \ldots, n$. Then the total kinetic energy of the system is

$$T = \sum_i \frac{1}{2} m_i \mathbf{x}_i \cdot \mathbf{x}_i,$$

and if the force field is conservative, then the total potential energy is the position function $V = V(\mathbf{x}_i) = V(q_i)$.

The Lagrangian (Function) L is defined as the difference between the kinetic and potential energies: $L = T - V$. This function contains the core dynamical information of the system. We will bypass much of the detail of the origin and derivation of this function here, and state that the equations of motion can be calculated from L via the Euler-Lagrange equations (sometimes called the *Lagrange Equations of the Second Kind*

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) = \frac{\partial L}{\partial q_j}, \quad j = 1, \ldots, m.$$

This is in general a system of n second-order ODEs, one for each generalized coordinate.

Example 5.15. For the Pendulum, we have

$$T(\mathbf{x}) = \frac{1}{2} m \mathbf{x}_i \cdot \mathbf{x}_i = \frac{1}{2} m \left(\ell^2 \dot{\theta}^2 \cos^2\theta + \ell^2 \dot{\theta}^2 \sin^2\theta\right) = \frac{1}{2} m \ell^2 \dot{\theta}^2 = \frac{1}{2} m \left(\ell\dot{\theta}\right)^2,$$

and

$$V(\mathbf{x}) = mg(\ell + y) = mg(\ell - \ell\cos\theta),$$

written this way to place the lowest potential energy at the low point on the circle and at 0. Then with

$$L = T - V = \frac{1}{2} m \left(\ell\dot{\theta}\right)^2 - mg(\ell - \ell\cos\theta),$$

we can derive the equations of motion as

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right) = m\ell^2\ddot{\theta} = -mg\ell\sin\theta = \frac{\partial L}{\partial \theta},$$

or $m\ell^2\ddot{\theta} + mg\ell\sin\theta = 0$.

The advantages of using the Lagrangian formulation to study Newtonian physics are many:

- The equations of motion can all be derived from a single function through simple calculus,
- The Lagrangian is simply a difference of energies, which as scalar fields are more easy to calculate than forces (vector fields),
- works well in all coordinate systems.

However, there are some shortcomings: Usually, the Lagrangian does not have a physical interpretation as some measurable quantity. And it can be quite difficult or impossible to actually solve the resulting n second-order differential equation system.

There is another formulation of classical mechanical systems that has been shown to be much more robust in generalizing to other areas of physics; the Hamiltonian formulation. Here, instead of using generalized coordinates and their velocities, one replaces the velocities with the corresponding coordinates' conjugate momenta. This was developed by William Rowan Hamilton in the 1830s. To start, for each choice of generalized coordinate q_i in the Lagrangian formulation, define a corresponding conjugate momentum p_i via

$$p_i = \frac{\partial L}{\partial q_i}.$$

Then, in the new coordinate system $(q_1, \dots, q_n, p_1, \dots, p_n)$, define the function $H(\mathbf{q}, \mathbf{p}, t) = \sum_i \dot{q}_i p_i - L(\mathbf{q}, \dot{\mathbf{q}}, t)$, the Legendre Transform of the Lagrangian. The function H is called the Hamiltonian of the system. As such, we can calculate its infinitessimal change via the differential

$$dH = \sum_i \left(\dot{q}_i \, dp_i + p_i \, d\dot{q}_i - \frac{\partial L}{\partial q_i} \, dq_i - \frac{\partial L}{\partial \dot{q}_i} \, d\dot{q}_i \right) - \frac{\partial L}{\partial t} \, dt \tag{5.1.2}$$

$$= \sum_i \left(\frac{\partial H}{\partial q_i} \, dq_i + \frac{\partial H}{\partial p_i} \, dp_i \right) + \frac{\partial H}{\partial t} \, dt. \tag{5.1.3}$$

But with the definition of conjugate momenta and the Euler-Lagrange equations,
****

Dynamically speaking, we can also play the game the other way: Let $n \in \mathbb{N}$ and endow $\mathbb{R}^{2n}$ with the coordinates $(q_1, \dots, q_n, p_1, \dots, p_n)$. Then ANY C^1-function $H : \mathbb{R}^{2n} \to \mathbb{R}$ can play the role of a total energy function (read: Hamiltonian) of some Newtonian system. The vector field can be given by $X_H = \begin{bmatrix} \frac{\partial H}{\partial p_i} \\ -\frac{\partial H}{\partial q_i} \end{bmatrix}$ so that the equations of motion are

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \quad \text{and} \quad \dot{p}_i = -\frac{\partial H}{\partial q_i}.$$

Notice that this vector field is automatically conservative, since

$$\text{dev} X_H = \sum_i \frac{\partial}{\partial q_i} \left(\frac{\partial H}{\partial p_i} \right) + \frac{\partial}{\partial p_i} \left(-\frac{\partial H}{\partial q_i} \right) = 0.$$

Exercise 142. Solve the planar system given by the function $H(x, y) = 3x^2 + 2xy + y^2$ and classify the equilibrium at the origin.

But this takes us directly back to exact ODEs:

- If we start with $H(x, y) = 3x^2 + 2xy + y^2$ and consider both x and y as dependent on time, then
$$\frac{\partial H}{\partial t} = 0 = (6x + 2y)\frac{dx}{dt} + (2x + 2y)\frac{dy}{dt}.$$
This is just an exact ODE $M \, dx + N \, dy = 0$, with $M = (6x + 2y)$ and $N = (2x + 2y)$ and $M_y = N_x = 2$.
- The connection is due to the fact that the integral curves are forced to "live" on the level sets of the function. Hence the value of the function is constant along the solution curves of the vector field generated by the ODE. This is no accident, and persists in higher dimensions with some subtleties.

For now, we leave you with a beautiful restatement of the classification theorem for the isolated equilibria of vector fields generated by functions:

Theorem 5.16.

Remark 5.17. We can generalize this Hamiltonian construction to many even-dimensional non-Euclidean spaces, as long as they satisfy certain conditions. We will not go into this in this text, but the spaces where this can be done are called symplectic, or more generally Poisson spaces. This leads to a mathematical version of Hamiltonian dynamics called symplectic geometry, a synthesis of differential geometry, differential topology and sometimes smooth dynamics.

5.2. Newtonian Systems of Classical Mechanics

Your previous work in ODEs suggested a general premise about systems of differential equations. If they are defined "nicely", then the present state of a mechanical system determines its future evolution through other states uniquely. One can place this in the language of dynamical systems to say that if a mathematical construction accurately models a mechanical system, than the construction determines a dynamical system on the space of all possible states of the system. The trick in many cases is to well understand what constitutes a state of a mechanical system. To start, given a mechanical system, the *configuration space* of the system is the set of all possible positions (value combinations of all of its variables) of the system. The *state space*, rather, is the set of all possible states the system can be in. This is usually much broader a description.

For example, consider the pendulum, a mass is attached to the free end of a massless rigid rod, while the other end of the rod is fixed. The set of all possible configurations of the pendulum is simply a copy of S^1. However, for each configuration, the pendulum is in a different state depending on what the mass' velocity is when it resides in a configuration. One can think of all possible states as the space $S^1 \times \mathbb{R}$. This reflects the data necessary to completely determine the future evolution of the pendulum by a knowledge of its position and velocity at a single moment, and the evolution equation which is a second-order, possibly non-linear and non-autonomous, ODE in the general form

$$\ddot{x} = f(t, x, \dot{x}).$$

In the case of a pendulum, time is not explicit on the right hand side, and the equation is autonomous. Under the standard practice of converting this ODE into a system of two first order ODEs, we can interpret the evolution as giving a vector field on the state space $S^1 \times \mathbb{R}$, with coordinates x and $\dot{x}$. This vector field determines a flow, which solves the ODE and determines the future evolution of the system based on knowledge of the state of the system at a particular moment in time.

Many systems behave in a way that their future states are completely determined by their present position and velocity, along with a notion of how they are changing. In classical (Newtonian) mechanics, Newton's Second Law of motion states roughly that the force acting on an object is proportional to how the velocity of the object is changing. The is the famous equation $f = ma$, where f is the total force acting on the object and a is its acceleration. As the velocity depends on the current position of an object, a good notion of how an object moves through a space under the influence of a force is completely determined by how its position and velocity are changing, at least when the force is static:

$$f(x) = ma = m\ddot{x} = m\frac{d^2x}{dt^2}.$$

This is a special case of the general second order ODE mentioned above.

Example 5.18. An object under the influence of only gravity satisfies Newton's Second Law and the differential equation is $\ddot{x} = -g$, where g is the gravitational constant. This is solved by integrating the "pure time" ODE twice

$$x(t) = -g\frac{t^2}{2} + v_0 t + s_0,$$

where v_0 and s_0 are the initial velocity and initial position, respectively (the two constants of the integrations).

EXAMPLE 5.19. Harmonic Oscillator. Recall Hooke's Law: the amount an object is deformed is linearly related to the force causing the deformation. This translates to $\ddot{x} = -kx$, which is autonomous. Solutions are given by

$$x(t) = a\sin\sqrt{k}t + b\cos\sqrt{k}t,$$

where a and b are related to the initial starting position and velocity of the mass.

As stated above, note that any ODE of the form $\ddot{x} = f(t,x,\dot{x})$ can be converted to a system of two first order (generally coupled) ODEs of the form

$$\begin{aligned} \dot{x} &= v \\ \dot{v} &= f(t,x,v) \end{aligned}$$

which defines a vector field (a static vector field if t does not appear explicitly in the equations) on the (x,v)-state space. In this autonomous case, we get for Newton's Equation, $\dot{x} = v$ and $\dot{v} = \frac{1}{m}f(x,\dot{x})$. Often, the model neglects the dependency of the vector field on the velocity component, as is in the case where friction is ignored. In this case, Newton's equation(s) reduce to $\dot{x} = v$ and $\dot{v} = \frac{1}{m}f(x)$. This is the case in the two examples above. We will treat this case presently.

Note: For the system defined by n coordinates and their velocities, we get the $2n$-system of first order equations defined as

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{v} \\ \dot{\mathbf{v}} &= \mathbf{f}(t,\mathbf{x},\mathbf{v}) \end{aligned}$$

The state space consists of the $2n$-dimensional vectors $\begin{bmatrix} \mathbf{x} \\ \mathbf{v} \end{bmatrix}$. Restricting to the case where time is not explicit in the ODEs and velocity-dependent effects are ignored, the vector field of this $2n$-system attaches the vector $V = \begin{bmatrix} \mathbf{v} \\ \frac{1}{m}\mathbf{f}(\mathbf{x}) \end{bmatrix}$ to each point $\begin{bmatrix} \mathbf{x} \\ \mathbf{v} \end{bmatrix}$. The divergence of this vector field is

$$\nabla \cdot V = \sum_{i=1}^{n} \frac{\partial}{\partial x_i}(v_i) + \sum_{i=1}^{n} \frac{\partial}{\partial v_i}\left(\frac{1}{m}f_i(\mathbf{x})\right) = 0.$$

Hence the flow preserves volume.

place around here all of the stuff regarding classical mechanics, with both Lagrange and Hamiltonian formulations and their symmetries.

REMARK 5.20. This fact is true for general autonomous Newtonian systems where the force is solely a function of position. One facilitating idea in Newtonian physics is to, in essence, factor out the mass. Define a new variable (a coordinate) $q := mx$, and then switch from the velocity coordinate to $p := mv$ as the (linear) momentum. Then $\dot{q} = m\dot{x} = mv$, and $f = ma = f = m\dot{v} = \dot{p}$, and the system becomes $\dot{q} = p$ and $\dot{p} = f\left(\frac{1}{m}q\right) = g(q)$. Not only does this make the system easier to work with, it exposes some hidden symmetries within the equations of conservative systems. This forms the basic framework for what are called Hamiltonian dynamics.

Now assume that the force $f(x)$ is a gradient field (this means that the force is the gradient of a function of position alone, or $f = -\nabla V$ for some $V(x)$. Then

$$f(x) = ma = m\dot{v} = -\nabla V.$$

Here, the function V is called the potential energy (energy of position), and the energy of motion, the kinetic energy is

$$K = \frac{1}{2}m\|v\|^2 = \frac{1}{2}m(v \cdot v).$$

The total energy $H = K + V$ satisfies

$$\frac{d}{dt}(H) = \frac{dK}{dt} + \frac{dV}{dt} = m\dot{v} \cdot v + \sum_{i=1} n \frac{\partial V}{\partial x_i} \cdot \frac{\partial x_i}{\partial t} = m\dot{v} \cdot v + \nabla V \cdot v = (\nabla V + m\dot{v}, v) = 0.$$

The conclusion is the total energy H is conserved as one evolves in a system like this. As H is a function defined on the state space given by the vectors x and mv, the solutions to the system of ODEs are confined to the level sets of this function. A system like this is called *conservative*, and is characterized by the idea that the force field is a gradient field. You have seen this before in a different guise:

5.2.1. Exact Differential Equations. Consider the nonlinear system of 2 first-order, linear, autonomous differential equations in 2 variables

$$\begin{aligned} \dot{x} &= 4 - 2y \\ \dot{y} &= 12 - 3x^2. \end{aligned} \tag{5.2.1}$$

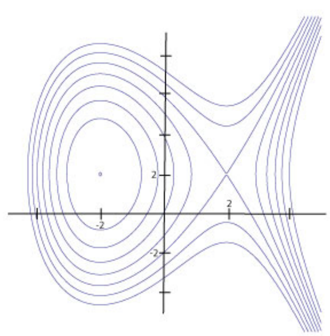
This system can also be written by the single differential equation

$$(12 - 3x^2)dx - (4 - 2y)dy = 0. \tag{5.2.2}$$

Note that this equation is exact, and separable, and upon integration, one obtains

$$4y - y^2 = 12x - x^3 + C.$$

this defines our solutions implicitly. In fact, we can use this directly.



Define a function $\varphi(x,y) = 4y - y^2 - 12x + x^3$. Then φ is conserved by the flow, and the flow must live along the constant level sets of φ (the sets that satisfy $\varphi(x,y) = C$.) These sets are given by the figure.

Now recall in Section 2.4 that an ODE $M\,dx + N\,dy = 0$ is exact if $M_y = N_x$ (this notation again refers to the first partial derivatives of the functions with respect to the subscripts). The reason is vector-calculus in nature: The solution is a function $\varphi(x,y) = C$ satisfying $\frac{\partial\varphi}{\partial x} = M$ and $\frac{\partial\varphi}{\partial y} = N$. The condition for exactness is simply the statement that for any C^2 function $\varphi(x,y)$, the mixed partials are equal:

$$M_y = \frac{\partial M}{\partial y} = \frac{\partial}{\partial y}\frac{\partial\varphi}{\partial x} = \frac{\partial^2\varphi}{\partial y\partial x} = \frac{\partial^2\varphi}{\partial x\partial y} = \frac{\partial}{\partial x}\frac{\partial\varphi}{\partial y} = \frac{\partial N}{\partial x} = N_x.$$

But we can now go even further: The vector field $\mathbf{F}(x, y) = \left(4 - 2y, 12 - 3x^2\right)$ of the system in Equation 5.2.1 corresponds to the exact ODE given by Equation 5.2.2 when $M = 12 - 3x^2$ and $N = -(4 - 2y)$, or $\mathbf{F}(x, y) = (-N, M)$. With this,

$$div(\mathbf{F}) = \frac{\partial}{\partial x}(-N) + \frac{\partial}{\partial y}(M) = -N_x + M_y = 0.$$

The vector field $\mathbf{F}$ is conservative, and the flow will preserve volume (area) in the plane. This is a general fact for vector fields of exact ODEs and leads directly to:

Proposition 5.21. *The flow of an exact ODE in $\mathbb{R}^2$ preserves volume in phase space.*

Corollary 5.22. *Equilibria of exact ODEs in $\mathbb{R}^2$ can only be saddles or centers.*

The repercussions of these facts are quite important: For instance, $\mathbf{p} = (-2, 2)$ is an equilibrium solution of Equation 5.2.1. What is its type and stability (forgetting the figure for a moment, that is) of $\mathbf{p}$? We can linearize this Almost Linear System (See Section 2.4) at $\mathbf{p}$:

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} \frac{\partial(-N)}{\partial x}(-2,2) & \frac{\partial(-N)}{\partial y}(-2,2) \\ \frac{\partial M}{\partial x}(-2,2) & \frac{\partial M}{\partial y}(-2,2) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 & -2 \\ 12 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

The eigenvalues $r = \pm\sqrt{24}$ ar epurely imaginary. Hence the linearlized equilibrium at the origin is a center. But centers are NOT *structurally* stable, in that a small perturbation in a center may result in a sink or a source, as well as a center (the eigenvalues may take on small real parts, either negative or positive). Hence we cannot by itself declare that $\mathbf{p}$ is in fact a center via the linearized system.

Somewhere around here place, talk mathematically about a classification theorem for equilibria of nonlinear, almost linear systems, both without and with phase volume preservation.

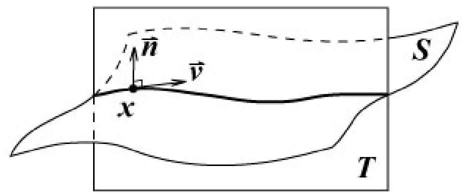
However, with the *additional* knowledge that $\mathbf{F}$ is conservative, then sinks and sources are not possible, and in fact, the point $\mathbf{p}$, an equilibrium of the nonlinear system, MUST be a center. Such is an import of phase volume preservation.

Exercise 143. Complete the phase diagram for this ODE by noting directions of motion along the level curves. Also, note the values of the level sets corresponding to the two equilibria solutions. Finally, show analytically that the equilibrium at $(2, 2)$ is unstable, while the equilibrium at $(-2, 2)$ is stable.

5.2.2. Newton's First Law. An object not under the influence of an external force will move linearly and with constant (maybe zero) velocity. How does this notion of linear (constant velocity) motion appear

- in Euclidean space?
- on $\mathbb{T}^n$?
- On S^2?
- On an arbitrary metric space? Here we must get a better understanding of just what a straight line is in a possibly curved space. We can use the metric to define a straight line as the path that is the shortest distance between two points. This path is called a *geodesic*.

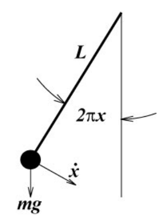
On a smooth surface $S \subset \mathbb{R}^3$, the Euclidean metric on $\mathbb{R}^3$ induces a metric on S. Choose a point $x \in S$. The surface has a well-defined tangent plane to S at x. With this tangent plane, we can choose a normal $\mathbf{n}$ to the surface at x, as well as a desired direction $\mathbf{v}$ in the tangent plane. Now, for a particle moving freely along the surface S in the direction of $\mathbf{v}$ at x, the ONLY force acting on the particle is the force keeping it on S. Thus, the acceleration vector of the particle is in the direction of $\mathbf{n}$. With no component of the force in the direction of motion, the speed $\|\mathbf{v}\|$ is constant along this intersection line.



QUESTION 5.23. *What do the geodesics look like on S^2?*

5.2.3. The Planar Pendulum. One can model the planar pendulum by the autonomous second order differential equation

$$2\pi mL\ddot{x} + mg\sin(2\pi x) = 0. \tag{5.2.3}$$



Some notes:

- This is the undamped pendulum as stated. If one were to consider damping, one can model this by adding a term involving $\dot{x}$. A common one is $cL\dot{x}$.
- This equation can be rewritten as
$$\frac{2\pi L}{g}\ddot{x} + \sin(2\pi x) = 0.$$
- To simplify even further, one can scale time by $\tau = \frac{t}{T}$, where $T = \sqrt{\frac{g}{2\pi L}}$. Then we get
$$\ddot{x} + \sin(2\pi x) = 0.$$

So the model becomes

$$\dot{x} = v \tag{5.2.4}$$
$$\dot{v} = -\sin 2\pi x, \tag{5.2.5}$$

which is Newtonian with $f(x) = -\sin 2\pi x$. Here the kinetic energy is $K = \frac{1}{2}v^2$, and V is the potential energy, where

$$f(x) = -\nabla V, \quad \text{and} \quad V = \int \sin 2\pi x\, dx = -\frac{1}{2\pi}\cos 2\pi x.$$

The total energy is $H = K + V = \frac{1}{2}v^2 - \frac{1}{2\pi}\cos 2\pi x$ and is conserved. Hence motion is along the level sets of H.

Some dynamical notes:

- For low energy values $H \in \left(-\frac{1}{2\pi}, \frac{1}{2\pi}\right)$, motion is periodic and all orbits are closed.

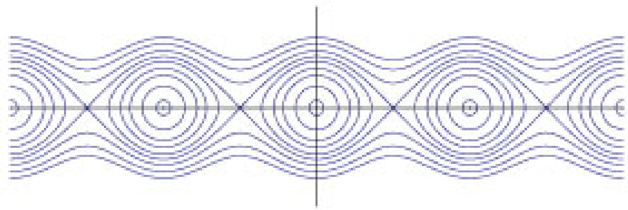


FIGURE 1. The phase plane of the planar pendulum.

