Nonlinear Systems

(. . . and chaos)

a brief introduction

http://csustan.csustan.edu/~ tom

Tom Carter

Computer Science
CSU Stanislaus

http://csustan.csustan.edu/~
tom/Lecture-Notes/Nonlinear-Systems/Nonlinear-Systems.pdf

September 24, 2013

Our general topics:

What are nonlinear systems? ←

- Let's start by adding another word to this, and ask the question, "What are nonlinear dynamical systems?" (and we'll go back to front . . .)

  A reasonable thing to say is that a *system* is a collection of entities that we can treat (for some purpose, in some context) as a unity of interacting parts or elements. At various times we will treat various collections of entities as systems, or subsystems. We may at times ignore certain elements that might otherwise be included. There will also often be times when we will engage in abstraction, and refer to a "system" when we are actually discussing, manipulating, or analyzing an abstraction from the real (physical) system we are interested in.

A *dynamical* system is one which changes over time. It is generally not unreasonable to assume that there is some form of energy flow involved in such a system.

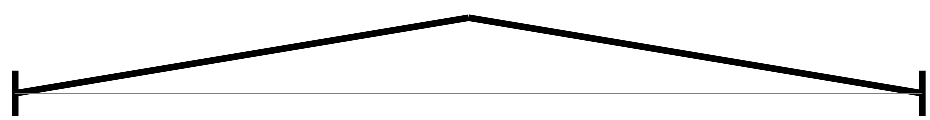
A *nonlinear* dynamical system is, as the name implies, a system whose best description (behavior) is not *linear*. There are various contexts and forms of description for the concept *linear*. We'll look at a variety of examples as we go along.

A Linear Example (string) ←

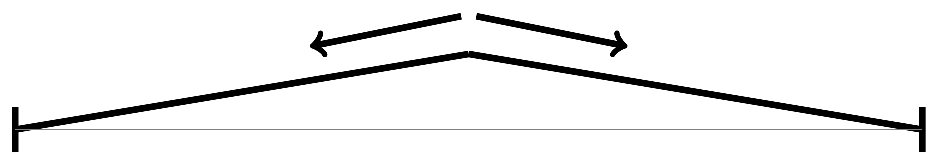
(a little string theory :-)

- It will be worth our while to have some very specific examples available, so let's start with this one – consider a string stretched tightly between two fixed endpoints:

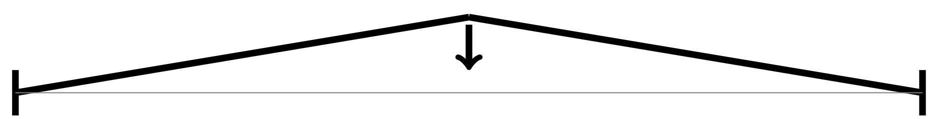
We can pull the center of the string up:



At the moment we release it, there will be forces acting on the string. Let's focus our attention on the center point of the string (where we had taken hold of it to pull it up). There will be forces pulling toward the two fixed ends:

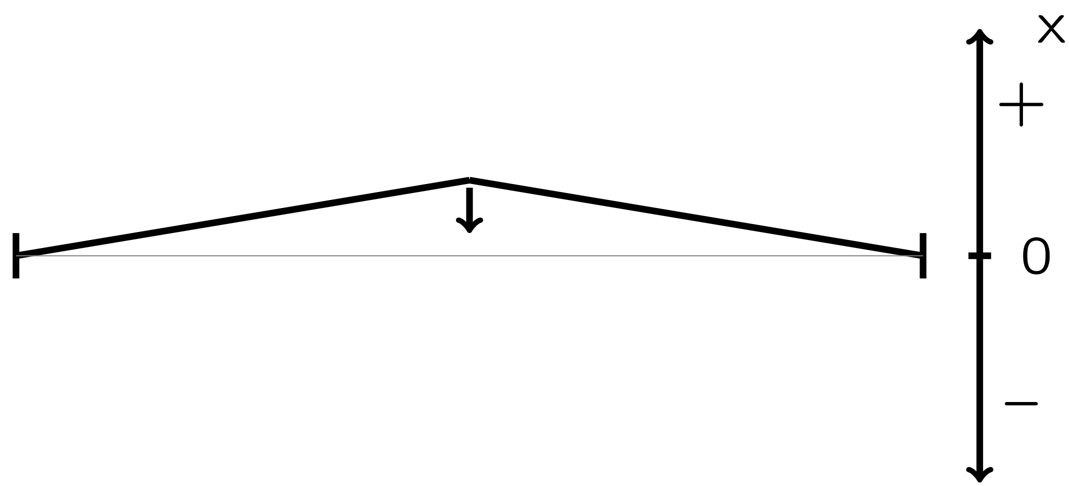


There will be a "net" force acting on the center point of the string:



Now let's put a "reference frame" on the system, so we can make things more explicit.

We will measure the vertical displacement of the center point of the string relative to the "resting position," with positive up and negative down:



At this point we'll make a set of simplifying assumptions, and in particular treat this as a *linear* system (in a sense to be made more explicit below).

- Among the assumptions we'll make are:
  - Everything is nicely symmetric (in particular, the force is exactly vertical).
  - The force is very simple (no complications from friction, etc.).
  - The force changes *linearly* with the displacement of the center point of the string.
  - The motion is always continuous and smooth (differentiable), and thus, also,
  - Newtonian mechanics (and all that that entails, including all the machinery of the calculus . . .)

- Now we'll set this up as a (Newtonian) differential system. The (vertical) displacement of the center point of the string will be denoted by x (which we will recall is actually a function of time t, but we will generally simplify the notation as x rather than $x(t)$). We will (often) denote the derivative of a function with respect to time as

$$\dot{x} = \frac{dx(t)}{dt}$$

and the second derivative as

$$\ddot{x} = \frac{d^2x(t)}{dt^2}.$$

As appropriate, we will refer to velocity and acceleration as

$$v = \dot{x}$$

and

$$a = \dot{v} = \ddot{x}.$$

For the time being, we will assume a very simple form for the force on the center point:

$$F = -x$$

(Note: these should really be vectors $\vec{F} = -\vec{x}$, but we'll keep the notation simple for now . . .). This is the place at which we are doing a *linear* approximation, and thus are working with a *linear dynamical system*.

Now we invoke Newton, and his fundamental equation of motion:

$$F = ma$$

(and we'll use units where $m = 1$, and thus write $F = a$).

- Putting the pieces together, our system is given by the linear ordinary differential equation:

$$a = F = -x$$

or

$$\ddot{x} = -x$$

or

$$\ddot{x} + x = 0.$$

We can use the standard machinery of calculus to solve this differential equation.

The characteristic polynomial of this differential equation is

$$z^2 + 1 = (z + i)(z - i)$$

with roots $z = i$ and $z = -i$. Hence the general form of the solution is

$$x = x(t) = b_0 e^{it} + b_1 e^{-it}.$$

Now we can remember (use) the definition

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!},$$

from which we get

$$
\begin{aligned}
e^{it} &= \sum_{n=0}^{\infty} \frac{(it)^n}{n!} \\
&= 1 + it - \frac{t^2}{2!} - i\frac{t^3}{3!} + \frac{t^4}{4!} + \cdots \\
&= 1 - \frac{t^2}{2!} + \frac{t^4}{4!} - \cdots \\
&\qquad + it - i\frac{t^3}{3!} + i\frac{t^5}{5!} - \cdots \\
&= \sum_{n=0}^{\infty} (-1)^n \frac{t^{(2n)}}{(2n)!} + i \sum_{n=0}^{\infty} (-1)^n \frac{t^{(2n+1)}}{(2n+1)!} \\
&= \cos(t) + i \sin(t).
\end{aligned}
$$

Then, using some standard techniques (and observing that $\cos(t)$ and $\sin(t)$ are linearly independent), we find that the general solution to our system can be written as:

$$x(t) = c_0 \cos(t) + c_1 \sin(t).$$

Of course, if we didn't want to be reminded of the general method of solving linear ordinary differential equations, or the definition of the exponential function, or Euler's formula ($e^{it} = \cos(t) + i \sin(t)$), we could have just observed that

$$\frac{d^2 \cos(t)}{dt^2} = -\cos(t)$$

and

$$\frac{d^2 \sin(t)}{dt^2} = -\sin(t)$$

and gone straight to our general solution :-)

Systems of Differential Equations ←

- There is another nice way to represent such a second order ordinary differential equation, as a system of first order differential equations.

  In this example, we start with the equation

  $$\ddot{x} + x = 0,$$

  then introduce a new variable $v = \dot{x}$ (which we have seen before – the velocity), and then express things as a system:

  $$\begin{aligned} \dot{x} &= v \\ \dot{v} &= -x \end{aligned}$$

  This is a very general approach, which can also be applied to nth order equations, resulting in a system n first order equations.

We can then rewrite the system in vector/matrix notation:

$$\begin{bmatrix} \dot{x} \\ \dot{v} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} x \\ v \end{bmatrix}$$

(note the convenient use of column notation for vectors).

Changing notation slightly, to simplify extension to nth order:

$$\begin{bmatrix} \dot{x_1} \\ \dot{x_2} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

and, writing $\mathbf{x}$ for $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, $\dot{\mathbf{x}}$ for $\begin{bmatrix} \dot{x_1} \\ \dot{x_2} \end{bmatrix}$, and $\mathbf{A}$ for $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$, we can write

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}.$$

Generalizing to nth order, writing $\mathbf{x}$ for $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, etc., and also allowing a constant b_i to be added to each row, we have a system

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{b}.$$

In this form, it is easy to see that these are *linear* systems. There is a well developed theory for solving such systems (using eigenvalues/eigenvectors, etc.), and even for generalizations where $\mathbf{A}$ and $\mathbf{b}$ can be time dependent:

$$\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x} + \mathbf{b}(t).$$

Discussion of this is available in various places, such as here: http://www.unf.edu/~ mzhan/chapter4.pdf.

- For future reference, here is an example of what a *nonlinear* system would like in this sort of notation (of course, the right hand sides won't be linear, and hence we won't get the nice simple matrix representation).

  These are the (famous) *Lorenz equations*, which give rise to the *Lorenz attractor*:

$$\begin{aligned}
\dot{x} &= \sigma(y - x) \\
\dot{y} &= x(\rho - z) - y \\
\dot{z} &= xy - \beta z
\end{aligned}$$

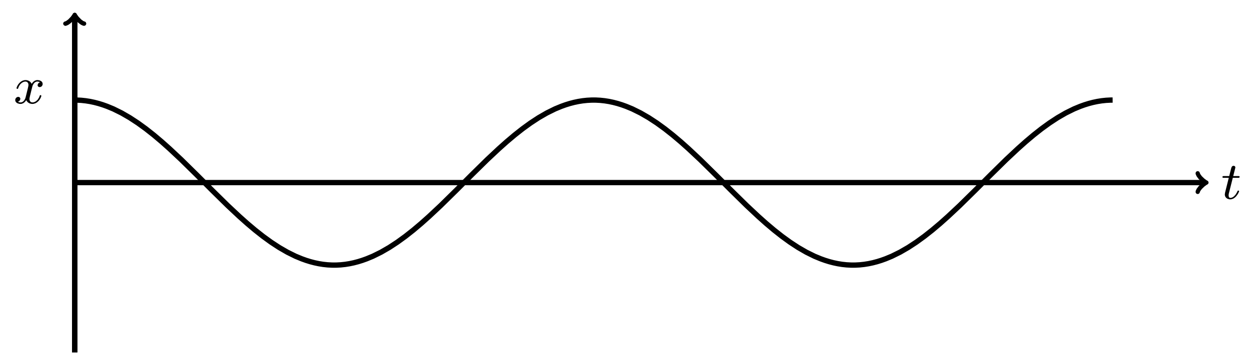
  where σ is the *Prandtl number*, ρ is the *Rayleigh number*, and β is another adjustable parameter. All of σ, ρ, and β are positive, with typical values $\sigma = 10$, $\beta = 8/3$, and ρ is varied. When $\rho = 28$, the system exhibits chaotic behavior. You can see the nonlinear (mixed) terms in the second and third equations (the xz and xy terms).

Simple Harmonic Oscillator $\leftarrow$

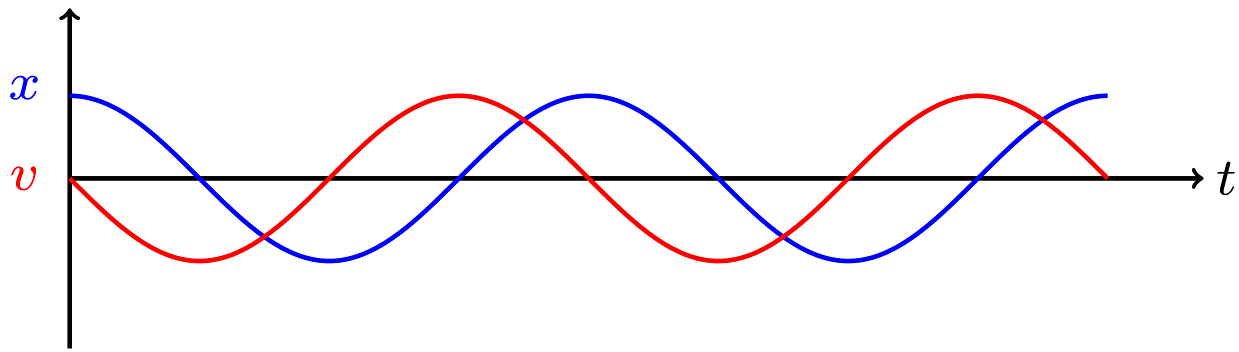
- Let's go back to our string example (i.e., our simple harmonic oscillator). For ease of reading, let's also go back to our notation of position x and velocity v, or, when we want to clarify the time dependence, $x(t)$ and $v(t)$:

$$\begin{aligned} \dot{x} &= v & \qquad \dot{x}(t) &= v(t) \\ \dot{v} &= -x & \qquad \dot{v}(t) &= -x(t) \end{aligned}$$

  The simplest solution to this system is $x(t) = \cos(t)$ (and, of course, $v(t) = -\sin(t)$). We can plot $x(t)$ vs. t:



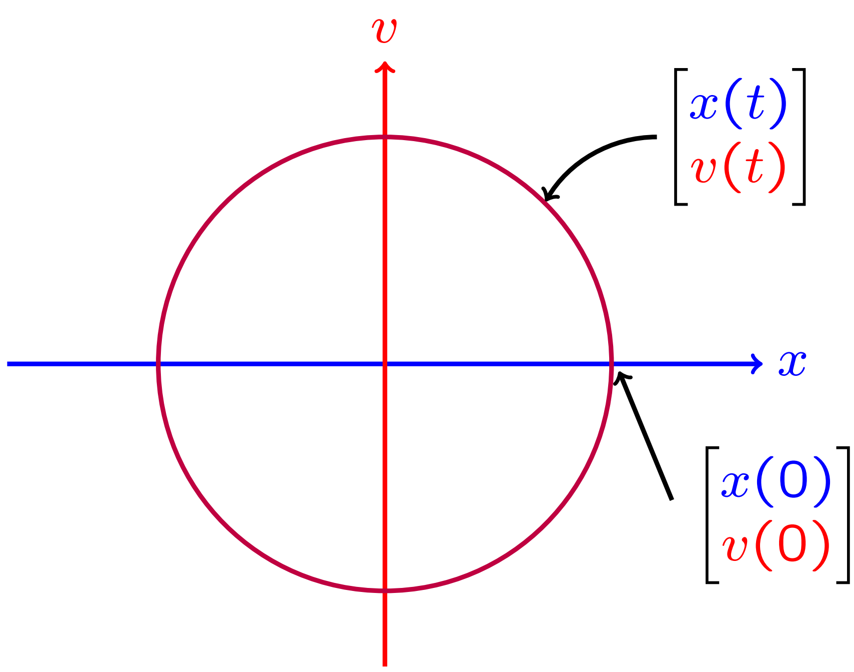
We can also try plotting both $x(t)$ and $v(t)$ on the same graph:



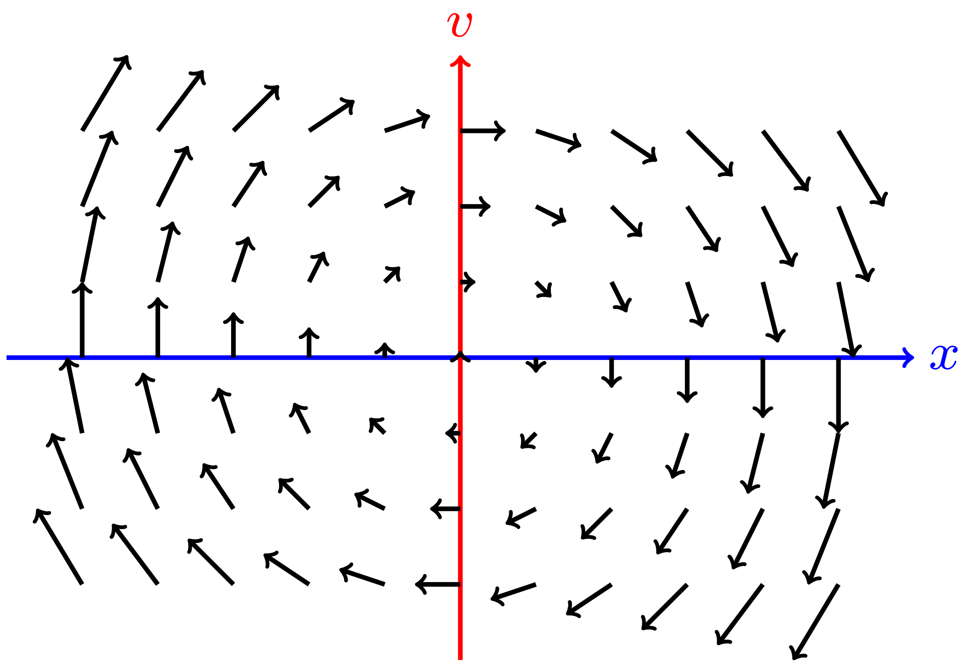
Before long, this can get fairly messy. We would like better ways to visualize the system.

One important point is that this is a fully deterministic system, and the state of the system is completely specified once we know x and v. We can thus visualize the system in *state space* (also called *phase space*), with one dimension for each variable. Each point in phase space represents the state of the system at a particular time, and over time the system will trace a *trajectory* through phase space.

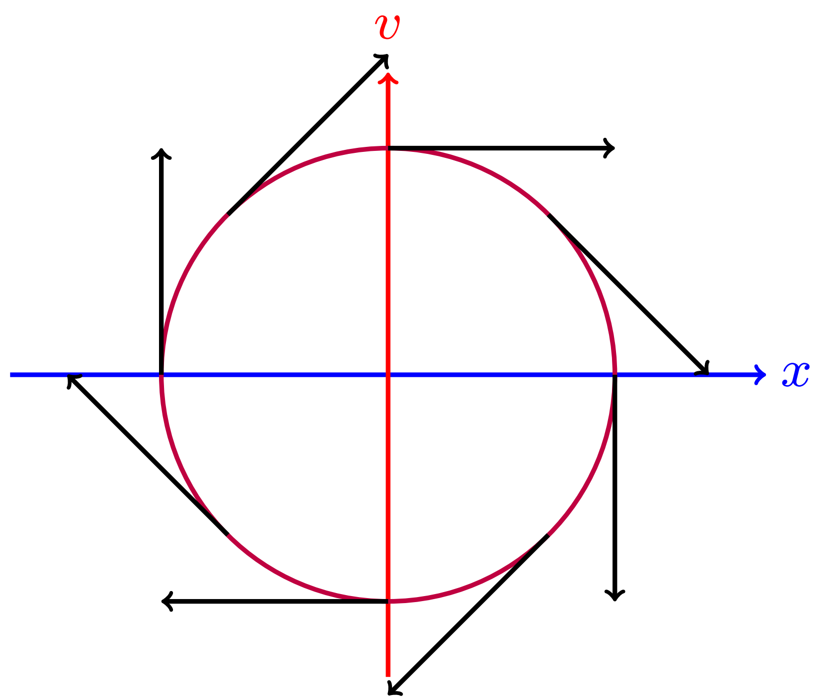
For our simple harmonic oscillator, the phase space trajectory takes a particularly nice form:



- We can also think about the vector field consisting of the derivatives. Associated with each point in phase space, there is the vector of the velocity/acceleration values. At the point $\begin{bmatrix} x(t) \\ v(t) \end{bmatrix}$ there is the derivative vector $\begin{bmatrix} v(t) \\ -x(t) \end{bmatrix}$:



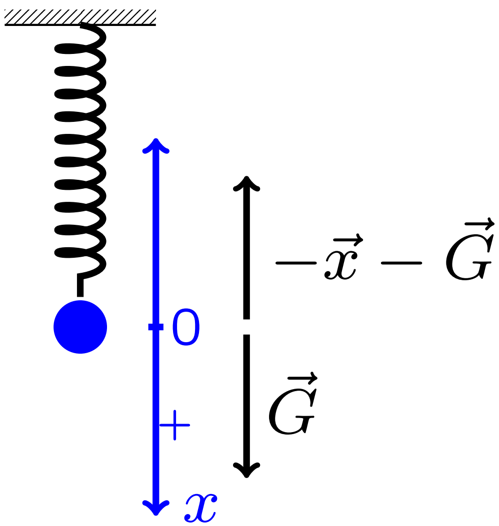
For a given trajectory in phase space, the derivative vectors along the trajectory will be tangent to the trajectory at each point:



Some Other Examples $\leftarrow$

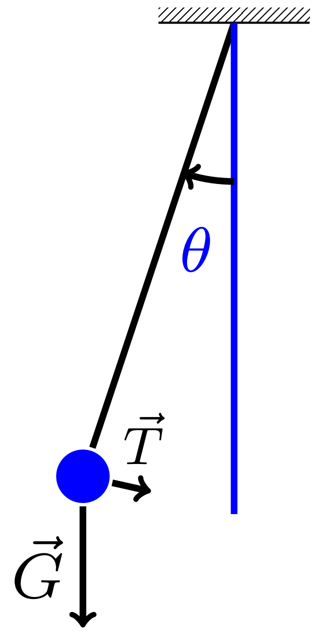
It is probably worth noting that there are various other examples we could have used to get here.

One example is a mass supported by a spring:



Exercise for the reader: make sense of this diagram . . .

Another example is a pendulum:



We measure the angle θ from the vertical. We have a tangential component of the gravitational force

$$|\vec{T}| = -\sin(\theta) \approx -\theta$$

(with the linear approximation, and etc. ...).

