The Frobenius-Perron Operator

A Magical Tool to Extract Information from Chaos

By Mark Walth

In this notebook, we explore the theory and application of the *Frobenius-Perron Operator*. At a high level, this operator provides a way to extract statistical information about the behavior of a dynamical system. What makes it remarkable is that even when a system exhibits *chaotic* dynamics —where precise long-term prediction is impossible — the Frobenius-Perron operator can still reveal the system's long-term distribution of behaviors. Seemingly paradoxically, by shifting our focus from individual trajectories to the evolution of an ensemble of trajectories, chaotic systems become more tractable.

Our treatment is based off of the book "Chaos, Fractals and Noise", by Lasota and Mackey.

Chapter 1: Introduction

Let's begin by looking at a motivating example. We will start with examples from the world of discrete dynamical systems to keep things simple.

The Logistic Map

Our first example is the **logistic map**, given by

$$\ell(x) = rx(1-x)$$

where $r \in [0, 4]$. The system evolves according to the recurrence relation:

$$x_{n+1} = \ell(x_n)$$

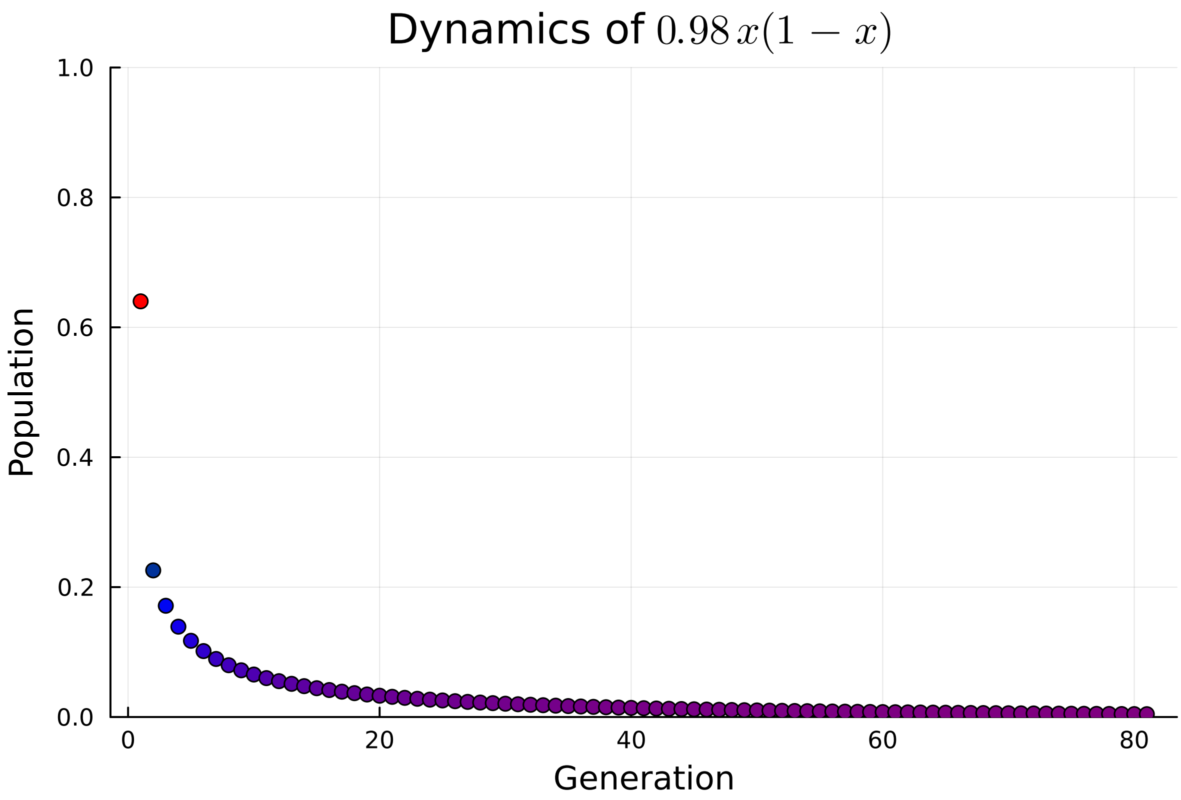
for $x \in [0, 1]$.

The logistic map appears in many mathematical and applied contexts, most notably in population dynamics. Here, x represents the size of a population as a fraction of the environment's carrying capacity. For example, $x = 0.5$ means the population is at half of its theoretical maximum, while $x = 1$ represents a fully saturated population at its environmental limit.

The parameter r governs the reproduction rate, determining how the population evolves over time. As we will see, varying r can lead to a surprising range of behaviors, from stable equilibrium to chaotic fluctuations.

Here's a little exploration of how $x_{n+1} = \ell(x_n)$ behaves for varying values of r. We start with an initial population of $x = .63$. For small values of r, the population eventually just settles down to some fixed value.

r value: 0.98



However, at $r = 3$, a curious thing happens. The stable fixed point splits into two values! The population oscillates each generation, bouncing between two different values, in a 'boom-bust' cycle.

But the story keeps getting stranger! Around $r = 3.45$, the boom-bust cycle itself splits into a new cycle, now with a period of 4 generations. (The exact value of r where this occurs is $1 + \sqrt{6}$.)