- For high energy values $H > \dfrac{1}{2\pi}$, motion looks unbounded. But is it really? What is the pendulum actually doing for these energy values? Recall the idea from before that the actual phase space is a cylinder (the horizontal coordinate – the position of the pendulum – is angular and hence periodic). Hence motion is still periodic and orbits are still closed.
- What happens for the energy value $H = \dfrac{1}{2\pi}$? What is the lowest energy value? What are the energy values of the equilibria solutions?

EXERCISE 144. On the phase plane above in Figure 1, complete the diagram by orienting the solutions curves (choose carefully your directions and justify by showing it is compatible with your choice of coordinates. Then create a Poincare section along the vertical axis of the phase diagram as an open interval runs from the top separatrix to the bottom separatrix. Compare the first-return map to any time-t map within the region bounded by the two separatrices.

Back to Poincare Recurrence. This system is conservative and hence exhibits phase volume preservation (incompressibility). What can we say about the recurrent points? Theorem 5.13 required a finite volume domain to establish the density of recurrent points. On the phase cylinder, we can create a finite volume domain simply by bounding the total energy $H < M$, for some $M > -\frac{1}{2\pi}$, so

$$X_M = \left\{ (x, v) \in S^1 \times \mathbb{R} \;\middle|\; H(x, v) < M \right\}.$$

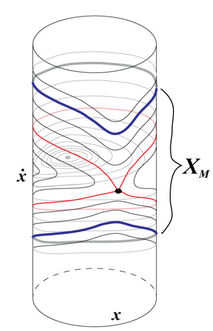


FIGURE 2. Phase cylinder of the planar pendulum.

Now by Theorem 5.13, almost all points on X_M are recurrent. Can you find points that are not recurrent in the phase space? Can you classify them? Look for points $\mathbf{x}$, where $\mathbf{x} \notin \omega(\mathbf{x})$. That the phase plane for the pendulum is actually a cylinder is an extremely important concept, if not for this reason alone.

The preceding two examples, that of the exact ODE and the Planar Pendulum, illustrate some very important phenomena. One important facet is that they are examples of non-linear, autonomous, first-order ODEs in the plane. And although the exact ODE system can be solved (the pendulum cannot), both can be effectively studied via an analysis of a few particular orbits. We take a moment here to expound on this.

Let's linearize the pendulum around the equilibrium at $(x,v) = (0,0)$. Here, in Equation 5.2.4, $f(x,v) = v$ and $g(x,v) = -\sin 2\pi x$, so the linearized system is

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x}(0,0) & \frac{\partial f}{\partial y}(0,0) \\ \frac{\partial g}{\partial x}(0,0) & \frac{\partial g}{\partial y}(0,0) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -2\pi & 0 \end{bmatrix}.$$

Note that this linear system is simply that of a harmonic oscillator $\ddot{x} - kx = 0$, with $k = -2\pi$.

One can see immediately the following:

- The eigenvalues of the matrix $\begin{bmatrix} 0 & 1 \\ -2\pi & 0 \end{bmatrix}$ are $\lambda = \pm\sqrt{2\pi}i$. Hence the linear system has a center at the origin. But according to the the Poincare-Lypaunov theorem, we cannot automatically use this to classify the origin of the non-linear system. This is true even though we do know that, in this case, at least, solutions of the undamped pendulum are in fact periodic.
- The total energy of this "classical" system is $H = \frac{1}{2}v^2 + \pi x^2$ and is conserved. Hence motion is along the level sets of H in the plane, which are concentric ellipses.
- If one solves the linear system, all of the periods of motion along the ellipses are the same. The question is, are the periods the same for the undamped pensulum?

Remark 5.24. As a mental exercise, draw an open interval Poincare section along the vertical axis of Figure 1, from the horizontal axis to the separatrix above it. Noting that near the origin, the system will look like the linear system solutions but that near the separatrix, all solutions will spend a lot of time moving slowly past the unstable equilirbria, one can reason that the periods are not the same along the concentric closed orbits. But what can you say happens as you approach the separatrix?

Exercise 145. Solve the associated linear systems at both equilibria explicitly and compare directly the linear system solutions to the nonlinear phase portraits.
.

Now linearize around the other equilibrium solution at $\left(\frac{1}{2}, 0\right)$. We get the linear system

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 2\pi & 0 \end{bmatrix}.$$

Here, the eigenvalues are $\lambda = \pm\sqrt{2\pi}$, real and distinct, and the POincare-Lyapunov Theorem classifies this equilibrium as a saddle, both for the linearized system as well as the original non-linear system. Questions: What do the level sets of total energy look like for the linear system here? Which level set corresponds to the solutions that limit to the equilibrium? Notice that in the phase cylinder (Figure 2), these are the homoclinic points of the pendulum. Can you describe just what a solution here looks like in terms of the actual mechanical device pendulum? In detail, what does a homoclinic solution look like physically? See Figure 3:

In any conservative system, the total energy H is called a *first integral* of the equations of motion. Sometimes also a *constant of the motion*. Motion is

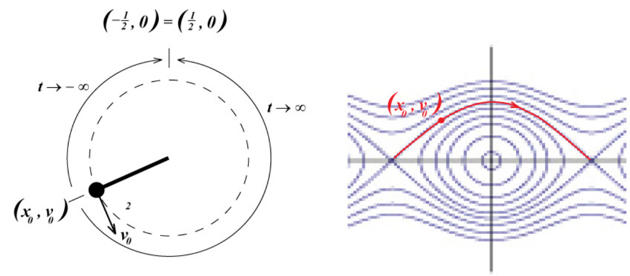


FIGURE 3. The separatrix orbit of the pendulum. It never stops advancing to the vertical position and never reaches it either.

confined to "live" on the level sets of H. If we go back to our exact ODE in Equation 5.2.2, and consider the solution function as a kind of total energy of the system, $\varphi(x,y) = H(x,y) = 4y - y^2 - 12x + x^3$. Then, as x and y vary with respect to t, so does H, and

$$\frac{\partial H}{\partial t} = \frac{\partial H}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial H}{\partial y}\frac{\partial y}{\partial t} = (-12 + 3x^2)(4 - 2y) + (4 - 2y)(12 - 3x^2) = 0.$$

So H is constant with respect to t along the flow.

So what happens in higher dimensions? For an ODE system with phase space $\mathbb{R}^n$ (or some subset), a first integral, or constant of the motion is a function $H : \mathbb{R}^n \to \mathbb{R}$. A regular level set of H (recall the notion of regular here means that the n-vector DH_x is not the 0-vector, is an $(n-1)$-dimensional subset of $\mathbb{R}^n$, called a hypersurface, given by the set of solutions to the equation $H(\mathbf{x}) = c \in \mathbb{R}$. Note that the hypersurface is regular as long as c is not a critical value of H, by the Implicit Function Theorem. The hypersurface is also called the inverse image of c in $\mathbb{R}^n$, and denoted $H^{-1}(c) \subset \mathbb{R}^n$ even though with $n > 1$ there is no possibility of the function H actually having an inverse.

Now, if one can find two such non-constant functions $G : \mathbb{R}^n \to \mathbb{R}$ and $H : \mathbb{R}^n \to \mathbb{R}$, and these two functions are "sufficiently different" from each other, then one can view solutions to the ODE system as living on the level sets of each of G and H, simultaneously. This means that solutions will live on the intersections of the two $(n-1)$-dimensional hypersurfaces. When the level sets not are tangent to each other, then motion is restricted to live on a smaller dimensional surface. This constrains the solutions further and is an effective tool for solving systems of ODEs in more than two variables. First, we need a good notion of what "sufficiently different" means here:

DEFINITION 5.25. For $n \geq m > 1$, a collection of C^1-functions $H_1, H_2, \ldots, H_m : \mathbb{R}^n \to \mathbb{R}$ are said to be *functionally independent* on some domain $U \subset \mathbb{R}^n$ if ∇H_1, ∇H_2, …, ∇H_m are linearly independent as vectors at each $x \in U$.

Note that, equivalently, one can create the function $H : \mathbb{R}^n \to \mathbb{R}^m$, $H(\mathbf{x}) = (H_1, H_2, \ldots, H_m)$. Then the Jacobian of H is just the derivative of H, denoted $DH(\mathbf{x})$. This is the $(m \times n)$-matrix whose ith row is the derivative of H_i, or the transpose of the gradient of H_i. The functions are functionally independent iff the Jacobian has full rank on U.

For example, Suppose $G : \mathbb{R}^3 \to \mathbb{R}$ and $H : \mathbb{R}^3 \to \mathbb{R}$ are two functionally independent first integrals of a system of first order autonomous ODEs in $\mathbb{R}^3$. Then the solution passing through $\mathbf{x}_0 \in \mathbb{R}^3$ must live on the intersection of the two level sets, or $\mathbf{x}(t) \subset G^{-1}(c) \cap H^{-1}(d)$ for all t where the solution is defined, and where $G(\mathbf{x}_0) = c$ and $H(\mathbf{x}_0) = d$. Figure 4 offers some examples of what this situation may look like.

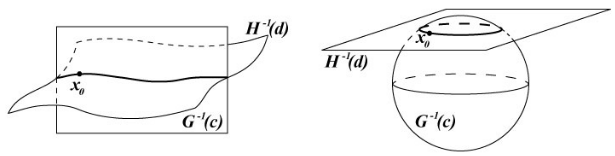


FIGURE 4. A solution in 3-space constrained by two first integrals.

5.3. Billiards

Nice treatment in www.personal.psu.edu/axk29/pub/Russian-bill.pdf.

We return to billiard maps now and present a more general situation. The 2-particle billiards is really a part of an entire field of study called convex billiards. To start, let D be a bounded, closed domain in the plane, where B is the boundary of D, so $B = \partial D$. Orbits of motion are line segments in D with endpoints in B, and adjacent line segments meet in B. When B, as a curve in the plane, is C^1, the angle which a line segment makes with the tangent to B at the end point is the same as the angle the adjacent line segment makes. This is what is meant by "angle of incidence equals angle of reflection". Should B also contain corners (points where B is not C^1), declare that an orbit entering the corner end there (this is sometimes referred to as "pocket" billiards. Motion is always considered with constant velocity on line segments, and collisions with B are specular (elastic).

Some dynamical criteria:

- Every orbit is completely determined by its starting point and direction.
- Recall for polygonal billiards, a billiard flow is continuous flow per unit time. It is certainly not a differentiable flow, as it fails at the collisions with B (Note: One can certainly define a smooth flow whose trajectory has corners. All that is necessary is for the flow to slow up and momentarily stop at the corner, to allow it to change direction smoothly. This is quite common for parameterized curves. Here, though, the flow does not slow up.)
- In the billiard flow on the triangle, we cured the non differentiable flow points by "unfolding" the table. Here, instead, we will analyze this situation by creating a completely different state space which collects only the relevant information from the actual billiard.

First, ignore the time between collisions of line segments with B, and consider orbits as simply a sequence of points on B, along with their angle of incidence. For each collision of an orbit with B, the point and the angle completely determine the next point and angle of collision. In the "space" of points of B and possible angles of collision, we get an assignment of the next point of collision and angle for each previous one. It turns out that this assignment is quite well defined. Call this assignment Φ, where $(x_1, \theta_1) \mapsto (x_2, \theta_2) \mapsto \cdots \mapsto (x_n, \theta_n) \mapsto \cdots$. For now, let B be C^1. Collect up all of the points of B, and you get a copy of S^1. Collect up all possible angles of incidence and you get the interval $[0, \pi]$ (really one gets the open interval, but one can limit to the orbits that simply run along the length of B. This is not such an important factor here. The state space is all of the points in B along with all of the incidence angles is a copy of $C = S^1 \times [0, \pi]$, the cylinder. The assignment takes $(x_1, \theta_1) \mapsto (x_2, \theta_2) \mapsto \cdots \mapsto (x_n, \theta_n)$. See the figure. The resulting cylinder, along with the evolution map Φ is called the billiard map.

EXAMPLE 5.26. Let

$$D = \left\{ (x, y) \in \mathbb{R}^2 \;\middle|\; x^2 + y^2 \leq 1 \right\}$$

be the unit disk in the plane. Here $B = \partial D = S^1$ is the unit circle, parameterized by the standard angular coordinate θ from polar coordinates in the plane (note that this parameter takes values in $[0, 2\pi)$ and is quite different from the parameterization we have been using for S^1 given by the exponential map $x \mapsto e^{2\pi i x}$). The state space is then $C = S^1 \times I$, where $I = [0, \pi]$. What are the dynamics? Go back to the light-ray in a circular mirrored room exercise from before. You will find that the initial angle of incidence never changes, and the evolution map is constant on the second coordinate.

EXERCISE 146. Show for S^1 the unit circle, that $\Phi(s, \theta) = (s + 2\theta, \theta)$.

EXERCISE 147. Show that this is not quite true for a billiard table whose radius in not 1.

Now do you recognize the evolution map on the state space in this dynamical system? This is basically the twist map on the cylinder, a map that you already showed was area preserving. And you already know the dynamics of this map. To continue our study, we can say more about the orbit structure within each invariant cross-section (constant θ section) of the cylinder: To each $\theta = \theta_0$ is associated a *caustic*:

- In optics, a caustic is the envelope of light rays reflected or refracted by a curved surface or object, or the projection of that envelope of rays on another surface.
- Or the caustic is a curve or surface to which each of the light rays is tangent, defining a boundary of an envelope of rays as a curve of concentrated light.
- In differential geometry and geometric optics (mathematics, in general), a caustic is the envelope of rays (directed line segments) either reflected or refracted by a manifold.

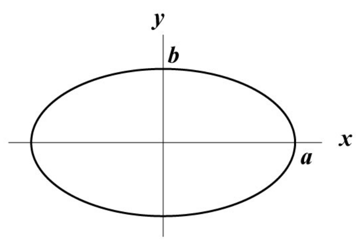
EXERCISE 148. For the circle billiard, let $\theta \notin \mathbb{Q}$. Then the caustic is the edge of the region basically filled with light. What shape is this caustic, and can you write the equation for this caustic as a function of the angle θ_0.

EXPERIMENT 1. Shine a light from a small hole horizontally into a circular mirrored room. Try to pass the light beam directly through the center of the room (force $\theta_0 = \frac{\pi}{2}$. What happens as you "focus" the light? How does the light fill the room as you approach $\frac{\pi}{2}$, and when you reach $\frac{\pi}{2}$?

Example 5.27. Let

$$D = \left\{ (x,y) \in \mathbb{R}^2 \,\middle|\, \frac{x^2}{a^2} + \frac{y^2}{b^2} \le 1 \right\}$$

be an ellipse, where the *diameter* is $2a$ and the *width* is $2b$.



Recall that a definition of an ellipse is the set of points in the plane whose combined distance from two reference points is a constant. The two reference points are the two foci of the ellipse, and in our case, the combined distance constant is the diameter $2a$. Written as above, the ellipse is centered at the origin in the plane, and the diameter and the width are along the horizontal and vertical axes, respectively. Upon inspection, one can see that it will not be the case, as with the circle, that there is a period-2 orbit passing through each point. There are four such points, though, and all four of these lie on one of the axes. Why is this true? We will see.

This billiard table has notable differences from the circular one, beyond the relative lack of period-2 orbits. To understand these difference better, we introduce a technique of study common in billiards: Generating Functions.

Parameterize the boundary by arc-length s and let p and p' be 2 points on $B = \partial D$. Now define a real-valued function on $B \times B$ by

$$H\left((s,s'\right) = -d\left(p,p'\right),$$

where d is the standard Euclidean metric in the plane. This function H is called the *generating Function* for the billiard:
Some notes:

- This function helps to identify points on the same orbit.
- Critical points of H determine period-2 orbits (think about what this means for the ellipse.)
- rarely can we find a good working expression for H in terms of s and s'. But we can discuss its properties and use them effectively.

Example 5.28. Let $a = b = 1$, and we are back at the circular billiard. Here $H(s,s') = -2\sin\frac{1}{2}(s-s')$.

Exercise 149. Derive this function using the geometry of the unit circle.

Exercise 150. For $a > b$, we do not have a good expression for H. However, we can surmise that the diameter boundary points are at a minimum for H (remember the minus sign), and the width boundary points are a saddle point for H. Why is this? Can you see it?

5.3.1. Dynamics of elliptic billiards. As in circular billiards, one way to discuss the orbit structure for an elliptic billiard is to try to describe any possible caustics (curves tangent to orbits, which help to define edges of envelopes of orbit regions. We have two results here:

Proposition 5.29. *Every smaller confocal (having the same foci) ellipse is a caustic.*

The proof here is constructive and can be found in the book. This family of ellipses works as a caustic for any orbit segment that does not pass between or meet the foci. Convince yourself that if an orbit segment does not meet or pass between the foci, then the entire orbit will not intersect the closed line segment connecting the foci. And once an orbit segment crosses that line, it will continue to cross that line both forward and backward for each line segment in the orbit. And if an orbit segment passes through a focus, where will it go next? Where will it go over the long term?

PROPOSITION 5.30. *There exists a caustic for every ray between the foci. The caustic of the orbit corresponding to this ray is both pieces of a hyperbola confocal to the ellipse.*

Note: Ellipses and hyperbolas are both conic sections, and related via their *eccentricity*, a nonnegative number that parameterizes conic sections via a ratio of their data. Indeed, along the major axis (the diameter) of a conic section, one can measure the distance from the curve to the origin (let's keep all conic section centered at eh origin for now). Call this the radius a. One can also measure the distance from the center to one of the foci. call this c. Then eccentricity e is the ration of these two numbers:

- For $e = \dfrac{c}{a} = 0$ (implying that $c = 0$), the section is a circle.
- For $0 < e = \dfrac{c}{a} < 1$, the section is an ellipse.
- For $e = \dfrac{c}{a} = 1$, the section is a parabola.
- For $e = \dfrac{c}{a} > 1$, the section is a hyperbola.

For the circle case, the equation is elliptical, with $a = b$, and we have $\dfrac{x^2}{a^2} + \dfrac{y^2}{a^2} = 1$, or $x^2 + y^2 = a^2$. For the hyperbolic case, we have $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1$.
And as a couple of final dynamic notes;

- An orbit that passes through one focus must pass through the other . What are the implications of this for the resulting orbit?
- There are tons of periodic orbits in elliptic billiards, of all periods. Can you draw some? Period-4 should be easy to see, as is period-2. How about period-3?

EXERCISE 151. Construct a period-4 orbit for an elliptic billiard and show analytically that it exists.

EXERCISE 152. Describe the long term behavior of ANY orbit that has a orbit segment that pass through one of the foci.

Going back to the generating functions, we can say more about orbits in general. Here are some properties. Recall for any convex billiard table, $H(s,s') = -d(p,p')$, where d is the standard Euclidean metric in $\mathbb{R}^2$.

LEMMA 5.31. $\dfrac{\partial}{\partial s'}H(s,s') = -\cos\theta'$, *and* $\dfrac{\partial}{\partial s}H(s,s') = \cos\theta$.

PROOF. This really is simply calculus. For the first result, fix and parameterize a small arc in the ellipse centered at s', $c(t)$, where $c(t_0) = p'$. Choose a parameterization such that the tangent vector is unit length. Then, noting that $d(p,p') = \|p - p'\|$, we have

$$\frac{\partial}{\partial s'} H(s,s') = \left.\frac{d}{dt}\right|_{t=t_0} - d\,(p, c(t)) = \frac{-1}{2d(p,c(t))}\left.\left(2\left(c'(t)\cdot(p-c(t))\right)\right)\right|_{t=t_0}$$
$$= \frac{-\|c'(t_0)\|\,\|p-c(t_0)\|\cos\theta'}{\|p-c(t_0)\|} = -\cos\theta'$$

by the cosine formula for the dot product of two vectors and since $\|c'(t_0)\| = 1$ by the parameterization. Hence $\dfrac{\partial H}{\partial s'} = -\cos\theta'$. The other result is similar. $\square$

Now apply this idea to any three points s_{-1}, s_0, and s_1 on the ellipse. can these three points lie successively on an orbit? The answer is , yes, if

$$\frac{\partial}{\partial s'} H(s,s') + \frac{\partial}{\partial s} H(s,s') = 0.$$

That is, if s_0 is a critical point of the assignment

$$s \longmapsto H(s_{-1}, s) + H(s, s_1).$$

This is a variational approach to the construction of orbits, and techniques like this form the content of our course 110.427 Introduction to the Calculus of Variations.

Experiment 2. Consider a convex billiard with one pocket (corner) p. Find all possible bank shots to sink a ball at p.

Refer back to the definition of convex (and strictly convex) back in the section n Contractions.

Recall the notion of a strictly convex domain, where $B = \partial D$ has non-zero curvature (where B is C^2 and where the second derivative is non-zero). Visually, this means that there are no straight-line segments on B, and certainly no inflection points (changes in concavity). It also means that we can effectively take the angle of incidence to be from the open interval $(0,\pi)$ instead of the closed interval. Thus the state space is the open cone.

Here are some quick results: First, switch from the angular coordinate θ to the rectilinear coordinate $r = -\cos\theta$, so that for $\theta \in (0,\pi)$, we have $r \in (-1,1)$.

Proposition 5.32. *For a convex billiard, the billiard map*

$$\Phi(s,r) = (\mathcal{S}(s,r), \mathcal{R}(s,r)) : C \to C$$

is area and orientation preserving.