It's not hard to see that each of these is described (in the linear approximation) by the same system, x and velocity v:

$$\begin{aligned} \dot{x} &= v \\ \dot{v} &= -x \end{aligned}$$

or, in the case of the pendulum, θ and $\dot{\theta}$:

$$\begin{aligned} \dot{\theta} &= v \\ \dot{v} &= -\theta. \end{aligned}$$

A Linear Approximation ←

- We can even use this to find (at least an approximation to) a trajectory, given an initial point (initial condition) in phase space. The basic idea comes from the definition of the derivative:

$$\frac{df(t)}{dt} = \lim_{\Delta t \to 0} \frac{f(t + \Delta t) - f(t)}{\Delta t}$$

which means that, for small Δt, we have the approximation

$$\frac{df(t)}{dt} \approx \frac{f(t + \Delta t) - f(t)}{\Delta t}$$

or,

$$\frac{df(t)}{dt} * \Delta t \approx f(t + \Delta t) - f(t)$$

and thus the approximation

$$f(t + \Delta t) \approx f(t) + \frac{df(t)}{dt} * \Delta t.$$

In the case of our simple harmonic oscillator, we have the approximation

$$\begin{bmatrix} x(t+\Delta t) \\ v(t+\Delta t) \end{bmatrix} = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} + \begin{bmatrix} \dot{x}(t) * \Delta t \\ \dot{v}(t) * \Delta t \end{bmatrix}$$
$$= \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} + \begin{bmatrix} v(t) * \Delta t \\ -x(t) * \Delta t \end{bmatrix}.$$

Working with our particular system, let's do a couple of steps, starting at $t = 0$, and with $\Delta t = 0.1$. We will have $\begin{bmatrix} x(0) \\ y(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, and $\begin{bmatrix} \dot{x}(0) \\ \dot{v}(0) \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$. Our first step of approximation will give us

$$\begin{bmatrix} x(0.1) \\ v(0.1) \end{bmatrix} \approx \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 * 0.1 \\ -1 * 0.1 \end{bmatrix}$$
$$= \begin{bmatrix} 1 \\ -0.1 \end{bmatrix}.$$

The next three steps of our approximation will be:

$$\begin{bmatrix} x(0.2) \\ v(0.2) \end{bmatrix} \approx \begin{bmatrix} 1 \\ -0.1 \end{bmatrix} + \begin{bmatrix} -0.1 * 0.1 \\ -1 * 0.1 \end{bmatrix}$$
$$= \begin{bmatrix} 0.99 \\ -0.2 \end{bmatrix}.$$

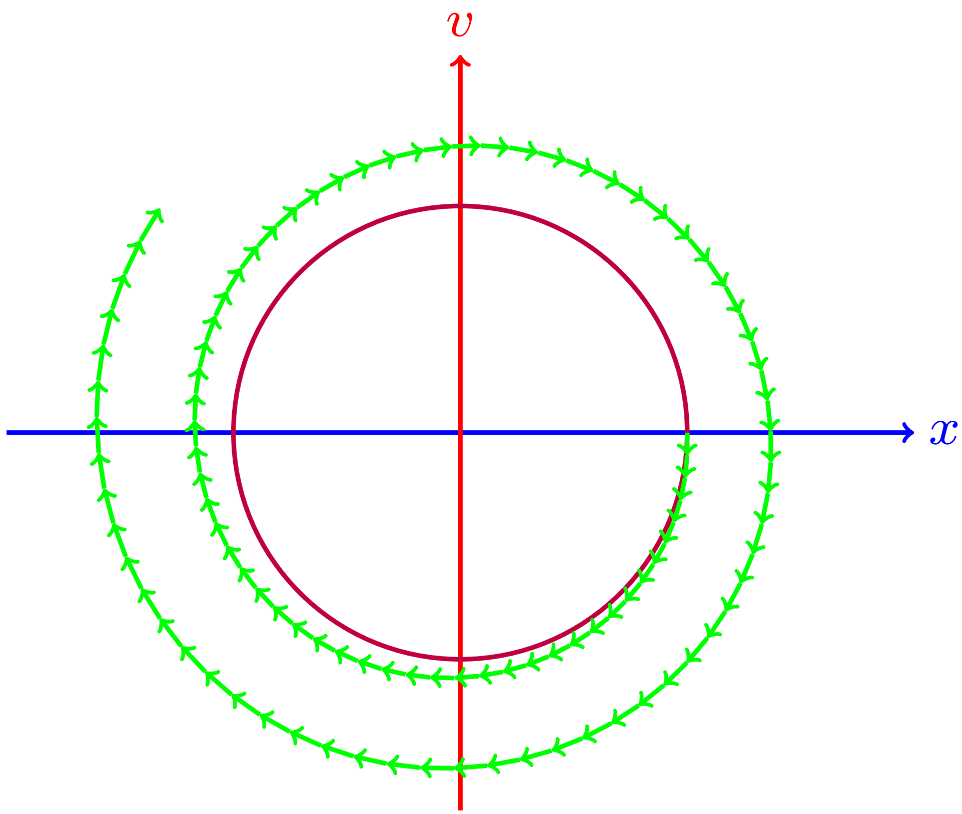
and

$$\begin{bmatrix} x(0.3) \\ v(0.3) \end{bmatrix} \approx \begin{bmatrix} 0.99 \\ -0.2 \end{bmatrix} + \begin{bmatrix} -0.2 * 0.1 \\ -0.99 * 0.1 \end{bmatrix}$$
$$= \begin{bmatrix} 0.97 \\ -0.299 \end{bmatrix}$$

and

$$\begin{bmatrix} x(0.4) \\ v(0.4) \end{bmatrix} \approx \begin{bmatrix} 0.97 \\ -0.299 \end{bmatrix} + \begin{bmatrix} -0.299 * 0.1 \\ -0.97 * 0.1 \end{bmatrix}$$
$$= \begin{bmatrix} 0.9401 \\ -0.396 \end{bmatrix}$$

This will look like (I have gone ahead and added several more steps):



The approximation starts out doing a reasonable job (but over time it does drift away from the real trajectory . . .).

- We'll keep in mind this approximation approach, and come back to it later. In particular, we'll need to think some about what cautions we should keep in mind when using approximations. But, let's briefly return now to the thrilling days of yesteryear, and do a little bit more with linear systems.

  Our linear approximation of the string system clearly leaves out a bunch of stuff. In particular, we know perfectly well that a vibrating string won't go on vibrating forever. Things like friction (both internal within the string itself, and external, like air resistance) will play a role in the dynamics. We can improve our system by adding a friction term. In keeping with our simplification of linearity, a reasonable approximation of the friction term is that it depends linearly on velocity. Keeping terms and constants simple, we can learn a reasonable amount by studying this system (the second term in the $\dot{v}$ line is friction):

$$\begin{bmatrix}\dot{x}\\ \dot{v}\end{bmatrix} = \begin{bmatrix} v \\ -x - bv\end{bmatrix}.$$

Using the "characteristic polynomial" approach, we can look at $\ddot{x} = -x - b\dot{x}$ or $\ddot{x} + b\dot{x} + x = 0$, and thus work with

$$r^2 + br + 1 = 0$$

giving us

$$r = \frac{-b \pm \sqrt{b^2 - 4}}{2}$$

or, assuming $|b| < 2$,

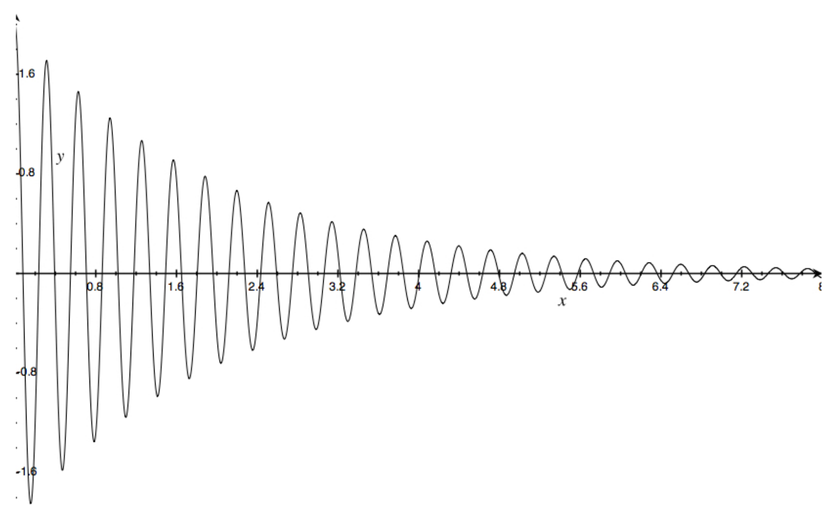
$$r = -\frac{b}{2} \pm i\frac{\sqrt{4 - b^2}}{2}$$

.

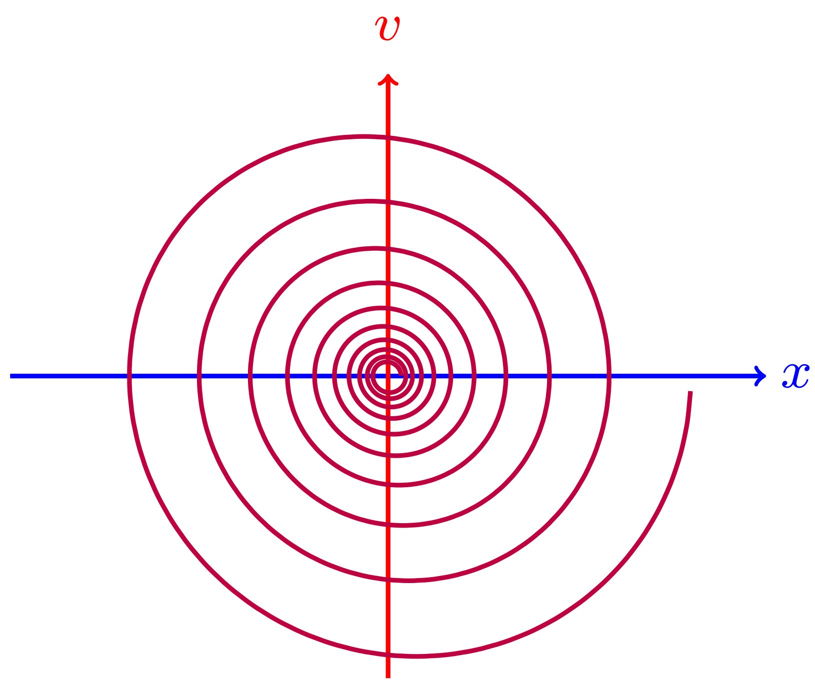
From this, we get a solution to the system in the form

$$
\begin{aligned}
x(t) &= e^{(-\frac{b}{2}+i\frac{\sqrt{4-b^2}}{2})t} + e^{(-\frac{b}{2}-i\frac{\sqrt{4-b^2}}{2})t} \\
&= e^{-\frac{b}{2}t}\left(e^{i\frac{\sqrt{4-b^2}}{2}t} + e^{-i\frac{\sqrt{4-b^2}}{2}t}\right) \\
&= e^{-\frac{b}{2}t} * 2\cos\left(\frac{\sqrt{4-b^2}}{2}t\right)
\end{aligned}
$$

which is decaying oscillations:

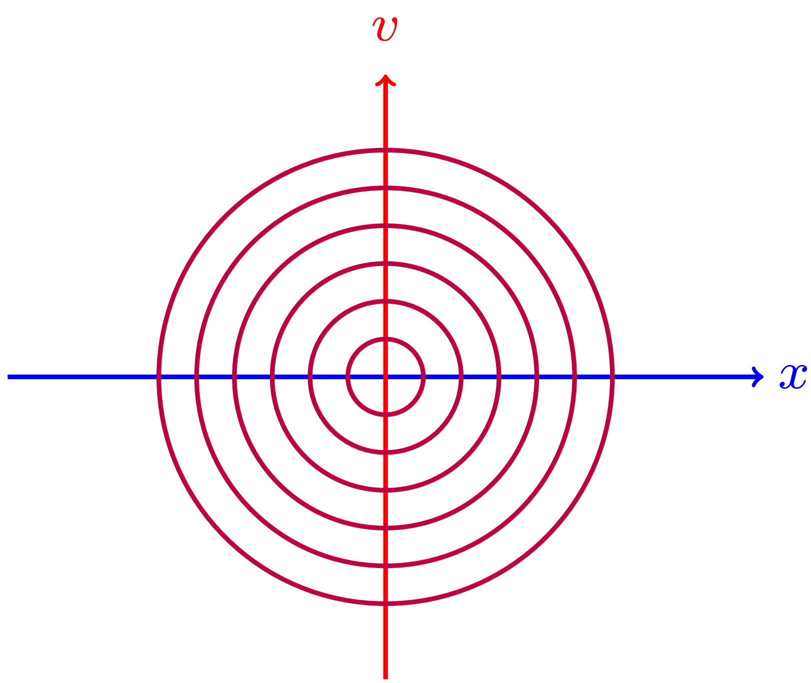


In phase space, this system spirals in toward $(0,0)$ (see appendix 1):



Trajectories ←

- We can characterize various points in phase space, and various trajectories. In the case of the simple harmonic oscillator, for any starting point (initial condition), the resulting trajectory is a simple closed trajectory – a *periodic orbit*.



In the special case of starting at $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$, the system is at a *fixed point*. This fixed point is, however, not *stable*, in the sense that if some noise jostles the system, it will follow a periodic orbit somewhat away from $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

On the other hand, in the case of of the damped harmonic oscillator, any initial condition will tend toward $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ as $t \to \infty$. $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ is a fixed point, and is an *attracting stable* fixed point, in the sense that if noise jostles the system, it will, as $t \to \infty$, return to $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

For any of these continuous systems (i.e., systems where t is a continuous variable), we can get at new system by replacing t with $-t$. In the case of the simple harmonic oscillator, there is complete symmetry, and the new system is indistinguishable from the original – in particular, all trajectories are closed periodic orbits.

In the case of the damped harmonic oscillator, there is a fixed point at $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$, but all other trajectories spiral out (in phase space) away from $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$. In this case, $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ is an *unstable, repelling* fixed point.

Discrete Linear Systems $\leftarrow$

- The systems we have been looking at all have time (t) as a continuous variable (and we have been limiting ourselves to smooth, i.e., differentiable, systems).

  Now I'd like to move to *discrete* systems, where time proceeds in steps rather than continuously. In other words, instead of looking at the variable $x(t)$, we will now take time to only take integral values, and we will be interested in systems with values $x_0, x_1, x_2, \dots$

  Instead of *differential* equations, we will be working with *difference* equations. The general form of a (one variable) difference equation is

$$x_{n+1} = f(x_n, x_{n-1}, \dots, x_0, c).$$

A very simple example is the Fibonacci sequence. This is given by the difference equation

$$x_{n+2} = x_{n+1} + x_n.$$

This is a second order difference equation. Once we specify initial conditions (e.g., $x_0 = 0$, $x_1 = 1$), we have the solution sequence $0, 1, 1, 2, 3, 5, 8, 13, 21, \ldots$

We can solve linear difference equations for the general solution for the nth term using a procedure similar to the approach for linear differential equations. For the Fibonacci sequence, we rewrite the equation as

$$x_{n+2} - x_{n+1} - x_n = 0,$$

and then work with the characteristic polynomial

$$r^2 - r - 1 = 0.$$

From this, we get

$$r = \frac{1 \pm \sqrt{1+4}}{2} = \frac{1 \pm \sqrt{5}}{2}.$$

The general form of the solution is then

$$x_n = a_1 \left(\frac{1+\sqrt{5}}{2}\right)^n + a_2 \left(\frac{1-\sqrt{5}}{2}\right)^n.$$

Once we specify the initial conditions, we can solve for a_1 and a_2. For example, when $x_0 = 0$ and $x_1 = 1$, we get

$$a_1 * 1 + a_2 * 1 = 0$$

$$a_1 * \left(\frac{1+\sqrt{5}}{2}\right) + a_2 * \left(\frac{1-\sqrt{5}}{2}\right) = 1,$$

from which $a_2 = -a_1$, and so

$$a_1 * \left(\frac{1+\sqrt{5}}{2}\right) - a_1 * \left(\frac{1-\sqrt{5}}{2}\right) = 1,$$

and hence

$$a_1 = \frac{1}{\sqrt{5}}$$

$$a_2 = -\frac{1}{\sqrt{5}}$$

and so

$$x_n = \frac{1}{\sqrt{5}} * \left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}} * \left(\frac{1-\sqrt{5}}{2}\right)^n.$$

This is a very general approach (with some minor subtleties when dealing with repeated roots) – and, of course, there are the problems of finding the roots for higher order polynomials, but this basically gets us going on general linear difference equations.

A Discrete Nonlinear System ←

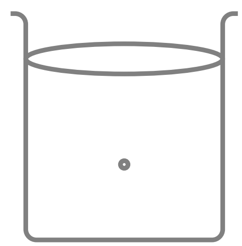
- Having spent some time looking at *linear* systems, now it's time to move on to *nonlinear* systems.

  We will start with a classic example, called the discrete *Logistic Equation*. This example is very simple to develop and specify, but in its details it reveals many of the characteristic properties of nonlinear dynamical systems.

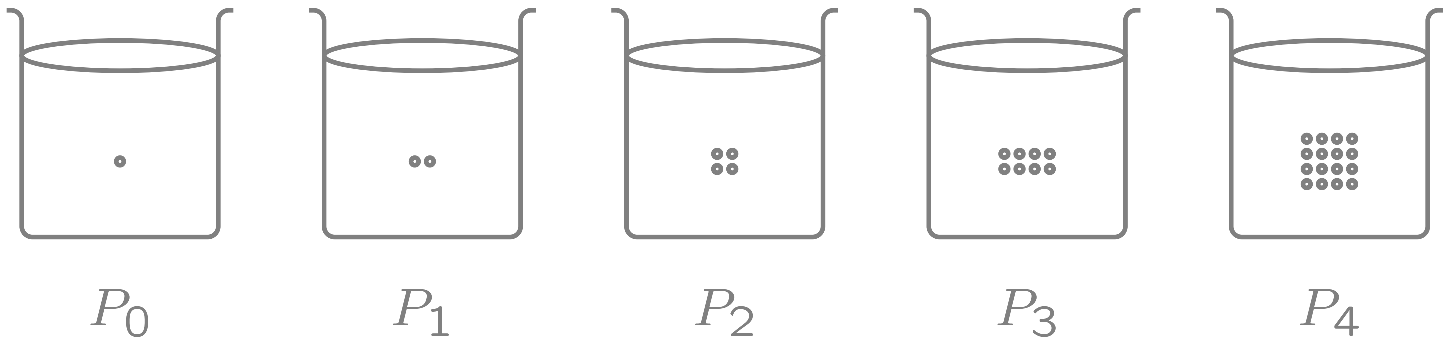
The Logistics Equation (derivation) ←

- Now it's time to get down to business, and start exploring a specific nonlinear example. We're going to look at a classic example from biology, concerning population growth.

  We'll start by considering a single species, living in an environment where it depends on a consumable resource. So, imagine a beaker of sugar water, into which we put a single bacterium:



As time passes, the bacteria reproduce (by dividing, time lapse pictures, bacteria greatly magnified . . .):



The population of bacteria at time $n+1$ will be given by

$$P_{n+1} = 2P_n \quad \text{(with } P_0 = 1\text{)},$$

and so in general

$$P_n = 2^n.$$

- Now we'll ask the traditional question at this point: If we put the first bacterium in the beaker at 9:00 am, the bacteria divide every 10 minutes, and the beaker is completely full of bacteria at 12:00 noon, at what time was the beaker exactly half full of bacteria?

  The answer is clearly "11:50 am." (Now we're supposed to think about natural resources like oil, and oil consumption doubling every 20 years – if in all the time of oil consumption up until now we have consumed half the total reserve of oil, in how many years will we have consumed all the oil? etc. . . .)

  But the question I really want to ask is, "What will the beaker look like at 10 minutes after 12:00 noon?"

  The first thought is that there will be sugar water and bacteria all over the table because the beaker will have overflowed from the next doubling (it was completely full at 12:00 noon) – but in the real world, that can't be right. In fact, there won't be any live bacteria in the beaker at 10 minutes after 12:00 noon, because they all will have died of starvation! The beaker was completely full of bacteria, so there was no sugar water left . . .

In fact, as the bacteria population gets closer to filling the beaker, there will be a downward pressure on population – there will be increasing competition for scarce resources. Therefore, a better model of the bacteria population at a given time $n+1$ will be

$$P_{n+1} = 2P_n(M - P_n)$$

where the 2 is from population growth by doubling, and M is the maximum number of bacteria that can fit in the beaker. Note that this corresponds with the idea that if the beaker is ever completely full of bacteria, in the next time step the population will go to 0 because there will be mass starvation . . .

- Now let's simplify the units – instead of keeping track of the total population P_n, we'll let $x_n = \frac{P_n}{M}$, that is, the proportion of the maximum population (also sometimes called the *carrying capacity*) that we have at time n. We will thus have $0 \le x_n \le 1$. Let's also generalize to other species that might have a net birth rate R other than 2, so that we will have the classic *logistics equation* for population of a single species in a resource limited environment:

$$x_{n+1} = R * x_n * (1 - x_n).$$

Note that this has an x_n^2 term, is thus nonlinear, and can't be "solved" in any straightforward way . . .

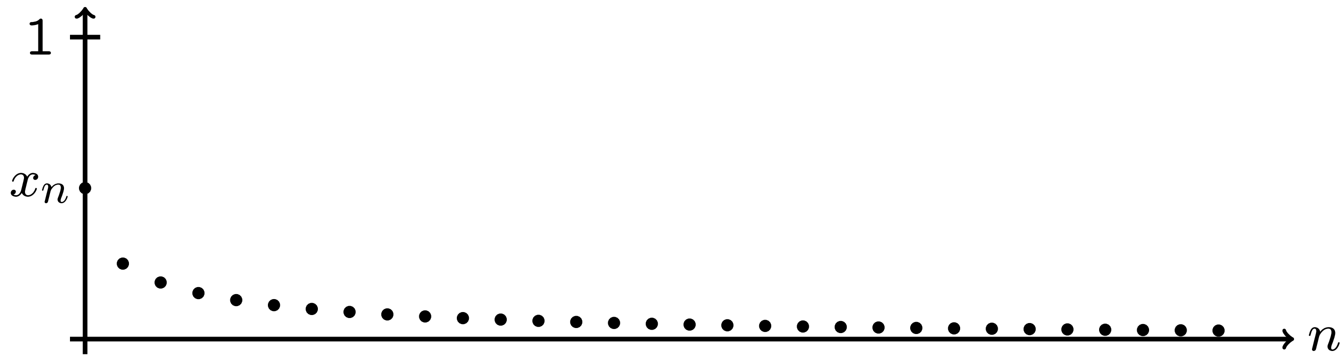
The Logistics Equation (analysis) $\leftarrow$

- The fact that we can't "solve" the logistics equation doesn't mean that we can't study its behavior, or analyze its characteristics. Let's do some work on that. We can start by just observing the system in action.