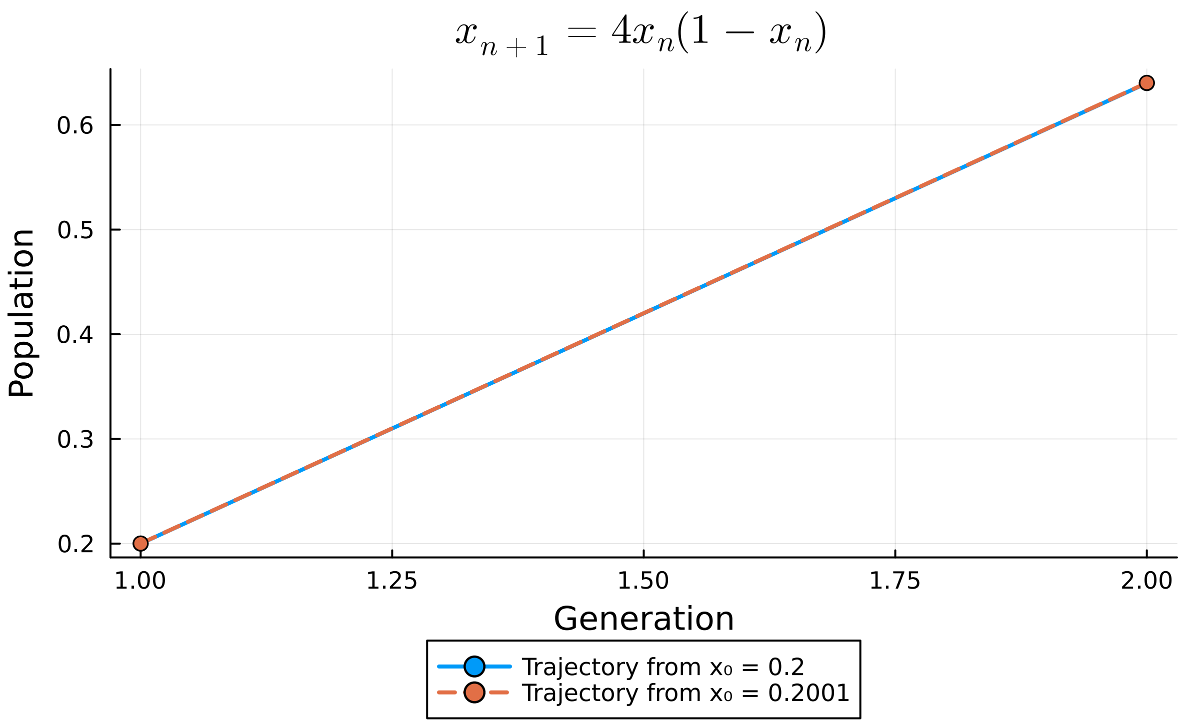
As r continues to increas, the trajectories get ever more complicated. At a critical value (which turns out to be roughly $r = 3.5699$) the population become *aperiodic* - it just bounces all over the place from year to year!

In fact, in this regime, the population is **chaotic**. There is no apparent pattern in the way the population changes from one year to the next. (You may notice that for certain parameter values, namely $r = 3.74$ and $r = 3.84$, the population is again periodic, with periods 6 and 3 respectively. These are known as 'windows of stability' amongst the chaos.)

A complete discussion of the dynamics of the logistic map would take us too far away from our main point, but its a beautiful story. We refer the reader to [Strogatz - Nonlinear Dynamics, Ch: 10] for further details.

For our purposes, the most important feature of the logistic map is its *sensitive dependence on initial conditions*. In the chaotic regime, if two different trajectories start off very very close to each other, they will eventually diverge. In the below demonstration, we take two initial conditions which are with 0.0001 of each other. Their trajectories look similar for the first 11 generations or so, but then they get totally scrambled and soon bear no resemblance to each other.

Number of iterations: 1



If we are scientists studying such a chaotic system, this seems to give us no chance at all of being able to make predictions about how the system will behave! A measurement error of 0.05% is enough to lead to totally different outcomes!

Ensemble Behavior of the Logistic Map

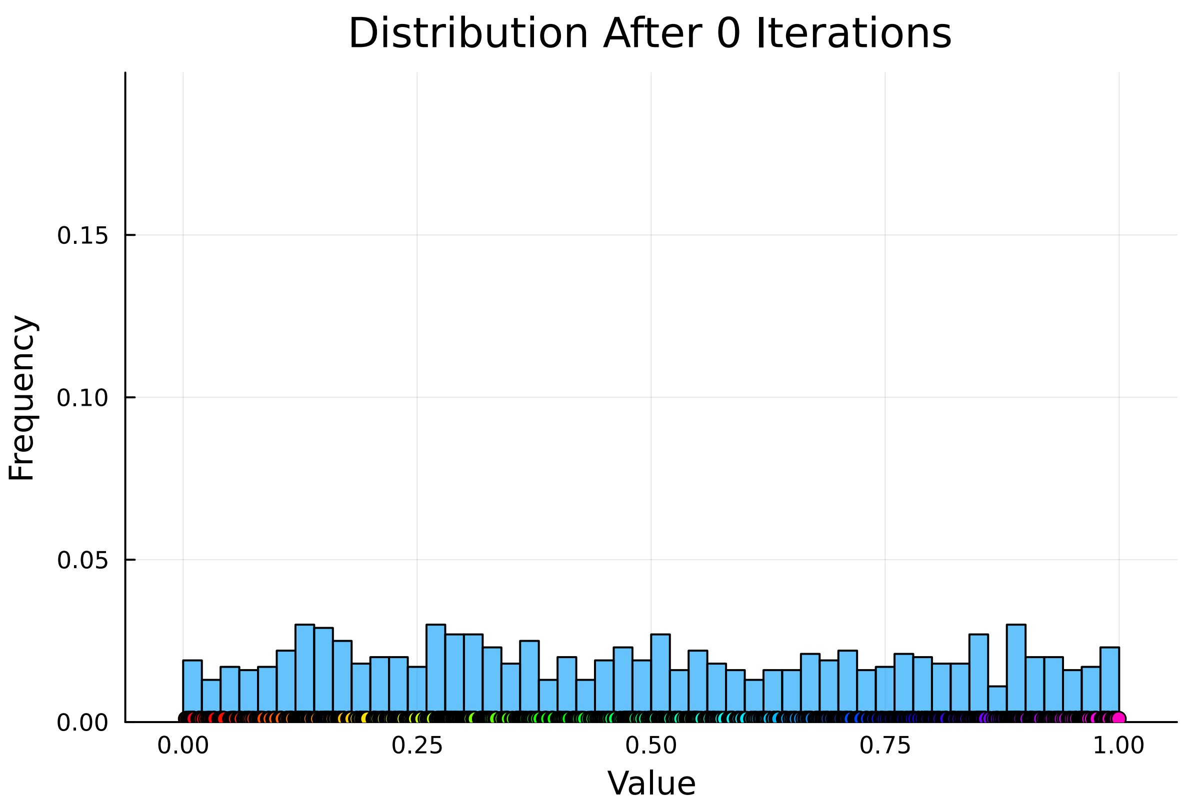
Now we are going to make a somewhat surprising move. Rather than just trying to understand the behavior of a *single* trajectory of the logistic map, let's study what happens to a *large ensemble* of trajectories. It seems at first like this would make the problem much harder - after all, why should studying 1000 trajectories be easier than studying 1 trajectory??