Proof. The proof is constructive and based on simply calculus. $\square$

Proposition 5.33. *If $B = \partial D$ is C^k (which means that the Euclidean coordinates are C^k functions of the length parameter), then both $\mathcal{S}$ and $\mathcal{R}$ are C^{k-1}.*

Proof. This is the Implicit Function Theorem. $\square$

Proposition 5.34. *For D strictly convex, the billiard map has at least two period-2 orbits; at the diameter and at the width.*

One can describe the width in the following way. Take two distinct vertical parallel lines tangent to the billiard table (necessarily on "opposite sides" of the table). As one rotates the table, the distance between these lines will change. When one reaches the diameter of the table (the largest possible Euclidean distance between 2 points on the boundary), the two points will lie along the line perpendicular to the vertical lines. This perpendicular line segment represents one of the period-2 orbits. The other comes comes at the point when the two vertical lines reach a local minimum distance (which is the minimum distance for a strictly convex table). At this point again, the line segment joining the two tangencies will be perpendicular to the vertical lines and represent another period-2 orbit. This is the width of the table.

One final note, finding these period-2 orbits using this method involves finding where the vertical lines reach a minimum and maximum distance from each other. But this is what the generating function H is also doing, and why the generating function is particular good at finding period-2 orbits. It is actually good at finding period-n orbits also, but this goes a bit beyond this course.

5.3.2. Application: Pitcher Problems. In the movie "Die Hard with a Vengeance", John McClain (Bruce WIllis) and Zeus (Samuel L. Jackson) are confronted with a puzzle in their ordeal to stop a terrorist's plot run by Peter Krieg (Jeremy Irons). The puzzle is one of the Pitcher Problems: Given two pitchers of a certain size, in integer gallons, how can one measure out some intermediate integer gallon amount through various fillings, emptying and transferring of fluid form one pitcher to another. In the movie, near a fountain stands two empty jugs, one a 5 gallon jug and the other a 3 gallon jug. There is also a bomb here, which can be disabled by weighing out exactly 4 gallons of fluid. The heroes figure out the process as a 6 stage one: Fill the 5 gallon jug; use it to fill the three gallon jug; empty the three gallon jug; put the remaining 2 gallons into the 3 gallon jug; fill the 5 gallon jug; and use it to refill the 3 gallon jug; What remains in the 5 gallon jug is 4 gallons. How they came up with this procedure? Watch the movie. Is there a systematic way to find the shortest procedure for doing this? Yes, and billiards is one way to work it out.

Construct a parallelogram billiard table with an acute angle of $\frac{\pi}{6}$ radians, whose side lengths are, in this case, 5 and 3 units long, as in Figure 5. The integer points in this parallelogram form a lattice. Define the corner collisions in the following way:

- if a point meets an obtuse corner coming in along an edge, then the collision looks like a collision with the other wall, with reflection angle $\frac{\pi}{6}$.
- any other corner collision ends the orbit (this is all that we will need.)

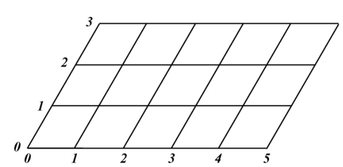


FIGURE 5. A 5×3-Parallelogram Billiard Table.

Now create an orbit that begins at the lattice point $(0,0)$ and runs along one of the edges. You have two choices. What is it's fate? Well, one result is that the orbit will always end at the obtuse corner opposite the first one encountered in the orbit. But even more interestingly, the orbit will meet every lattice point along the boundary, except for the opposite acute corner. In fact, for a table like this with side lengths $p \leq q$, this is true iff p and q are relatively prime (have no factors in common except for 1. Why is this important?

Follow the orbit. Every horizontal segment corresponds to either filling or emptying one of the jugs and every vertical segment (defined along the other edge) corresponds to either filling or emptying the other jug. The diagonal segments (diagonal with respect to the sides) corresponds to transferring fluid from one jug to the other. So each lattice point (a, b) corresponds to the current state of the fluid in each jug and at leat one of entries in each edge collision is either 0 or full. The result is that each lattice point in an orbit corresponds to a "next move". If and when the desired entry k appears as a part of an orbit, you can count how many moves it takes to create that desired amount of fluid. There are only two possible starting moves, so there is a shortest path (possibly two of them?) There is a beautiful number theory result concerning just how long these shortest orbits can be to reach any desired intermediate measure.

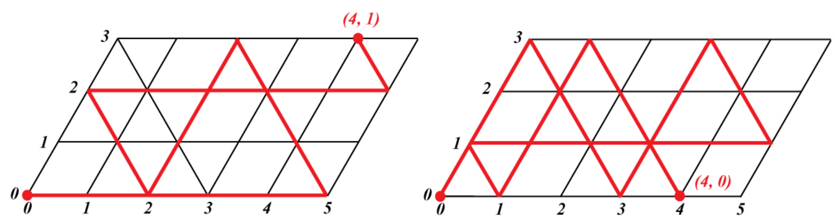


Figure 6. The two orbits from $(0, 0)$ to the first instance of a boundary lattice point with a 4.

In general, for any two pitchers of size $p < q \in \mathbb{N}$ units, the number of fillings, emptyings and transfers needs to measure out any $r < p \in \mathbb{N}$ is at most $q + 1$. Finding it, however, by simply exploring possibilities, may take a while. With a parallelogram billiard, just take a shot!

Exercise 153. Find a way to measure out 1 gallon of fluid with only two jugs of sizes 8 and 11 gallons. What is the shortest number of fillings, emptyings and transfers to do so.

Exercise 154. Given two jugs of sizes 6 gallons and 9 gallons, determine precisely which intermediate measurements are NOT possible.

Exercise 155. Find a way to cook a perfectly timed 11 minute boiled egg, using only a 5 minute egg-timer and a 9 minute egg-timer.

CHAPTER 6

Complicated Orbit Structure

6.1. Counting Periodic Orbits

It seems like we spend a lot of time in our study of dynamical systems on the classification and counting of periodic orbits of a map $f : X \to X$. To understand why, consider

- n-periodic points are the fixed points of the map f^n.
- Periodic points, like fixed points, have stability features.
- There are existence theorems for periodic points.
- many times, we can "solve" for them, without actually solving the dynamical systems.

Recall Definition 2.22 for the set of n-periodic points of a map $f : X \to X$: $\mathrm{Per}_n(f) := \left\{x \in X \,\middle|\, f^n(x) = x\right\}$. Here, we are interested in the cardinality of this set.

Definition 6.1. For $f : X \to X$ a map, let

$$P_n(f) := \#\left\{x \in X \,\middle|\, f^n(x) = x\right\}$$

be the number of all n-periodic points of f. And let

$$P(f) := \bigcup_{n\in\mathbb{N}} P_n(f).$$

Note that $P_n(f)$ also includes all m-periodic point when $m|n$. In particular, the 1-periodic points are the fixed points and these are counted in $P_n(f)$ for all $n \in \mathbb{N}$.

As a sequence, $\{P_n(f)\}_{n\in\mathbb{N}}$ can say a lot about f.

Consider the map $E_2 : S^1 \to S^1$, $E_2(z) = z^2$, where $z = e^{2\pi i x} \in \mathbb{C}$ is a complex number restricted to the unit modulus complex numbers. Another way to see this map is $E_2(s) = (2s \mod 1)$, for $s \in S^1$, depending on your model for S^1.

Proposition 6.2. *$P_n(E_2) = 2^n - 1$, and all periodic points are dense in S^1 (i.e., $\overline{P(E_2)} = S^1$).*

Proof. Using the model $E_2(z) = z^2$, we find that z is an n-periodic point if

$$\left(\cdots\left(\left(z^2\right)^2\right)\cdots\right)^2 = z \text{ or } z^{2^n} = z \text{ or } z^{2^n-1} = 1.$$

Thus every periodic point is an order-$(2^n - 1)$ root of unity (and vice versa). And there are exactly $2^n - 1$ of these, uniformly spaced around the circle. In fact, to any rational $\frac{p}{q} \in \mathbb{Q}$, the point $e^{2\pi i\left(\frac{p}{q}\right)}$ is a qth root of unity. If $q = 2^n - 1$, for $n \in \mathbb{N}$, then $e^{2\pi i\left(\frac{p}{q}\right)}$ is an order-n fixed point. Now as n goes to ∞, the spacing between order-$(2^n - 1)$ roots of unity goes to 0. Hence any point $x \in S^1$ can be written as the limit of a sequence of these points. Hence will be in the closure of $P(E_2)$. $\square$

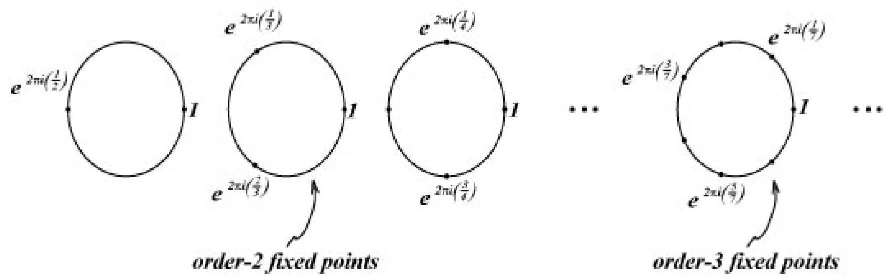


FIGURE 1. Order-n fixed points of E_2 spaced evenly in S^1

EXAMPLE 6.3. z is 2-periodic if $z^{2^2-1} = z^3 = 1$. These are the points $z = e^{2\pi i(\frac{k}{3})}$, for $k = 1, 2$. To see how this works,

$$E_2(e^{2\pi i(\frac{1}{3})}) = \left(e^{2\pi i(\frac{1}{3})}\right)^2 = e^{2\pi i(\frac{1}{3})*2} = e^{2\pi i(\frac{2}{3})}, \text{ while}$$
$$E_2(e^{2\pi i(\frac{2}{3})}) = \left(e^{2\pi i(\frac{2}{3})}\right)^2 = e^{2\pi i(\frac{2}{3})*2} = e^{2\pi i(\frac{4}{3})} = e^{2\pi i(\frac{1}{3})}.$$

We can calculate the growth rate of $P_n(E_2)$ in the obvious way: Define the *truncated* natural logarithm

$$\ln_+ x = \begin{cases} \ln x & x \geq 1 \\ 0 & \text{otherwise} \end{cases}.$$

Then define $p(f) = \overline{\lim_{n\to\infty}} \dfrac{\ln_+ P_n(f)}{n}$ as the relative logarithmic growth of the number of n-periodic points of f with respect to n.

For our case, then, where $E_2(z) = z^2$,

$$p(E_2) = \overline{\lim_{n\to\infty}} \frac{\ln_+(2^n - 1)}{n} = \overline{\lim_{n\to\infty}} \frac{\ln_+ 2^n(1 - 2^{-n})}{n}$$
$$= \overline{\lim_{n\to\infty}} \frac{\ln_+ 2^n + \ln_+(1 - 2^{-n})}{n} = \ln 2.$$

This is the exponential growth rate of the periodic points of the map E_2. Note that the growth factor is 2 at each stage, hence the exponential growth rate is the exponent of e which corresponds to the growth factor. Here $2 = e^{\ln 2}$.

PROPOSITION 6.4. *For $f : S^1 \to S^1$, $f(z) = z^m$, where $m \in \mathbb{Z}$ and $|m| > 1$,*

$$P_n(f) = |m^n - 1|,$$

the set of all periodic points is dense in S^1, and $p(f) = \ln |m|$.

EXERCISE 156. Show this for $m = -3$.

Here is an interesting fact: Let $f(z) = z^2$. The image of any small arc in S^1 is twice as long as the original arc. However, there are actually 2 disjoint pre-images of each small arc, and each is exactly half the size. Combined, the sum of the lengths of these two pre-images exactly matches the length of the image. Thus this expanding map on S^1 actually preserves length! Some notes about this:

- This is actually true for all of the expanding maps $E_m : S^1 \to S^1$, $E_m(z) = z^m$, where $m \in \mathbb{Z}$, and $|m| > 1$.
- This is a somewhat broadening of the idea of area preservation for a map. When the map is onto but not 1-1 (in this case, the map is 2-1), the relationship between pre-image and image is more intricate, and care is needed to understand the relationship well.

6.1.1. The quadratic family. For $\lambda \in \mathbb{R}$, let $f_\lambda : \mathbb{R} \to \mathbb{R}$, where $f_\lambda(x) = \lambda x(1-x)$ is also called the logistic map on $\mathbb{R}$. For $\lambda \in [0,4]$, we can restrict to $I = [0,1]$, and $f_\lambda : I \to I$ is the interval map family we partially studied already. In fact, we can summarize our results so far: For $\lambda \in [0,3]$, the dynamics are quite simple. There are only fixed points, and no nontrivial periodic points, and all other points are asymptotic to them. The fixed points are at $x = 0$ and $x = 1 - \frac{1}{\lambda}$.

Some new facts:

(1) for $\lambda \in [3,4]$, a LOT happens! (we will get to this later in the course.)
(2) for $\lambda > 4$, I is not invariant.
(3) since f_λ is quadratic, f_λ^n is at most of degree 2^n. Thus the set of n-periodic points must be solutions to the equation $f_\lambda^n(x) = x$. Bringing x to the other side of the equation, the set $P_n(f_\lambda)$ must consists of the roots of an (at most) 2^n-degree polynomial. Hence
$$P_n(f_\lambda) \leq 2^n, \text{ for all } \lambda \in \mathbb{R}.$$
(4) For $\lambda > 4$, many points escape the interval I. However, as we will see, many points have orbits which do not. We can still talk about the map on the set of all of these points....

Let $\lambda > 4$, and consider the first iterate of f_λ. Notice (see the figure), that the intervals I_1 and I_2 are both mapped onto $[0,1]$ and that each contains exactly one fixed point. Under the second iterate of the map, f_λ^2, only points in the 4 intervals J_i, $i = 1,2,3,4$ remain in $[0,1]$. Here there are 4 fixed points (again one in each interval). But notice that only two of them are new, y_1 and y_2. These two new ones are period-2 points that are not fixed points. See in the cobwebbed figure the period-2 orbit on the right of the figure.

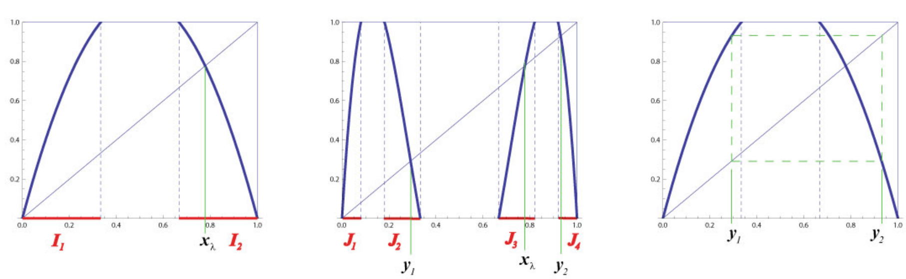


FIGURE 2. The map f_λ, f_λ^2, and the period-2 orbit

Continue iterating in this fashion, and one can see that there will be

- 2^n intervals of points that remain in I after n-iterates.
- The next iterate of f_λ maps each of these 2^n intervals onto $[0,1]$, creating a single fixed point in each interval (of f_λ^n).

- You can see then (and one can prove this by induction) that $P_n(f_\lambda) = 2^n$, when $\lambda > 4$.

EXERCISE 157. Show that for f_λ where $\lambda > 4$, if $\mathcal{O}_x^+ \notin I$, then $\mathcal{O}_x \longrightarrow -\infty$. Also show that once an iterate of x under f_λ leaves I, it never returns.

So what about the points whose orbits stay in I? We can construct this as follows: For $x \in I$, call $\mathcal{O}_x^n$ the nth partial (positive) orbit of x, where

$$\mathcal{O}_x^n = \left\{ y \in I \;\middle|\; y = f^i(x),\ i = 0, \dots, n-1 \right\}.$$

Then define

$$C_n = \left\{ x \in I \;\middle|\; \mathcal{O}_x^n \in I \right\}.$$

Then $C_0 = I$, $C_1 = I_1 \cup I_2$, $C_2 = J_1 \cup J_2 \cup J_3 \cup J_4$, and $C_n \subset C_{n-1}$ for all $n \in \mathbb{N}$. Finally, define

$$C = \bigcap_{n=0}^{\infty} C_n,$$

Then $f_\lambda : C \to C$ is a discrete dynamical system.

What does this set C look like? For starters, it seems quite similar in construction to our Canter Ternary Set. Be careful here, though. The connected subintervals of C_n will not always be the same length in C_n. You can see this in the above figure, but should also check specifically for C_2. It should be certain that if $x \in I$ is n-periodic, then $x \in C$. But are these the only points whose entire orbit lies in I? What about a point $y \in [0,1]$ which is well-approximated by periodic points? This means that there is a sequence of periodic points in I which converges to y. Is that enough to ensure that $\mathcal{O}_y \in [0,1]$? This is an important question (which should be yes, by continuity.) It turns out that there are a lot of non-periodic points in C. In fact, there are an uncountable number. In fact, a Cantor's set-worth! To see this, we need a better definition of a Cantor Set than what comes from our Canter Ternary Set above.

DEFINITION 6.5. A non-empty subset of I is called a *Cantor Set* if it is a closed, totally-disconnected, perfect subset of I.

DEFINITION 6.6. A non-empty subset $C \subset I$ is *perfect* if, for every point $x \in C$, there exists a sequence of points $x_i \in C$, $i \in \mathbb{N}$, where $\{x_i\}_{i \in \mathbb{N}} \longrightarrow x$.

DEFINITION 6.7. A non-empty subset $C \subset I$ is *totally-disconnected* if, for every $x, y \in C$, the closed interval $[x, y] \notin C$.

Roughly, there are no isolated points in a perfect set. And there are no closed, positive-length intervals in a totally disconnected subset of an interval.

PROPOSITION 6.8. *Let $f_\lambda : I \to \mathbb{R}$ be defined by$f_\lambda(x) = \lambda x(1-x)$, where $\lambda > 4$ and let*

$$C = \left\{ x \in I \;\middle|\; \mathcal{O}_x \in I \right\}.$$

Then C is a Cantor Set and $f_\lambda\big|_C$ is a discrete dynamical system.

PROOF. By the exercise above, we already know that all periodic points are in C. For the moment, let's consider only the case that $\lambda > 2 + \sqrt{5} > 4$. In this case, we are assured that $|f_\lambda'(x)| > \mu > 1$, $\forall x \in C_1$ and some number μ.

EXERCISE 158. Verify this fact.

And hence, by the Chain Rule, we have $|f'_\lambda(x)| > 1$, $\forall x \in C$. By Definition **??**, we know f_λ is expanding. Since C is an arbitrary intersection of closed sets, it is certainly closed.

As for totally-discontinuous, let's assume that for $x, y \in C$, where $x \neq y$, the interval $[x, y] \in C$. Then the orbit of the entire interval lies completely in C. But since f_λ is expanding, $|f(x) - f(y)| > \mu|x - y|$. And for each $n \in \mathbb{N}$, $|f^n(x) - f^n(y)| > \mu^n|x - y|$. Choose $n > -\frac{\ln|x-y|}{\ln\mu}$. Then $|f^n(x) - f^n(y)| > 1$. But then $f^{n+1}([x,y]) \notin C$. This contradiction means that no positive-length intervals exist in C, and establishes that C is totally discontinuous.

To see that C is perfect, assume for a minute that there exists an isolated point $z \in C$. Being isolated means that there is a small open interval $U(z) \subset I$, where for all $x \in U(z)$, where $x \neq z$, we have $x \notin C$. Now, since $z \in C$, it is in a subinterval of every C_n. For any choice of $n \in \mathbb{N}$, call the interval $[x_n, y_n] \in C_n$ where $z \in [x_n, y_n]$. Create a sequence of nested closed intervals $\{[x_i, y_i]\}_{i \in \mathbb{N}}$, where for every i, $z \in [x_i, y_i] \subset C_i$. Each endpoint x_i is eventually fixed and hence $x_i \in C$ for all $i \in \mathbb{N}$. But C is totally disconnected. Hence the intersection

$$\bigcap_{i=1}^{\infty}[x_i, y_i]$$

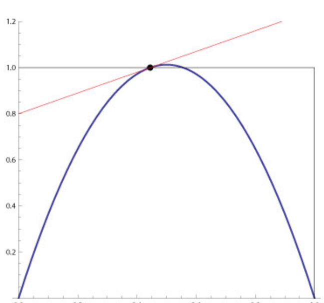
can only consist of one point, and z is in this set. Thus, as a sequence $\{x_i\}_{i\in\mathbb{N}} \longrightarrow z$, and z is NOT isolated in C. Hence C is perfect, and hence C is a Canter Set.