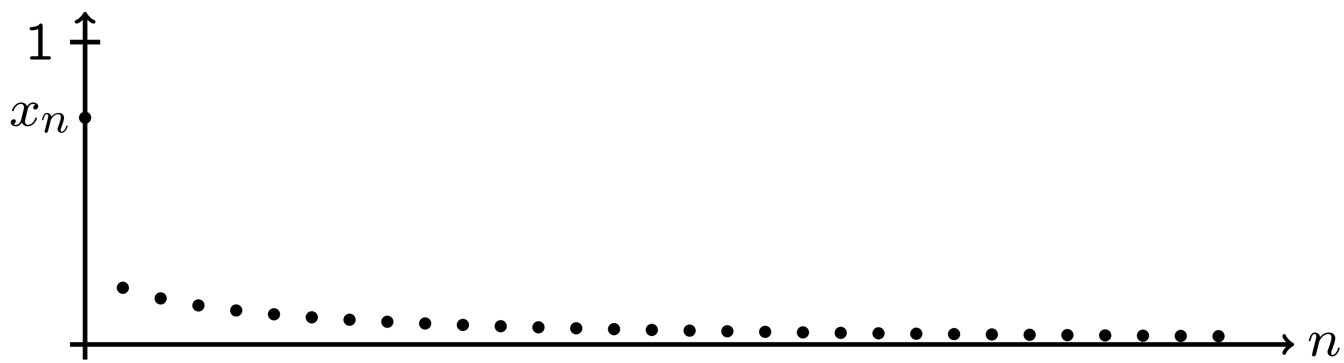
  For our first example, let's look at what happens when $R = 1$. We'll choose a starting value of $x_0 = \frac{1}{2}$. We will have

$$
\begin{aligned}
x_0 &= \frac{1}{2} \\
x_1 &= 1 * \frac{1}{2} * (1 - \frac{1}{2}) = \frac{1}{4} \\
x_2 &= 1 * \frac{1}{4}(1 - \frac{1}{4}) = \frac{1}{4} * \frac{3}{4} = \frac{3}{16} \\
&\ldots
\end{aligned}
$$

Plotting x_n vs. n, we see (first $x_0 = \frac{1}{2}$, then $x_0 = 0.75$):

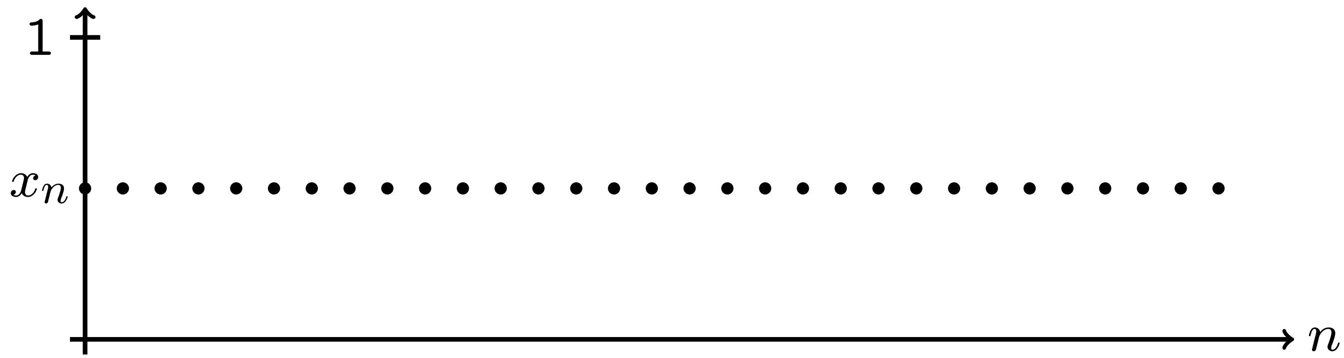


$$x_{n+1} = 1.0 * x_n * (1 - x_n), \quad x_0 = \frac{1}{2}$$

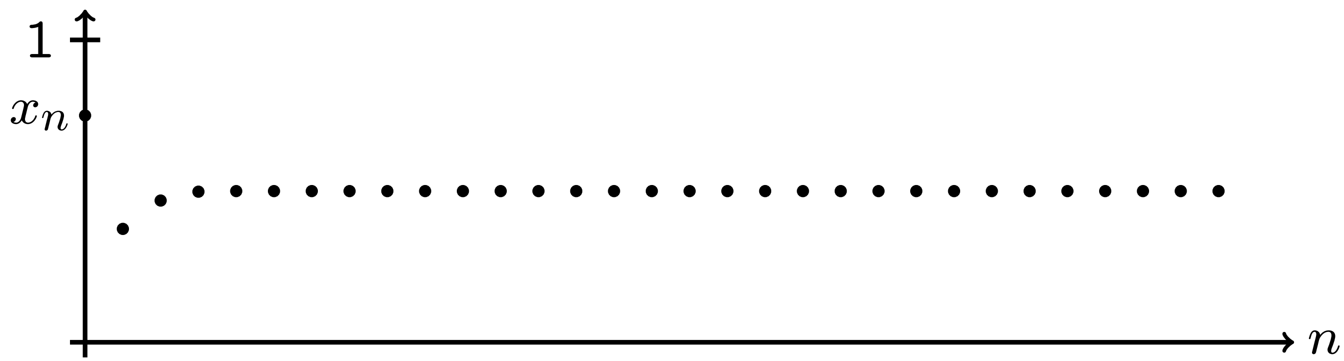


$$x_{n+1} = 1.0 * x_n * (1 - x_n), \quad x_0 = 0.75$$

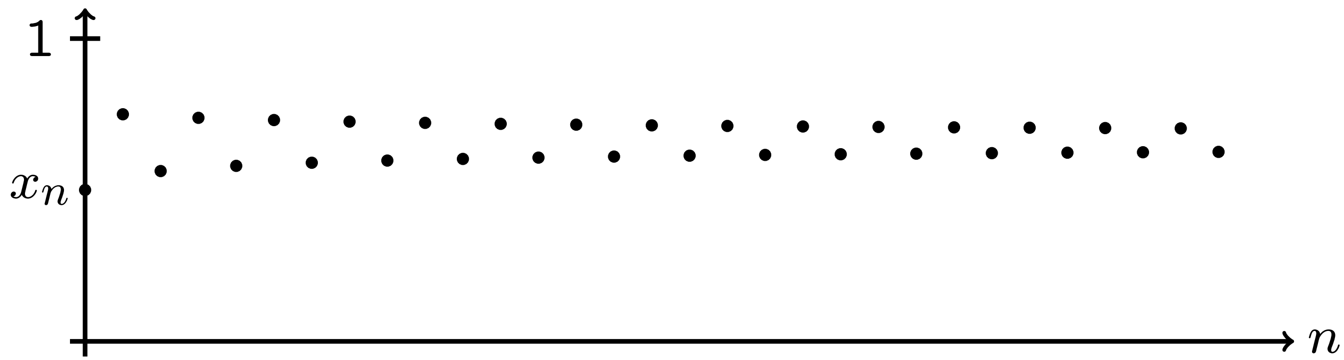
Let's look at some other values of R (and values of x_0):



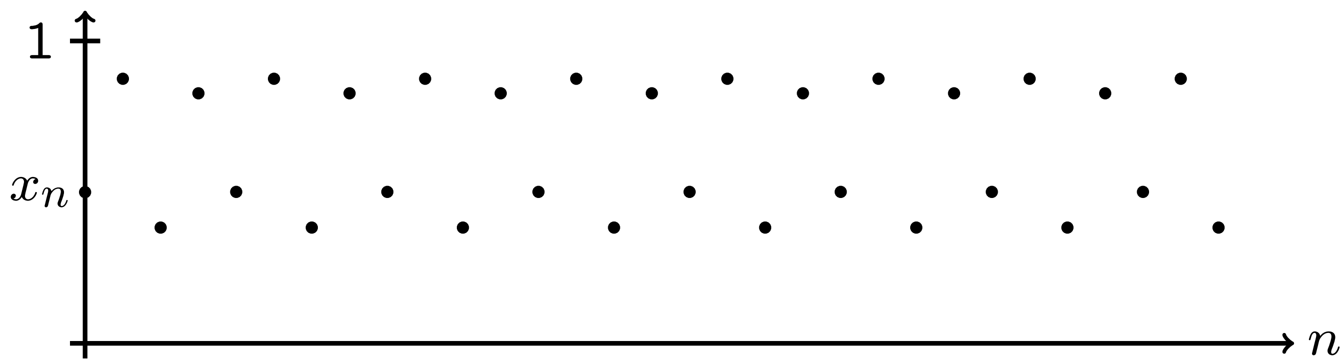
$$x_{n+1} = 2.0 * x_n * (1 - x_n), \quad x_0 = 0.5$$



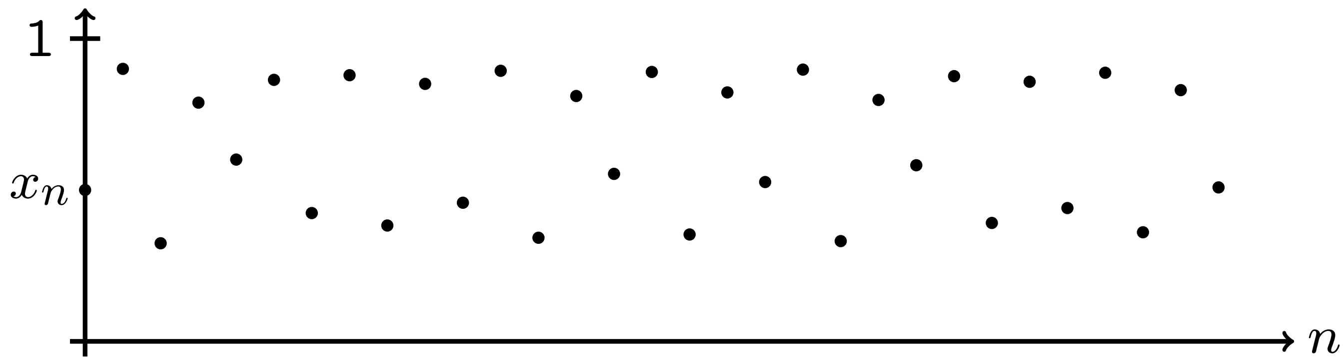
$$x_{n+1} = 2.0 * x_n * (1 - x_n), \quad x_0 = 0.75$$



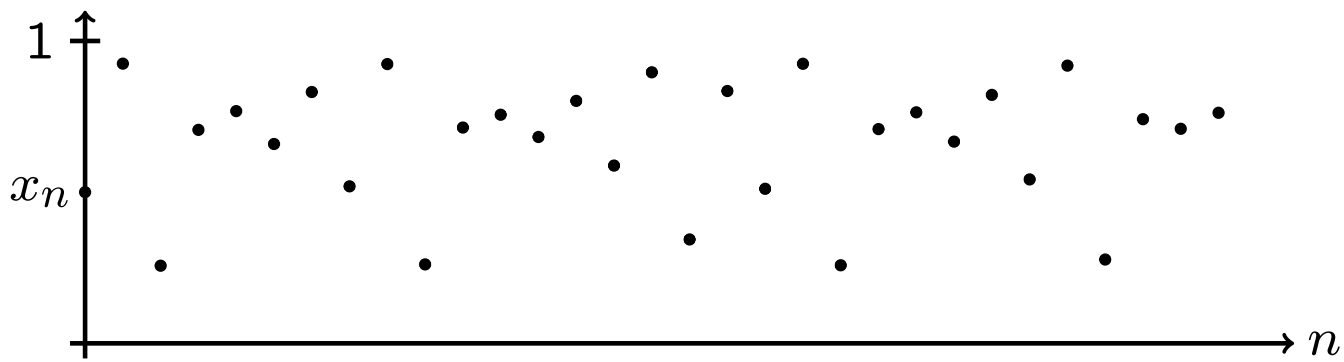
$$x_{n+1} = 3.0 * x_n * (1 - x_n), \quad x_0 = 0.5$$



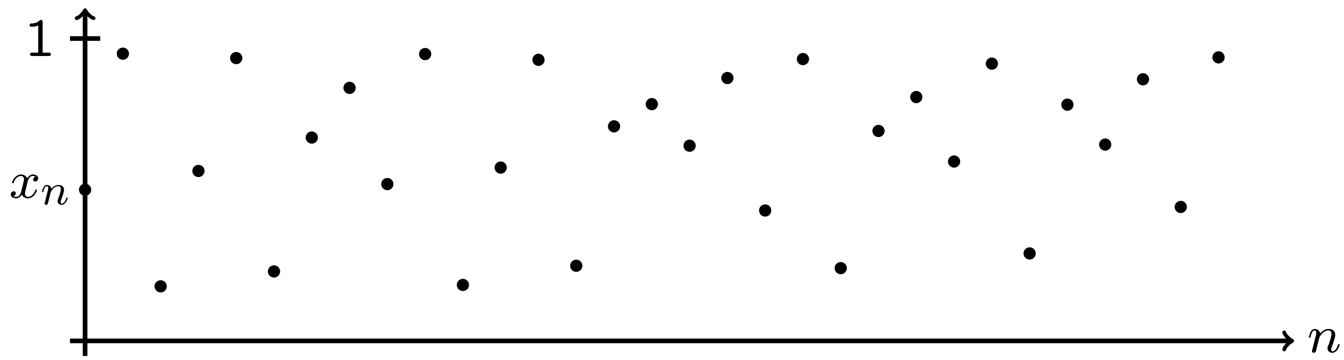
$$x_{n+1} = 3.5 * x_n * (1 - x_n), \quad x_0 = 0.5$$



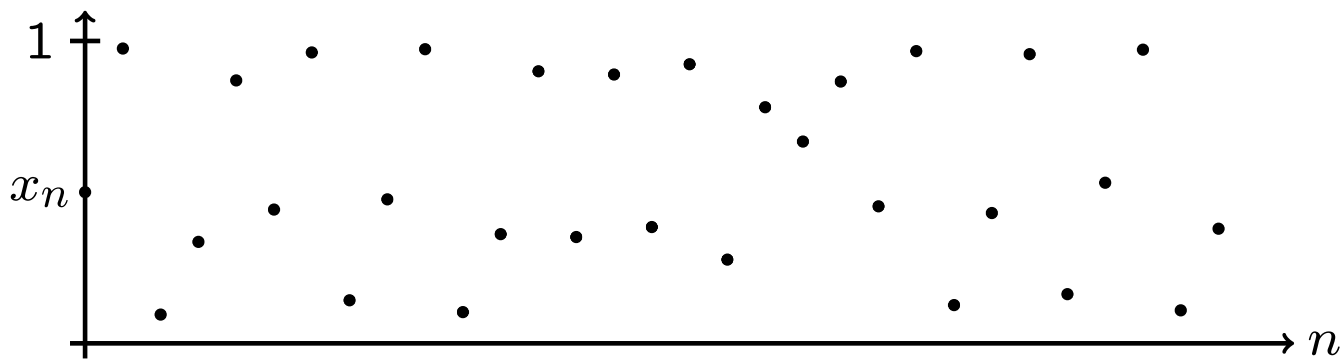
$$x_{n+1} = 3.6 * x_n * (1 - x_n), \quad x_0 = 0.5$$



$$x_{n+1} = 3.7 * x_n * (1 - x_n), \quad x_0 = 0.5$$

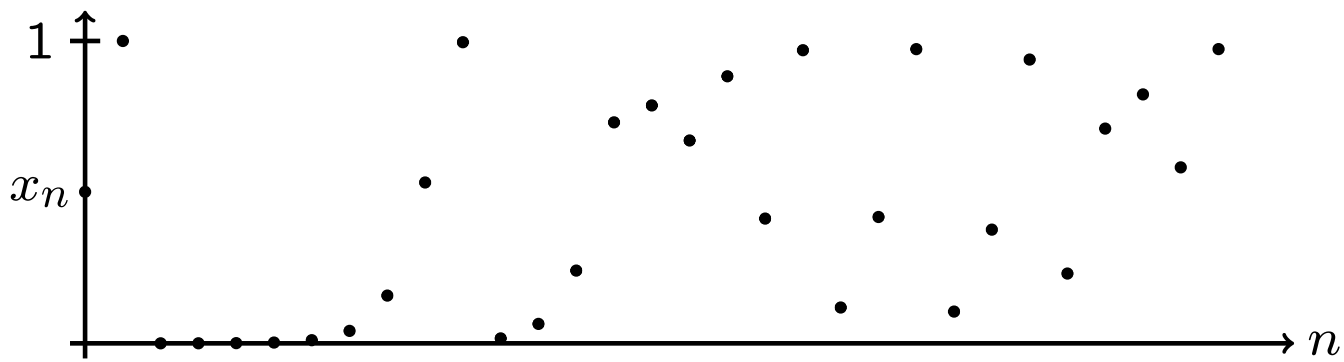


$$x_{n+1} = 3.8 * x_n * (1 - x_n), \quad x_0 = 0.5$$

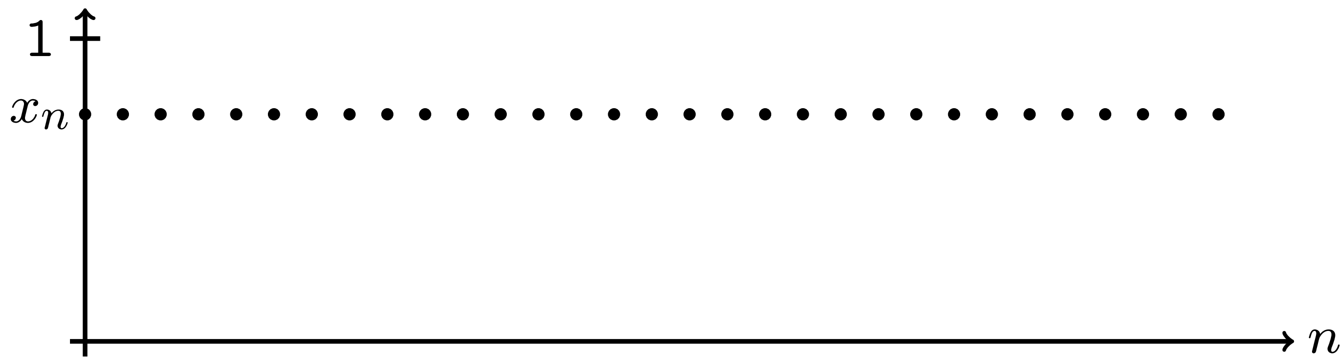


$$x_{n+1} = 3.9 * x_n * (1 - x_n), \quad x_0 = 0.5$$

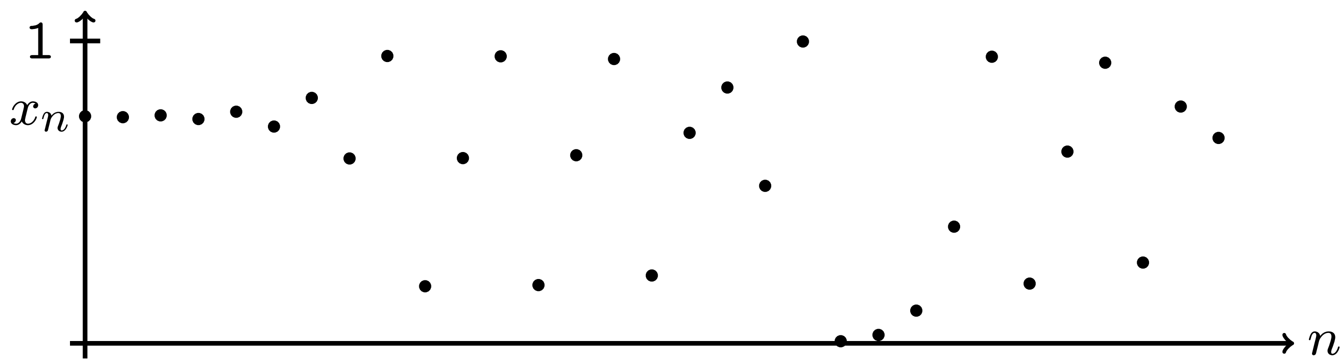
$$x_{n+1} = 4.0 * x_n * (1 - x_n), \quad x_0 = 0.5$$



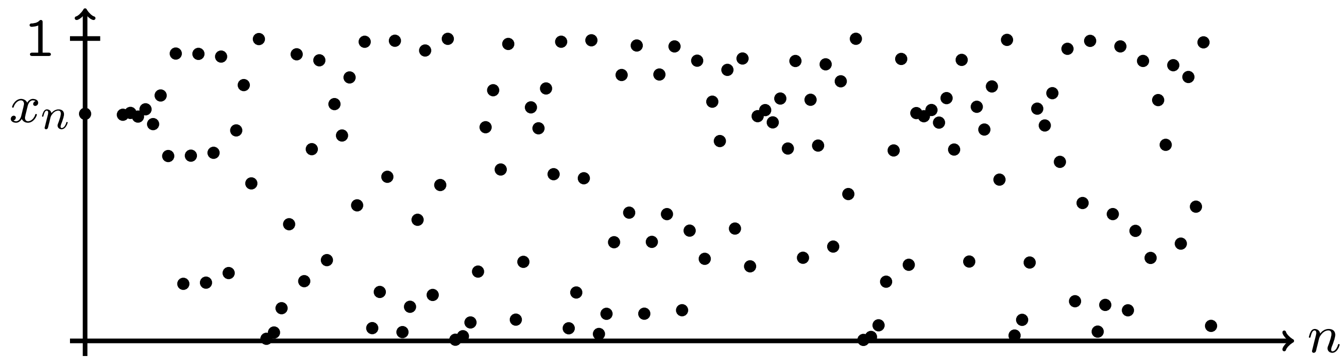
$$x_{n+1} = 4.0 * x_n * (1 - x_n), \quad x_0 = 0.501$$



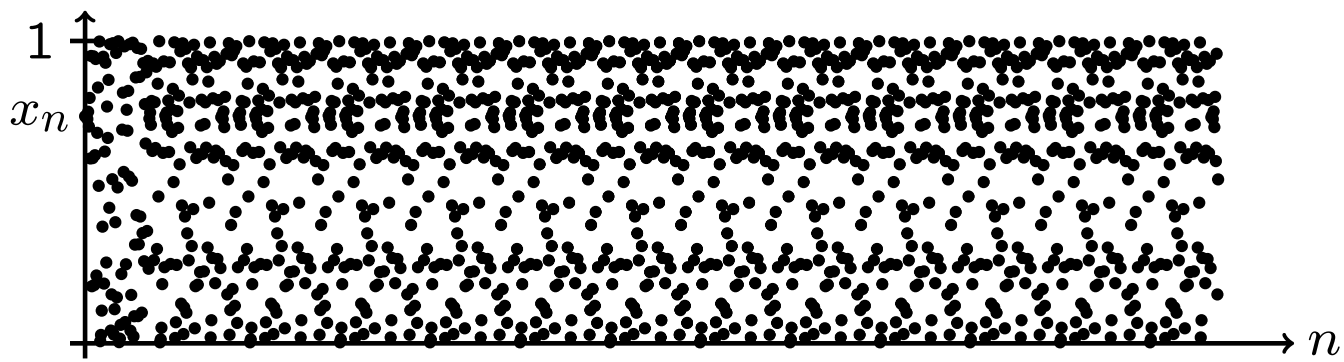
$$x_{n+1} = 4.0 * x_n * (1 - x_n), \quad x_0 = 0.75$$



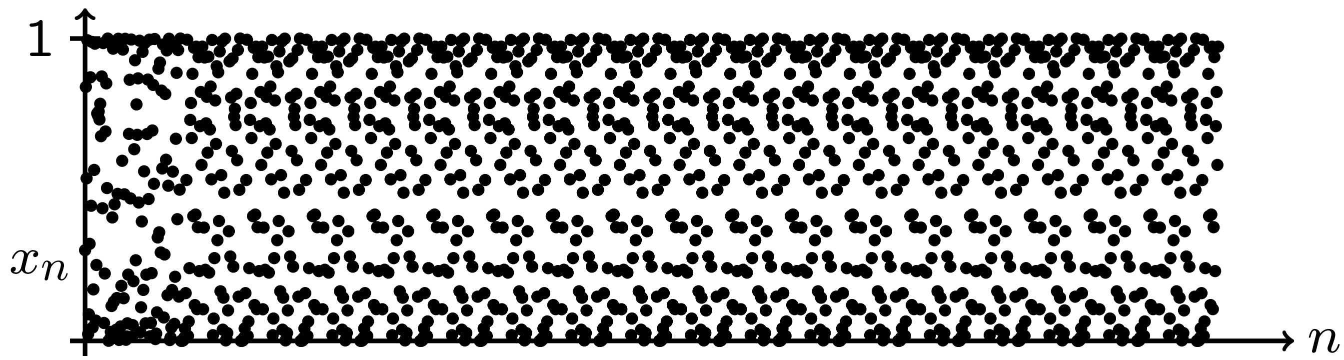
$$x_{n+1} = 4.0 * x_n * (1 - x_n), \quad x_0 = 0.751$$