To begin, take a look what happens if we take 100 uniformly randomly distributed points, and apply the map $\ell(x) = 4x(1-x)$ to them repeatedly. Points are color coded by their starting position.

Number of iterations: 0

Clearly, after just a few iterations, the points get all scrambled. To really see what's going on, let's get a sense of where the points are ending up after each iteration. To that end, let's plot a histogram of how many points are in each interval after each iteration. To make the effect stronger, we'll use 1000 points this time.

Number of iterations: 0



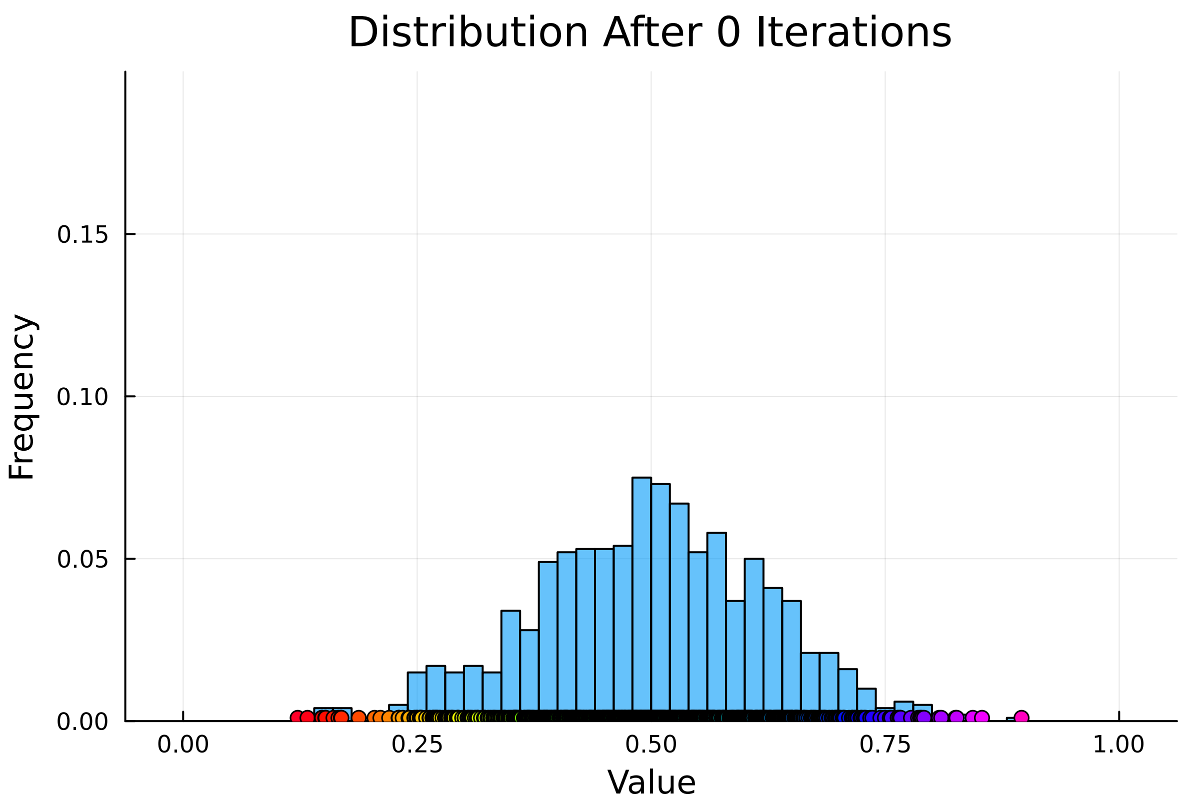
How intersesting! After about 4 iterations, the points seems to converge to some U-shaped distribution, and then stay there! Note that the individual points are still moving all over the place each iteration, but the *distribution* appears (more or less) fixed!

What's amazing is that this fixed distribution seems to be independent of the initial choice of distribution. Below you can play around with what happens if you change the initial distribution of points. Eventually, they all lead to the same final **stationary** distribution.

Initial Distribution: Normal

Number of iterations: 0

Histogram when the initial distribution is Normal



Now we are lead to the following natural

Question: What is that stationary distribution?

Answering this question will lead us to our goal, the **Frobenius-Perron Operator**.

As a first step towards answering this question, we can ask: Given an initial distribution of data, what will be the distribution after applying our map f one time?