As a final note, we will relegate a discussion of why C is still a Canter set when $4 < \lambda < 2 + \sqrt{5}$ to the following remark, noting that the proof requires a subtle bit of finesse not totally germane to the current discussion. □

Remark 6.9. When $4 < \lambda < 2 + \sqrt{5}$, the map f_λ is not expanding on C_1. Indeed, for $\epsilon > 0$, let $\lambda = 4 + \epsilon$. Then the first intersection of the graph of f_λ and the $y = 1$ line is at $x_1 = \frac{1}{2}\left(1 - \sqrt{1 - \frac{4}{\lambda}}\right)$. The derivative of f_λ at this crossing is $f'_\lambda(x_1) = \sqrt{\lambda^2 - 4\lambda}$, which evaluates (when $\lambda = 4 + \epsilon$, see the figure below, to

$$f'_\lambda(x_1) = \sqrt{\lambda^2 - 4\lambda} = \sqrt{4\epsilon + \epsilon^2} > 2\sqrt{\epsilon}.$$

The derivative of the square of the map at x_1 has a much higher derivative since the derivative of the image of x_1 is $-\lambda = -(4 + \epsilon)$ at the image point $f_\lambda(x_1) = 1$. Hence the derivative of the square of this map is greater than $8\sqrt{\epsilon}$. This happens all though the interval, and the map can be said to be *eventually expanding*, in that $\exists N \in \mathbb{N}$ where for all $n > N$ the map $f^n_\lambda(x)\big|_{C_n}$ is expanding. Then the proof above holds. Thus the proposition is true for all $\lambda > 4$.



This quadratic family is an example of a *unimodal* map: A continuous map defined on an interval that is increasing to the left of an interior point and decreasing thereafter.

Proposition 6.10. *Let $f : [0,1] \to \mathbb{R}$ be continuous with $f(0) = f(1) = 0$ and suppose there exists $c \in (0,1)$ such that $f(c) > 1$. Then $P_n(f) \geq 2^n$. If, in addition, f is unimodal and expanding, then $P_n(f) = 2^n$.*

DEFINITION 6.11. A map $f : [0,1] \to [0,1]$ is *expanding* if

$$|f(x) - f(y)| > |x - y|$$

on each interval of $f^{-1}([0,1])$.

Examples of expanding maps include the logistic map for suitable values of $\lambda > 4$, and the circle maps E_m, where $m \in \mathbb{Z}$ and $|m| > 1$ (you should modify the definition here to include maps of S^1). Note here:

- In the Proposition, the condition $f(0) = f(1) = 0$ and continuity ensure that the map will "fold" the image over the domain, and
- the condition $f(c) > 1$ ensures the folding will be complicated, with lots of points escaping, while lots of points will not.

6.1.2. Expanding Maps. Here is a better definition of an expanding map (albeit limited now to circle maps):

DEFINITION 6.12. A C^1-map $f : S^1 \to S^1$ is expanding if $|f'(x)| > 1$, $\forall x \in S^1$.

EXAMPLE 6.13. It should be obvious by this definition that the map E_m, where $m \in \mathbb{Z}$ and $|m| > 1$ is expanding, since $E_m(x) = mx \mod 1$ is differentiable and $|E'_m(x)| = |m| > 1$ for all $x \in S^1$.

Recall that the degree of a circle map is a well defined property that measures how many times the image of a map of S^1 winds itself around S^1.

LEMMA 6.14. *Let $f, g : S^1 \to S^1$ be continuous. Then*

$$\deg(g \circ f) = \deg(g)\deg(f).$$

PROOF. Degree is defined via a choice of lift: Given lifts $F, G : \mathbb{R} \to \mathbb{R}$ of these two maps, we have for $s \in S^1$ and $k \in \mathbb{Z}$,

$$G(s+k) = G(s+k-1) + \deg(g) = G(s+k-2) + 2\deg(g) = \cdots = G(s) + k\deg(g).$$

But this means

$$G(F(s+1)) = G(F(s) + \deg(f)) = G(F(s)) + \deg(f)\deg(g).$$

□

EXAMPLE 6.15. $\deg(f^n) = (\deg(f))^n$.

Hence we can use this to show:

PROPOSITION 6.16. *If $f : S^1 \to S^1$ is expanding, then $|\deg(f)| > 1$ and $P_n(f) = \left|(\deg(f))^n - 1\right|$.*

6.1.3. Hyperbolic Toral Automorphisms. Here is a 2-dimensional version of periodic point growth. Let $L : \mathbb{R}^2 \to \mathbb{R}^2$, $L(x,y) = (2x+y, x+y)$. We can also write L as the linear vector map

$$L(\mathbf{x}) = A\mathbf{x}, \text{ where } A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}.$$

REMARK 6.17. Note here that when vector spaces are finite dimensional and bases are given or understood, then any linear map can be represented by the matrix defining the action. Hence we will simply call L the transformation of the plane given by the matrix A, and refer to L as either the map or the matrix $L = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$. In mathematics, this is known as an abuse of notation. And when little added confusion results, it can be an effective way to reduce the number of objects involved in a discussion.

We know that since A has integer entries, it takes integer vectors to integer vectors, and hence descends to a map on the two torus $\mathbb{T}^2$. Indeed, if $\mathbf{x}_1 = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$ and $\mathbf{x}_2 = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$ satisfy $\mathbf{x}_1 - \mathbf{x}_2 \in \mathbb{Z}^2$, then

$$L(\mathbf{x}_1 - \mathbf{x}_2) = L(\mathbf{x}_1) - L(\mathbf{x}_2) \in \mathbb{Z}^2.$$

But then $L(\mathbf{x}_1) - L(\mathbf{x}_2) = \mathbf{0} \mod 1$, which means $L(\mathbf{x}_1) = L(\mathbf{x}_2) \mod 1$. Hence the map L induces a map on $\mathbb{T}^2$ which assigns

$$(x,y) \longmapsto (2x + y \mod 1, x + y \mod 1).$$

We will call this new *induced map* on the torus $F_L : \mathbb{T}^2 \to \mathbb{T}^2$, where

$$F_L(\mathbf{x}) = A\mathbf{x}, \quad A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}, \quad \mathbf{x} \in \mathbb{T}^2.$$

Some notes:

- This map is an *automorphism* of $\mathbb{T}^2$: A homeomorphism that preserves also the ability of points on the torus to be added together (multiplied, if one defines the multiplication correctly).
- F_L is also invertible since it is an integer matrix of determinant 1. The inverse map $F_L^{-1} : \mathbb{T}^2 \to \mathbb{T}^2$ is given by the matrix $A^{-1} = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$.
- The eigenvalues of F_L (really the eigenvalues of A) are the solutions to the quadratic equation $\lambda^2 - 3\lambda + 1 = 0$, or
$$\lambda = \frac{3 \pm \sqrt{5}}{2}.$$
Note that
$$\lambda_1 = \frac{3 + \sqrt{5}}{2} > 1, \qquad \text{and } \lambda_2 = \lambda_1^{-1} = \frac{3 - \sqrt{5}}{2} < 1,$$
so that the matrix defining F_L is a hyperbolic matrix (determinant-1 with eigenvalues off the unit circle in $\mathbb{C}$). Hence L here is a hyperbolic map of the plane, given the classification in Section 3.3. Than F_L is is an example of a *hyperbolic toral automorphism*.

Generalize all of this to hyperbolic toral automorphisms: Let $L : \mathbb{R}^2 \to \mathbb{R}^2$, $L(x,y) = (ax + by, cx + dy)$ be a linear map of $\mathbb{R}^2$, where $ad - bc = 1$ and $|a + c| > 2$. These are the hyperbolic planar maps that descend to hyperbolic toral automorphisms F_L.

EXERCISE 159. Show that a determinant-1, 2×2-matrix with integer entries is hyperbolic iff the trace has magnitude greater than 2.

QUESTION 6.18. *How does F_L act on $\mathbb{T}^2$?*

Really the answer to this question relies on how L acts on $\mathbb{R}^2$. Watching the model of $\mathbb{T}^2$ as the unit square in $\mathbb{R}^2$ as it is acted on by L provides the means to study the F_L action on $\mathbb{T}^2$. This is the two dimensional version of studying a lift of a circle map on $\mathbb{R}$ as a means of studying the circle map.

Linear maps of the plane take lines to lines. Hence they take polygons to polygons, and, in this case, they take parallelograms to parallelograms. The image of the unit square can be found by simply finding the images of the four corners of the square and constructing the parallelogram determined by those points by connecting corresponding adjacent point via lines. See the left side of Figure 3. But there is more. L is area preserving. Hence the image of the parallelogram will also have area 1 (remember the discussion around Equation 5.1.1.) And due to the equivalence relation given by the exponential map on $\mathbb{R}^2$, every point in the image of the unit square has a representative within its equivalence class INSIDE the unit square. We can reconstruct the unit square by translating back all of these outside points back into the square. This becomes the image of points on the torus back into the torus, as seen in Figure 3 on the right.

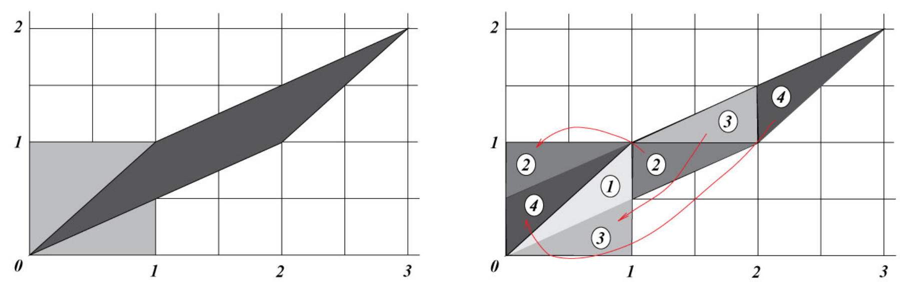


FIGURE 3. The map F_L, for

Notice in the figure that

$$L\left(\begin{bmatrix}1\\1\end{bmatrix}\right)=\begin{bmatrix}3\\2\end{bmatrix},\quad L\left(\begin{bmatrix}1\\0\end{bmatrix}\right)=\begin{bmatrix}2\\1\end{bmatrix},\quad \text{and}\quad L\left(\begin{bmatrix}0\\1\end{bmatrix}\right)=\begin{bmatrix}1\\1\end{bmatrix}.$$

EXERCISE 160. Draw the torus in its representation in $\mathbb{R}^3$ (as the surface of a doughnut) with its two canonical loops that correspond to the edges of the unit square in $\mathbb{R}^2$, viewed as a fundamental domain. Then carefully draw the images of these two curves under the following hyperbolic linear toral maps given by the following matrices. You may want to draw the images of the edges of the fundamental domain in $\mathbb{R}^2$ first.

(a) $\begin{bmatrix}2 & 1\\1 & 1\end{bmatrix}$.

(b) $\begin{bmatrix}3 & 2\\1 & 1\end{bmatrix}$.

QUESTION 6.19. *How are any periodic points distributed?*

we have the following proposition:

PROPOSITION 6.20. *The set of all periodic points of* $F_L : \mathbb{T}^2 \to \mathbb{T}^2$ *is dense in* $\mathbb{T}^2$ *and* $P_n(F_L) = \lambda_1^n + \lambda_1^{-n} - 2$.

PROOF. The first claim we will make to prove this result is the following: *Every rational point in* $\mathbb{T}^2$ *is periodic.* To see this, note that every rational point in $\mathbb{T}^2$ is a point in the unit square with coordinates $x = \frac{s}{q}$, and $y = \frac{t}{q}$, for some $q, s, t \in \mathbb{Z}$. For every point like this, $F_L(x,y)$ is also rational with the same denominator (neglecting fraction reduction, do you see why?) But there are only q^2 distinct points in $\mathbb{T}^2$ which are rational and which have q as the common denominator. Hence, at some point, $\mathcal{O}_{(x,y)}$ will start repeating itself. Hence this claim is proved. Now notice that the set of all rational points in $\mathbb{T}^2$ is dense in $\mathbb{T}^2$, or

$$\overline{\mathbb{Q} \cap [0,1]} \times \overline{\mathbb{Q} \cap [0,1]} = [0,1]^2.$$

Hence the periodic points are dense in $\mathbb{T}^2$.

The next claim is: *Only rational point are periodic.* To see this, assume $F_L\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x \\ y \end{bmatrix}$. Then

$$F_L^n\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} \mod 1,$$

and this forces the system of equations

$$\begin{array}{c} ax + by = x + k \\ cx + dy = y + \ell \end{array}, \text{ for } k, l \in \mathbb{Z}.$$

Simply solve this system for x and y and you get that $x, y \in \mathbb{Q}$.

EXERCISE 161. Solve this system for x and y.

The number of periodic points can be found by creating a new linear map. Define

$$G_n\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = F_L^n\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) - \begin{bmatrix} x \\ y \end{bmatrix} = (F_L^n - I_2)\begin{bmatrix} x \\ y \end{bmatrix}.$$

The n-periodic points are precisely the kernel of this linear map:

$$Per_n(F_L) = \ker(G_n) = \left\{\begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 \,\middle|\, G_n\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}\right\}.$$

We can easily count these now to calculate $P_n(F_L)$. They are precisely the pre-images of integer vectors!

CLAIM. All pre-images of $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ under the map $G_n = F_L^n - I_2$ are given by $\mathbb{Z}^2 \cap (L^n - I_2)([0,1) \times [0,1))$.

- Since F_L is to be understood as simply the matrix L where images are taken modulo 1, the map G_n is simply the map $L^n - I_2$ where images are taken modulo 1. Hence we can study the effect of G_n by looking at the image of $L^n - I_2$.

- To avoid over-counting points, we modify our unit square, eliminating twice-counted points (on the edges) and quadruply-counted points (the corners). Consider the "half-open box" $[0,1)^2$ as our model of $\mathbb{T}^2$. In this model, every point lives in its own equivalence class.

We try a few early iterates:

EXAMPLE 6.21. $G_1 = F_L - I_2$. Then the corresponding map on $\mathbb{R}^2$ is

$$L - I_2 = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}.$$

This map is a shear on $[0,1)^2$ and we see that the only integer vector in the image is the origin: The red dot in Figure 4. Thus

$$P_1(F_L) = \lambda^1 + \lambda^{-1} - 2 = \frac{3+\sqrt{5}}{2} + \frac{3-\sqrt{5}}{2} - 2 = 3 - 2 = 1.$$

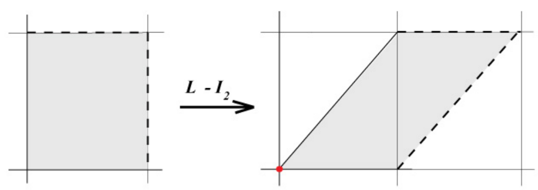


FIGURE 4. G_1-action on $\mathbb{T}^2$ as seen through the map $L - I_2$ on $[0,1)^2$.

EXAMPLE 6.22. $G_2 = L^2 - I_2$. Here

$$G_2 = L^2 - I_2 = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 3 \\ 3 & 1 \end{bmatrix}.$$

This map is a little more complicated, and we see in Figure 5 that there are a few more integer vectors in the image, namely the points $(2,1)$, $(3,2)$, $(4,2)$, and $(5,3)$. And since

$$P_2(F_L) = \lambda^2 + \lambda^{-2} - 2 = 7 - 2 = 5,$$

we see that the formula continues to hold.

EXERCISE 162. What were the original points in $[0,1)^2$ that correspond to these 5 integer vectors under G_2?

EXERCISE 163. Draw the image of $[0,1)^2$ in $\mathbb{R}^2$ under the linear map corresponding to G_3 for F_L above. Calculate $P_3(F_L)$ via the formula and verify by marking the integer points in $(L^3 - I_2)\big([0,1)^2\big)$. Choose two non-zero integer vectors in the image and identify the original 3-periodic points in $\mathbb{T}^2$ that correspond to them.

This proof ends by appealing to a strikingly beautiful result by Georg Alexander Pick in 1899, which we now call Pick's Theorem:

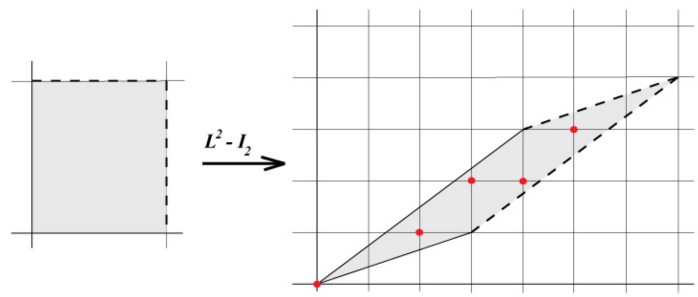


FIGURE 5. G_2-action on $\mathbb{T}^2$ as seen through the map $L^2 - I_2$ on $[0,1)^2$.

THEOREM 6.23 (Pick's Theorem). *Let $A \in \mathbb{R}^2$ be a polygon in the plane whose vertices are integer points (points with integer coordinates). Then for I the number of integer-points in the interior of A and B the number of integer-points on A, we have*

$$area(A) = I + \frac{1}{2}B - 1.$$

Here, the polygonal integer points are edge points and vertex points, so that $\frac{1}{2}B - 1 = \frac{B-2}{2}$. For a parallelogram, as in our case, this amounts to collecting up all integer-points on the polygon, and counting each as a half and also counting all four vertices as 1. Sort of like double counting edge points and quadruple counting vertex points. The unit square is the fundamental domain of $\mathbb{T}^2$. Hence opposite edges are identified and all four corners are the same point in the quotient. Hence Pick's Formula, applied to our square $[0,1)^2$ with its only two edges and o vertex, reduces to a simple counting of interior points and the remaining edge points (See Figure 6. In sum: The area of $G_n\left([0,1)^2\right)$ is precisely equal the number of integer-vectors in the image. And the latter is given by

$$|\det(G_n)| = |\det L^n - I_2| = |(\lambda^n - 1)(\lambda^{-n} - 1)| = \lambda^n + \lambda^{-n} - 2,$$

where λ is the largest eigenvalue (in magnitude) of L. $\square$

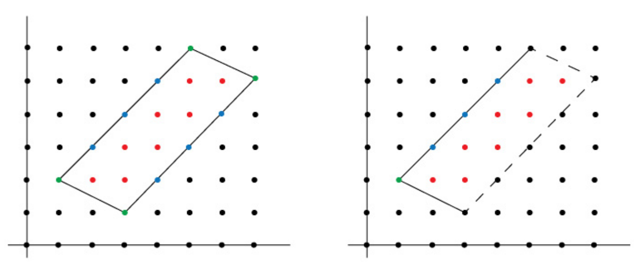


FIGURE 6. Pick's Formula for parallelograms.

Note: G_2 on $\mathbb{R}^2$ is NOT area preserving! In fact,

$$\det(G_2) = \begin{vmatrix} 4 & 3 \\ 3 & 1 \end{vmatrix} = 5.$$

Note that the map F_L above was area-preserving on the torus. It is also invertible (any determinant one matrix with integer coefficients is invertible, and the inverse is also of determinant one with integer entries!) However, area-preserving does NOT ensure invertibility of the map. The prime example is the circle map $E_m : S^1 \to S^1$, where $E_m(z) = z^n$. The map is area-preserving, if we sum all of the lengths of the disjoint pre-images of small sets. But it is also of degree m. And if $|m| > 1$, the map is m to 1. Invertibility is a very desirable quality for a map, as it allows us to work both forwards and backwards in constructing orbits. Fortunately, there are ways to study non-invertible maps by encoding their information in a (different) invertible dynamical system. We will introduce this concept here, but not spend a lot of time on it for now:

6.1.4. Inverse Limit Spaces. One issue with many-to-one maps like E_2 above is precisely that the map is not invertible: Each point in the range of the function has two distinct pre-images. For example,

$$E_2^{-1}\left(\frac{1}{4}\right) = \left\{\frac{1}{8}, \frac{5}{8}\right\}.$$

That the map is volume preserving is well-hidden by this property as we have seen. One can account for the non-injectivity via a "choice" of pre-image for a particular point, but one cannot move backward along an orbit in general. There is a way to account for all orbits pre-origins, however, which we do now.

For X a metric space with $f : X \to X$ a surjective map, let

$$(X, f) = \left\{ \underline{\mathrm{x}} = (\dots, x_i, \dots, x_{-2}, x_{-1}, x_0) \;\middle|\; x_i \in X, i \in -\mathbb{N}, f(x_i) = x_{i+1} \right\}$$

be the set of all sequences of points in X that serve as orbits leading up to x_0, for all choices of $x_0 \in X$. The set (X, f) may seem a bit unwieldy and complicated, but it has some nice properties. For example, one can use the topology of X to endow this set with its own topology. Indeed, (X, f) is a subset the bigger set of all infinite sequences of X, called sometimes X^∞, a space given the (infinite) product topology. By a famous theorem of Tychanoff, if X is compact, then so is (X, f). And, as a topological space, (X, f) also can be made a metric space with the metric

$$\underline{\mathrm{d}}(\underline{\mathrm{x}}, \underline{\mathrm{y}}) = \sum_{i \le 0} 2^i d x_i - y_i),$$

where d is the metric on X.

Exercise 164. Show that this is a metric.

Now, define the map $T_f : (X, f) \to (X, f)$ by

$$T_f (\dots, x_{-3}, x_{-2}, x_{-1}, x_0) = (\dots, x_{-2}, x_{-1}, x_0, f(x_0)).$$

It turns out that this map is a homeomorphism: It is continuous in the product topology (we will not show this), as is its inverse (drop off the final term), and we have the exercise:

Exercise 165. Show T_f is one-to-one and onto.