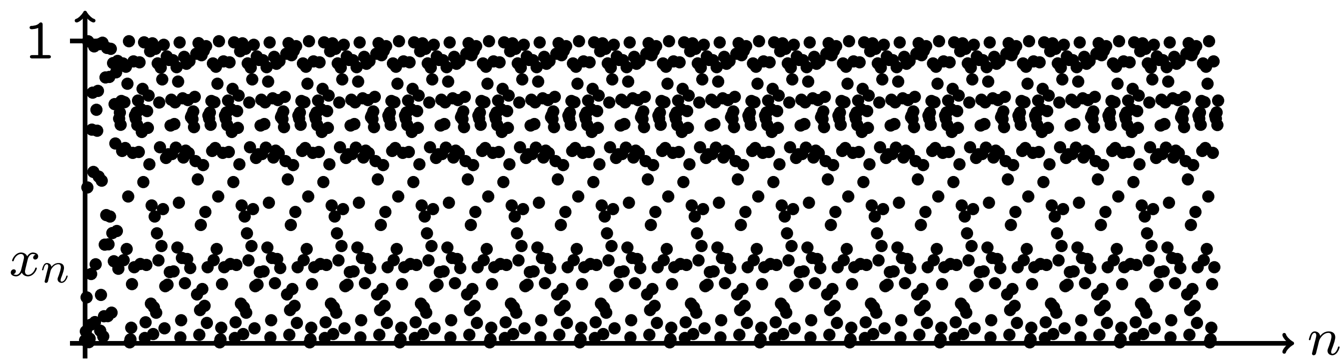
$$x_{n+1} = 4.0 * x_n * (1 - x_n), \quad x_0 = 0.751$$



$$x_{n+1} = 4.0 * x_n * (1 - x_n), \quad x_0 = 0.751$$



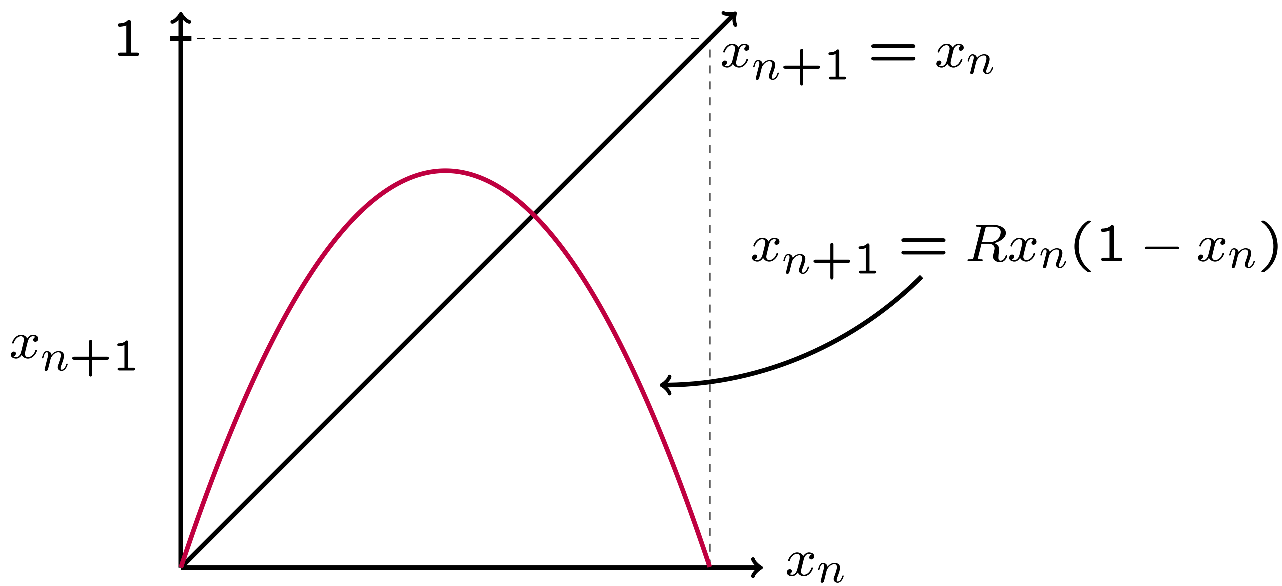
$$x_{n+1} = 4.0 * x_n * (1 - x_n), \quad x_0 = 0.3$$



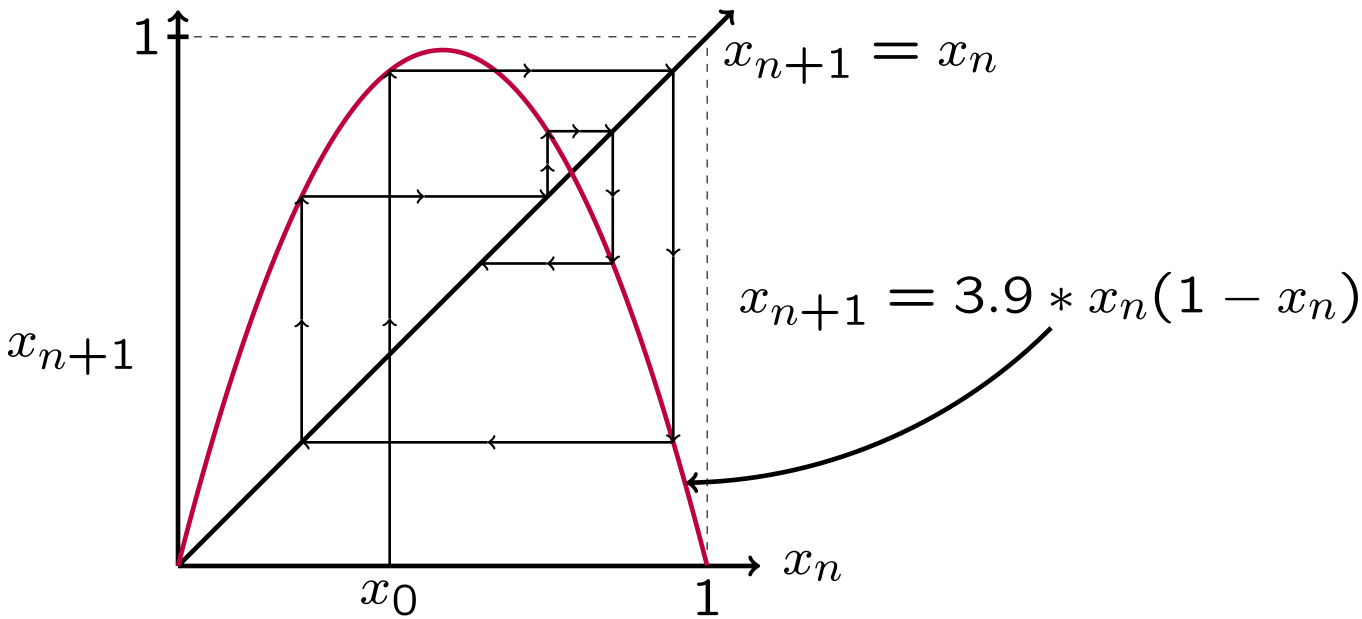
$$x_{n+1} = 4.0 * x_n * (1 - x_n), \quad x_0 = 0.01$$

Cobweb Diagrams $\leftarrow$

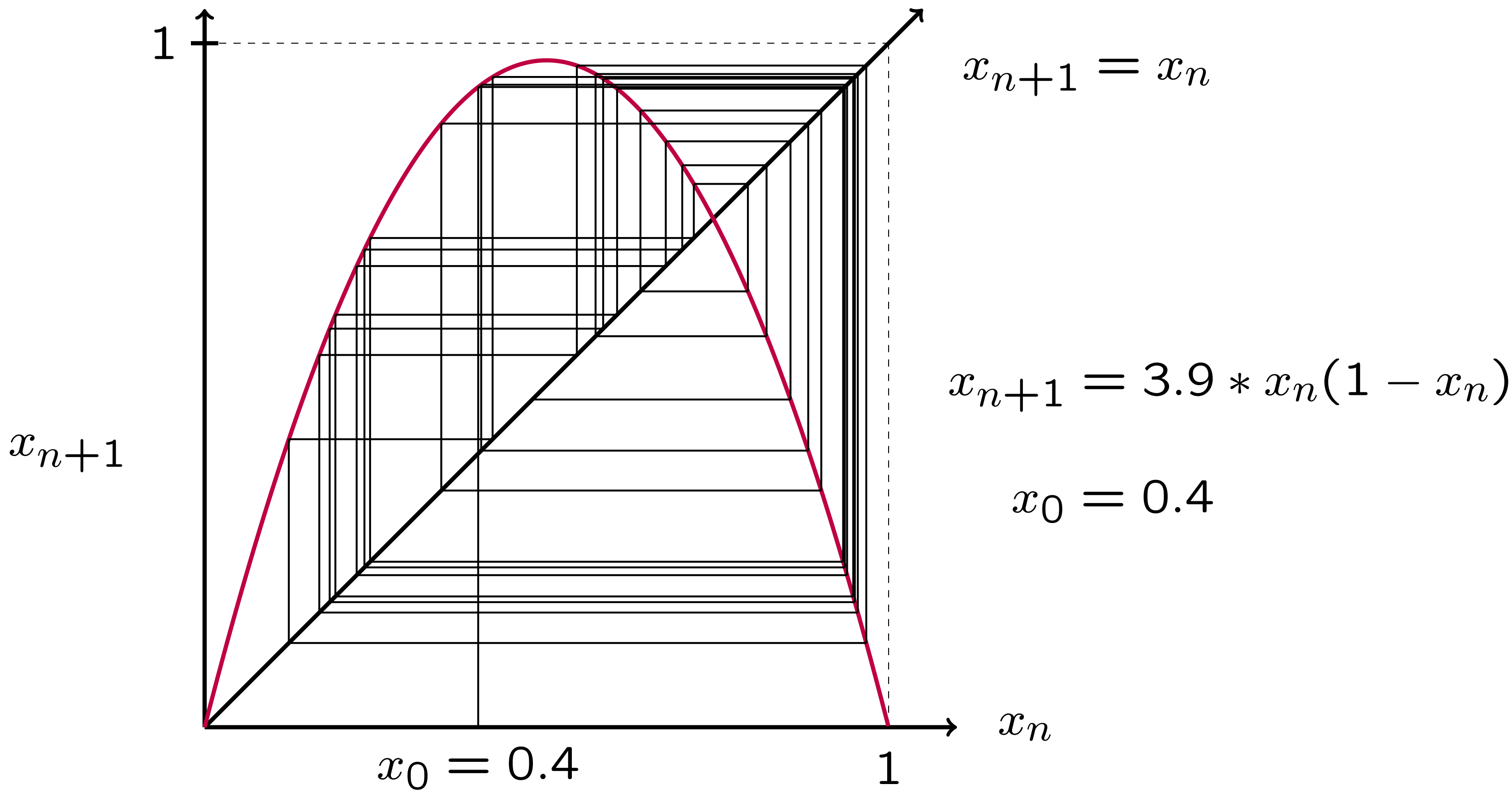
- We can view the system in a somewhat different way. This is a form of *phase space* for the system, where we will plot x_{n+1} against x_n. For any value of $R > 0$, the right hand side is a parabola opening down, with roots at $x_n = 0$ and $x_n = 1$. The maximum value of the parabola occurs at $x_n = \frac{1}{2}$, and the maximum value is $\frac{R}{4}$.



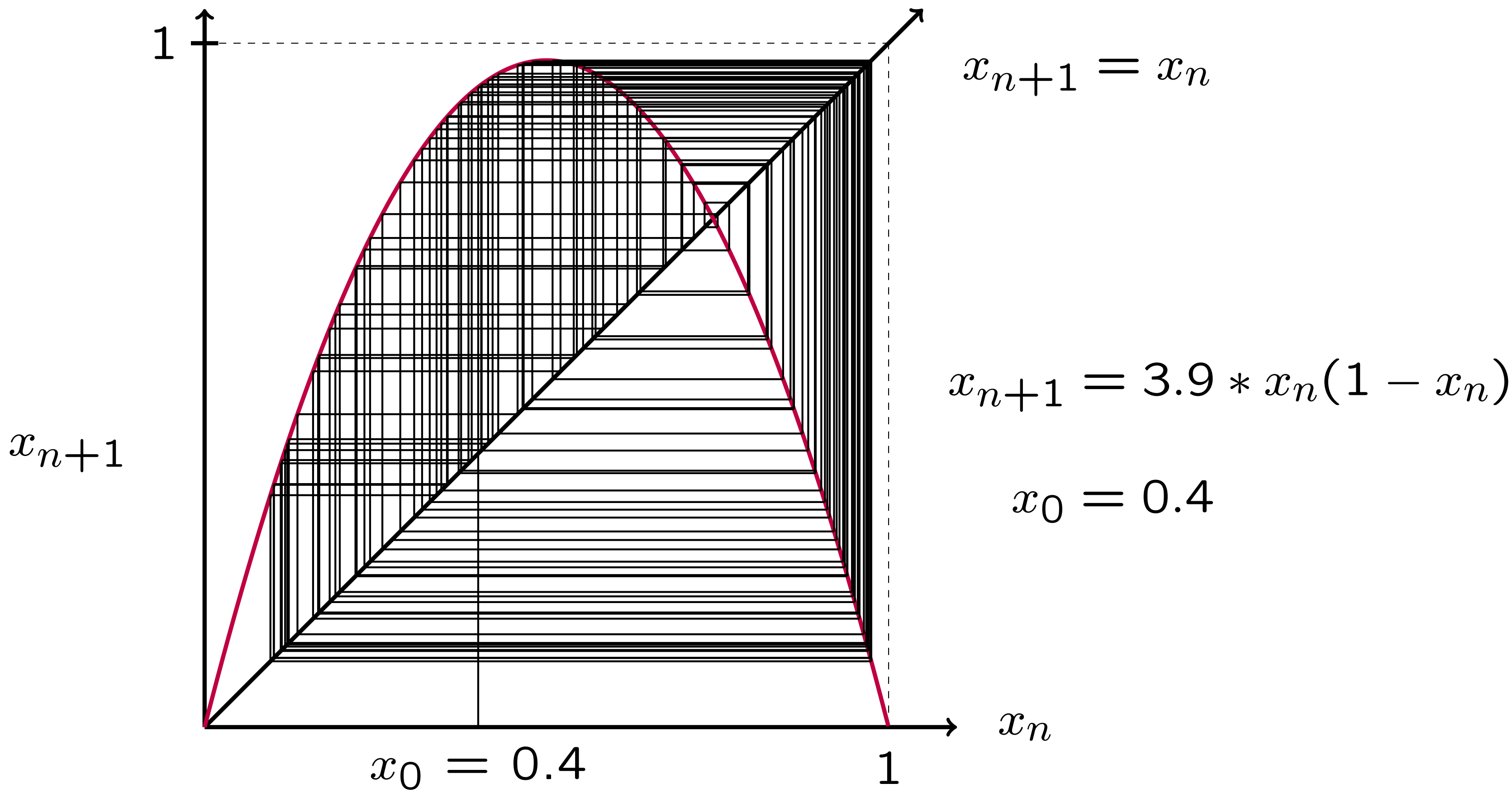
Now we'll follow the trajectory of the system. We'll visualize this by drawing a sequence of lines (often called a cobweb diagram). The coordinates of the endpoints of the lines will be $(x_0, 0) -> (x_0, x_1)$, $(x_0, x_1) -> (x_1, x_1)$, $(x_1, x_1) -> (x_1, x_2)$, $(x_1, x_2) -> (x_2, x_2)$, etc.



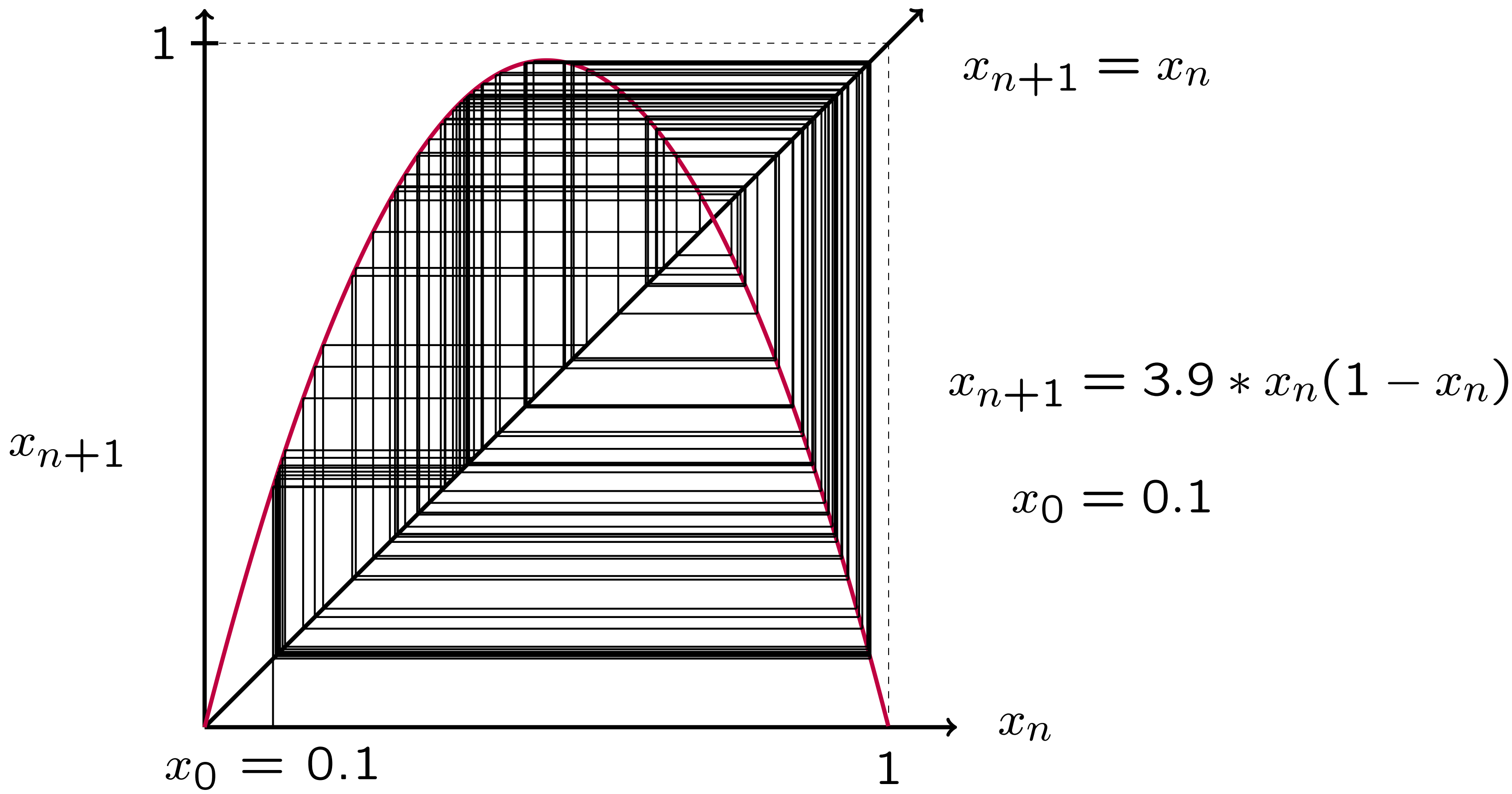
Let's let that same example run for a while:



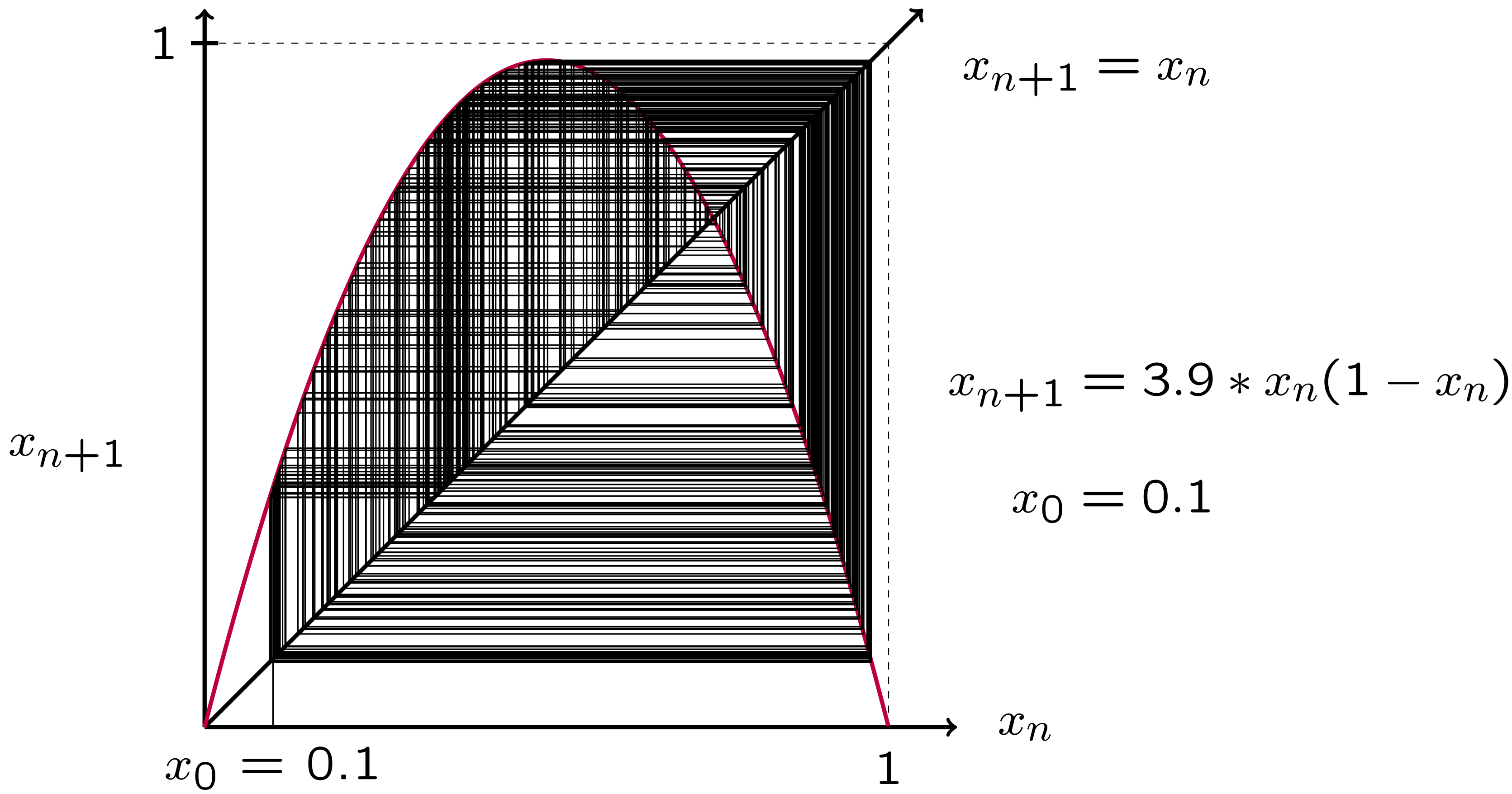
And even longer:



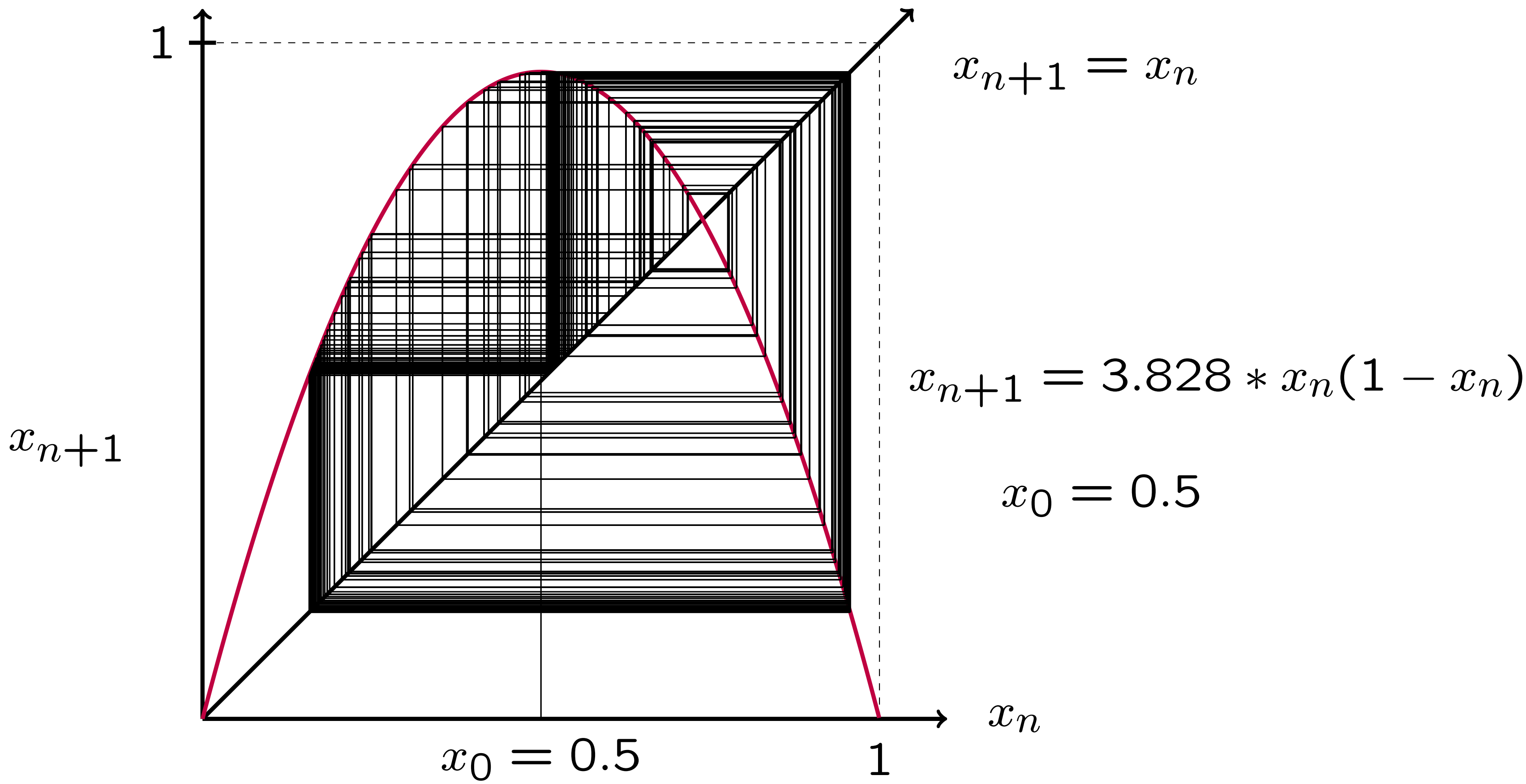
With a different starting value:



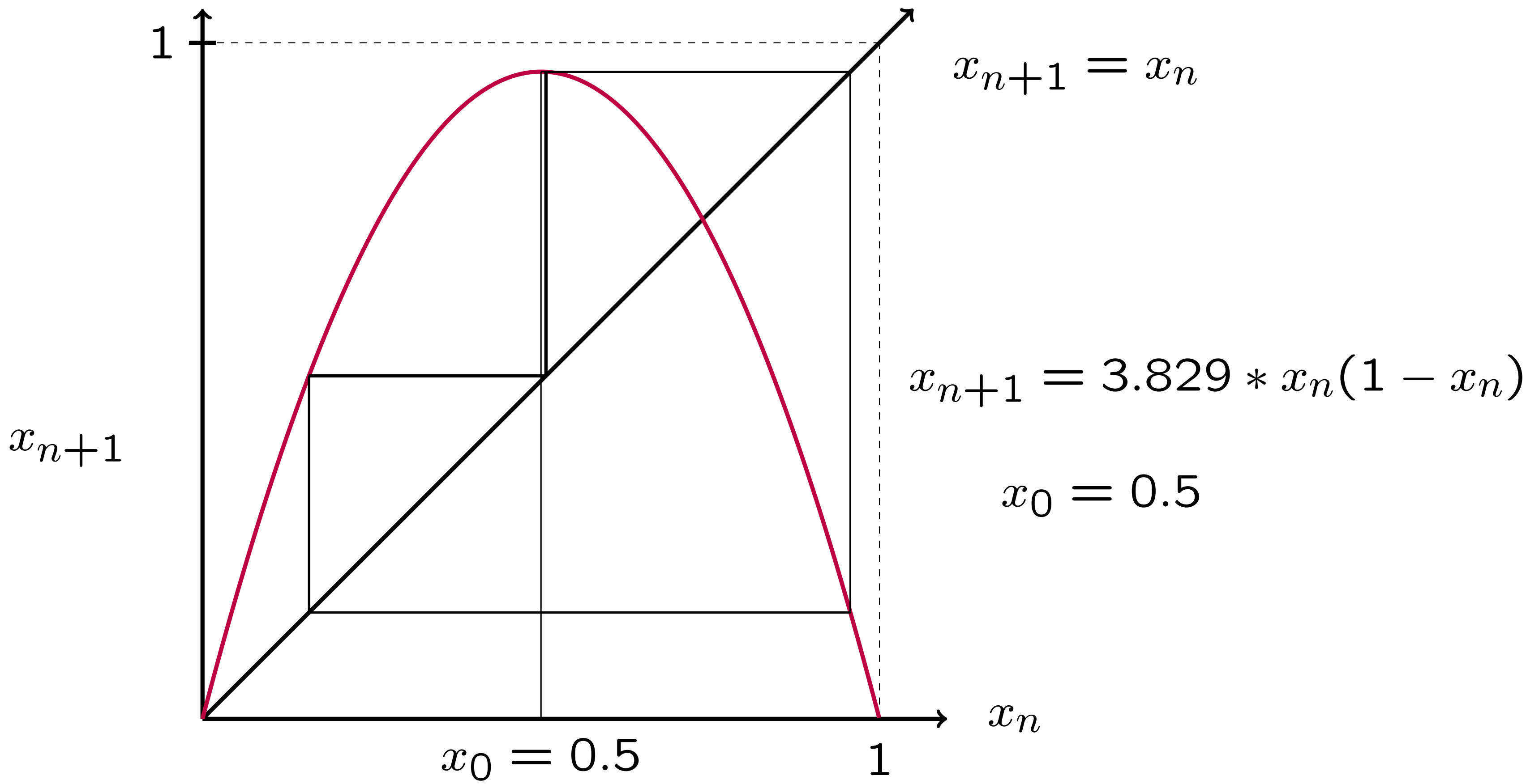
And even longer:



For other values of R (320 iterations):

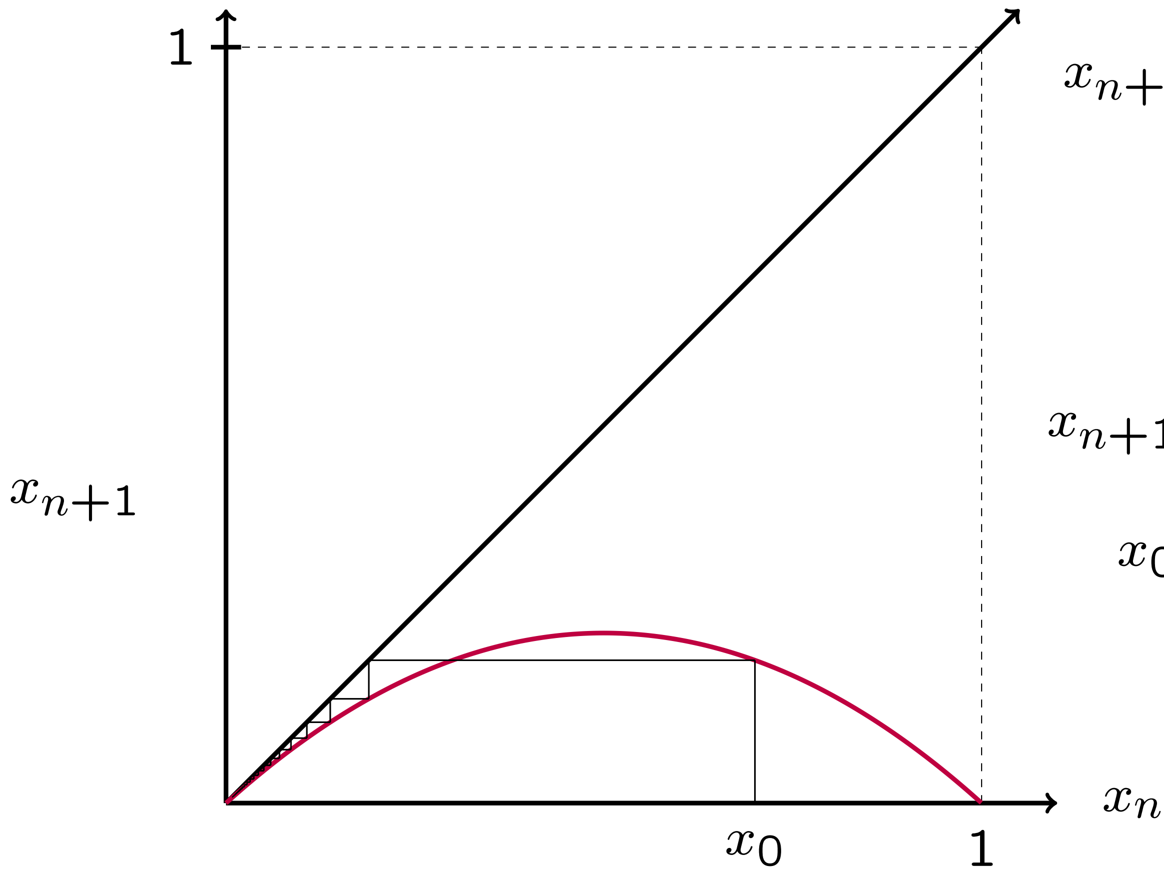


And slight changes (also 320 iterations):



- Now we'll be a little more systematic in our analysis. We'll start with small values of R, where the dynamics are relatively simple. Let's look at the slope of the tangent line to the parabola at 0. We are working with the parabola $P(x) = R * x * (1 - x)$, or $P(x) = -Rx^2 + Rx$, so the derivative is $P'(x) = -2Rx + R$. We thus have $P'(0) = R$, and hence when $R \leq 1$, the entire parabola (for $x > 0$) lies below the line $x_{n+1} = x_n$. In this case, the dynamics just die away to zero (i.e., $\lim_{n\to\infty}(x_n) = 0$). In particular, 0 is an *attracting fixed point* of the system. (Note that for any value of R, 0 is a fixed point of the system.)

  For $R \leq 1$, the dynamics don't depend in any significant way on the starting value – almost immediately, the system begins decaying to 0, and continues directly there.



$$x_{n+1} = x_n$$

$$x_{n+1} = 0.9 * x_n(1 - x_n)$$

$$x_0 = 0.7$$

For $R > 1$, we have a new feature – the parabola crosses the line $x_{n+1} = x_n$ at some point $x > 0$. In particular, we solve for that crossing point by setting $x_n = x_{n+1}$ – in other words, we have

$$x_n = R * x_n * (1 - x_n),$$

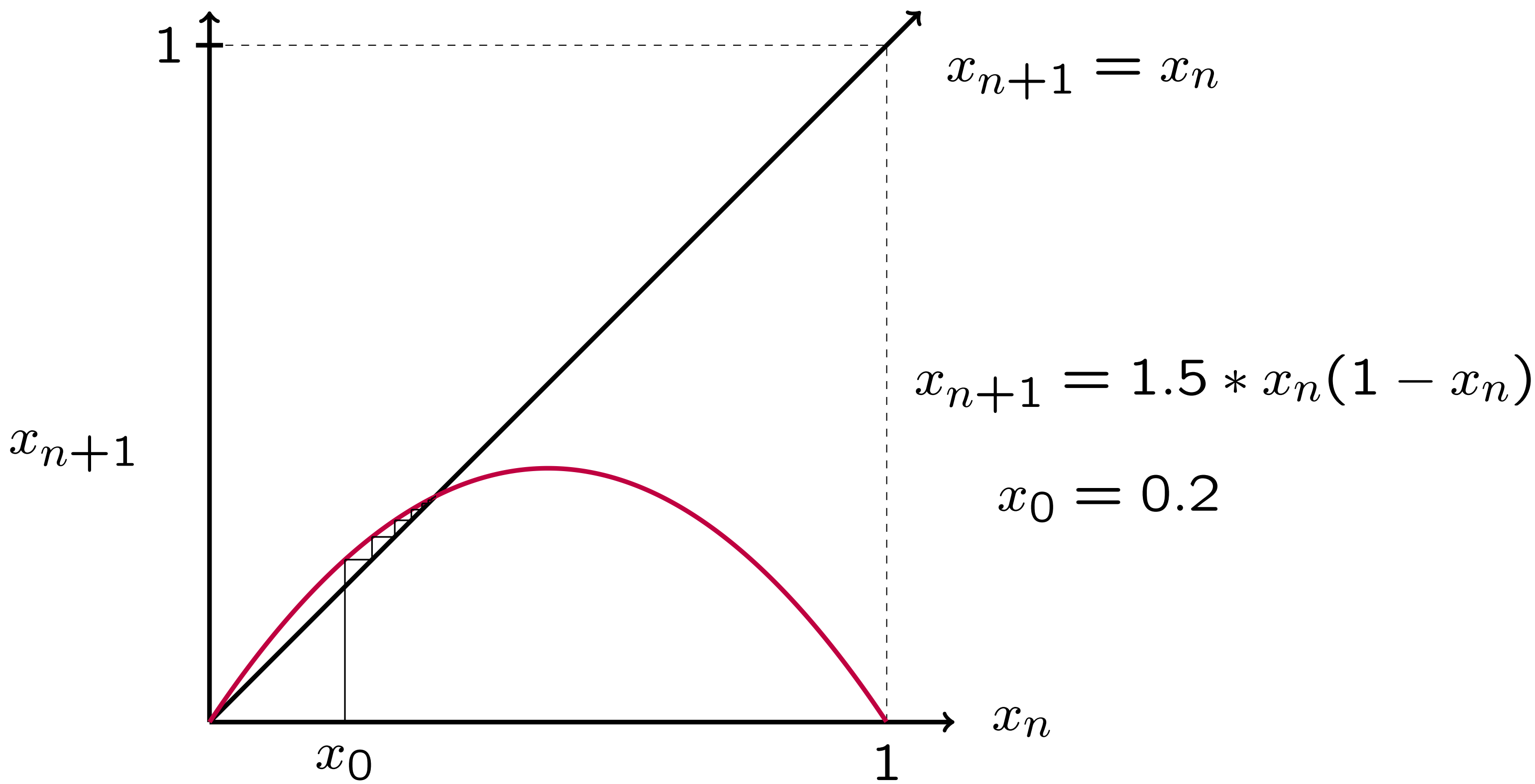
and so

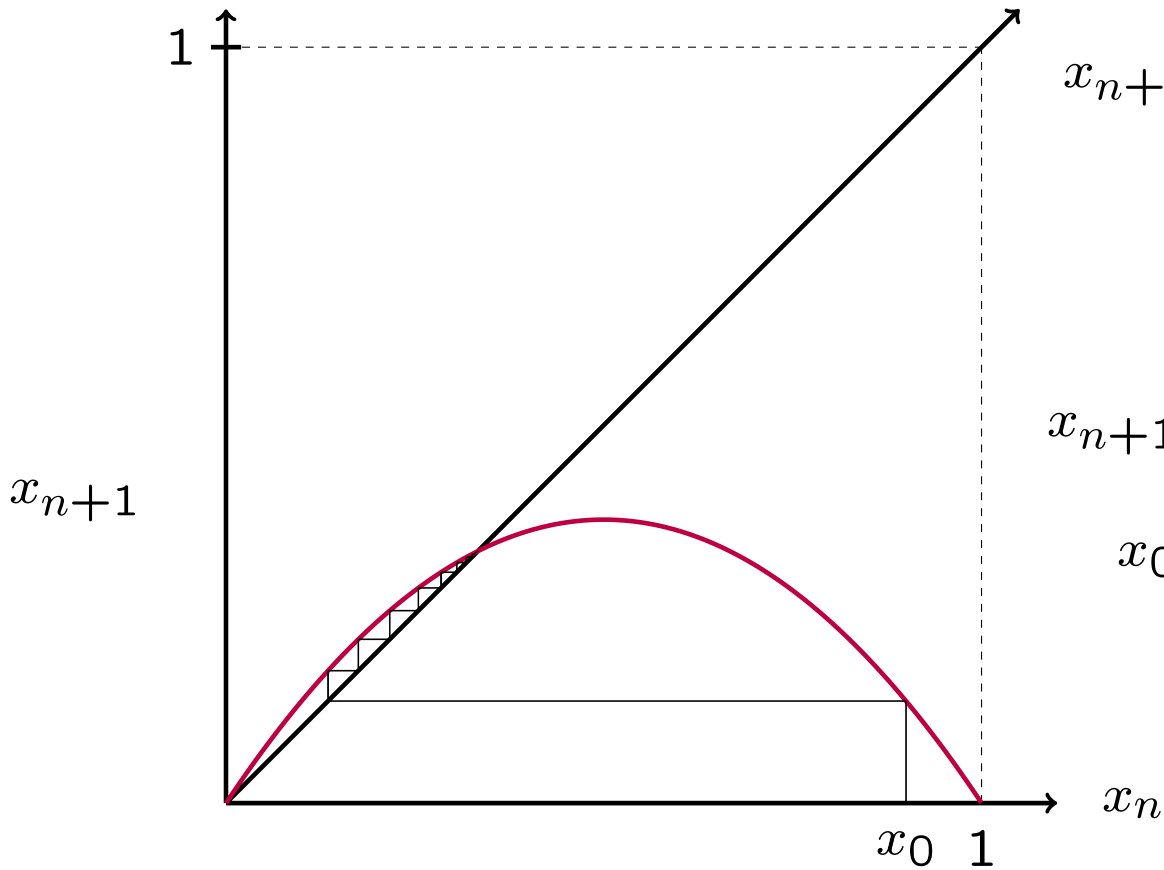
$$\begin{aligned} x_n - Rx_n(1 - x_n) &= 0 \\ x_n - Rx_n + Rx_n^2 &= 0 \\ Rx_n^2 + (1 - R)x_n &= 0 \\ x_n(Rx_n + 1 - R) &= 0, \end{aligned}$$

and hence the crossing point occurs at

$$x_n = \frac{R - 1}{R}.$$

For all values of $R > 1$, this crossing point is a fixed point of the system. For values $1 < R < 3$, this is an attracting fixed point for any starting value x_0 (except $x_0 = 0$ or 1, and 0 is now a *repelling* fixed point):