Derivation of Frobenius-Perron Operator

To begin to answer this question, we need to introduce some background concepts. We'll work in a more general setting, and then get back to our example of the logistic map.

Let's start by fixing some map, which we'll call S, which takes $[0, 1]$ onto itself, $S : [0, 1] \to [0, 1]$. Given some large initial collection of states, $x_1^0, x_2^0, \ldots, x_N^0$, we'd like to understand something about the distribution of states after applying S. Define

$$x_1^1 = S(x_1^0), x_2^1 = S(x_2^0), \ldots, x_N^1 = S(x_N^0).$$

In what follows, it is useful to define the notion of a *characteristic function*. Let Δ be some subinterval of $[0, 1]$. We'll define the characteristic function 1_Δ by

$$1_\Delta(x) = \begin{cases} 1, & \text{if } x \in \Delta, \\ 0, & \text{if } x \notin \Delta. \end{cases}$$

Now, to talk about the distribution of our states, we're going to borrow the idea of a *density function* from probabibilty theory. Loosely speaking, a density function tells you what percentage of your values lie in a given interval. Symbolically, we can say this as follows, though this definition is still rough. A density function f_0 for the initial states $x_1^0, \ldots, x_N^0$ is one such that for any interval Δ,

$$\int_\Delta f_0(u)du \approx \frac{1}{N}\sum_{j=1}^N 1_\Delta(x_j^0).$$

That is, if we integrate our density over an interval, its about equal to what we'd get if we count what percentage of our data points are in that interval.

Now, what we would like to understand is, given an initial density f_0 for the initial state $x_1^0, \ldots, x_N^0$, can we find a density f_1 for the state one step later, after applying our map S to all the data? Such a density would satisfy

$$\begin{aligned} \int_\Delta f_1(u)du &\approx \frac{1}{N}\sum_{j=1}^N 1_\Delta(x_j^1) \\ &= \frac{1}{N}\sum_{j=1}^N 1_\Delta(S(x_j^0)) \end{aligned}.$$

Now, we make the key observation. The indicator functions that appear in the equation above provide us the way forward. When is the indicator $1_\Delta(S(x_j^0)) = 1$? In other words, when is $S(x_j) \in \Delta$? This happens precisely when $x_j \in S^{-1}(\Delta)$, where by S^{-1} we mean the preimage. This observation gives us the key relationship

$$1_\Delta(S(x)) = 1_{S^{-1}(\Delta)}(x).$$

This may all feel a bit technical. Bear with it, we're almost there.

Putting the pieces together, we want our new density function f_1 to satisfy

$$\int_\Delta f_1(u)du = \frac{1}{N}\sum_{j=1}^N 1_{S^{-1}(\Delta)}(x_j^0),$$

but notice that this right hand side is just exactly equal to $\int_{S^{-1}(\Delta)} f_0(u)du$! Hence, we have arrived at how our densities will evolve as a result of the dynamics S:

$$\int_\Delta f_1(u)du = \int_{S^{-1}(\Delta)} f_0(u)du.$$

To make this observation actionable, we make the concrete choice $\Delta = [a, x]$. Then the above equation becomes

$$\int_a^x f_1(u)du = \int_{S^{-1}([a,x])} f_0(u)du$$

. Finally, differentiating both sides with respect to x, we obtain

$$f_1(x) = \frac{d}{dx}\int_{S^{-1}([a,x])} f_0(u)du$$

.

This inspires the following definition. Given a density of states f and some dynamics S, we define

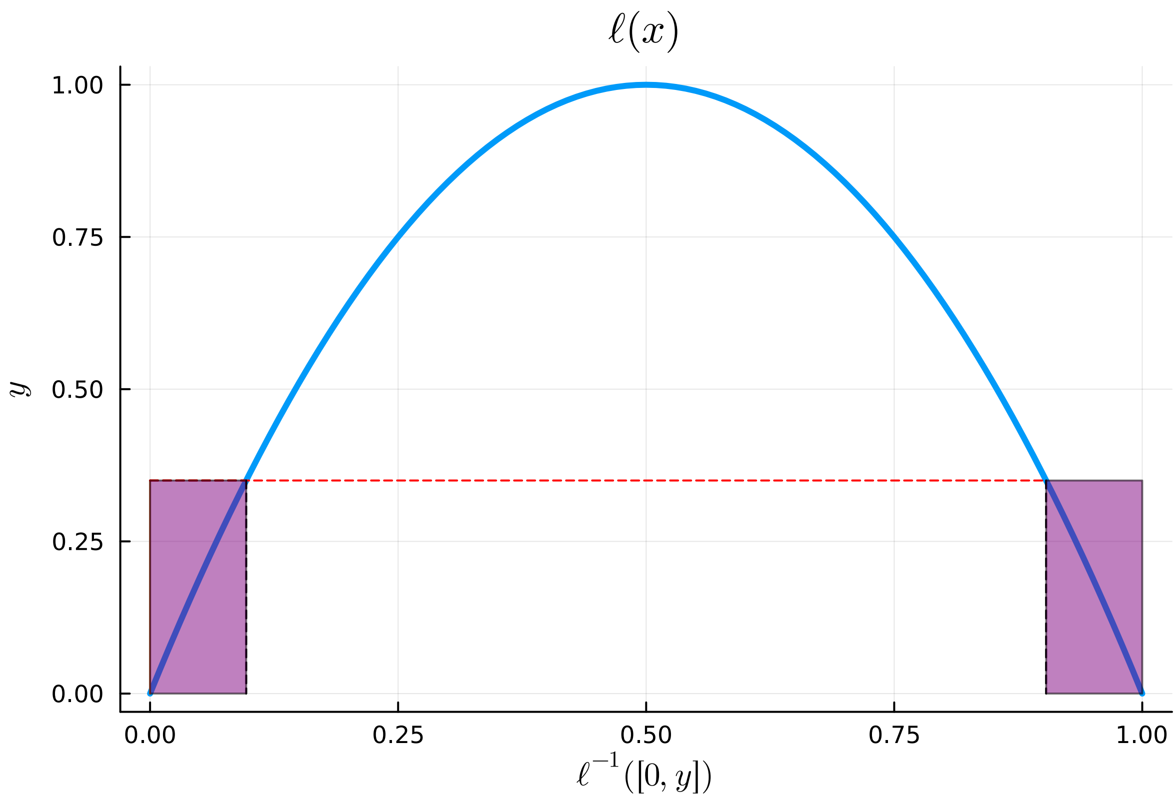
$$Pf(x) = \frac{d}{dx}\int_{S^{-1}([a,x])} f_0(u)du.$$