In fact,

$$T_f(\underline{\mathrm{x}}) = \underline{\mathrm{f(x)}}.$$

Here, (X, f) is called the *inverse limit space* of X, given f. It catalogs every orbit that leads to x_0. Now create the space of bi-infinite sequences (or a two-sided sequence space)

$$X' = \left\{\left(\dots, x_{-3}, x_{-2}, x_{-1}, x_0, \mathcal{O}^+_{x_0}\right) \,\middle|\, x_i \in X, i \in \mathbb{Z}, f(x_i) = x_{i+1}\right\}$$

and extend T_f to X'. Then T_f is just the left shift map on X'. It remains a homeomorphism, X' is also compact when X is, and every point in X' constitutes an entire $\mathbb{Z}$-orbit. The inverse, T_f^{-1} is the right shift.

EXAMPLE 6.24. For $E_2(x) = 2x \mod 1$ on S^1, some of these sequences which correspond to $x_0 = 1$ look like

$$\begin{gathered}
0^{\text{th}} \text{ place} \\
\downarrow \\
\left\{\dots, \frac{1}{8}, \frac{1}{4}, \frac{1}{2}, \ \mathbf{1}\ , 1, 1, \dots\right\} \\
\left\{\dots, \frac{1}{4}, \frac{1}{2}, 1, \ \mathbf{1}\ , 1, 1, \dots\right\} \\
\left\{\dots, \frac{3}{8}, \frac{3}{4}, \frac{1}{2}, \ \mathbf{1}\ , 1, 1, \dots\right\} \\
\left\{\dots, \frac{7}{8}, \frac{3}{4}, \frac{1}{2}, \ \mathbf{1}\ , 1, 1, \dots\right\}
\end{gathered}$$

DEFINITION 6.25. For X a metric space and $f : X \to X$ continuous, the *inverse limit* is defined on the space of sequences

$$X' = \left\{ \{x_n\}_{n\in\mathbb{Z}} \,\middle|\, x_n \in X,\ f(x_n) = x_{n+1},\ \forall n \in \mathbb{Z} \right\}$$

by $F\left(\{x_n\}_{n\in\mathbb{Z}}\right) = \{x_{n+1}\}_{n\in\mathbb{Z}}$.

This is a new dynamical system defined by the map F on the *inverse limit space* X'. Note that since this map takes entire sequences to sequences, it is 1-1, and hence we can go backwards. On sequences, this map is invertible, since the entire history of a point is already in the "point" (read: sequence).

EXAMPLE 6.26. Back to the map E_2 on S^1, the limit space is

$$\mathbb{S} = \left\{ \{x_n\}_{n\in\mathbb{Z}} \,\middle|\, x_n \in S^1,\ E_2(x_n) = x_{n+1},\ \forall n \in \mathbb{Z} \right\}$$

with the map $F\left(\{x_n\}_{n\in\mathbb{Z}}\right) = \{2x_n \mod 1\}_{n\in\mathbb{Z}}$. The space $\mathbb{S}$ is called a solenoid, and a picture of it site on the cover of the book.

6.2. Symbolic Dynamics

Thinking along the lines of the inverse limit spaces we saw recently, we can adapt that construction to create a new type of dynamical system that serves as a beautiful and important model for many of the concepts we will see soon. To start, let $\mathcal{M} = \{0, 1, 2, \dots, n-1\}$ be a finite, discrete set of n symbols (we use numbers here, but there is no natural reason why.) We can topologize $\mathcal{M}$ with the discrete topology (every subset is considered “open”) so that $\mathcal{M}$ is a topological space. It is even a metric space, as one can define $d_{\mathcal{M}}(x,y) = 1 - \delta_{ij}$ on $\mathcal{M}$ as a metric where every element is distance-1 from every other element.

Exercise 166. Show that this is a metric.

Now construct two other sets,

$$\mathcal{M}^{\mathbb{N}} = \left\{ \mathbf{x} = \{x_i\} \;\middle|\; i \in \mathbb{N}, x_i \in \mathcal{M} \right\},$$
$$\mathcal{M}^{\mathbb{Z}} = \left\{ \mathbf{x} = \{x_i\} \;\middle|\; i \in \mathbb{Z}, x_i \in \mathcal{M} \right\},$$

respectively the set of one-sided and two-sided infinite sequences of elements in $\mathcal{M}$. It is easy to see that these sets can also be made into spaces via the product topology. In this topology, one can construct open sets via a metric. Let

$$d(\mathbf{x},\mathbf{y}) = \sum_{i \in \mathbb{K}} \frac{x_i - y_i}{2^{|i|}}.$$

Then the $\frac{1}{2^n}$-neighborhoods of a particular sequence are the precisely the set of all sequences that agree with the particular sequence at every position from 0 to n in the case of $\mathcal{M}^{\mathbb{N}}$. And in the case of $\mathcal{M}^{\mathbb{Z}}$, that agree on the positions from $-(n+1)$ to $n+1$. These are called open cylinders in the product topology.

Definition 6.27. Let $(X, \mathcal{P})$ be a partitioned space. We say a map $f : X \to X$ respects $\mathcal{P}$, if $\forall i \in \{1, 2, \ldots, n\}$,

$$f(\mathcal{P}_i) = \bigcup_{j=1}^{n} \delta_{ij} \mathcal{P}_j, \text{ where } \delta_{ij} = \begin{cases} 1 & \exists x \in \mathring{P}_i \text{ such that } f(x) \in \mathring{P}_j \\ 0 & \text{otherwise.} \end{cases}$$

Here, when $(X, \mathcal{P})$ is a partitioned space and f respects the partition, we say that f has the *Markov condition* and will call $\mathcal{P}$ a *Markov partition* for f on X.

Definition 6.28. For $\mathcal{P}$ a Markov partition for f on X, the *transition matrix* is an $n \times n$ matrix A where $a_{ij} = \delta_{ij}$.

Markov partitions are way of encoding information about orbits without tracking the precise points in the orbit. In essence, one divides a domain into a finite number of pieces, and then records only the piece an orbit visits at each natural number or integer. Now this only works when the map takes each partition element onto a union of other partition elements; partition elements must map to partition elements. But this course record of orbit behavior can illustrate a lot of dynamical information, as we will see.

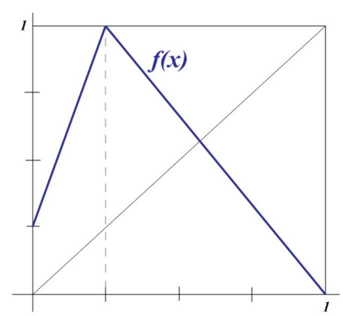


Figure 7. The graph of $f(x)$.

Example 6.29. Let $I = [0,1]$ and $f : I \to I$ be defined as

$$f(x) = \begin{cases} 3x + \frac{1}{4} & 0 \le x \le \frac{1}{4} \\ \\ \frac{4}{3} - \frac{4}{3}x & \frac{1}{4} \le x \le 1. \end{cases}$$

See Figure 7. Here we create $\mathcal{P} = \{\mathcal{P}_1, \mathcal{P}_2\}$, where $\mathcal{P}_1 = \left[0, \frac{1}{4}\right]$ and $\mathcal{P}_2 = \left[\frac{1}{4}, 1\right]$. Then $\mathcal{P}$ is a Markov partition for f on I, since $f(\mathcal{P}_1) = \mathcal{P}_2$ and $f(cP_2) = \mathcal{P}_1 \bigcup \mathcal{P}_2$. The transition matrix is then $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$. And f on I is a subshift of finite type. One can say that, symbolically, f corresponds to the one-sided left shift map on the sequence space $\mathcal{M}^{\mathbb{N}}$, where $\mathcal{M} = \{0, 1\}$,where the invariant subset of allowable sequences are the ones with no consecutive 0's in them.

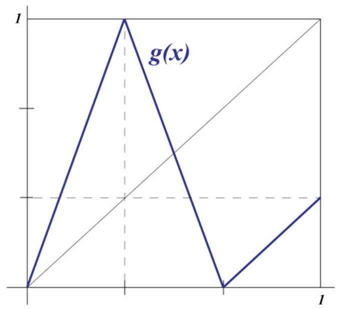


FIGURE 8. The graph of $g(x)$.

The number of elements in the partition dictates the "size" of the sequence space, and accordingly the size of the transition matrix. That the map f above is not injective will play a crucial role later, as we will see. For now, contrast this with another example:

EXAMPLE 6.30. Let $g : I \to I$ be given as in Figure 8. Then a Markov partition for g is $\mathcal{P} = \{\mathcal{P}_1, \mathcal{P}_2, \mathcal{P}_3\}$, where $\mathcal{P}_1 = \left[0, \frac{1}{3}\right]$, $\mathcal{P}_2 = \left[\frac{1}{3}, \frac{2}{3}\right]$ and $\mathcal{P}_3 = \left[\frac{2}{3}, 1\right]$. The transition matrix here is $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix}$. Here, g corresponds to the left shift map on the subshift of finite type on $\{0, 1, 2\}^{\mathbb{N}}$ characterized by the forbidden blocks $\mathcal{F} = \{23, 33\}$.

EXERCISE 167. Create a piecewise linear C^0 map on I where $\mathcal{F} = \{11, 22, 23\}$, and write out the transition matrix A.

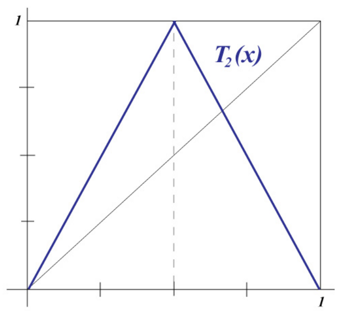


FIGURE 9. The graph of the full tent map $T_2(x)$.

When each element of a partition maps to the union of all of the element of the partition, there are no forbidden blocks. Here, the subshift is a full shift on $\mathcal{M}$.

EXAMPLE 6.31. Consider the map on I at left, given by $T_2(x) = \begin{cases} 2x & 0 \le x \le \frac{1}{2} \\ 2(1-x) & \frac{1}{2} \le x \le 1 \end{cases}$. The graph of $T_2(x)$, the notation will become apparent soon) is often called the *full tent map* on the unit interval I, as seen in Figure 9. Given the obvious partition, T_2 corresponds to the full left shift on the sequence space $\{0, 1\}^{\mathbb{N}}$. As both partition intervals are mapped onto I, there are no forbidden blocks here. A slight alteration of notation reveals another interesting fact. Write each sequence as a binary expansion:

$$\mathbf{x} = \{x_i\}_{i \in \mathbb{N}} = (.x_1 x_2 x_3 \ldots)_2 ,$$

where the decimal simply indicates a starting point for the sequence. Then the shift map is just the map $2x \mod 1$ on these binary numbers, since

$$\begin{aligned} f(\mathbf{x}) = f(\{x_i\}) &= 2 \cdot (.x_1 x_2 x_3 \ldots)_2 \mod 1 \\ &= (x_1.x_2 x_3 x_4 \ldots)_2 \mod 1 = (.x_2 x_3 x_4 \ldots)_2 = \{x_{i+1}\} . \end{aligned}$$

6.3. Chaos and Mixing

Recall a map $f : X \to X$ on a metric space is *topologically transitive* if there exists a dense orbit. Some examples that we looked at included the irrational rotations of S^1 and the irrational linear flows on the two torus $\mathbb{T}^2$. Note that these examples had no periodic points at all, and all orbits were dense. Contrast that with the idea some dynamical systems seemed to be full of periodic points. Think of rational rotations of S^1 and rational linear flows on $\mathbb{T}^2$. Again, on these examples, all points were periodic, and none of these maps are topologically transitive.

These properties seem to be mutually exclusive, and are when the dynamics are relatively simple to describe. However, for dynamical systems which possess both a dense supply of periodic orbits as well as a dense orbit, the dynamics can be labeled quite complex. How complex?

Definition 6.32. A continuous map $f : X \to X$ of a metric space is said to be *chaotic* if

- f is topologically transitive,
- $\overline{Per(f)} = X$.

Notes:

- There are many definitions of chaos floating around in this area, as efforts to finally pin down the concept continue. This definition really is one of the better universal definitions we have for the concept. That said, there is still a slight problem even with this definition. For details now, see Theorem 6.52 and and adjacent Example 6.53.
- Either one of these properties without the other means that the dynamics are relatively simple to describe.

Some examples that we were recently playing with:

(1) Let $E_m : S^1 \to S^1$ be the linear expanding map of S^1, for $|m| > 1$.
(2) Let $f_\lambda : C \to C$ be the logistic map for $\lambda > 4$, restricted to the Cantor set of point whose orbit lies completely within the unit interval.
(3) Let $F_L : \mathbb{T}^2 \to \mathbb{T}^2$ be the linear hyperbolic automorphism of the two-torus given by the linear automorphism of the plane determined by the hyperbolic matrix L.

In the first and third cases, we showed that the periodic points are dense in the respective spaces. Hence the dynamical systems are chaotic if we can show there actually exists a dense orbit. The same holds for the Cantor map, although we did not actually show that the periodic points are dense. However, showing directly that there exists a dense orbit is not easy. We will instead construct a bit more machinery, and show that these maps possess some stronger properties that transitivity. In this way, we can study the maps in more detail, and gain some additional insight into the structure of these dynamical systems. To start:

Proposition 6.33. *Let X be a complete separable metric space with no isolated points. For $f : X \to X$ continuous, the following are equivalent:*

(1) f has a dense orbit and is topologically transitive,
(2) f has a dense positive semiorbit,
(3) if $U, V \subset X$ are open and nonempty, $\exists N \in \mathbb{Z}$ such that $f^N(U) \cap V \neq \varnothing$,
(4) if $U, V \subset X$ are open and nonempty, $\exists N \in \mathbb{N}$ such that $f^N(U) \cap V \neq \varnothing$.

Remark 6.34. Recall that a metric space is complete if all Cauchy sequences converge. And X is separable if there exists a countable dense subset. These properties, along with the "no isolated points" condition, are technical in nature and while necessary, should not keep you from well understanding how this proposition works on the nice spaces we are used to. So don't worry too much at this point about these technicalities.

Proof. Obviously $4 \Rightarrow 3$ and $2 \Rightarrow 1$. If we can show that $3 \Rightarrow 2$ and $1 \Rightarrow 4$, we would be done. We will not do this, however. The real point of this exposition is to understand the relationship between 1 and 3. To this end, we will prove the statement $1 \Rightarrow 4$.

Let f be topologically transitive, with a dense orbit given by $\mathcal{O}_x$, $x \in X$. Then for any choice of nonempty, open sets $U, V \subset X$, $\exists n, m \in \mathbb{Z}$ such that $f^n(x) \in U$, and $f^m(x) \in V$. If we suppose for a minute that $m > n$, then we would get that $f^{m-n}(U) \cap V \neq \varnothing$, with $N = m - n > 0$. In the case that f is invertible, this makes sense, since $f^{-n}(U)$ would be a neighborhood of x, so that $f^m(x) \in f^m(f^{-n}(U))$. With $f^m(x) \in V$, the result follows. However, this works even in the case where f is not invertible. Simply think of $f^{-n}(U)$ as being the inverse (set theoretic) image of U (the set of all things that go to U under f^n). See the picture.

If $m - n < 0$, then we can either increase the value of m if f has a dense positive semiorbit, or decrease n if f has a dense negative semiorbit (it must have at least one), via a construction similar to that of Corollary 5.7 of Theorem 5.6. Indeed, suppose f has a dense positive semiorbit. Then, since X has no isolated points, $y = f^m(x) \in U$ is not isolated, and there must be a $m' > m$ where $f^{m'}(x) \in U$ (close enough to y to be in U.) Indeed, there will be a sequence $m_k \to \infty$, where $f^{m_k}(x) \in U$ and $f^{m_k}(x) \to y$. Choose k such that $m_k > n$. Then let $N = m_k - n$. Decreasing the value of n is similar. $\square$

Corollary 6.35. *A continuous, open map f of a complete metric space is topologically transitive iff there does not exist two disjoint, open f-invariant sets.*

It will help in understanding this last statement to understand the notion of an open map. Roughly speaking, a map is open if it takes open sets to open sets, something that is not generally true for continuous maps (think of a constant map). We will expound on this and topology in general shortly. However, the idea in the previous corollary and discussion is that finding a dense orbit is equivalent to the notion that the orbit of ANY open set in X must eventually intersect any other open set in X actually provides a method of discovery for dense orbits. A set $V \subset X$ is f-invariant, if $f(V) \subset V$. Now assume that you have such a set V which is open. Now take any other open set U. Whether it is invariant or not, its entire orbit $\mathcal{O}_U$ is a union of all of its images and is hence open in X if the map f is open. The Corollary says that an open map is topologically transitive iff we cannot divide the space into two disjoint open sets which are each invariant under f. Put this way, the two notions look very much alike.

To better get some of these ideas, lets go over a bit of topology:

Definition 6.36. A *topology* on a set X is a well-defined notion what constitutes an open subset of X.

What well-defined means is: A topology on X is a collection $\mathcal{T}_X$ of subsets of X that satisfy

- $\varnothing$ and X are in $\mathcal{T}_X$,
- the union of the elements of any subcollection of $\mathcal{T}_X$ is in $\mathcal{T}_X$, and
- the intersection of the elements in any finite subcollection of $\mathcal{T}_X$ is in $\mathcal{T}_X$.

For a topology $\mathcal{T}_X$ on X, the elements of $\mathcal{T}_X$ are called *open*. Also, any set that is given a topology, is called a *topological space*.

EXAMPLE 6.37. The set of all open intervals $(a,b) \subset \mathbb{R}$ constitutes a topology on $\mathbb{R}$, called the *standard* topology $\mathcal{T}_{\mathbb{R}}$, for $-\infty \leq a \leq b \leq \infty$. It should be obvious that this allows all of $\mathbb{R}$ to be in $\mathcal{T}_{\mathbb{R}}$, and if we let $a = b$, then the element $(b,b) = \varnothing$ is also in $\mathcal{T}_{\mathbb{R}}$. The union of any collection of open intervals is certainly open also. Now, without the last condition, however, we would have a problem: Suppose we allowed that the intersection of any subcollection of $\mathcal{T}_{\mathbb{R}}$ to be in $\mathcal{T}_{\mathbb{R}}$. Then the set

$$\bigcap_{n=1}^{\infty} \left(-\frac{1}{n}, \frac{1}{n}\right) = \{0\}$$

would have to be open. But then all individual points would also be open, and thus by the middle constraint, any subset of X would be open! You can see why the third provision is necessary. Incidentally, there is a topology on $\mathbb{R}$ (or any other set), where each of the points is considered open. It is called the *trivial* topology on the set, and although it works via the definitions, it does not describe well the actual set as a space.

Some facts:

DEFINITION 6.38. For $f : X \to Y$ a (not necessarily continuous) map between two topological spaces, f is *continuous* if whenever $V \subset Y$ is open (an element of its topology $\mathcal{T}_Y$), then $f^{-1}(V) \subset X$ is open (an element of $\mathcal{T}_X$).

This allows us to talk about maps being continuous between arbitrary topological spaces, in a way that is entirely compatible with what you already learned as the definition of continuity between spaces like subsets of $\mathbb{R}$ in Calculus I, or subsets of $\mathbb{R}^n$ in Calculus III. Then, we simply assumed the standard topologies on Euclidean space, and the notions of "nearness" which is at the center of continuity comes out of the little ϵ-balls used to define continuity.

We can alter this definition to better fit the notion you are already familiar with:

DEFINITION 6.39. A function $f : X \to Y$ is continuous at some point $x \in X$ if and only if for any neighborhood V of $f(x)$, there is a neighborhood U of x such that $f(U) \in V$.

In topology, any open set in X (a member of $\mathcal{T}_X$) is called a neighborhood of any of its points. Hence again, this definition depends on the topologies of X and Y. It basically says that no matter how "small" we choose an open set V containing $f(x)$, we can always find an open set U, containing x, where $f(U)$ sits entirely inside V.

Now, for a minute, choose $Y = X$, a metric space (so that we can talk about ϵ-balls around points given a metric). Then we have

DEFINITION 6.40. For x a metric space, a function $f : X \to X$ is continuous at a point $x = a \in X$, if given $\epsilon > 0$, there is a $\delta > 0$ such that when $|x - a| < \delta$, (really, when $x \in B_\delta(a)$), then $|f(x) - f(a)| < \epsilon$ (really, $f(x) \in B_\epsilon(f(a))$.)

Now think about this last definition and the definition you saw in Calculus (either Calculus I or Calculus III). In Calculus I, the standard ϵ-δ definition of a limit, altered for continuity by switching the L for $f(a)$, is exactly the last definition above, as the graph at right shows.

Back to the task at hand, we go one step farther with another definition:

DEFINITION 6.41. $f : X \to Y$ is called an *open map*, if it is continuous and if whenever $U \subset X$ is open, then $f(U) \subset Y$ is open also.

While continuity is common among maps, "openness" is not, and is kind of a special property. When the map f has a continuous inverse, then f is open. But this is not that common a property.

EXERCISE 168. Prove Corollary 6.35 in detail.

Now, we know expanding maps of S^1 and hyperbolic automorphisms of $\mathbb{T}^2$ look messy dynamically. The question is: How messy are they?