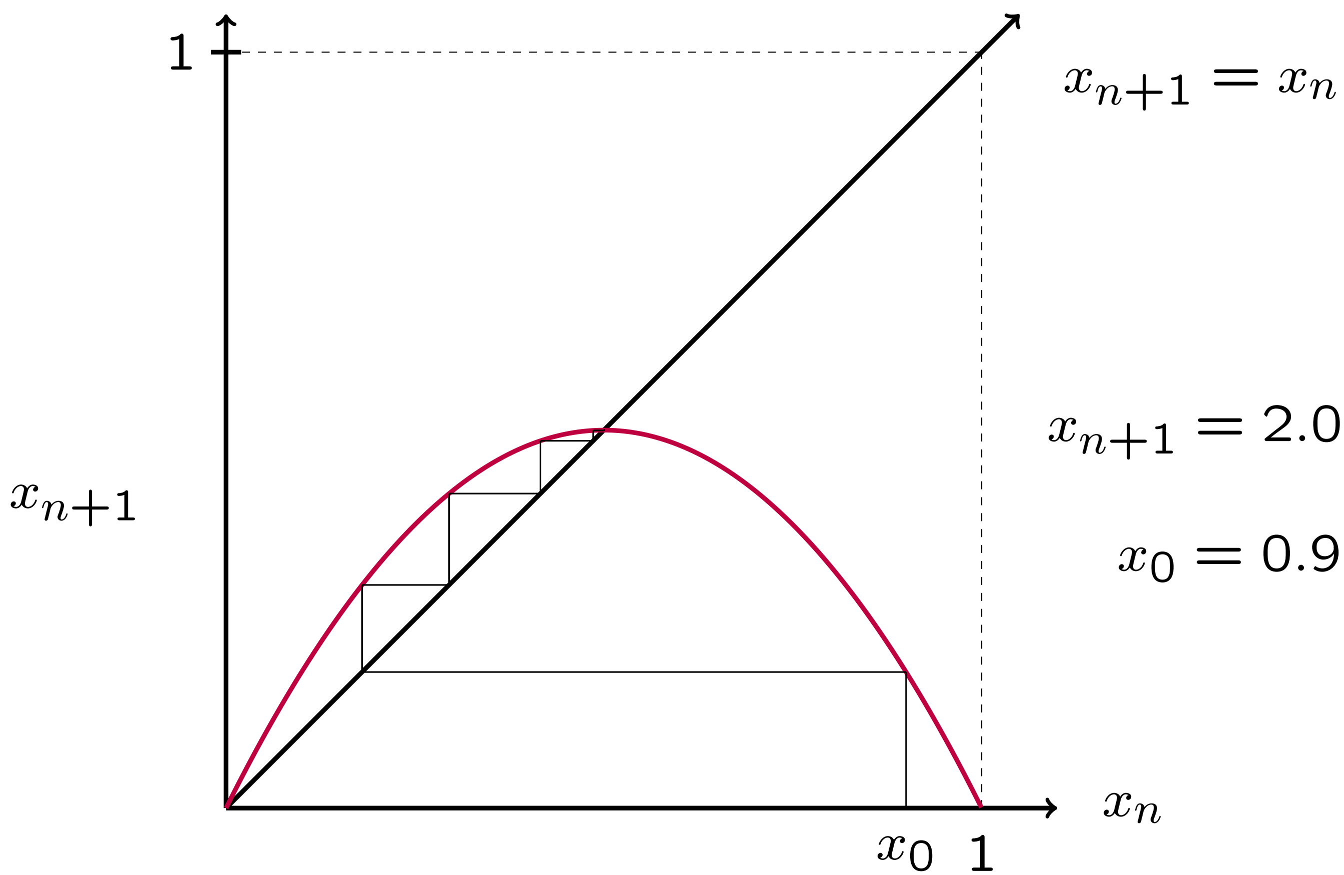




$x_{n+1} = x_n$

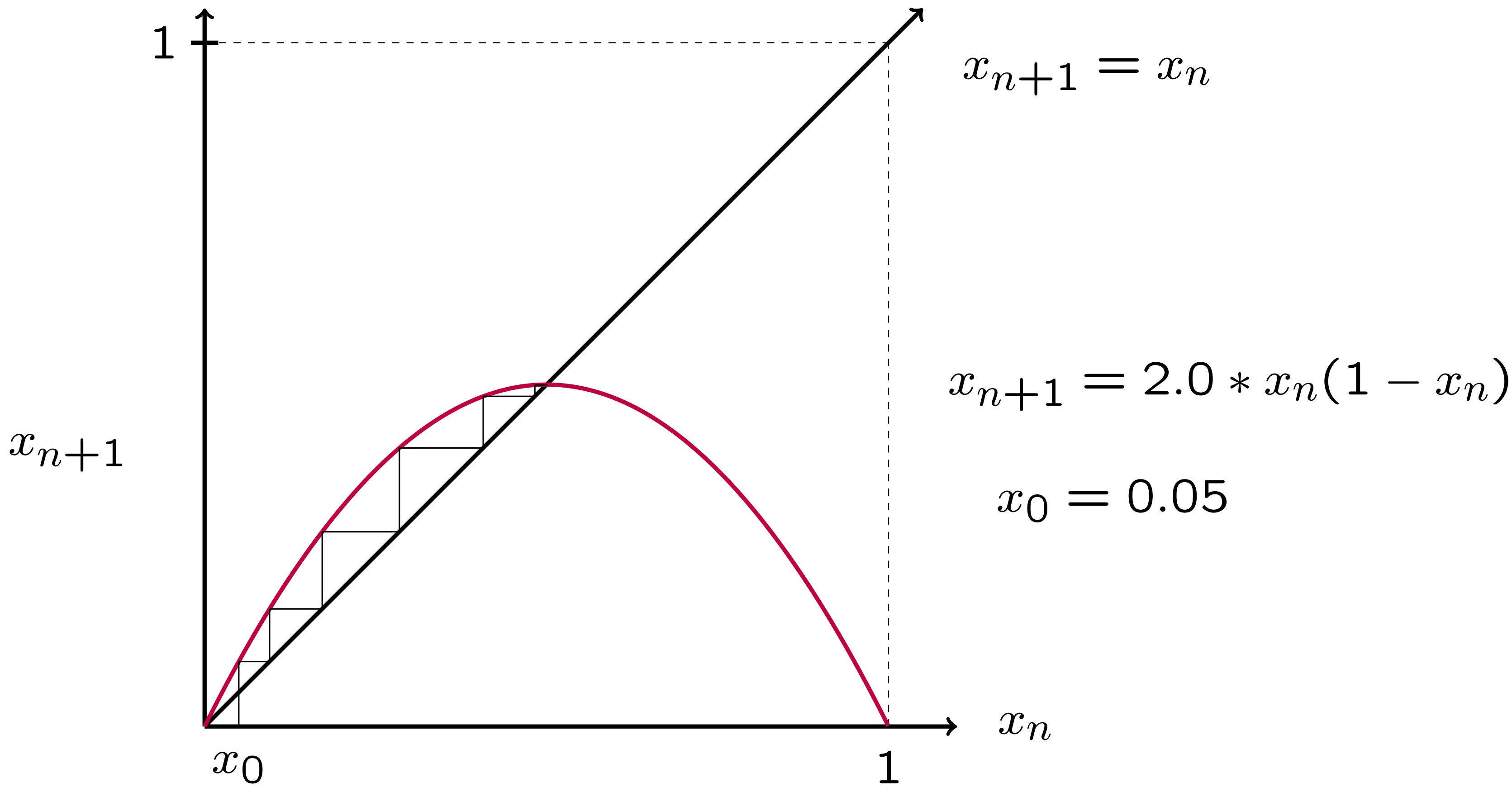
$x_{n+1} = 1.5 * x_n(1 - x_n)$

$x_0 = 0.9$

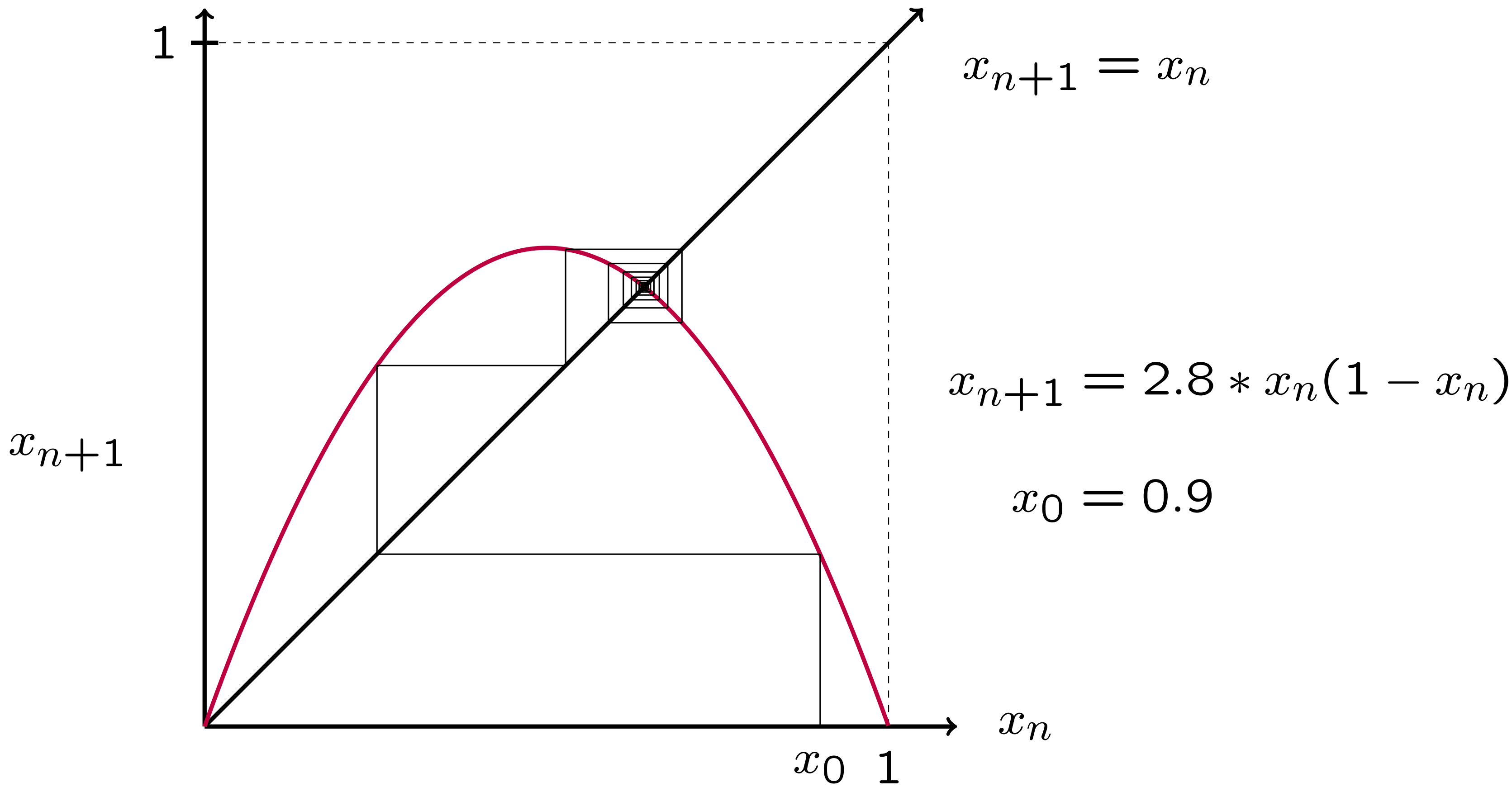


$$x_{n+1} = 2.0 * x_n(1 - x_n)$$

$$x_0 = 0.9$$

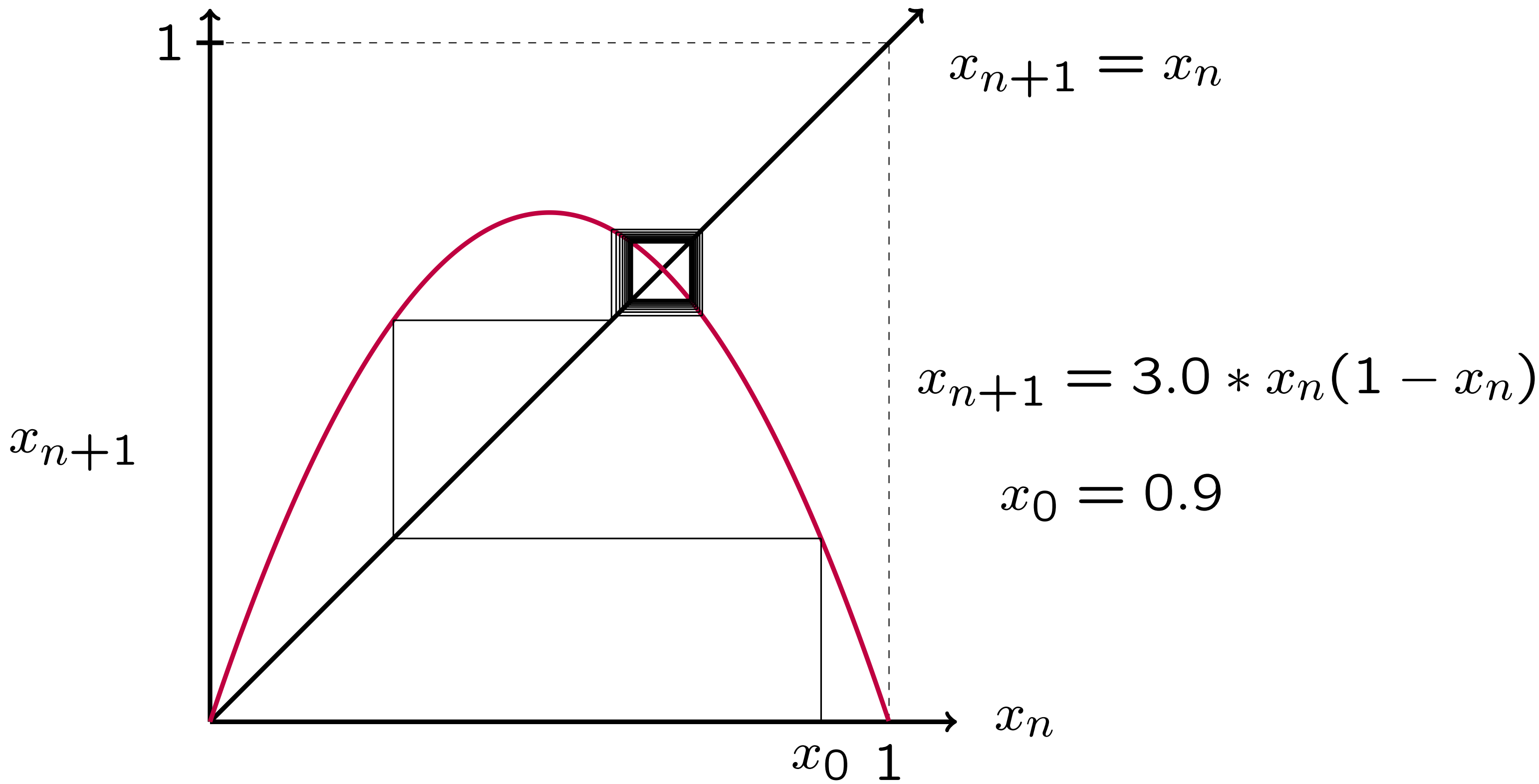


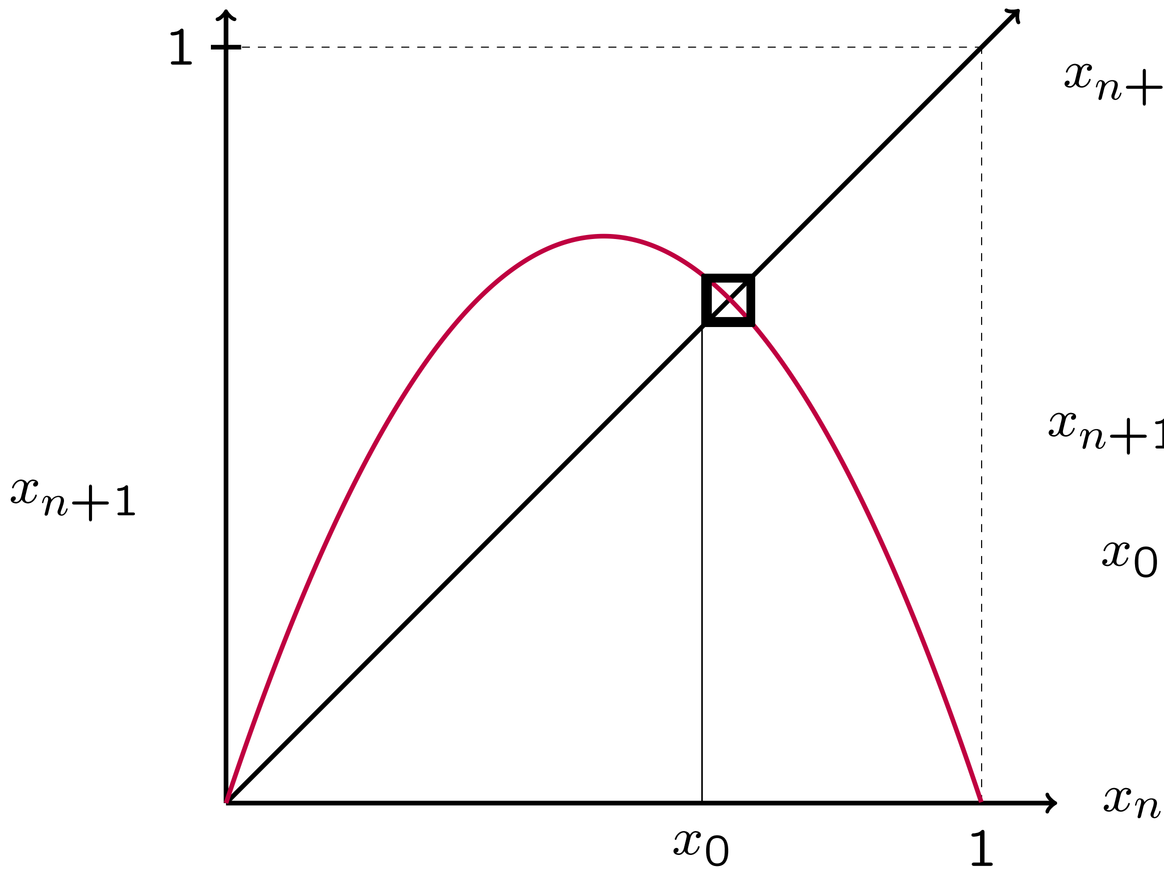
1
$x_{n+1} = x_n$
$x_{n+1} = 2.0 * x_n(1 - x_n)$
$x_0 = 0.05$
x_{n+1}
x_n
x_0
1



Note that for $2 < R < 3$, the parabola is coming back down at the fixed point $\frac{R-1}{R}$, and so the trajectory spirals in . . .

When $R = 3$, we have a new phenomenon. We still have a fixed point at $\frac{R-1}{R}$, but it is no longer and attracting fixed point:

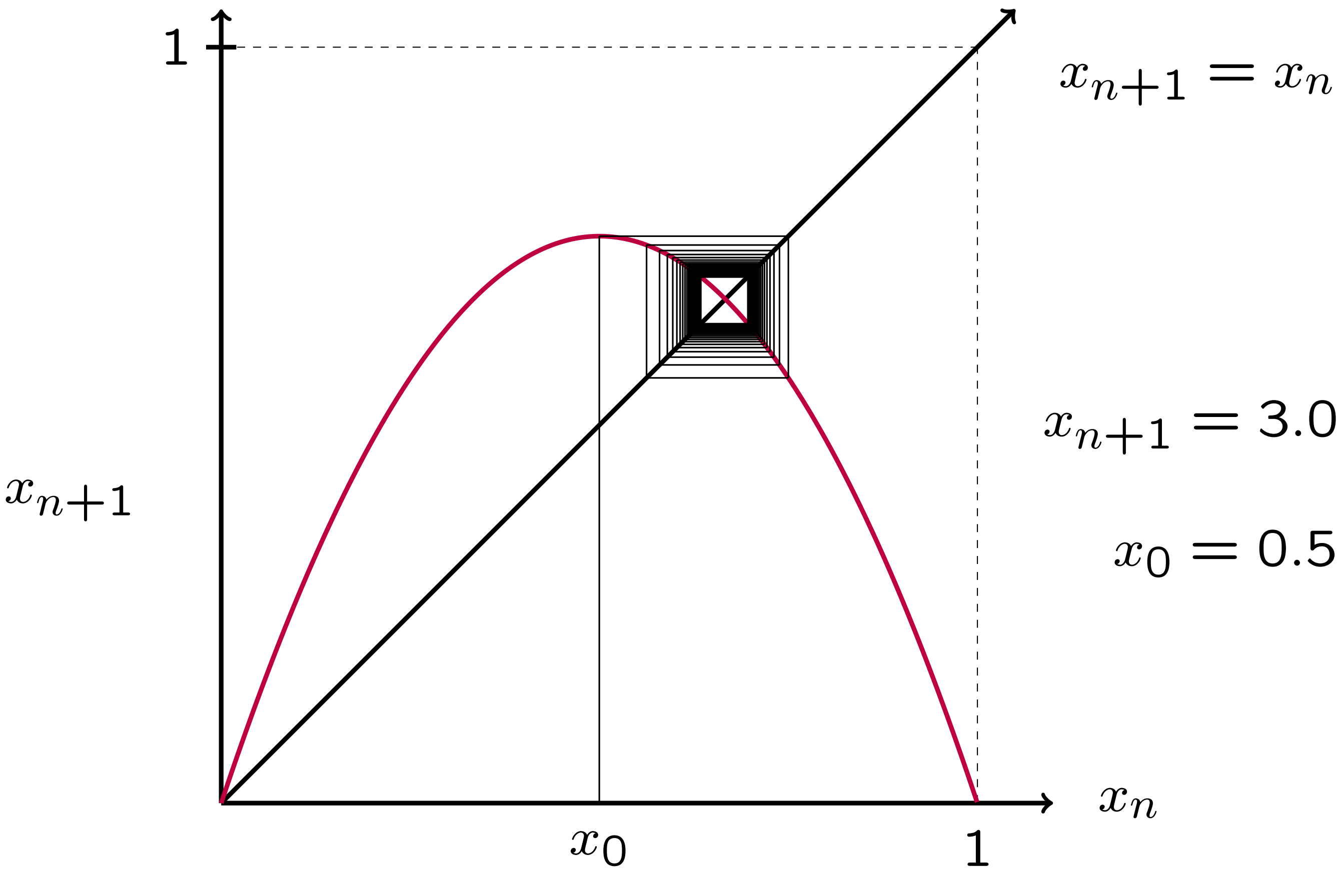




$x_{n+1} = x_n$

$x_{n+1} = 3.0 * x_n(1 - x_n)$

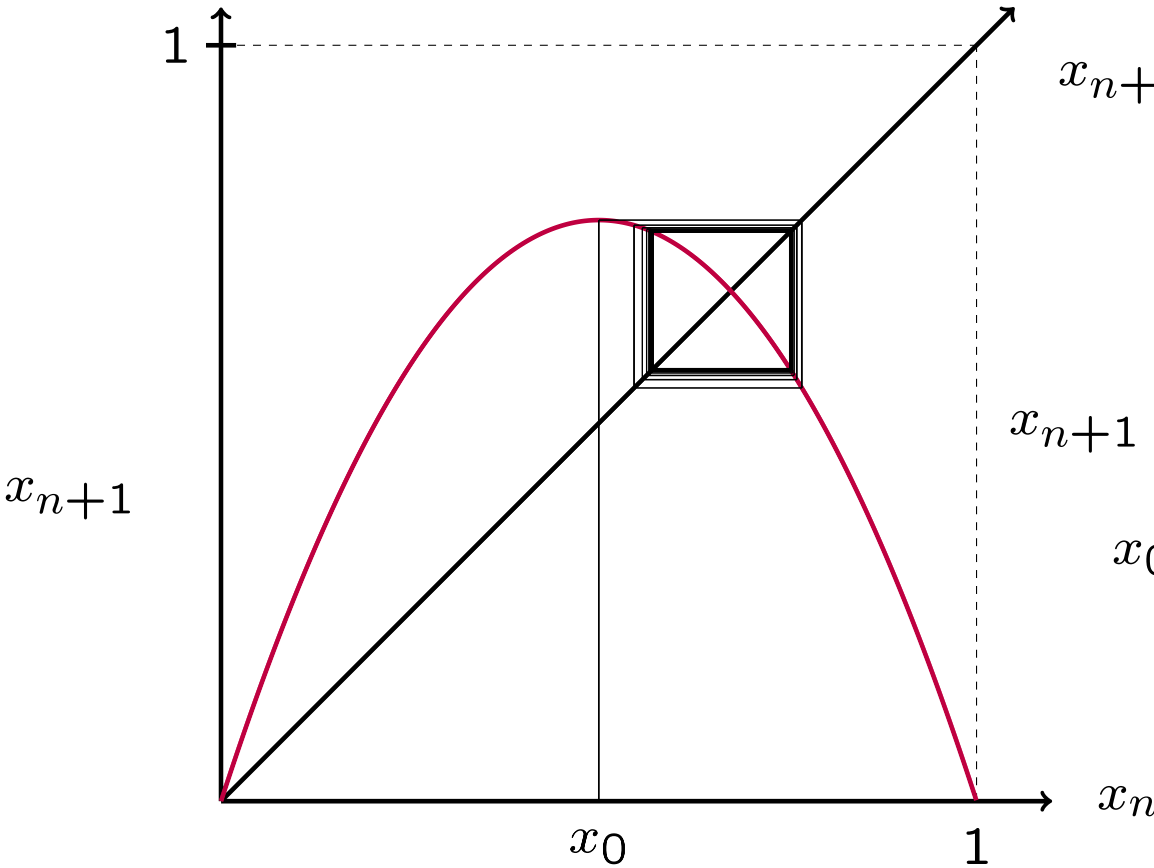
$x_0 = 0.63$



$$x_{n+1} = 3.0 * x_n(1 - x_n)$$

$$x_0 = 0.5$$

For R slightly bigger than 3, we have new behavior:

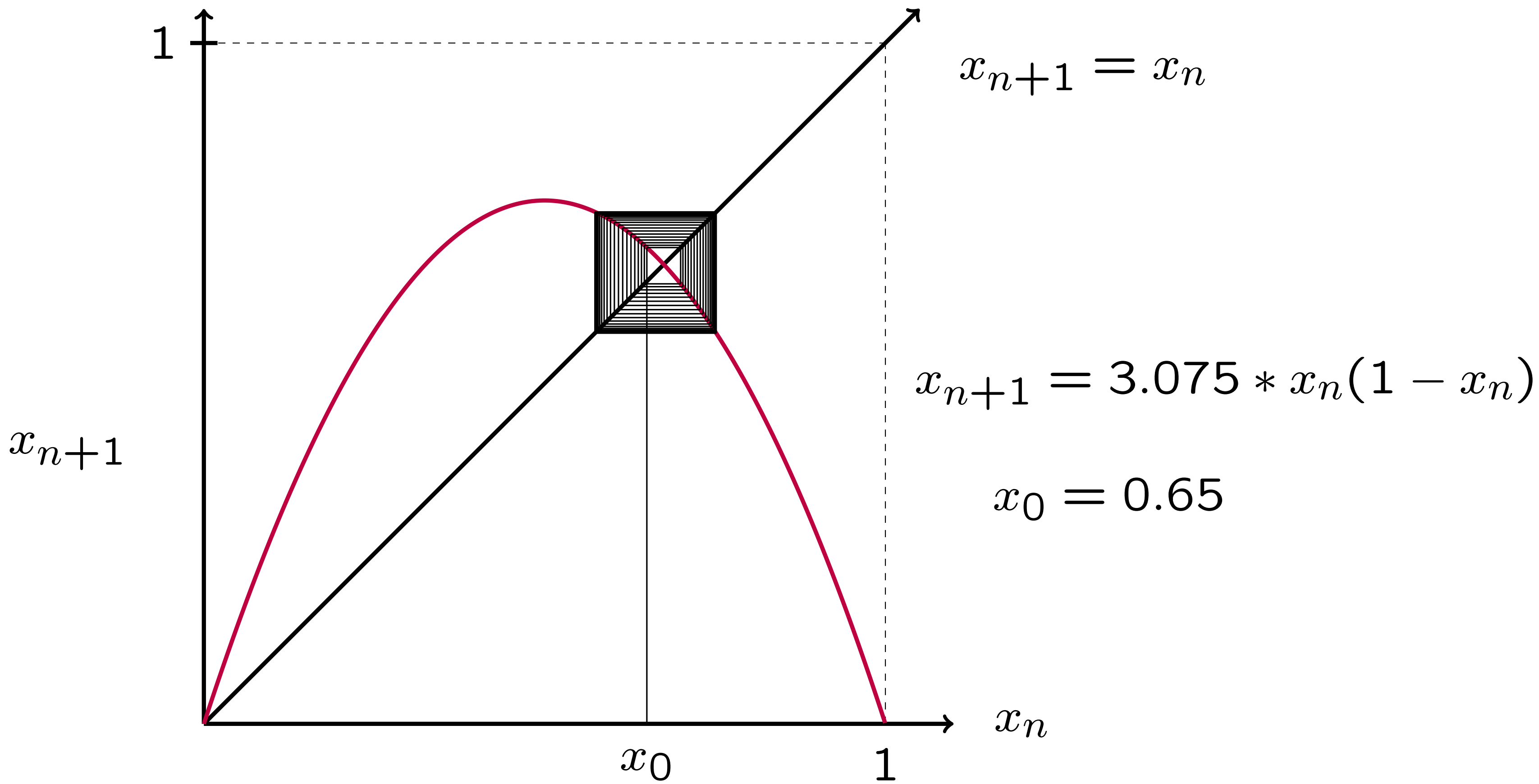


$$x_{n+1} = x_n$$

$$x_{n+1} = 3.075 * x_n(1 - x_n)$$

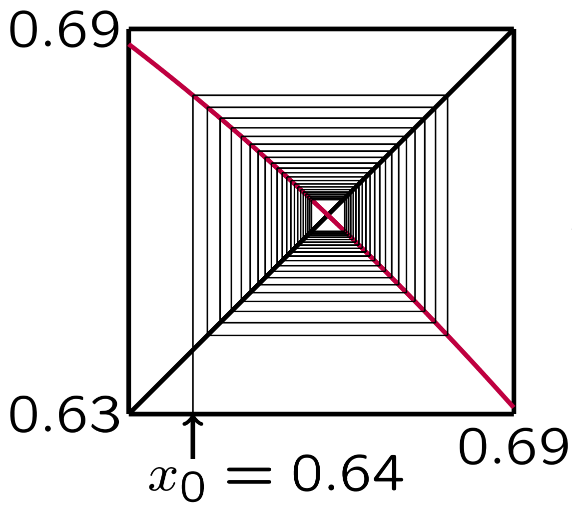
$$x_0 = 0.5$$

Or, starting closer to the fixed point:

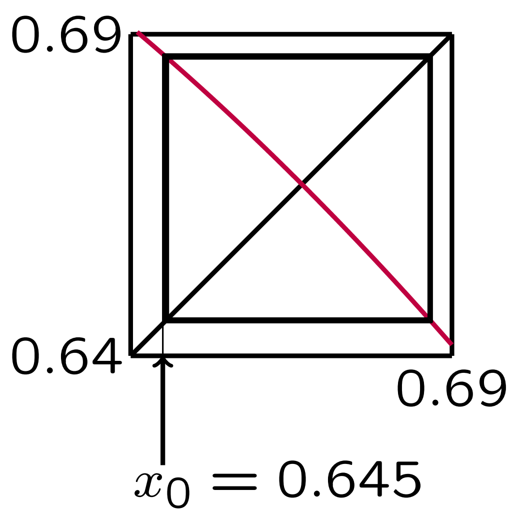


As R passes through 3, the fixed point goes from being an attracting fixed to point to being an *unstable fixed point* (note that we actually already observed this as R passed through 1, where 0 became an unstable fixed point).

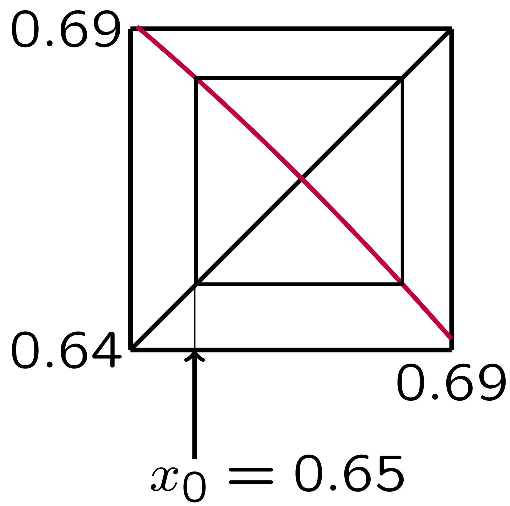
By looking more closely at the region of the fixed point as R passes through 3, we can understand better what is happening. In particular, the critical feature of the system is that as R goes through 3, the slope of the tangent to the parabola goes from above -1 to below -1 (i.e., becomes more steep).



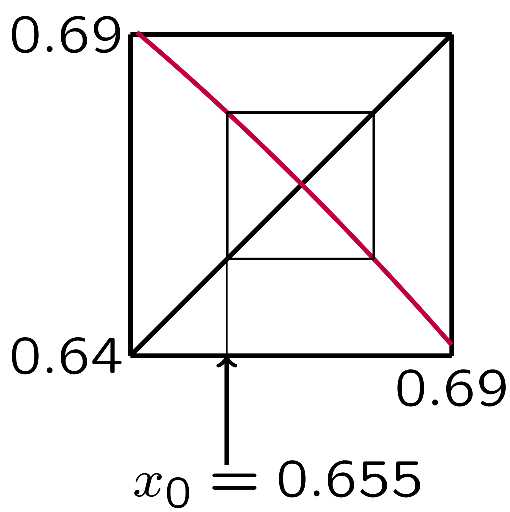
$$x_{n+1} = 2.95 * x_n(1 - x_n)$$



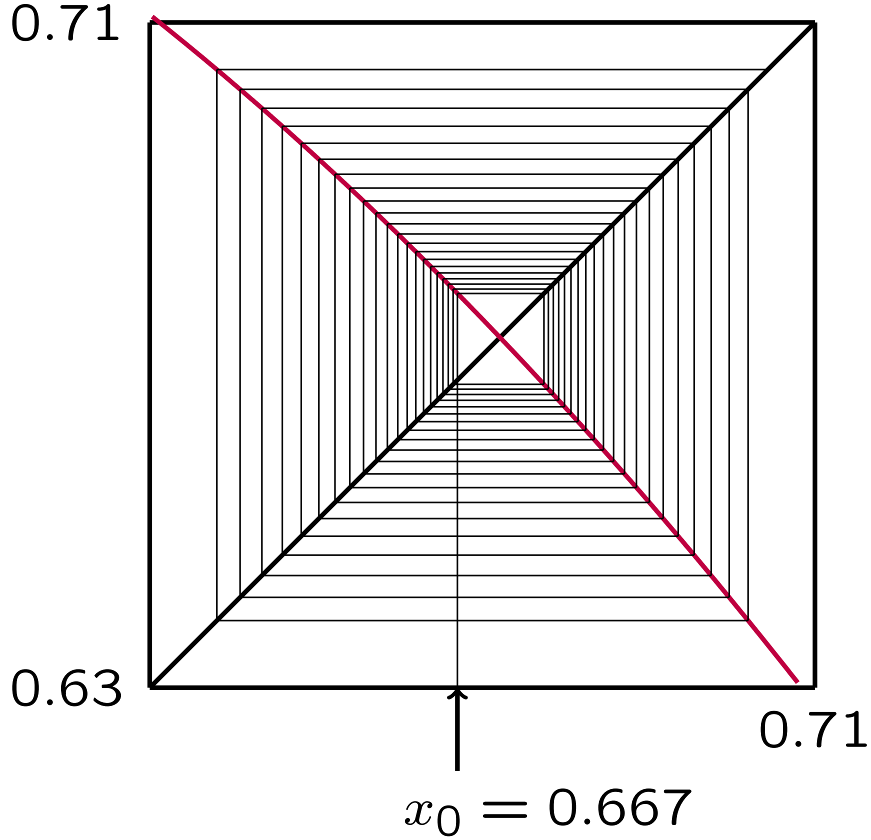
$$x_{n+1} = 3.0 * x_n(1 - x_n)$$



$$x_{n+1} = 3.0 * x_n(1 - x_n)$$

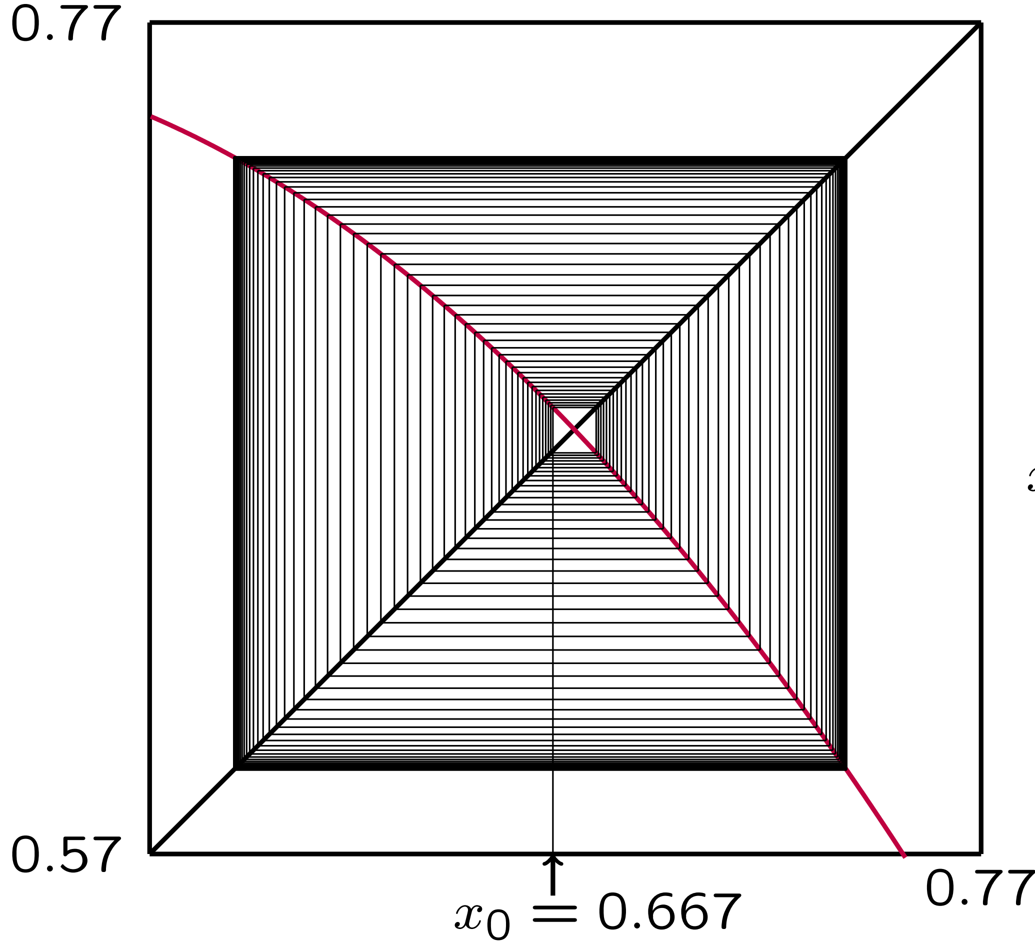


$$x_{n+1} = 3.0 * x_n(1 - x_n)$$

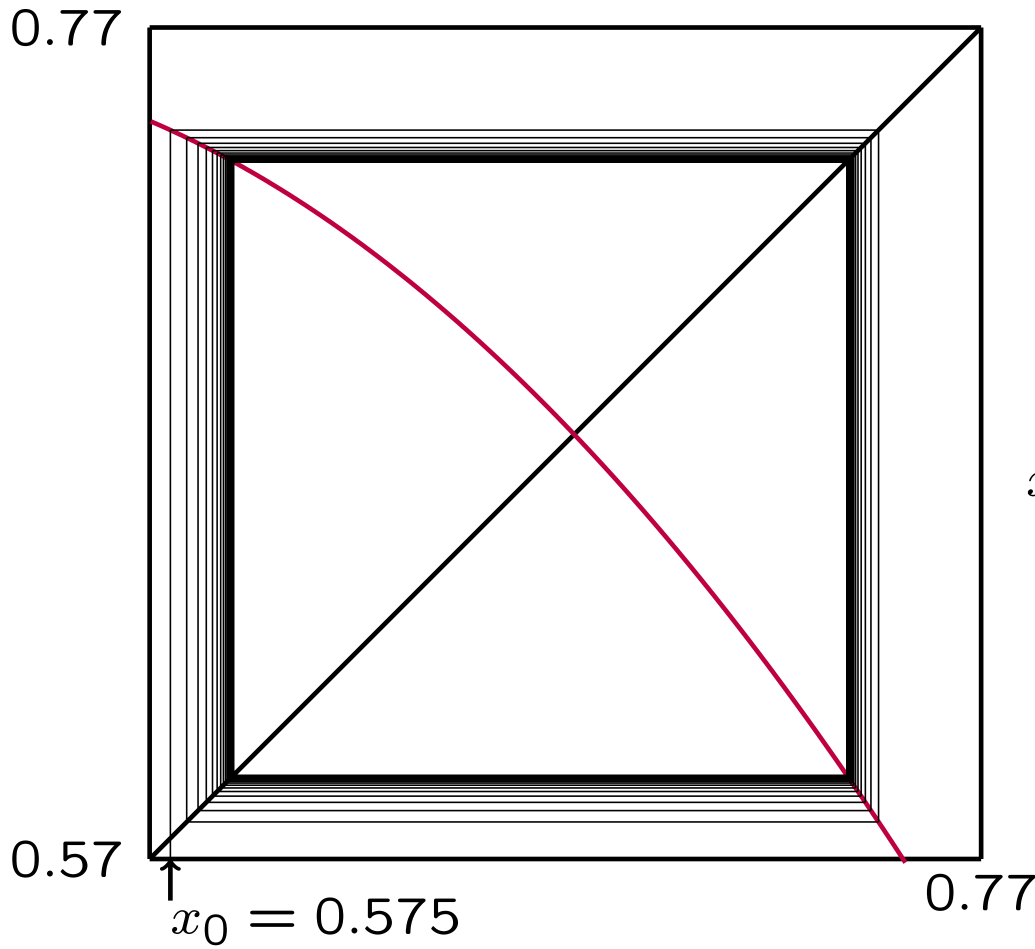


$$x_{n+1} = 3.05 * x_n(1 - x_n)$$

If we zoom out from this, we see another interesting phenomenon arise . . .



$$x_{n+1} = 3.05 * x_n(1 - x_n)$$



$$x_{n+1} = 3.05 * x_n(1 - x_n)$$

- We are seeing here an orbit of period 2 – the system bounces back and forth between two values. For values of $R < 3$ we had an attracting fixed point. The attracting fixed point is at 0 for $0 \leq R \leq 1$, and at $\frac{R-1}{R}$ for $1 < R < 3$. At $R = 3$ we still have the fixed point (and in fact that fixed point remains for all $R \geq 1$), but it is no longer a stable fixed point. For $R > 3$, it is a *repelling fixed point* – values near, but not exactly on, the fixed point will, in successive iterations, move further away.

  At $R = 1$ and $R = 3$, changes in R result in significant changes in the dynamics of the system. These significant changes in the dynamics of the system resulting from changes in a controlling parameter are called *bifurcations*. At $R = 1$, the stable (attracting) fixed point at 0 becomes unstable, and becomes a repelling fixed point. We acquire a new stable (attracting) fixed point at $\frac{R-1}{R}$.

At $R = 3$, the fixed point at $\frac{R-1}{R}$ becomes unstable, and we acquire a new attracting stable orbit of period 2. These is the first of a sequence of bifurcations in the dynamics of the logistics equation. At each of these bifurcations, an orbit of period 2^n becomes unstable (although it continues to exist in the dynamics), and a new stable orbit of period 2^{n+1} arises. This process is called a *period doubling bifurcation cascade*.

We can calculate the x values of this period 2 orbit by looking at two steps of the difference equation:

$$\begin{aligned} x_{n+2} &= Rx_{n+1}(1 - x_{n+1}) \\ &= R(Rx_n(1 - x_n))(1 - Rx_n(1 - x_n)) \\ &= R^2 x_n(1 - x_n)(1 - Rx_n + Rx_n^2) \\ &= R^2 x_n(-Rx_n^3 + 2Rx_n^2 - (R+1)x_n + 1) \end{aligned}$$

We are looking for period 2 orbits – in other words, values where $x_{n+2} = x_n$. We therefore want

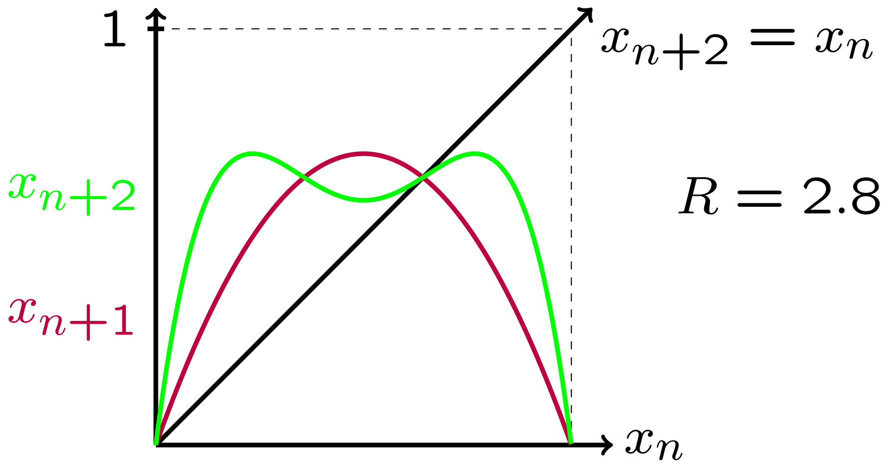
$$x_n = R^2 x_n(-Rx_n^3 + 2Rx_n^2 - (R+1)x_n + 1)$$

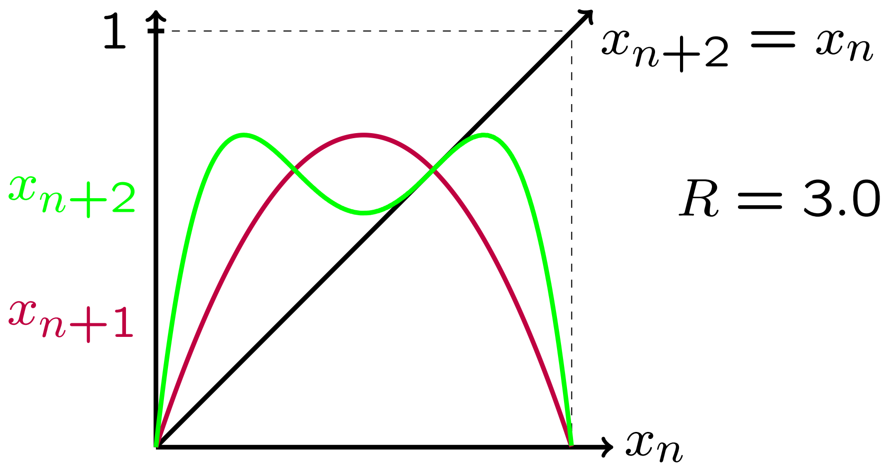
or

$$R^2 x_n(-Rx_n^3 + 2Rx_n^2 - (R+1)x_n + 1) - x_n = 0.$$

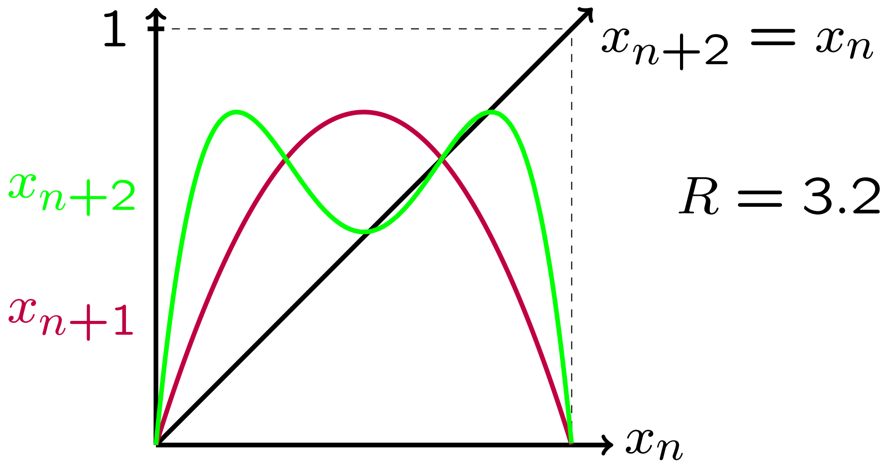
This is a 4th degree polynomial is x_n, which, for $R > 3$, has 4 roots. It is easy to see that $x_n = 0$ is a root (but we already knew that, because 0 is a fixed point of the original equation, and hence also repeats itself every 2 steps). Similarly, $x_n = \frac{R-1}{R}$ is a root, because it too is a fixed point. We could proceed with the (somewhat messy) algebra to find the other two roots, but instead let's look at the cobweb diagram for x_{n+2} vs. x_n (in these diagrams, we're also seeing the original parabolas, for reference . . .).