P is called the **Frobenius-Perron operator.** It allows us to compute how the distribution of states will evolve as we continue to apply our dynamics: if we start with an initial distribution f_0, then after n iterations, our distribution will be given by $P^n f_0$.

Frobenius-Perron and The Logistic Map

What's amazing is that in our logistic map example, we can actually explicitly compute the Frobenius-Perron operator. As before, we'll let $\ell(x) = 4x(1-x)$. To get an explicit form for Pf with this map, we will need to consider what is $\ell^{-1}([a,y])$. For simplicity, we'll set $a = 0$. A picture is helpful.

value of y : 0.35



The picture above shows that for a given y, the inverse image $\ell^{-1}([0,y])$ will generally have two components, visualized as where the purple rectangles touch the x axis. Knowing this, it's not hard to compute that $\ell^{-1}([0,y]) = [0, \frac{1}{2} - \frac{1}{2}\sqrt{1-y}] \cup [\frac{1}{2} + \frac{1}{2}\sqrt{1-y}, 1]$.

Therefore, combining this with the definition for $Pf(x)$, we have that

$$\begin{aligned} Pf(x) &= \frac{d}{dx}\int_0^{1/2-1/2\sqrt{1-x}} f(u)du + \frac{d}{dx}\int_{1/2+1/2\sqrt{1-x}}^{1} f(u)du \\ &= \frac{1}{4\sqrt{1-x}}\left(f\left(\frac{1}{2} - \frac{1}{2}\sqrt{1-x}\right) + f\left(\frac{1}{2} + \frac{1}{2}\sqrt{1-x}\right)\right) \end{aligned}.$$

We now have an *explicit formula* for how the distribution of states will evolve with our logistic function dynamics.

Let's do a reality check. Suppose we start off with a uniform distribution, $f_0(x) = 1$. Then it's not too hard to calculate that

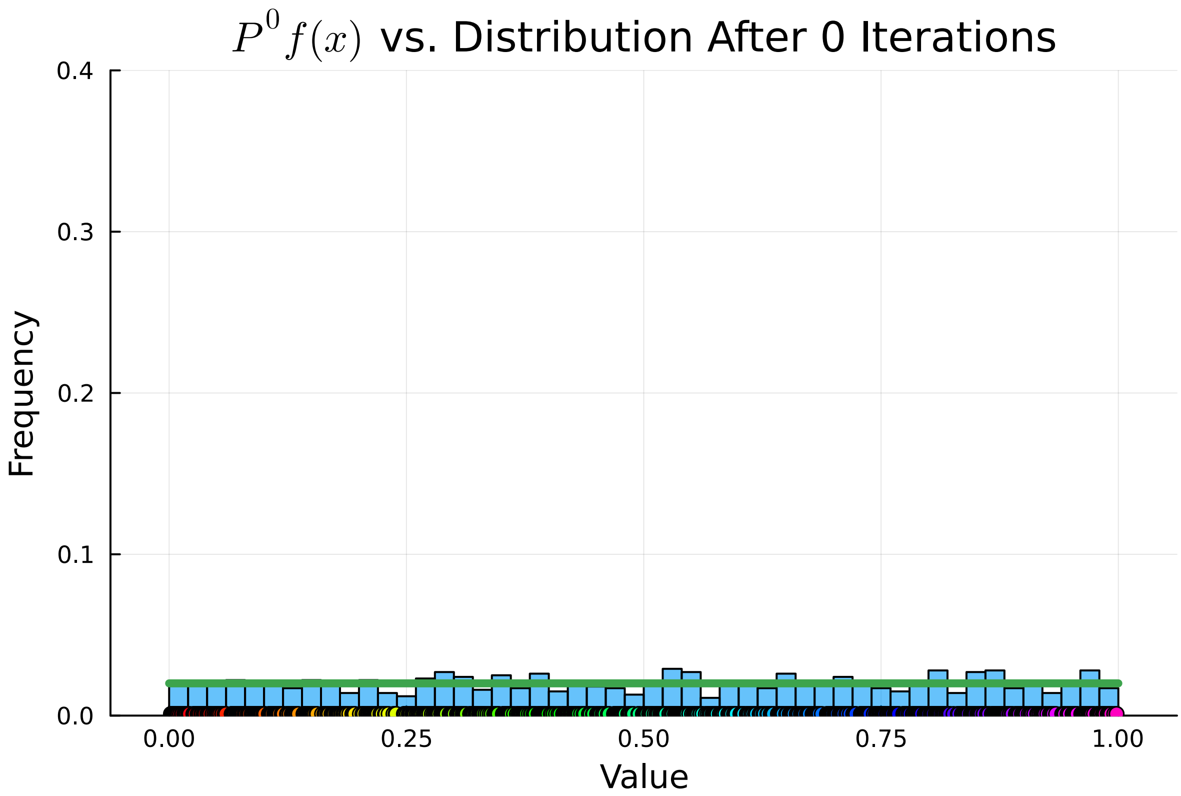
$$\begin{aligned} Pf_0(x) &= \frac{1}{2\sqrt{1-x}} \\ P^2 f_0(x) &= \frac{\sqrt{2}}{8\sqrt{1-x}}\left(\frac{1}{\sqrt{1+\sqrt{1-x}}} + \frac{1}{\sqrt{1-\sqrt{1-x}}}\right) \end{aligned}$$

but computing anything beyond that by hand is clearly going to be a hassle!