DEFINITION 6.42. A continuous map $f : X \to X$ is said to be *topologically mixing* if, for any two nonempty, open set $U, V \subset X$, $\exists N \in \mathbb{N}$, such that $f^n(U) \cap V \neq \varnothing$, $\forall n > N$.

Notes:

- Do you see how much stronger (more restricting) this is to topological transitivity? For instance, (topologically mixing)$\Rightarrow$(topologically transitive), but not vice-versa. To see why, think of the irrational rotations of the circle. The orbit of a small open interval will eventually intersect any other small open interval. But, depending on the rotation, will most likely leave again for a while before returning. This is not mixing!

  EXERCISE 169. Show that topological mixing implies topological transitivity.

- Actually, the problem with irrational circle rotations is a bit deeper; they are isometries:

  LEMMA 6.43. *Isometries are not topologically mixing.*

  PROOF. Under an isometry, the diameter of a set $U \subset X$, $diam(U)$ is preserved. Let $U = B_\delta(x) \subset X$ be a small δ-ball about a point $x \in X$. Here $diam(U) = 2\delta$ and $\forall n \in \mathbb{N}$, $diam(f^n(U)) = 2\delta$. Now choose $v_1, v_2 \subset X$, such that the distance between v_1 and v_2 is greater than 4δ. Let $V_1 = B_\delta(v_1)$ and $V_2 = B_\delta(v_2)$ (so that the minimal distance between these two balls is greater than 2δ). If we assume that the isometry $f : X \to X$ is top. mixing, then there will be a $k \in \mathbb{N}$, such that both $f^n(U) \cap V_1 \neq \varnothing$, and $f^n(U) \cap V_2 \neq \varnothing$. $\forall n > k$. But this is impossible since V_1 and V_2 are too far apart to both have nonempty intersection with an iterate of U. Hence f cannot be mixing. $\square$

PROPOSITION 6.44. *Expanding maps on S^1 are topologically mixing.*

PROOF. for now, suppose that the expanding map is C^1. Differentiable expanding maps have the property that for $f : S^1 \to S^1$, $|f'(x)| \geq \lambda > 1$, $\forall x \in S^1$. Let $F : \mathbb{R} \to \mathbb{R}$ be a lift. it is an exercise to show that the lift also shares the derivative property, $|F'(x)| \geq \lambda$, $\forall x \in \mathbb{R}$. So choose a small closed interval $[a,b] \subset \mathbb{R}$, where $b > a$. Then, by the Mean Value Theorem, $\exists c \in (a,b)$, such that

$$|F(b) - F(a)| = |F'(c)||b-a| \geq \lambda(b-a).$$

Hence, the length of the iterate of the interval is greater by a factor of λ than the interval. This continues at each iterate of F, so that $\exists n \in \mathbb{N}$, such that $\|F^n([a,b])\| > 1$. But then $\pi(F^n([a,b])) = S^1$.

Now simply grab the open interval (a,b), noting that $\pi((a,b))$ will also be open (on small intervals, π is a homeomorphism), and let $U = \pi((a,b))$. With V be any other open set in S^1, we are done. $\square$

Corollary 6.45. *Linear expanding maps of S^1 are chaotic.*

Exercise 170. Without using Proposition 6.44 and topological mixing, show expanding maps are chaotic.

Somewhere around here, place the Smale Horseshoe as a means to describe a planar chaotic set.

Proposition 6.46. *$F_L : \mathbb{T}^2 \to \mathbb{T}^2$, the linear hyperbolic automorphism of the two torus given by the hyperbolic matrix L is topologically mixing.*

Corollary 6.47. *F_L is chaotic.*

For a brief idea why the previous proposition is true, recall for F_L given by the matrix $L = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$, the eigenvalues were $\lambda = \frac{3 \pm \sqrt{5}}{2}$, and the eigenvalue greater than 1 (the "expending" eigenvalue) has eigendirection given by the vector $\begin{bmatrix} 1 \\ \frac{-1+\sqrt{5}}{2} \end{bmatrix}$. Choose a small open line segment T along the line $y = \left(\frac{-1+\sqrt{5}}{2}\right)x + c$ within the box representing the torus. As we iterate the map, the orientation of the line stays the same, while the length of the line grows by a factor of $\lambda = \frac{3+\sqrt{5}}{2}$ at each iterate. For $N >> 1$, we would find that the length of $F_L(T)$ will be huge, and wrap around the torus quite densely. In fact, we can choose this N so that $F_L^N(T)$ will intersect ANY ball of radius ϵ in $\mathbb{T}^2$. Hence choose any ϵ-ball V and any other ϵ-ball U, and take as our T the diameter of U in the direction of the line $y = \left(-1 + \sqrt{5}2\right)x + c$. Then after the above chosen N, we would have $F_L^n(U) \cap V \neq \varnothing$, for all $n > N$. Hence F_L is topologically mixing on $\mathbb{T}^2$.

And finally, we will not prove this explicitly, but we have the following:

Proposition 6.48. *The logistic map $f_\lambda : C \to C$, where $\lambda > 2 + \sqrt{5} > 4$, and where $C \in [0,1]$ is the Canter set of points whose entire orbits stay within $[0,1]$ is expanding.*

Proposition 6.49. *f_λ as above, is topologically mixing.*

Corollary 6.50. *f_λ as above, is chaotic on C.*

6.4. Sensitive Dependence on Initial Conditions

So the next questions is: What information does chaos, as a property, convey about the dynamical system? Flippantly speaking, it tells us that the orbit structure is quite complicated. It tells us that arbitrarily close to a periodic point, are non-periodic points whose orbits are dense in the space. On the other hand, it tells us that arbitrarily close to a point whose orbit is dense in the space, are periodic points of arbitrarily high period. Hence simply being very close to a point of a certain type does not mean that the orbits will be similar. This means that one cannot rely on estimates or precision to help determine orbit behavior. Mathematically, it means the following:

DEFINITION 6.51. A map $f : X \to X$ of a metric space is said to exhibit a *sensitive dependence on initial conditions* if $\exists \Delta > 0$ (called a sensitivity constant), where $\forall x \in X$ and $\forall \epsilon > 0$, $\exists$ a point $y \in X$ where $d(x,y) < \epsilon$ and $dA(f^N(x), f^N(y)) > \Delta$ for some $N \in \mathbb{N}$.

There are lots of notes to say on this topic:

- The idea here is, for certain constants, no matter how small a neighborhood of a chosen point x you start, there will always be a point y in this neighborhood that after a time, its neighborhood will be far away from the orbit of x.
- The existence of at least one point in each neighborhood of an arbitrary point x whose orbit veers away from the orbit of x is the notion that everywhere there is an expanding direction (think of a differentiable map whose derivative everywhere, as a matrix, has at least one eigenvalue of modulus greater than 1). This is like the hyperbolic action on the torus.
- This idea was quite profound: Early developers of classical mechanics tended to believe that eventually we would understand the universe completely. Given the universes state in an instant, we should be able to predict its state at any future moment. This was the thinking around the early 1800's of people like Laplace.
- Poincare, in the late 1800's, saw this phenomenon of a sensitive dependence on initial conditions in the classical three-body problem. He understood immediately that the earlier reasoning was flawed. Indeed, knowing the precise state of all things in the universe was impossible. And with the presence of a sensitive dependence on initial conditions (even in the simplistic three-body problem), a reasonable approximation to the universe's state in an instant would never be good enough to make good long term predictions.
- Edward Lorentz, studying early climate models on a computers like at MIT in the 1960's, saw his deterministic (though nonlinear!) computer model make wildly divergent predictions given the exact same input values in redundant runs of his program. Puzzled as to why this was the case, it became clear that the model was fine. It was the assumption that any number is known to infinite precision in a computer. For example, zero is not zero on a computer. Setting a variable to 0 on a computer makes the number 0 only to within a certain precision (it stores the number in a certain number of bytes). For example, 0 in single precision, is only 0 down to 10^-7. If in the model this number is multiplied by a very large number, any variance from true 0 would be multiplied into the realm where it will change the calculations. Do this same run twice and you would get two different values for the result. This small variance is like trying to grab a point like x above and instead getting a nearby number like y instead. In the calculations, the resulting orbits would veer away from each other,

and the results would be different. Eventually, this was the discovery that Lorentz had made. Incidentally, the Lorentz Butterfly is an example of what is called a "strange attractor", and came out of the puzzle Lorentz created.

- Isometries cannot exhibit a sensitive dependence on initial conditions. Why not?
- For $f(x) = 2x \mod 1$, distance between nearby points grow exponentially by 2^n. This is a sensitive dependence on initial conditions. Eventually, this distance is larger than 1, and at this point, future iterates of each orbit tend to look unrelated to each other.
- for a map exhibiting a sensitive dependence on initial conditions in a compact space (closed and bounded), one can see how the orbit structure can be complicated. If all orbits are moving away from each other, and yet cannot go beyond the boundaries of the space, they just wind up mixing around each other. Think of smoke rising form a hot cup of coffee, or rising from a cigarette, and you can see just how complicated the orbits can be in this case.

EXERCISE 171. Show isometries cannot exhibit a sensitive dependence on initial conditions.

THEOREM 6.52. *Chaotic maps exhibit a sensitive dependence on initial conditions, except when the entire space consists of one periodic orbit.*

EXAMPLE 6.53. Let

$$X = \left\{ 0\,,\ \frac{1}{5}\,,\ \frac{2}{5}\,,\ \frac{3}{5}\,,\ \frac{4}{5} \right\},$$

and $f : X \to X$, $f(x) = x + \frac{1}{5} \mod 1$. Here, f is continuous with respect to the topology X inherits from $\mathbb{R}$ (this is called the subspace topology: Write $X \subset \mathbb{R}$ in the obvious way. Then declare any subset of X to be open if it can be written as an intersection of X with open set of $\mathbb{R}$. This is really the trivial topology of X as it is a finite, discrete subset of $\mathbb{R}$, so each point of X is open). Here, X certainly has a dense orbit (every point of X live in the orbit of $\frac{1}{5}$). And the set of all periodic points of X are dense in X (ALL points of X are periodic). Hence f is chaotic on X. But there certainly is not a sensitive dependence on initial conditions here.

EXAMPLE 6.54. The twist map on the cylinder does have a sensitive dependence on initial conditions. To see this, recall that each horizontal circle is invariant, and has a different rotation along it which is a linear function of height. Now take any point x, and any small neighborhood of x. This small neighborhood will include points on horizontal circles different from that of x. Choose any one of these points. Eventually, x and this other point will wind up pretty much on opposite sides of the cylinder. So what is the sensitivity constant (the largest such Δ)?

EXERCISE 172. For the twist map and the standard parameterization (and metric) of the circle given by the exponential map $f(x) = e^{2\pi i x}$, show the sensitivity constant is $\frac{1}{2}$.

PROPOSITION 6.55. *A topological mixing map (on a non-trivial space) exhibits a sensitive dependence on initial conditions.*

Remark 6.56. Perhaps a better definition of chaos is one which requires a sensitive dependence on initial conditions as a third condition along with the other two. This would discount the "chaotic" map in the above example (which is hardly chaotic in a non-mathematical sense), while not restricting in any detrimental way the intent of the property. In fact, this is a fairly widely accepted set of conditions for a map to be chaotic. Note also that the condition of sensitive dependence on initial conditions cannot by itself constitute a chaotic system. The twist map is an example of a system hardly in a chaotic state. And even the *star node*, the equilibrium at the origin of the map $\dot{\mathbf{x}} = I_2\mathbf{x}$ exhibits a sensitive dependence on initial conditions. Again, hardly chaotic, with neither of the other two conditions satisfied.

6.5. Topological Conjugacy

Place here the example of the flow on three space with one attractive cycle as an example of how the time-t maps of a flow may not be topologically conjugate. Also place here Denjoy's Theorem. And then work out the idea of equivalency of flows and do that also.

Definition 6.57. Suppose $g : X \to X$ and $f : Y \to Y$ are maps of metric spaces and there exists a surjective map $h : X \to Y$ such that

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

Then f is called a factor of g under h and f is said to be *topologically semiconjugate* to g via the *semiconjugacy* h. Furthermore, if h is a homeomorphism, then h is a conjugacy and f is topologically conjugate to g. We say in this case that $f \sim_h g$.

Remark 6.58. Homeomorphism defines an equivalence relation on the set of all topological spaces, in the sense that if two spaces are homeomorphic, they are for all intents and purposes equivalent. Indeed, via Definition 2.82, if $h : X \to Y$ is a homeomorphism, then $X \sim_h Y$. With $X \sim X$, using the identity map, $Y \sim X$ using h^{-1}, also a homeomorphism when h is, and for $g : Y \to Z$ a homeomorphism, we have $X \sim Z$ using the homeomorphism $g \circ h$. Now for maps on each of these spaces, we can construct the same properties: $f \sim f$, $f \sim g$ iff $g \sim f$ and when $f \sim g$ and $g \sim j$, then $f \sim j$. And in a conjugacy, we can see directly that $g^m = h^{-1} \circ f^m \circ h$, so that orbits go to orbits via h. Thus the orbit structure of g and that of f are the same. This becomes an isomorphism for dynamical systems; as we will see, the existence of a conjugacy allows us to study hard-to-study dynamical systems by instead establishing a conjugacy between them and easy to study ones.

We can generalize the full tent map from Example 6.31 to the *tent map* $T_r : [0,1] \to [0,1]$ in Figure 10; a continuous, piece-wise linear, unimodular interval map given by

$$T_r(x) = \begin{cases} rx & \text{if } \ 0 \le x \le \frac{1}{2} \\ r(1-x) & \text{if } \ \frac{1}{2} \le x \le 1. \end{cases} \tag{6.5.1}$$

This is also sometimes called the sawtooth function. Its height, at $x = \frac{1}{2}$, is $\frac{r}{2}$.

In contrast, the linear expanding map E_2 on S^1 has the graph at left. As a map

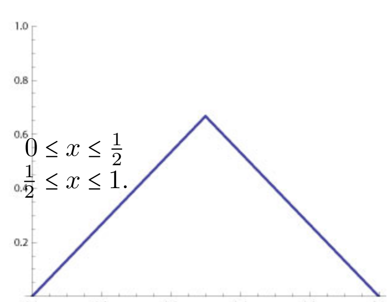
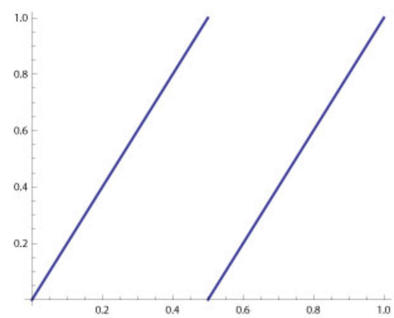


Figure 10. The tent map T_r.

on S^1, it is certainly continuous (here, the point 0 is the same as 1 in both the domain and the range. Hence the map can run off the top of the graph and reappear at the bottom and still be continuous). As a graph in the unit square displays much of the same information as the tent map when the peak is precisely at 1. In fact, we can define E_2 as an interval map via

$$E_2(x) = \begin{cases} 2x & \text{if} \quad 0 \le x \le \frac{1}{2} \\ 2x - 1 & \text{if} \quad \frac{1}{2} \le x \le 1. \end{cases}$$

PROPOSITION 6.59. *The logistic map $f_4(x) = 4x(1-x)$ on $[0,1]$ is topologically semi-conjugate to $E_2(x) = 2x \mod 1$ on S^1 via $h_1(x) = \sin^2 \pi x$, and topologically conjugate to the tent map $T_2 : [0,1] \to [0,1]$ via the conjugacy $h_2(x) = \sin^2 \frac{\pi}{2} x$.*



EXERCISE 173. Show that $h_2(x)$ is a homeomorphism, while $h_1(x)$ is surjective but cannot be a homeomorphism.

PROOF. Here, we will explicitly show the conjugacies. First, we show $h_1 \circ E_2 = f_4 \circ h_1$. This semi-conjugacy condition needs to be parsed along the linear pieces of E_2. Hence we want

$$h_1(2x) = f_4\left(\sin^2 \pi x\right) \text{ for } 0 \le x \le \frac{1}{2} \text{ and} \tag{6.5.2}$$

$$h_1(2x-1) = f_4\left(\sin^2 \pi x\right) \text{ for } \frac{1}{2} \le x \le 1. \tag{6.5.3}$$

As for the left hand sides of these two equations, in Equation 6.5.2, we get $h_1(2x) = \sin^2 \pi(2x) = \sin^2 2\pi x$.. And in Equation 6.5.3, on the left, we also have

$$h_1(2x-1) = \sin^2 \pi(2x-1) = \sin^2(2\pi x - \pi) = \sin^2 2\pi x$$

since $\sin(x - \pi) = -\sin x$. On the right hand side of each, we see

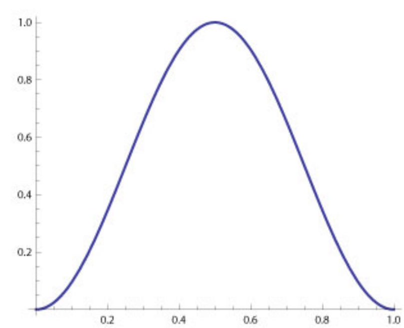
$$\begin{aligned} f_4\left(\sin^2 \pi x\right) &= 4\left(\sin^2 \pi x\right)\left(1 - \sin^2 \pi x\right) \\ &= 4\left(\sin^2 \pi x\right)\left(\cos^2 \pi x\right) \\ &= 4\left(\frac{1}{2} - \frac{1}{2}\cos 2\pi x\right)\left(\frac{1}{2} + \frac{1}{2}\cos 2\pi x\right) \\ &= 4\left(\frac{1}{4} - \frac{1}{4}\cos^2 2\pi x\right) \\ &= 4\left(\frac{1}{4}\sin^2 2\pi x\right) = \sin^2 2\pi x. \end{aligned}$$

As for the conjugacy $h_2(x)$, we need to show that $h_2 \circ T_2 = f_4 \circ h_2$. Again, we would need to parse this condition along the two linear pieces of T_2. The two resulting equations are almost identical to the previous case. In fact, Equation 6.5.2 is precisely the same with all of the factors π replaced by $\frac{\pi}{2}$ (thereby replacing h_1 with h_2). And for Equation 6.5.3, this time we get

$$h_2(2-2x) = \sin^2 \frac{\pi}{2}(2-2x) = \sin^2 \pi(1-x) = \sin^2 \pi - \pi x = \sin^2 2\pi x$$

since $\sin(\pi - x) = \sin x$. $\square$

Notes:



- The maps h_1 and h_2 are truly related, and come from the relationship between S^1 and I. Conjugacies are really all about maps that take orbits to orbits, and any map that satisfies this condition will transfer the dynamics of one system to the other. In this case, both the tent map and the expanding circle map have a certain symmetry about them; $E_2\left(x + \frac{1}{2}\right) = E_2(x)$ on I, while $T_2(x) = T_2(1-x)$. f_4 shares the latter property with T_2, and $T_2(x) = 1 - E_2(x)$ on the interval $\frac{1}{2} \le x \le 1$. The sine function has the appropriate property that $\sin \pi x = \sin \pi(1-x)$. The sine function is also a beautiful way to map S^1 down onto an interval. Indeed, view points of S^1 as $e^{2\pi i x}$, for $x \in I$, and the real part of $z = e^{2\pi i x} \in S^1$ is $\cos 2\pi x$. We can scale this as a "tent-like" map on I as the function
$$x \mapsto \frac{1 - \cos 2\pi x}{2} = \frac{1}{2} - \frac{1}{2}\cos 2\pi x = \sin^2 \pi x.$$
This is precisely h_1 above. For h_2, halving the angle makes h_2 1-1 on I.
- Once a (semi)-conjugacy is specified, ALL of the interesting dynamics of the logistic map for $\lambda = 4$ are present in the tent map for $r = 2$, as well as the linear expanding map E_2 on S^1.

Example 6.60. The map $f_4(x) = 4x(1-x)$ on $[0,1]$ is not topologically conjugate to $f_\lambda(x) = \lambda x(1-x)$, for $\lambda \in [0,1]$ since for each choice of λ, the latter family does not have a nontrivial period-2 point, while $x_* = \frac{5+\sqrt{5}}{8}$ satisfies $f_4(x_*) \ne x_*$, but $f_4^2(x_*) = x_*$. Check this.

Exercise 174. Show that the map $h : [0,1] \to [-2,2]$, $h(x) = 2\cos \pi x$, establishes a conjugacy between the Tent map T_2 and the map $g(x) = 2 - x^2$.

Here is another beautiful family of examples:

Example 6.61. For $\alpha > 0$ a real number, let $\varphi_\alpha : \mathbb{R} \to \mathbb{R}$ be defined by

$$\varphi_\alpha(x) = \begin{cases} x^\alpha & x \ge 0 \\ -|x|^\alpha & x < 0. \end{cases}$$

EXERCISE 175. Show φ_α is a homeomorphism on $\mathbb{R}$, and $\varphi_\alpha^{-1} = \varphi_{\frac{1}{\alpha}}$.

This family has the following property: Let $g_\lambda : \mathbb{R} \to \mathbb{R}$, $g_\lambda(x) = \lambda x$ be a family of linear maps for $\lambda \in \mathbb{R}$. Then for $\lambda' = \varphi_\alpha(\lambda)$, we have $g_\lambda = \varphi_\alpha^{-1} \circ g_{\lambda'} \circ \varphi_\alpha$, so that g_λ and $g_{\lambda'}$ are topologically conjugate.

EXERCISE 176. Verify that for $\lambda \in \mathbb{R}$, and $\alpha > 0$ real, $g_\lambda = \varphi_\alpha^{-1} \circ g_{\lambda'} \circ \varphi_\alpha$, when $\lambda' = \varphi_\alpha(\lambda)$.

This leads to the following. Partition $\mathbb{R}$ into the following intervals:

$$(-\infty, -1), \{-1\}, (-1, 0), \{0\}, (0, 1), \{1\}, (1, \infty)$$

and define an equivalence relation R on $\mathbb{R}$ with these equivalence classes. We have the following:

PROPOSITION 6.62. *For the family of linear maps $g_\lambda : \mathbb{R} \to \mathbb{R}$, $g_\lambda(x) = \lambda x$, $\lambda \in \mathbb{R}$,*

$$g_\lambda \sim g_{\lambda'} \quad \textit{iff} \quad \lambda \sim_R \lambda'.$$

EXERCISE 177. Prove Proposition 6.62.

Some notes about conjugacy:

- The Grobman-Hartman Theorem is a statement on the local conjugacy of two flows: In a neighborhood of a hyperbolic equilibrium solution, the flow of $\dot{\mathbf{x}} = f(\mathbf{x})$ is topologically conjugate to a linear flow W[GHTheorem].
- In general, the homeomorphism h establishing the conjugacy (or the surjective map establishing the semi-conjugacy) is difficult, if not impossible, to find. And even in the case where both f and g are smooth (C^∞), h need not be differentiable at all! As an example, $f(x) = 2x$ and $g(x) = 4x$ are topologically conjugate, as in Example 6.61. But the conjugacy $\varphi_2(x)$ has an inverse which is not differentiable. However, showing two maps are not conjugate may in fact be quite easy. For example, if one map has more fixed or n-periodic points then the other, they are not conjugate.

Here are a couple of interesting examples:

EXAMPLE 6.63. *[Arrowsmith & Price]* Let $f : \mathbb{R} \to \mathbb{R}$ be a diffeomorphism (a C^1-homeomorphism with a C^1-inverse), where $Df(x) > 0$ for some $x \in \mathbb{R}$. Then $f \simeq \varphi^1$, where φ^1 is the time-1 map of the flow $\varphi^t : \mathbb{R}^3 \to \mathbb{R}^2$ of the differential equation $\dot{x} = f(x) - x$. Indeed, f and φ^1 share many of the same properties as diffeomorphisms:

- Both are strictly increasing functions on $\mathbb{R}$ (do you see why?).
- Both have precisely the same fixed points, which correspond to the equilibria of the ODE.

  EXERCISE 178. Show this.

  NOTE. *f can have an arbitrary finite number of fixed points (even zero), a countably infinite number, or even a continuum. However, $Fix(f)$ is always a closed subset of $\mathbb{R}$.*

- If there exist gaps between successive fixed points for f, these gaps form open intervals without fixed points. There can only be a countable number of such fixed point gaps (why?)

Suppose $f \neq Id_{\mathbb{R}}$ and choose one such gap I_0. Its endpoints, if they exist, will be fixed points. Suppose, for the sake of argument, it has both, so that we can label them $I_0 = (x_0^*, x_1^*)$ (we will leave it to the reader to adapt this argument for I_0 an infinite length interval.) Now index the rest of the gap intervals using a subset of $\mathbb{Z}$ compatible with the ordering on $\mathbb{R}$.

- f has no nontrivial periodic points and hence each gap I_i on the index set is invariant: If $x \in I_i$, then $\mathcal{O}_x \subset I_i$.

Figure **??** is an example of this construction. Here, we explicitly construct $h : \mathbb{R} \to \mathbb{R}$ so that $h \circ \varphi^1 = f \circ h$. Choose I_i and $x_0, y_0 \in I_i$. Construct the orbits

$$\mathcal{O}_{x_0,f} = \{P_n\}_{n\in\mathbb{Z}} = \{f^n(x_0)\}_{n\in\mathbb{Z}}\,, \text{ and} \tag{6.5.4}$$

$$= \mathcal{O}_{y_0,\varphi^1} \qquad = \{Q_n\}_{n\in\mathbb{Z}} = \{\varphi^n(y_0)\}_{n\in\mathbb{Z}}\,. \tag{6.5.5}$$

Then $f : [P_n, P_{n+1}] \to [P_{n+1}, P_{n+2}]$ and $\varphi^1 : [Q_n, Q_{n+1}] \to [Q_{n+1}, Q_{n+2}]$ are orientation-preserving diffeomorphisms. Construct the homeomorphism (any will do, but we will use the linear one)

$$h_0(y) = x_0 + (y - y_0)\left(\frac{f(x_0) - x_0}{\varphi^1(y_0) - y_0}\right).$$

Extend this homeomorphism to all of the closed interval $\overline{I}_i = [x_0^*, x_1^*]$ via

$$h_n : [Q_n, Q_{n+1}] \to [P_n, P_{n+1}], \quad h_n(y) = f^n \circ h_0 \circ \varphi^{-n}(y).$$

Here, on the successive interval edges, we have $h_n(Q_{n+1}) = h_{n+1}(Q_{n+1}) = P_{n+1}$. Then $h_{\overline{I}_i} : \overline{I}_i \to \overline{I}_i$ is defined by

$$h_{\overline{I}_i}(y) = \begin{cases} x_i^* & y = x_i^* \\ h_n(y) & y \in [Q_n, Q_{n+1}] \quad n \in \mathbb{Z} \\ x_{i+1}^* & y = x_{x+1}^* \quad . \end{cases}$$

Do this on each interval gap I_i, and extend to all of $\mathbb{R}$ via

$$h_(y) = \begin{cases} h_{\overline{I}_i}(y) & y \in \overline{I}_i \\ y & \text{otherwise.} \end{cases}$$

Exercise 179. Verify that this h is indeed a homeomorphism.

It is also a conjugacy, since if $y \in Fix(f)$, then it is obvious that $h \circ \varphi^1(y) = f \circ g(y)$. But also if $y \notin Fix(f)$, then $y \in [Q_n, Q_{n+1}] \subset I_i$ for some $n, i \in \mathbb{Z}$. And then

$$h \circ \varphi^1(y) = h_{n+1} \circ \varphi^1(y) = f^{n+1} \circ h_0 \circ \varphi^{-(n+1)} \circ \varphi^1(y) \tag{6.5.6}$$

$$= f\left(f^n \circ h_0 \circ \varphi^{-n}(y)\right) = f \circ h_n(y) = f \circ h(y). \tag{6.5.7}$$

Example 6.64. Let $f, g : \mathbb{R} \to \mathbb{R}$ both satisfy $f(x) > x$, $g(x) > x$ $\forall x \in \mathbb{R}$, so that both are strictly increasing functions. Then $f \sim g$. Indeed, let $x_0 \in \mathbb{R}$ and consider $\mathcal{O}_{x_0} = \left\{f^i(x_0)\right\}_{i\in\mathbb{Z}} \subset \mathbb{R}$. Here, of course, $\mathcal{O}_{x_0}$, for any x_0 is monotonically increasing and defines a partition of $\mathbb{R}$,

$$P_{x_0} = \bigcup_{i\in\mathbb{Z}} [x_i, x_{i+1}]\,.$$

Then for $y_0 \in (x_0, x_1)$, we have automatically $y_i = f^i(y_0) \in (x_i, x_{i+1})$, $\forall i \in \mathbb{Z}$. Hence for all $x \in \mathbb{R}$, $\mathcal{O}_x$ has a unique iterate in $[x_0, x_1)$. f defines an equivalence relation on $\mathbb{R}$ and the set of equivalence classes can be represented by the interval $[x_0, x_1)$.

Now choose any $x'_0 \in \mathbb{R}$ and under g, construct another equivalence relation on $\mathbb{R}$ in the same manner. Let h_0 be any orientation-preserving bijection from $[x_0, x_1]$ to $[x'_0, x'_1]$. Then h_0 defines a unique homeomorphism $h : \mathbb{R} \to \mathbb{R}$ where $h\Big|_{[x_0,x_1]} = h_0$ which takes $\mathcal{O}_{x_0,f}$ to $\mathcal{O}_{x'_0,g}$ and indeed takes every orbit of f to a unique orbit of g.

EXERCISE 180. Formally construct h from h_0 and the two maps f and g.

Now if we very g, the new map will still be conjugate to the original f as long as the condition of strictly increasing holds. To lose this condition would be to introduce a fixed point for the new g. Since f has no fixed points, the two maps would at this point not be conjugate.

There is a tendency to identify maps that seem to do the same thing even though their descriptions are different. This is an abuse of logic that we try to get away with when we value understanding over formal accuracy. For example, recall that S^1 had multiple geometric interpretations.

EXAMPLE 6.65. Let $X = \mathbb{R}/\mathbb{Z}$ and define the rotation $R_\alpha : X \to X$ by $R_\alpha[x] = [x+\alpha]$, where again $[\cdot]$ denotes the fractional part of any real number (Mod 1). Then let $Y = \left\{ z \in \mathbb{C} \,\middle|\, |z| = 1 \right\}$, and define the map $r_\alpha(z) = z_\alpha z$, where $z_\alpha = e^{2\pi i \alpha}$. Then R_α and r_α are really just topologically conjugate. SO what is the homeomorphism that takes X to Y? It is the exponential map $h : x \mapsto e^{2\pi i x}$. Now show that it is a homeomorphism (on S^1.) And what is its inverse?

CHAPTER 7

Dynamical Invariants

Recall our indicators of dynamical complexity from before: topological transitivity, minimality, density of periodic orbits, chaos, growth rates of periodic orbits, etc. These properties are called *dynamical invariants* under conjugacy:

- Allows you to study a new dynamical system by establishing a conjugacy with a known one.
- Allows you to classify dynamical systems (everything conjugate to a known dynamical system has the same dynamical invariants.

THEOREM 7.1. *Any two degree-2, expanding maps of S^1 are conjugate.*

Thus, the only degree-2 expanding map to study is the map $E_2 : S^1 \to S^1$, $E_2(x) = 2x \mod 1$.

EXAMPLE 7.2. Rotation number classifies circle homeomorphisms

Herein we define a new dynamical invariant, called the *topological entropy* of a map. Roughly speaking, it is the exponential growth rate of the number of orbit segments distinguishable with arbitrary precision. To motivate this discussion, we can say that entropy is a sort of analytical version of Lyapunov exponents.

Do a historical treatment and give a better description. W[Lyapunov Exponent]

DEFINITION 7.3. *Lyapunov Exponents* are numbers which represent the exponential rate of divergence of nearby trajectories.

The idea here is to take two trajectories of initial separation $\delta_0 > 0$. If after time-t, the separation is $\delta_t = e^{\lambda t}\delta_0$, then the Lyapunov exponent is λ. Note that this depends largely on the direction of the measurement. Different directions of travel will result in different separation rates. Of interest typically is the largest:

- For C^1-dynamical systems, the exponents are related directly to the eigenvalues of the Jacobian matrix, a local linearization of the system.
- for C^0-systems, there is no Jacobian matrix to work with. However, one can still calculate the maximum exponent via

$$\lambda = \lim_{t\to 0} \frac{1}{t} \lim_{\delta_0 \to 0} \log \frac{\delta_t}{\delta_0}.$$

- Calculations of Lyapunov exponents are usually done numerically and only locally. Only rarely can they be calculated analytically or over the entire space.

Now to actually define topological entropy, we will need some more machinery:

7.1. Box Dimension.

7.1.1. Capacity. Let X be a metric space. A set $E \subset X$ is called *r-dense* if, using the metric,

$$X \subset \bigcup_{x \in E} B_r(x).$$

That is, if X can be covered by a set of r-balls all of whose centers lie in E. Then, in the case that X is a compact metric space (both closed and bounded), the *r-capacity* of X, with metric d is the minimal cardinality of any r-dense set. Denote the r-capacity of a set X by $S_{X,d}(r)$ (or simply $S_d(r)$ when the space X is either understood or not necessary to be explicit about).

Some Notes:

- This is simply a way of denoting the "thickness" of sets which have no actual volume by how they sit inside X (think cantor sets sitting inside an interval).
- It does not really matter ultimately, but we will mostly consider closed balls in these calculations.
- Some examples:

  Example 7.4. $\mathbb{Z}$ is r-dense in $\mathbb{R}$ if $r > \frac{1}{2}$ if the balls are open, and $r \geq \frac{1}{2}$ is the balls are closed.

  Example 7.5. $\mathbb{Z}^2$ is r-dense in $\mathbb{R}^2$ if $r > \frac{\sqrt{2}}{2}$ if the balls are open, and $r \geq \frac{\sqrt{2}}{2}$ is the balls are closed. Can you visualize this?

  Example 7.6. Let $I = [0,1]$ be the unit interval. Using open balls here, the $\frac{1}{2}$-capacity of I is 2. The $\frac{1}{4}$-capacity is 3. The $\frac{1}{8}$-capacity is 5, and the $\frac{1}{16}$-capacity is 9. One can show that $S_d\left(\frac{1}{2^n}\right) = 2^{n-1} + 1$.

  Exercise 181. Show this.

  Exercise 182. Determine a bound on r for which $\mathbb{Z}^3$ is r-dense in $\mathbb{R}^3$.

- These calculations work well with Cantor Sets. Studying how $S_d(r)$ changes as r changes (really, it is the order of magnitude of $S_d(r)$) leads to a generalized notion of dimension.

A rough notion of dimension for a topological space would be how many coordinates it would take to completely determine a point in the space (in relation to the other points). For example, the common description of the two-sphere S^2 is as the unit-sphere in $\mathbb{R}^3$; the set of all unit-length vectors in $\mathbb{R}^3$. However, using spherical coordinates (ρ, θ, ϕ) (see the connection), all of these points have coordinate $\rho = 1$, and hence each point on the sphere only requires two coordinates to differentiate between them. Hence, in a way, S^2 is two-dimensional as a space. This notion is not mathematically precise, however, as there do exist curves (1-dimensional lines) that can "fill" a two-dimensional space (Peano curves, some examples are called). Hence is this curve 1-dimensional, or 2 dimensional? Here, we will explore one mathematically precise notion of dimension (there are many), which will be useful in our definition of topological entropy.

Definition 7.7. A metric space X is called *totally bounded* if $\forall r > 0$, X can be covered by a finite set of r-balls all of whose centers are in X.

Really, this definition is technical, and is meant to account for the general metric space aspect of this discussion. That the centers need to be within X really only is a factor when the metric space X is a subspace of another space Y (otherwise there is no "outside" of X. And in Euclidean space, the notion of totally bounded is just the common notion of bounded that you are used to.

DEFINITION 7.8. For X totally bounded,

$$\mathrm{bdim}(X) := \lim_{r\to 0} \frac{-\log S_{(X,d)}(r)}{\log r}$$

is called the *box dimension* of X.

Notes:

- This concept is also called the Minkowski-Bouligard dimension, or the entropy dimension or the Kolmogorov dimension.
- This is an example of the idea of fractional dimension; some sets may look bigger then 0-dimensional, yet smaller than 1-dimensional, for example.
- In the case where this limit may not exist (I cannot think of an example where it wouldn't for a totally bounded set), certainly one can use the limit superior or the limit inferior to gain insight as to the "size" of a set.
- To calculate, really simply find a sequence of r-sizes going to 0, and calculate the r-capacities for this sequence. If the limit exists, then ANY sequence of r's going to 0, with their associated r-capacities will determine the same box dimension (Why?).

See W[Box Dimension]. Do a bit of history. Compare to Hausdorff dimension? Perhaps do a treatment of that also? Nice fractals satisfy the Open Set Condition (OSC). BoxD and Hausdorff are equal here.

EXAMPLE 7.9. Calculate $\mathrm{bdim}(I)$, for $I[0,1]$ with the metric d that I inherits from $\mathbb{R}$. Recall that if we were to use closed balls, then the $\frac{1}{2^n}$-capacity for I is $S_{(X,d)}\left(\frac{1}{2^n}\right) = 2^{n-1}$. But for open balls, we have $S_{(X,d)}\left(\frac{1}{2^n}\right) = 2^{n-1}+1$. The box dimension should be the same for both. Indeed, it is: For the harder one,

$$\begin{aligned}\mathrm{bdim}\,(I) &= \lim_{r\to 0} \frac{-\log S_{(X,d)}(r)}{\log r} = \lim_{n\to\infty} \frac{-\log\left(2^{n-1}+1\right)}{\log\left(\frac{1}{2^n}\right)} = \lim_{n\to\infty} \frac{\log\left(2^{n-1}+1\right)}{\log 2^n} \\ &\geq \lim_{n\to\infty} \frac{\log 2^{n-1}}{\log 2^n} = \lim_{n\to\infty} \frac{n-1}{n} = 1,\end{aligned}$$

and

$$\begin{aligned}\mathrm{bdim}\,(I) &= \lim_{r\to 0} \frac{-\log S_{(X,d)}(r)}{\log r} = \lim_{n\to\infty} \frac{-\log\left(2^{n-1}+1\right)}{\log\left(\frac{1}{2^n}\right)} = \lim_{n\to\infty} \frac{\log\left(2^{n-1}+1\right)}{\log 2^n} \\ &\leq \lim_{n\to\infty} \frac{\log 2^{n-1}\cdot n}{\log 2^n} = \lim_{n\to\infty} \frac{\log 2^{n-1}}{\log 2^n} + \lim_{n\to\infty} \frac{\log n}{\log 2^n} \\ &= \lim_{n\to\infty} \frac{n-1}{n} + \lim_{n\to\infty} \frac{\log n}{n} = 1 + 0 = 1.\end{aligned}$$

Hence $\mathrm{bdim}\,(I) = 1$. Using the closed ball construction is even easier.

EXAMPLE 7.10. Let C be the Ternary Cantor Set. Show $\mathrm{bdim}(C) = \frac{\log 2}{\log 3}$. Here, assume that C sits inside I from the previous example, and again inherits its metric d from I. And since we can choose our sequence of r's going to zero, we will choose

$r = \frac{1}{3^n}$, and consider only closed balls. Then one can show that $S_{(C,d)}\left(\frac{1}{3^n}\right) = 2^{n+1}$. (Think about this: At each stage, we remove the middle third of the remaining intervals. That means that at each stage we can cover each interval by a closed ball of radius $\frac{1}{3^n}$. But the mid-point is NOT in C, Hence we have to shift over a bit to find a point in C. Which means that we will need another ball to cover the remainder on this side. This over covers the interval, but is not enough to cover two adjacent intervals. And since at each stage there are 2^{n+1} intervals, we are done. See the figure.

The calculation is now easy:

$$\begin{aligned} \operatorname{bdim}(C) = & \lim_{r\to 0} \frac{-\log S_{(C,d)}(r)}{\log r} = \lim_{n\to\infty} \frac{-\log\left(2^{n+1}\right)}{\log\left(\frac{1}{3^n}\right)} = \lim_{n\to\infty} \frac{\log\left(2^{n+1}\right)}{\log 3^n} \\ = & \lim_{n\to\infty} \frac{n+1}{n} \cdot \frac{\log 2}{\log 3} = \frac{\log 2}{\log 3}. \end{aligned}$$

Exercise 183. By construction, calculate the r-capacity and hence the box dimension of the Cantor set formed by removing the middle half of each subinterval of the unit interval at each stage.

Exercise 184. Let $B = \left\{0, 1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots, \frac{1}{n}, \ldots\right\}$. Calculate $\operatorname{bdim}(B)$.

In fact, we have the following:

Theorem 7.11. *Let $C \subset I$ be the Cantor set formed by removing the middle interval of relative length $1 - \frac{2}{\alpha}$ at each stage. Then*

$$bdim(C) = \frac{\log 2}{\log \alpha}.$$

A special note: All Cantor sets are homeomorphic. Yet, if we change the size of a removed interval at each stage, we effectively change the box dimension. This means that box dimension is NOT a topological invariant (remains the same under topological equivalence). Since a homeomorphism here would also act as a conjugacy between two dynamical systems on Cantor Sets, this also means that box dimension is also NOT a dynamical invariant. Here we define the Bowen-Dinaburg (metric) topological entropy. W[Top Ent].

For $f: X \to X$, a continuous map on a metric space (X, d), consider a sequence of new metrics on X indexed by $n \in \mathbb{N}$:

$$d_n^f(x,y) := \max_{0\le i\le n} d\left(f^i(x), f^i(y)\right).$$

Here, with $d_0^f = d$, the new metrics d_n^f actually measure a "distance" between orbit segments

$$\begin{aligned} \mathcal{O}_{x,n} &= \left\{x, f(x), \ldots, f^n(x)\right) \\ \mathcal{O}_{y,n} &= \left\{y, f(y), \ldots, f^n(y)\right) \end{aligned}$$

as the farthest that these two sets diverge along the orbit segment, and assigns this distance to the pair x and y.

Exercise 185. Show for a given n that d_n^f actually defines a metric on X.