Before we go any further, let's just see how our calculations align with the simulations we ran earlier.

Initial Distribution: Uniform

Number of Iterations: 0



The dark green line shows the calculated value of the distribution, using the F.P. operator we calculated earlier. Notice how well it lines up! Also, notice that no matter which initial distribution of the data we use, the calculated value always seems to go to the same U shaped curve! Amazing!

Recap

So far, we have computed an operator P which tells you how a large ensemble of states will evolve under a given dynamics. This means that **even for chaotic systems** we have a means of extracting long term statistical information about how the system will behave.

Question: But what *is* that stationary distribution??

Ah, you've notice that we still haven't answered the earlier question! Very astute. Well it turns out its hard.

Clearly, a stationary distribution f^* must satisfy $Pf^* = f^*$, as this is what it means to be stationary. However, if you write down what that means in our case, you'll see it's not so obvious how to find f^*. It was Jon von Neumann who first solved this problem:

Answer: $f^*(x) = \frac{1}{\pi\sqrt{x(1-x)}}$

Exercise: (Easy)

Verify that the above $f^*(x)$ is indeed stationary.

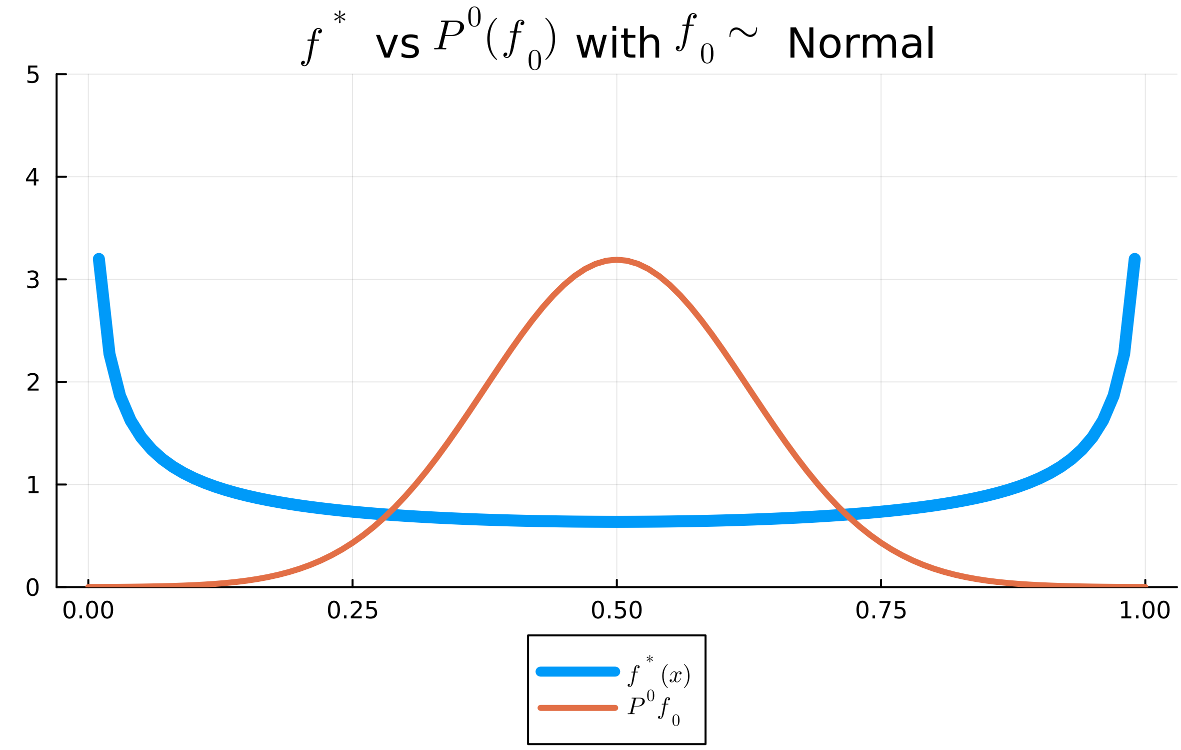
Exercise: (Much harder)

Show that $P^n f_0(x) \to f^*$ if you start with the initial distribution $f_0(x) = 1$.

Just to give a nice visual proof that $f^*(x)$ is indeed the stationary distribution, take a look at the below demonstration

Initial Distribution Normal

Number of Iterations: 0



Chapter 2: Extensions

In this chapter, we will touch on a few extensions of what we have discussed so far. The Frobenius-Perron operator can be defined in much greater generality than what we did in Chapter 1. Here, we'll sketch some of those extensions, and offer a few applications in some more complicated systems.

Extension 1: Higher Dimensional Systems

It should be no surprise that the Frobenius-Perron Operator can be defined in higher dimensional dynamical systems.

Let $S : \mathbb{R}^n \to \mathbb{R}^n$ be a differentiable map with differentiable inverse (i.e. a diffeomorphism). Again, we are interested in how an initial distribution of states f will evolve with each application of the dynamics S. We skip the derivation this time. It turns out we can define the F.P. operator with respect to the dynamics S by

$$P_S f(x) = f(S^{-1}(x))J^{-1}(x)$$

where $J^{-1}(x) = \left| \frac{dS^{-1}(x)}{dx} \right|$ is the determinant of the Jacobian of S^{-1}. This again makes it feasible to actually explicitlly compute P_S for certain mappings S.