Now, using the metric d_n^f, we can define an r-ball as the set of all neighbor points y whose nth orbit-segment $\mathcal{O}_{y,n}$ stays within r distance of $\mathcal{O}_{x,n}$:

$$B_r(x,n,f) = \left\{ y \in X \middle| d_n^f(x,y) < r \right\}.$$

Convince yourself that as we increase n, the orbit segment is getting longer, and more and more neighbors y will have orbit segments that move away from $\mathcal{O}_{x,n}$. Thus the r-ball will get smaller as n increases. But by continuity, the r-balls for any n will always be open sets in X that have x as an interior point. Also, as r goes to 0, the r-balls will also get smaller, right?

Now define the r-capacity of X, using the metric d_n^f and the new r-balls $B_r(x,n,f)$, denoted $S_{(X,d)}(r,n,f)$ (this is the SAME notion of r-capacity as the one we used for the box dimension! We are only changing the metric on X to d_n^f. But the actual calculations of the r-capacity depend on the choice of metric). As before, as r goes to 0, the r-balls shrink, and hence the r-capacity grows. And also, as n goes to ∞, we use the different d_n^f to measure ultimately the distances between entire positive orbits. This also forces the r-balls to shrink, and hence the r-capacity to grow. What is the exponential growth rate of the r-capacity as $r \to 0$? This is the notion of topological entropy:

Definition 7.12. Let $h_d(r,f) := \overline{\lim}_{n\to\infty} \dfrac{\log S_d(r,n,f)}{n}$. Then

$$h_d(f) := \lim_{r\to 0} h_d(r,f)$$

is called the *topological entropy* of the map f on X.

There are many things to say about this. To start:

- Topological entropy is a measure of the tendency of orbits to diverge from each other. It will always be a non-negative number, and the higher it is, the faster orbits are diverging. In Euclidean space, maybe this is not so special (think of the linear map on $\mathbb{R}^2$ given by the matrix λI_2, with $\lambda > 1$. All orbits diverge, but the dynamics is not very interesting), but in a compact space with all orbits diverging, the resulting messy nature of the dynamics can be quite interesting. Thus, topological entropy is a measure of the orbit complexity, and the higher the number, the more interesting (read messy) the dynamical structure.
- Another common notation for topological entropy is $h_{top}(f)$ or $h_T(f)$ or even $h(f)$. These are, in a sense, more accurate since it turns out that the topological entropy of a map does not actually depend on the metric d, at least up to equivalence, chosen for use in its definition. It is possible, however, that inequivalent metrics may lead to either the same or a different entropy. We will use the notation $h(f)$ in our subsequent discussion.
- Contractions and isometries have no entropy:

  Proposition 7.13. *Let f be either a contraction or an isometry. Then $h(f) = 0$.*

  Proof. In the case of f an isometry, for any $n \in \mathbb{N}$, $d_n^f = d$, since distances between iterates of a map are the same as the original distances between the initial points. Hence the r-capacity $S_{(X,d)}(r,n,f) = S_{(X,d)}(r,f)$

does not depend on n, and hence $h(r,f) = 0$. For a contraction, the iterates of two distinct points are always closer together than the original points. Hence also here $d_n^f = d$. This leads to the same conclusion. □

- Topological entropy is a dynamical invariant (invariant under conjugacy). This means that if f is (semi-)conjugate to g, then $h(f) = h(g)$. However, it is also useful to use the contrapositive: If one has two maps where $h(f) \neq h(g)$, then it is not possible that f is (semi-)conjugate to g.
- topological entropy measures, in a way, the exponential growth rate of the number of trajectories that are r-separable after n iterations. Suppose this number is proportional to e^{nh}. Then h would be the growth rate for a fixed r, and as $r \to 0$, this h would tend to the entropy.
- defining the topological entropy for a flow is simply a matter of replacing the $n \in \mathbb{N}$ with $t \in \mathbb{R}$ in all of the definitions for the invariant. we can relate the two in a way: The topological entropy of a flow is equal to the topological entropy of its time-1 map (really, its time-t for any choice of t, since the flow provides the conjugacy of any t-map with any other).
- In practice, topological entropy is quite hard to calculate. However, in many cases, and in response to the last bullet point, the entropy is directly related to the largest Lyapunov exponent of the system, at least for C^1 systems.

PROPOSITION 7.14. *For the expanding map* $E_m : S^1 \to S^1$, *where* $E_m(x) = mx \mod 1$, *and* $|m| \geq 1$, $h(E_m) = \log|m|$.

PROPOSITION 7.15. *For* $f : \mathbb{T}^2 \to \mathbb{T}^2$, *given by* $\vec{x} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \vec{x}$ *(this was the map* F_L *from before),* $h(f) = \log \frac{3+\sqrt{5}}{2}$.

Note: In both of these cases, the topological entropy of the map IS the maximum positive Lyapunov exponent of the system.

EXAMPLE 7.16. Show that $h(E_2) = \log 2$.

To do this calculation, we will need to quantify the r-capacity of S^1 under this map. This amounts to calculating $S_{(S^1,d)}(r,n,E_2)$ for a fixed r and as a function of the iterate number n. Hence we start with a good idea of what constitutes the actual size of an r-ball $B_r(x,n,E_2)$ for a choice of n. Note first that by its definition, $B_r(x,n,E_2)$ is the set of points whose distance away from x is less than r after n iterates of E_2. As the map is expanding by a factor of 2 (locally), distances double after each iterate (see the figure). Hence we will have to get closer to x when we start iterating to remain within r as we iterate. Hence $B_r(x,n,E_2)$ will shrink in size as n increases. How will it shrink?

Suppose for a minute that $r = \frac{1}{4}$. Choose an $x \in S^1$, and recall that

$$B_{\frac{1}{4}}(x,0,E_2) = \left\{ y \in S^1 \,\middle|\, d_0^{E_2}(x,y) = d(x,y) = |x-y| < \frac{1}{4} \right\}.$$

The radius of $B_r(x,n,E_2)$ is $\frac{1}{4}$ here. After one iterate, however,

$$B_{\frac{1}{4}}(x,1,E_2) = \left\{ y \in S^1 \,\middle|\, d_1^{E_2}(x,y) = \max\{|x-y|, |2x-2y|\} < \frac{1}{4} \right\}.$$

Here, it is obvious that the condition that $d_1^{E_2}(x,y) = |2x-2y| = 2|x-y| < \frac{1}{4}$ means that the actual distance between x and y would have to be $|x-y| < \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{8}$. Hence the radius of $B_{\frac{1}{4}}(x,1,E_2)$ is only $\frac{1}{8}$. Similarly, the radius of $B_{\frac{1}{4}}(x,2,E_2)$ is only $\frac{1}{16}$, and in general we have that

$$\text{radius}\left(B_{\frac{1}{4}}(x,n,E_2)\right) = \frac{1}{4}\cdot\frac{1}{2^n}.$$

But, really, the initial size of r does not determine the relative sizes of the r-balls with respect to each other. Hence, we can say that, for any choice of $r>0$, we have

$$\text{radius}\left(B_r(x,n,E_2)\right) = r\cdot\frac{1}{2^n}.$$

Recall that the r-capacity, $S_{(S^1,d)}(r,n,E_2)$ is the minimum number of the r-balls $B_r(x,n,E_2)$ it takes to cover S^1. Think of S^1 as being parameterized by the unit interval $[0,1]$ with the identification of 0 and 1. Then we really only need to find out how many r-balls we need for a given iterate n to cover an interval of length 1. Call this number K_n. Hence, we solve the equation (really, it is an inequality, but since adding one more ball to each quantity will not change the limit, this is an okay simplification)

$$\#\left(B_r(x,n,E_2)\right)\cdot 2\cdot\text{radius}\left(B_r(x,n,E_2)\right) = K_n\cdot 2\cdot r\cdot\frac{1}{2^n} = 1.$$

Which is solved by $K_n = \frac{1}{r}\cdot 2^{n-1}$. This is $S_{(S^1,d)}(r,n,E_2)$.

We now calculate

$$\begin{aligned}
h(E_2,r) &= \overline{\lim_{n\to\infty}}\frac{\log S_{(S^1,d)}(r,n,E_2)}{n} \\
&= \lim_{n\to\infty}\frac{\log\frac{1}{r}\cdot 2^{n-1}}{n} \\
&= \lim_{n\to\infty}\left(\frac{\log\frac{1}{r}}{n}+\frac{\log 2^{n-1}}{n}\right) \\
&= 0+\log 2\cdot\left(\lim_{n\to\infty}\frac{n-1}{n}\right) = \log 2.
\end{aligned}$$

Here again, the r-topological entropy does not depend on r at all, so that

$$h(E_2) = \lim_{r\to 0} h(E_2,r) = \lim_{r\to 0}\log 2 = \log 2.$$

7.2. Quadratic Maps (revisited)

We begin today by going back to quadratic maps:

Let $I=[0,1]$ and $f_\lambda: I\to I$, $f_\lambda(x)=\lambda x(1-x)$, but this time let $\lambda\in[3,4]$.

Definition 7.17. Let $x\in X$ be fixed for the map $f:X\to X$. The *basin of attraction* of x is

$$B(x) = \left\{ y\in X \,\middle|\, \mathcal{O}_y \to x\right\}.$$

- Sometimes the basin of attraction is easy to describe:

  Example 7.18. Let $\dot r = r(r-1)$, $\dot\theta = 1$ be the planar ODE system. It should be obvious now that the only equilibrium solution is at the origin of the plane, and the only other "interesting" behavior is the unstable limit cycle given by the equation $r(t)\equiv 1$. Since solutions are unique on all of $\mathbb{R}^2$ (and hence cannot cross), what starts inside the unit circle stays inside. And since the limit cycle is repelling, and there are no other limit cycles or equilibria inside the unit circle, it must be the case that the origin is attracting (you can also see this directly by noting that $\dot{r} < 0$, $\forall r \in (0,1)$). hence the basin of attraction of the origin is the open unit disk

$$B((0,0)) = \left\{ (r, \theta \in \mathbb{R}^2 \middle| r < 1 \right\}.$$

- Sometimes, it is not:

Example 7.19. Let $f : \mathbb{C} \to \mathbb{C}$, $f(z) = z^2 + c$, for $c \in \mathbb{C}$ a constant. For $c = 0$, we get a rather plain model. $\mathcal{O}_z \to 0$ $\forall |z| < 1$, and $\mathcal{O}_z \to \infty$ $\forall |(|z) > 1$. Do you recognize the map on the unit circle $|z| = 1$? It is the expanding (and chaotic) map $E_2 : S^1 \to S^1$ from before.

Definition 7.20. For $P : \mathbb{C} \to \mathbb{C}$ a polynomial map, the *Julia Set* is the closure of the set of repelling periodic points of P.

Keep this in mind. For the map E_2 in the circle, recall that the periodic points are dense in S^1 (this was a feature of chaos). And since the map is expanding, you can show that all of these periodic points are actually repelling (simultaneously!). The resulting mess is actually what a "sensitive dependence on initial conditions" is all about. Here again, the origin in $\mathbb{C}$ is an attracting fixed point, and its basin of attraction is everything inside the unit circle.

Now, though, let c be small and non-zero. There will still be two fixed points, right? (think of solving the equation $z = f(z) = z^2 + c$. The solutions will be $z = \frac{1 \pm \sqrt{1-4c}}{2}$. For $z \in \mathbb{C}$, this always has two solutions!) The one near the origin will still be attracting, while the one near the unit circle will still be a part of a set of repelling periodic points whose closure will form a (typically) fractal structure. This is again the Julia Set for this value of c, and can be highly bizarre looking. I showed you a few examples in class.

In sum, for general $c \in \mathbb{C}$, the Julia set is not a smooth curve. For example, let $c < -2$ be real. Then $f_c(z) = z^2 + c$ is topologically conjugate to a map of the form $x \mapsto \lambda x(1-x)$ for $\lambda > 4$ (this conjugacy is really just a change of variables. Can you find it?) The ramifications of this being that 1) the dynamics outside of the Julia Set are rather simple (think that outside of those interesting orbits of f_λ that stay within I forever, all orbits basically go to $-\infty$). But this implies that that Julia Set is conjugate to a Cantor Set. But this also means that the Cantor Set of points whose orbits stay within I under f_λ, $\lambda > 4$, consists of the closure of a set of repelling periodic points.

Definition 7.21. An m-periodic point p is called *attracting* under a continuous map f if $\exists \epsilon > 0$ such that $\forall x \in X$, where $d(x,p) < \epsilon$, then $d(f^n(x), f^n(p)) \overset{n\to\infty}{\longrightarrow} 0$.

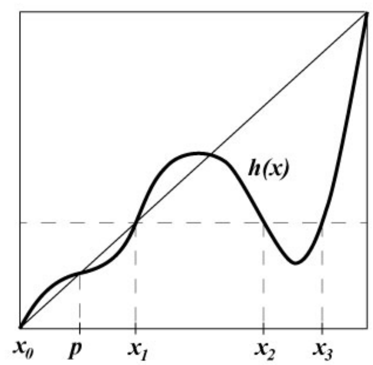
Exercise 186. Show that for an attracting m-periodic point p, each distinct point in its orbit is also attracting.

Call the basin of attraction for an m-periodic point p the union of the basins of attraction for each point of $\mathcal{O}_p$. That is, for $\mathcal{O}_p = \left\{p.f(p), f^2(p), \ldots, f^{m-1}(p)\right\}$, the basin of attraction of p is

$$B(p) = \left\{ x \in X \,\middle|\, d(f^n(x), f^{n+k}(p)) \overset{n\to\infty}{\longrightarrow} 0 \text{ for some } k \in \mathbb{N} \right\}.$$

DEFINITION 7.22. The *immediate basin of attraction* of an m-periodic point p is the largest interval $IB(p)$ containing p such that $\forall x \in IB(p)$, $\mathcal{O}_x \longrightarrow \mathcal{O}_p$. The immediate basin of attraction of a periodic orbit is the union of the immediate basins of attraction of each point in the orbit.

The basin of attraction of a periodic point will in general not consist of a single contiguous interval. However, the immediate basin always is. For $h(x)$ in the figure, for example, $B(p) = (x_0, x_1) \cup (x_2, x_3)$, while $IB(p) = (x_0, x_1)$ (draw some mental cobwebs to convince yourself of this). Back to our discussion of the logistic map, we see that the structure of the graph of $f_\lambda(x)$ on $[0,1]$ says a lot about the dynamical structure of the map:



PROPOSITION 7.23. *Let $f : [a,b] \to [a,b]$ be C^2 and concave down, where $f(a) = f(b) = a$. Then f has at most one attracting periodic orbit.*

We will not prove this here, but the idea rests on three important facts:

- The structure of f (twice-differentiable, concave down with images of endpoints equal) implies that it has a unique critical point $x_0 \in (a,b)$;
- the immediate basin of attraction of any attracting periodic orbit must contain x_0 (this is the non-trivial part of the proof); and,
- basins of attraction cannot overlap.

This proves very useful in our analysis of the logistic map.

EXAMPLE 7.24. In all of our examples of $f_\lambda : I \to I$, where $\lambda \in [0,3]$ there was always an attracting fixed point. However, for $\lambda = 3.1$, for example, the fixed point at $x = 0$ is repelling, and there is an attracting period-2 orbit (can you find the numeric values for this orbit?)

THEOREM 7.25. *If f_λ has an attracting periodic orbit, then the set outside of the basin of attraction (called the universal repeller) is a nowhere-dense null set.*

Some notes:

- A nowhere dense null set in a metric space is a set that can be covered by balls whose total volume is less than ϵ.
- What can lie within the universal repeller? First, any repelling fixed or periodic points, of course. But since the logistic map is a two-to-one map, the pre-image of a fixed point consists of two points, and includes a point that was not previously fixed.

  EXAMPLE 7.26. Let $\lambda = 3.1 = \frac{31}{10}$. It can be shown that $f_{3.1}$ has an attracting period-2 orbit. And $x_\lambda = 1 - \frac{1}{\lambda} = 1 - \frac{1}{\frac{31}{10}} = \frac{21}{31}$ is fixed under

$f_{3.1}$ and repelling (check this!) But the point $1 - x_\lambda = \frac{10}{31}$ also maps to $\frac{21}{31}$. In fact, the point $1 - x_\lambda$ is always the pre-image of the fixed point x_λ due to where it sits on the graph of f_λ. Both of these points are NOT in the basic of attraction of any periodic orbit. But also, $1 - x_\lambda$ is NOT a periodic point. It is an eventually fixed point, but that is different. Now the point $1 - x_\lambda$ also has two pre-images (find them: cobwebbing them is easy. Calculating them?)), and these two pre-images also have two pre-images. In fact, there are a countable number of pre-images that eventually get mapped onto x_λ. All of this set lies outside of the basic of attraction of any attracting periodic orbit, when x_λ is repelling. These points also give a sense for the difference between the basin of attraction and the immediate basin of attraction of an attracting periodic or fixed point. This gives you an idea of what is considered part of the universal repeller. Now think about how this set of pre-images of x_λ sit inside the interval! If you think about it correctly, you start to see just how fractals are born.

Example 7.27. For $\lambda \in \left[3, 1 + \sqrt{6}\right]$, there exists an attracting, period-2 orbit. The basin of attraction is everything except for the points 0, and $x_\lambda = 1 - \frac{1}{\lambda}$ and ALL of their pre-images.

Let's work out the situation: For $\lambda \in (1, 3]$, 0 is a repelling fixed point, x_λ is an attracting fixed point, and there are no other periodic points. In contrast, for $\lambda \in \left(3, 1 + \sqrt{6}\right)$, both $x = 0$, and x_λ are repelling fixed point, and there now exist an attracting period-2 orbit. This means that we have reached a bifurcation value for λ at $\lambda = 3$. This type of bifurcation is called a *period-doubling bifurcation*, and is visually a "pitchfork" bifurcation for the map f_λ^2. See the figure. Analytically, what happens is that the value of $|f'(x_\lambda)| < 1$ for $\lambda < 3$ and $|f'(x_\lambda)| > 1$ for $\lambda > 3$. But these derivatives are negative, right? for the map f_λ^2, this means that the same thing happens but the derivative are positive!. Geometrically, this determines how the graph of f_λ^2 crosses the line $y = x$, and the crossing changes from over/under to under/over as we pass through the value $\lambda = 3$. And when the graph of f_λ^2 passes to the under/over configuration, it creates two new fixed points (for the f_λ^2 map). You do not see these new crossings in the original map f_λ because they are only period two points. You can cobweb them to see that they are there, though. The under/over crossing means that the derivative is greater than 1, and hence the map is expanding near the fixed point (repelling). In contrast, the two new fixed points are over/under crossings, with a derivative less than 1, and hence are attracting. Again, see the figure. it is all in there.

Finally, what happens when $\lambda = 1 + \sqrt{6}$. Basically, the same thing, except that the period-2 orbit becomes a repelling orbit and a period-4 orbit is born and is attracting! Another period-doubling bifurcation.

Theorem 7.28. *There exists a monotonic sequence of parameter values*

$$\lambda_1 = 3, \quad \lambda_2 = 1 + \sqrt{6}, \quad \lambda_3 = \ldots, \quad \text{such that } \forall \lambda \in (\lambda_n, \lambda_{n+1}),$$

the quadratic map $f_\lambda(x) = \lambda x(1 - x)$ has an attracting period-2^n orbit, two repelling fixed points at $x = 0, x_\lambda$, and one repelling period-2^k orbit for each $k = 1, \ldots, n - 1$.

Notes:

- This is called a *period-doubling cascade*.

- At every new λ_n, the previous attracting periodic orbit becomes repelling, and adds (with all of its pre-images) to the universal repeller.
- The length of the intervals $(\lambda_n, \lambda_{n+1})$ decrease exponentially as n increases, and go to 0 somewhere before $\lambda = 4$.
- In fact, one can calculate the exponential decay of these interval lengths:

$$\delta = \lim_{n\to\infty} \frac{\text{length}\,(\lambda_{n-1}, \lambda_n)}{\text{length}\,(\lambda_n, \lambda_{n+1})} = \lim_{n\to\infty} \frac{\lambda_n - \lambda_{n-1}}{\lambda_{n+1} - \lambda_n} \cong 4.6992016010\cdots.$$

  This number has a universal quality to it, as it is always the exponential decay rate of the lengths between bifurcation values in period-doubling cascades. It is called the *Feigenbaum Number*.
- The full bifurcation diagram looks like the figure. At the back edge of the cascade is a place called the transition to chaos. At this point, there are a countable number of repelling periodic points. This collection along with all of their pre-images wind up being dense in the interval, and hence cause a sensitive dependence on initial conditions, commonly found in chaotic systems. This is the Julia set, which in a chaotic system is the entire set.
- Note the self-similar structure of the bifurcation diagram. it is not a fractal, really, but it is related to many of them in interesting ways.
- Look carefully at the bifurcation diagram. Even after the transition to chaos, there seem to be regions of calm periodic behavior. These are not artifacts. In fact, there exists an attracting period-3 orbit 9can you see it?) for a small band of values of λ. This attracting period-3 orbit eventually becomes a repeller, and starts another period doubling cascade (period-6 to period-12, etc.). In fact, there exists a period doubling cascade within this diagram for each prime number n. Look carefully and check in the book in chapter 11 for more details on this fascinatingly simple complicated map.