If S does not have a differentiable inverse, you can still write down $P_S f$, but it's more combersome. For example, if $S : \mathbb{R}^2 \to \mathbb{R}^2$, then

$$P_S f(x, y) = \frac{\partial^2}{\partial x \partial y} \int\int_{S^{-1}([a,x]\times[c,y])} f(s, t) \text{ ds dt}$$

Let's do some examples. We'll look at several examples. The first is so called the **"Baker's Map,"**

$$B : \mathbb{R}^2 \to \mathbb{R}^2$$

$$B(x, y) = \begin{cases} (2x, \frac{y}{2}), & \text{if } x < 0.5 \\ (2x - 1, \frac{y}{2} + \frac{1}{2}), & \text{if } x \geq 0.5. \end{cases}$$

The Baker's Map is so called because it takes the unit square $[0, 1] \times [0, 1]$, stretches it, and folds it back onto itself, much like a baker kneeding a piece of dough. Below we see a visualization of what the Baker map does to an initial collection of points. Notice how a distribution which is initially quite square gets spread out into thin strips, like a baker folding filo dough.

Our next example is a variant on the famous Arnold Cat Map, given by

$$C(x, y) = (3x + y, x + 3y) \mod 1.$$

While the end behavior of the Cat map and the Baker map look quite similar, looking at the first few iterates, it looks like somehow the Cat map is somehow *more* irregular than the Baker map.

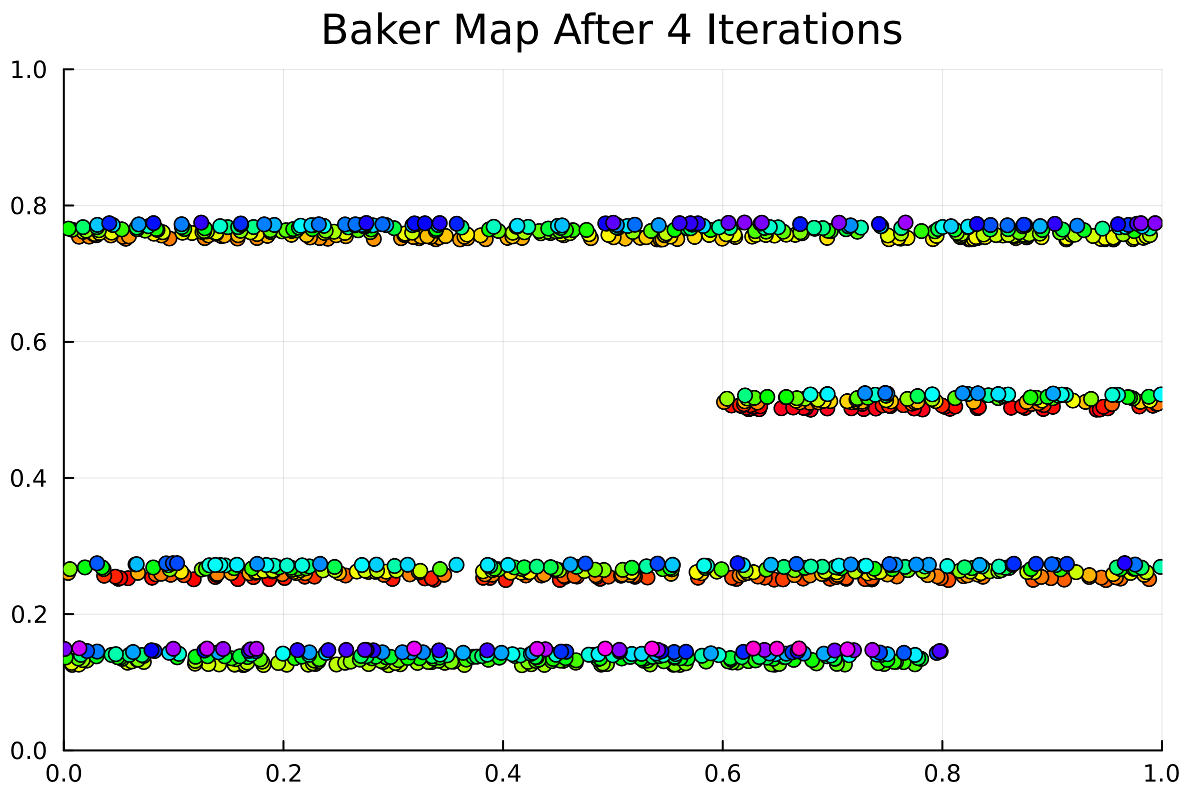
As a final example, we consider the map Erg given by

$$Erg(x, y) = (\sqrt{2} + x, \sqrt{3} + y).$$

Looking at several iterates of Erg, it appears to be the most "regular" of the three maps.

Map: Baker

Number of Iterations: 4



Notice that after about 7 iterations, both the Baker's distribute the points fairly uniformly across the space. For the Cat map, it only takes about 3 iterations. In both cases, this is similar to what happened with the logistic map, how eventually we saw that initial conditions get scattered throughout the space. This is an indication that Baker and Cat both have something like chaotic behavior going on. For now, we will call these behaviors "irregular."

The different varieties of behavior can be fit into a general scheme. The map Erg is "ergodic," the Baker map is "mixing," and the Cat map - seemingly the most irregular - is "exact." It turns out that

$$\text{Exact} \implies \text{Mixing} \implies \text{Ergodic}$$

so that in some sense, being `exact' is the strongest form of irregularity.

The precise definitions are below:

- **Ergodic:** A map S is *ergodic* if

$$S^{-1}(A) = A \iff \mu(A) \text{ or } \mu(X \backslash A) = 0$$

Where μ is some measure on the phase space. Intuitively, this means that the phase space of a system can't be split the phase space into non-interacting regions. The map Erg is ergodic because it eventually sends every point in the phase space near to every other point in the phase space.

- **Mixing:** S is *mixing* if

$$\lim_{n\to\infty} \mu(A \cap S^{-n}(B)) = \mu(A)\mu(B)$$

for all measurable A, B. Intuitively, this says that every part of every set `gets close' to every other set. This happens with the Baker's map.

- **Exact:** S is *exact* if

$$\lim_{n\to\infty} \mu(S^n(A)) = 1$$

for all measurable A with $\mu(A) > 0$. Intuitively, this says that every set blows up to fill up the whole phase space. This is what happens with the Cat map.

Levels of Irregularity and the Frobenius-Perron Operator

We bring up the different levels of irregularity because they can also be completely characterized in terms of the convergence of the Frobenius operator. The following beautiful theorems capture this behavior. (We are not stating the theorems completely precisely, just enough to give an idea of the connection. For precise statements, see [Lasota and Maceky - Chaos Fractals and Noise].)

Theorem

1. The map S is ergodic if and only if $\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} P_S^f = 1$ for all distributions f.
2. The map S is mixing if and only if $P_S^n f$ is weakly convergent to 1 for all distributions f.
3. The map S is exact if and only if $P_S^n f$ is strongly convergent to 1 for all distributions f.

This hopefully gives the reader a taste of how the Frobenius-Perron operator connects with the deep subjects of ergodic theory and measure preserving dynamcis. We hope that this introduction piques your interest.

The Frobenius-Perron Operator of the Baker's Map

To wrap up this section, let's actually compute the Frobenius Perron Operator for the Baker's Map. Using

$$P_S f(x,y) = \frac{\partial^2}{\partial x \partial y} \int\int_{S^{-1}([a,x]\times[c,y])} f(s,t)\, ds\, dt$$

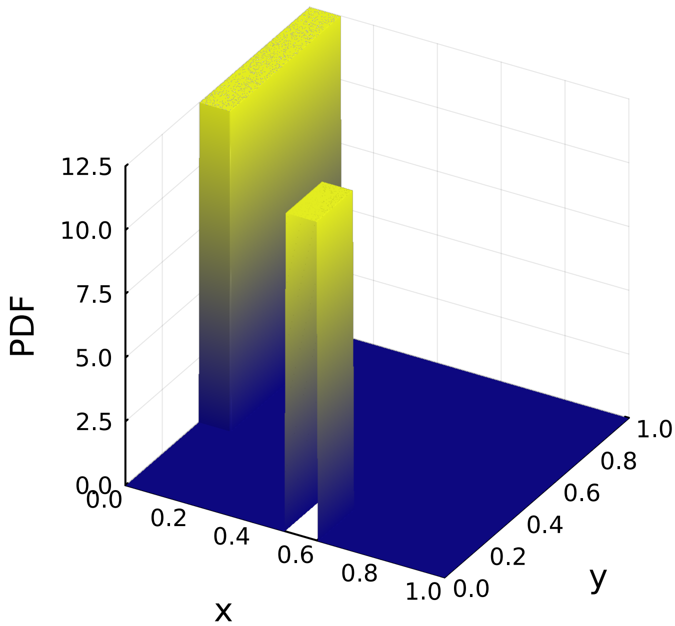
we leave it as an exercise to show that for the Baker's map,

$$P_B f(x,y) = \begin{cases} f\left(\frac{1}{2}x, 2y\right), & \text{if } y \leq \frac{1}{2} \\ f\left(\frac{1}{2} + \frac{1}{2}x, 2y-1\right), & \text{if } y > \frac{1}{2} \end{cases}$$

Let's take a look at what P_B does to an initial density of states. We'll take f_0 to be the uniform distribution on the rectangle $[0.1, 0.3] \times [0, 0.4]$, which is how the initial states were distributed in the above animation. Let's see how the distribution changes under repeated applications of P_B.

Number of Iterations: 2

$P_B^2(f_0)$, where f_0 is Uniform on $[0.1, 0.3] \times [0, 0.4]$



The visualization above helps to see that the uniform density is an invariant density for the Baker's map. This fact is easy to verify directly. This also jives with the theorem above, that for exact maps, $\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} P_S^f = 1$ for all distributions f

Another Generalization: Continuous Time

We note one further generalization of the ideas we've discussed so far. We have only discussed discrete time dynamical systems, but we note that it is completely possible to define a Frobenius-Perron operator in continuous time dynamical systems as well. We note it here for completeness.

Let $\frac{dx}{dt} = F(x)$ and let $S_t(x_0)$ be the associated time t flow map. Then there is an associated Frobenius-Perron operator $P_t : L^1 \to L^1$ for each time t which satisfies

$$\int_A P_t f(x) dx = \int_{S_t^{-1}(A)} f(x) dx$$

for all measurable sets A and distributions f.

Conclusion

In this notebook, we have explored the (unintuitive!) idea that by studying large ensembles of states, we can sometimes extract more information than we can by studying trajectories of individual states. The nature of this information is inherently statistical and probabilistic rather than deterministic, but considering how limited our ability is to make predictions at all in the chaotic case, this represents a huge win. The technique which allows us to do this is the Frobenius-Perron operator. We introduced the operator in a few elementary cases, and hopefully showed enough extensions to pique the reader's curiosity to learn more. We cannot recommend the text "Chaos Fractals and Noise" by Lasota and Mackey highly enough for the reader who wants to dive deeper.