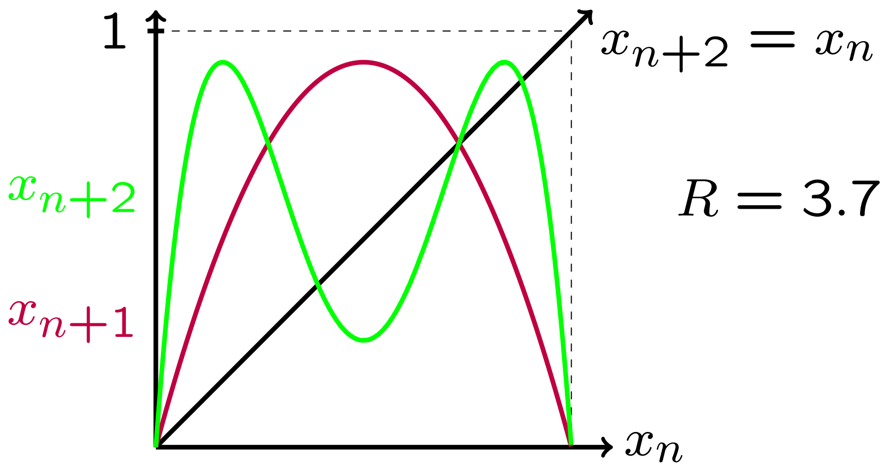
You can see the bifurcation happen as R reaches 3.0:



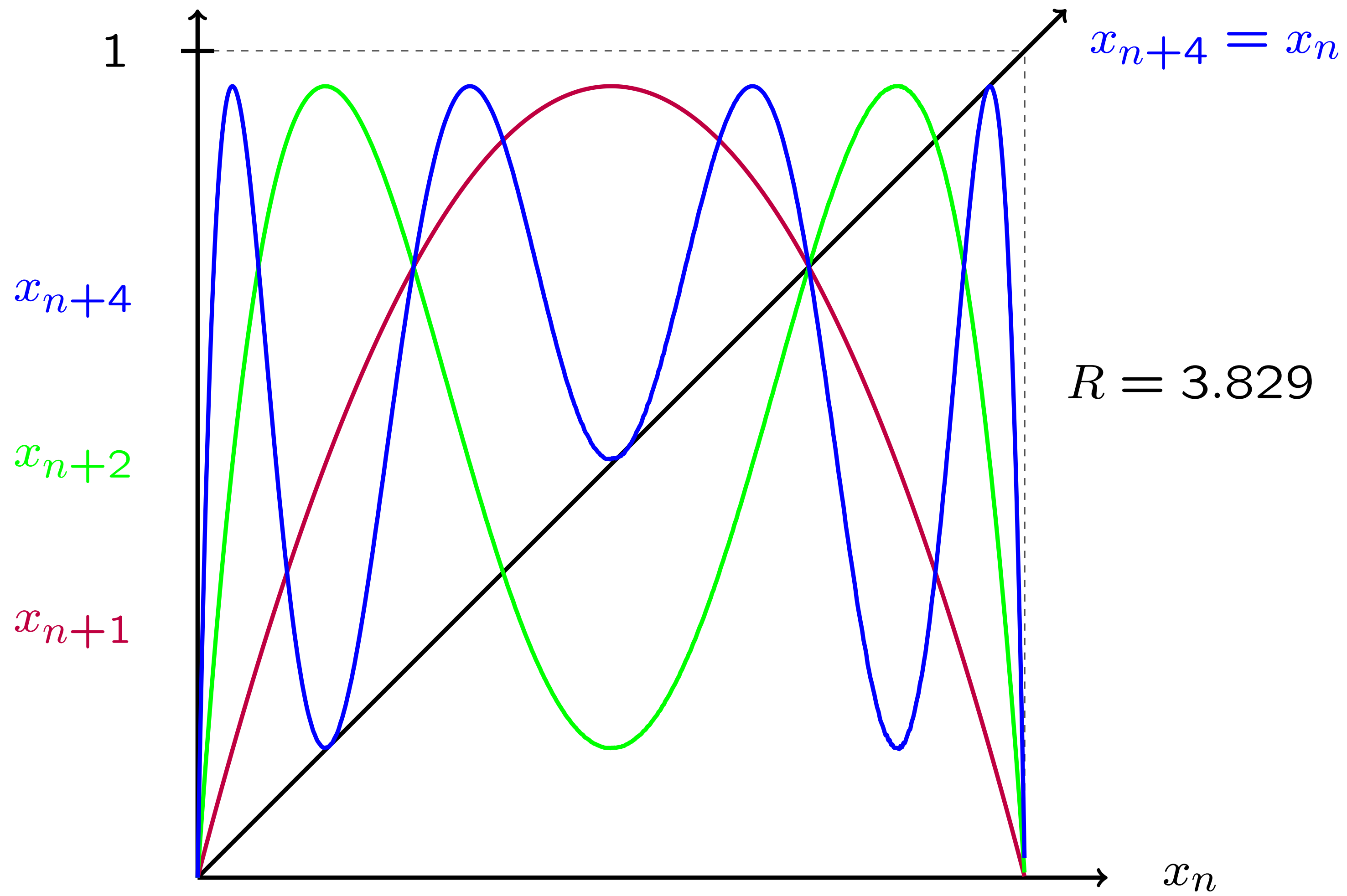


and moves on past . . .

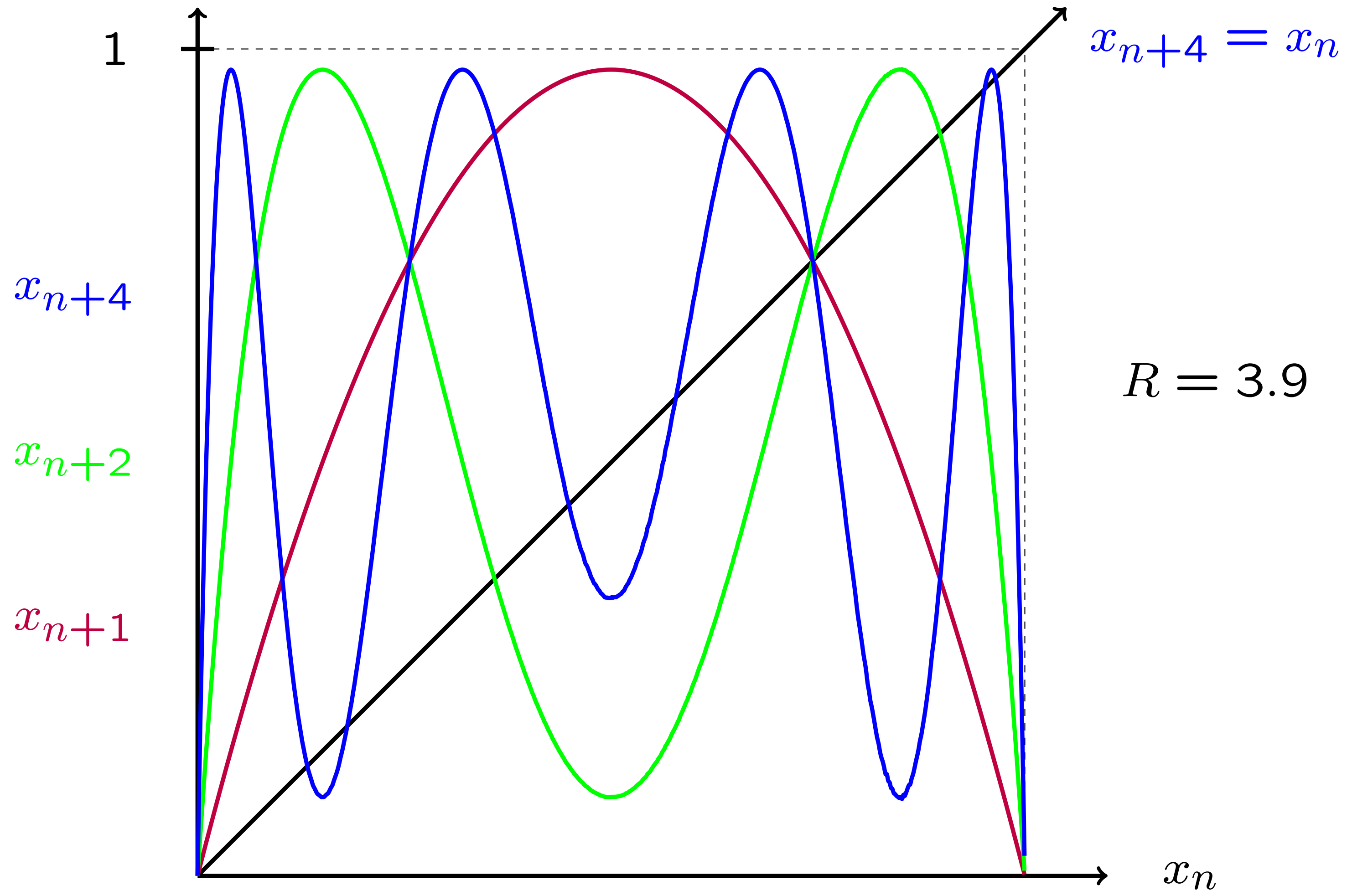




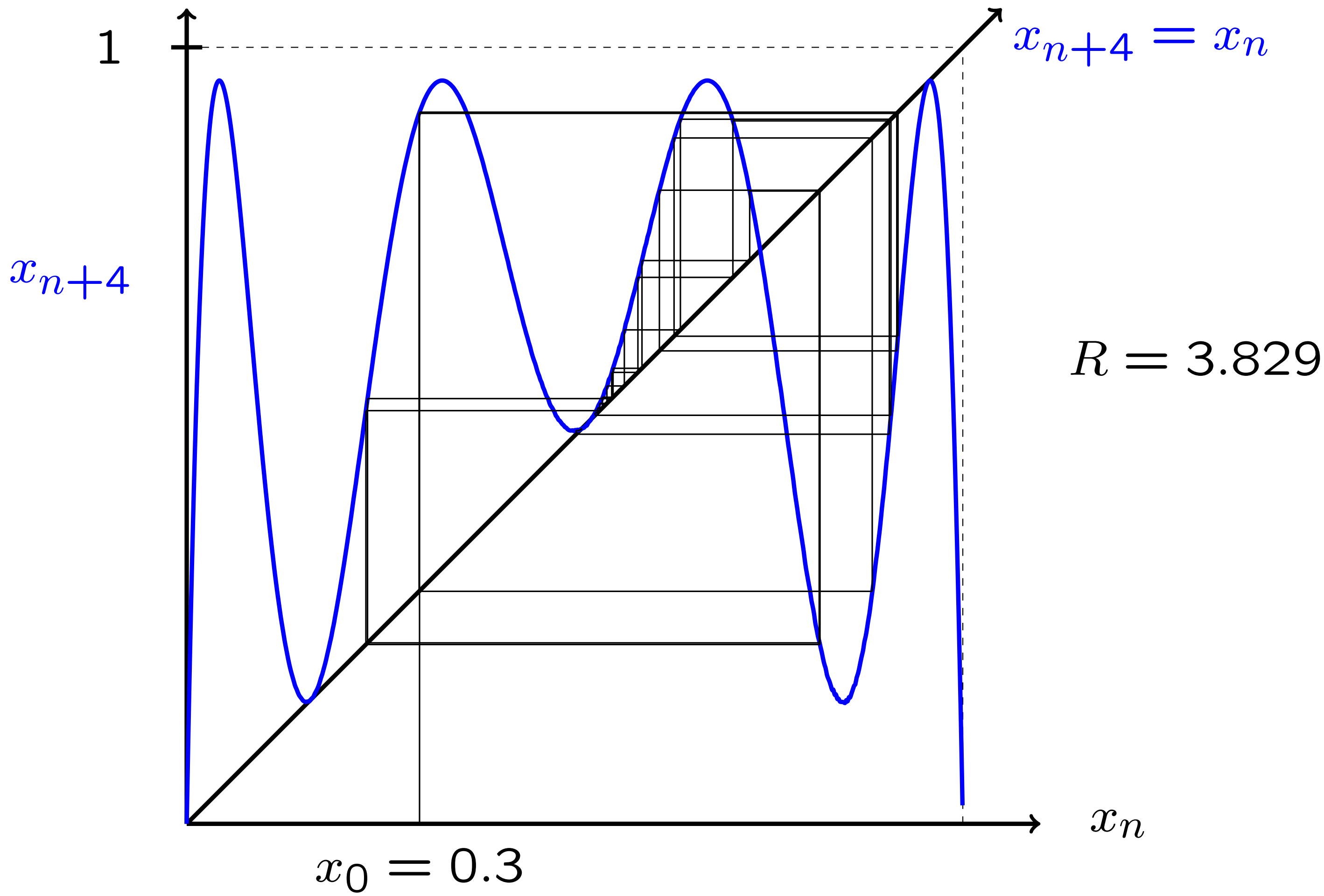
We can also look at higher order iterates:



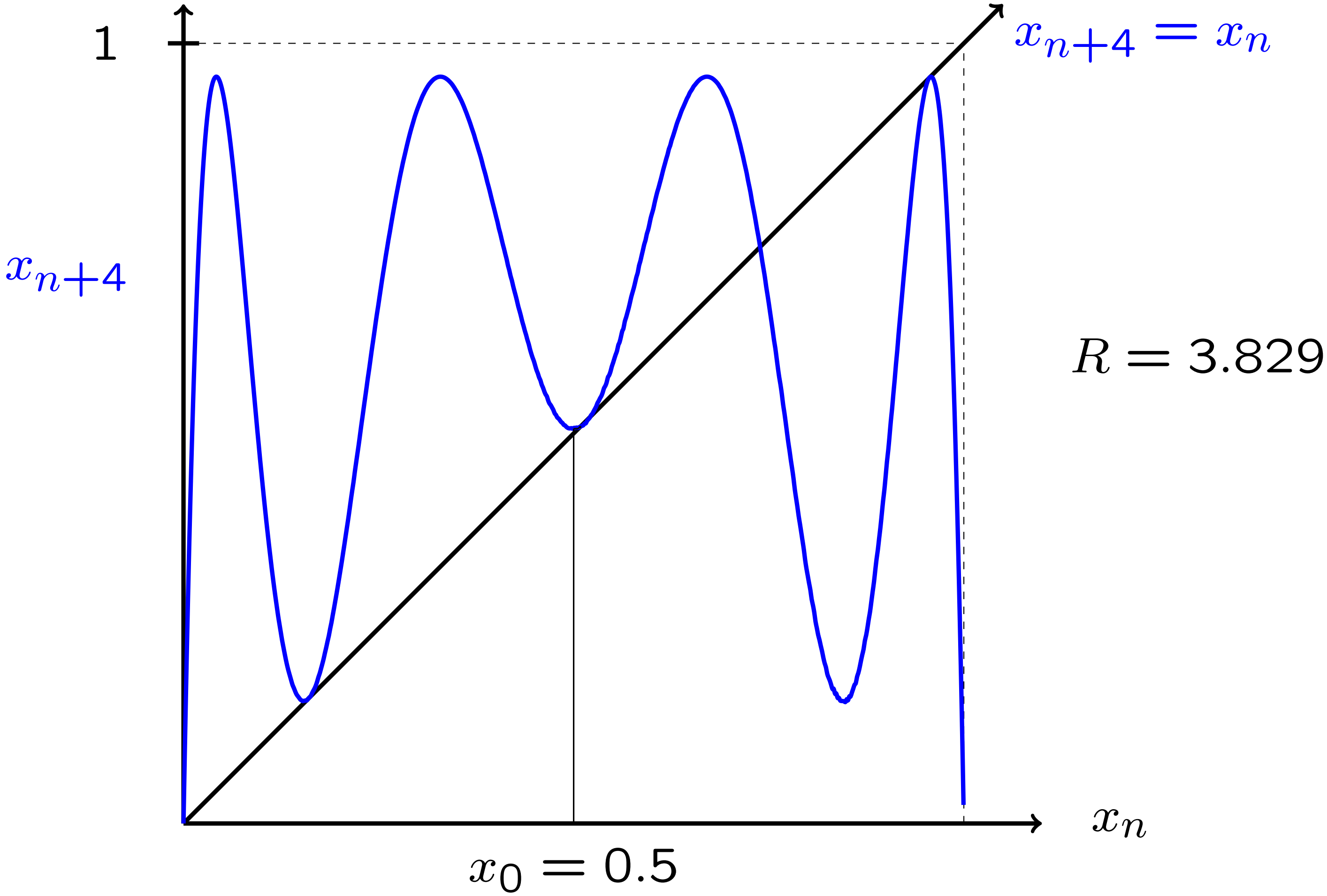
For various values of R:



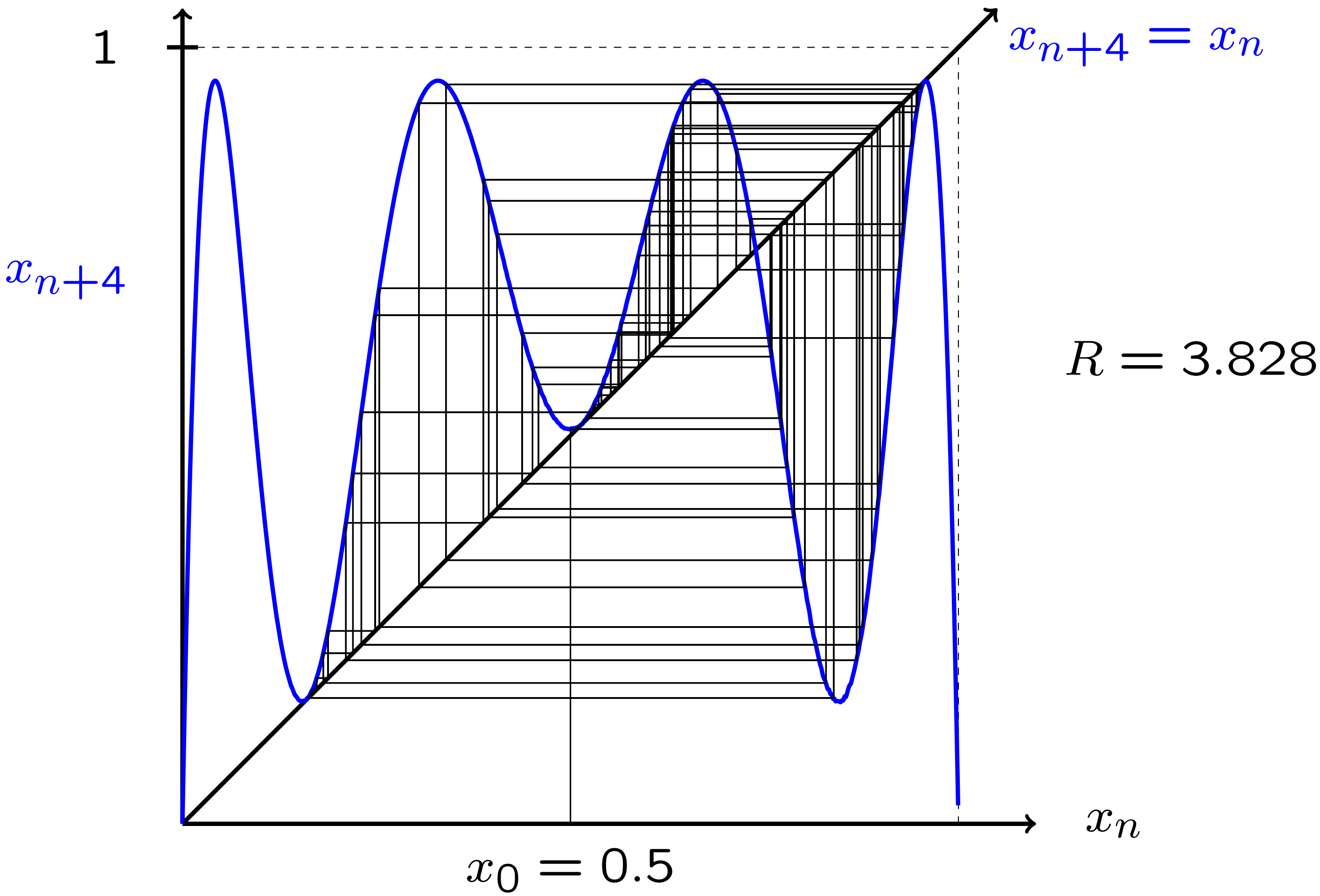
We can do cobweb diagrams also (320 iterations):



With changing starting values (still 320 iterations):



And changing values of R (also 320 iterations):

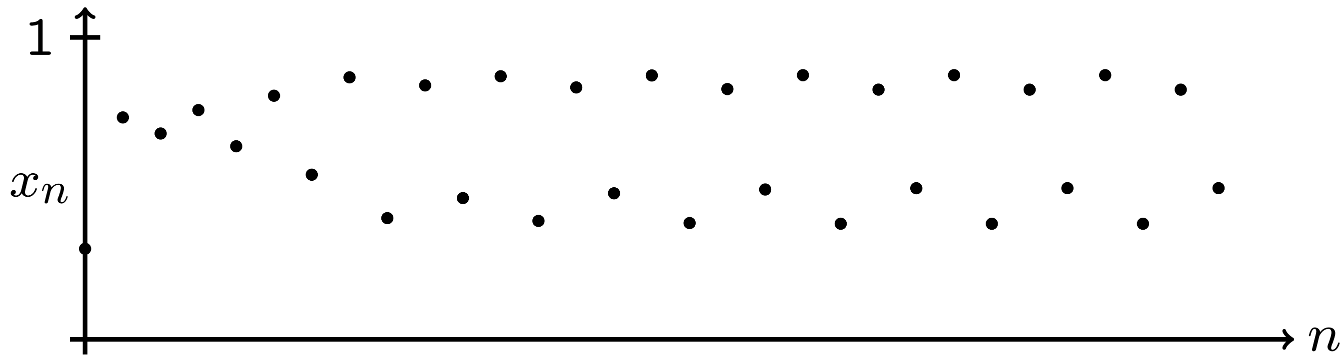


Bifurcation Diagrams ←

- There is another way of viewing the behavior of the logistics equation, which makes it easier to see the changes in behavior as R changes. We will display trajectories for various values of R all at once in a single two dimensional picture. On our two dimensional image, the horizontal axis will be for various values of R. For each value of R, the vertical direction will show a sequence of values of x_n.

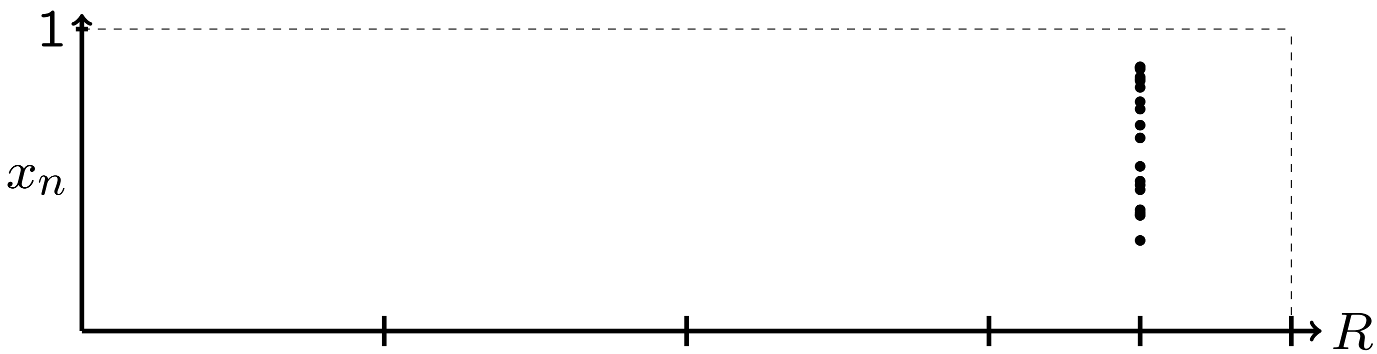
  This approach will allow us to see overall changes in the dynamics as we vary the controlling parameter R. As we will see, there are values of R at which we will have significant changes in the dynamics. As indicated above, these changes are called *bifurcations*, and hence this view is often called a *bifurcation diagram*.

Thus, for example, when $R = 3.5$ (and $x_0 = 0.5$):



$$x_{n+1} = 3.5 * x_n * (1 - x_n), \quad x_0 = 0.5$$

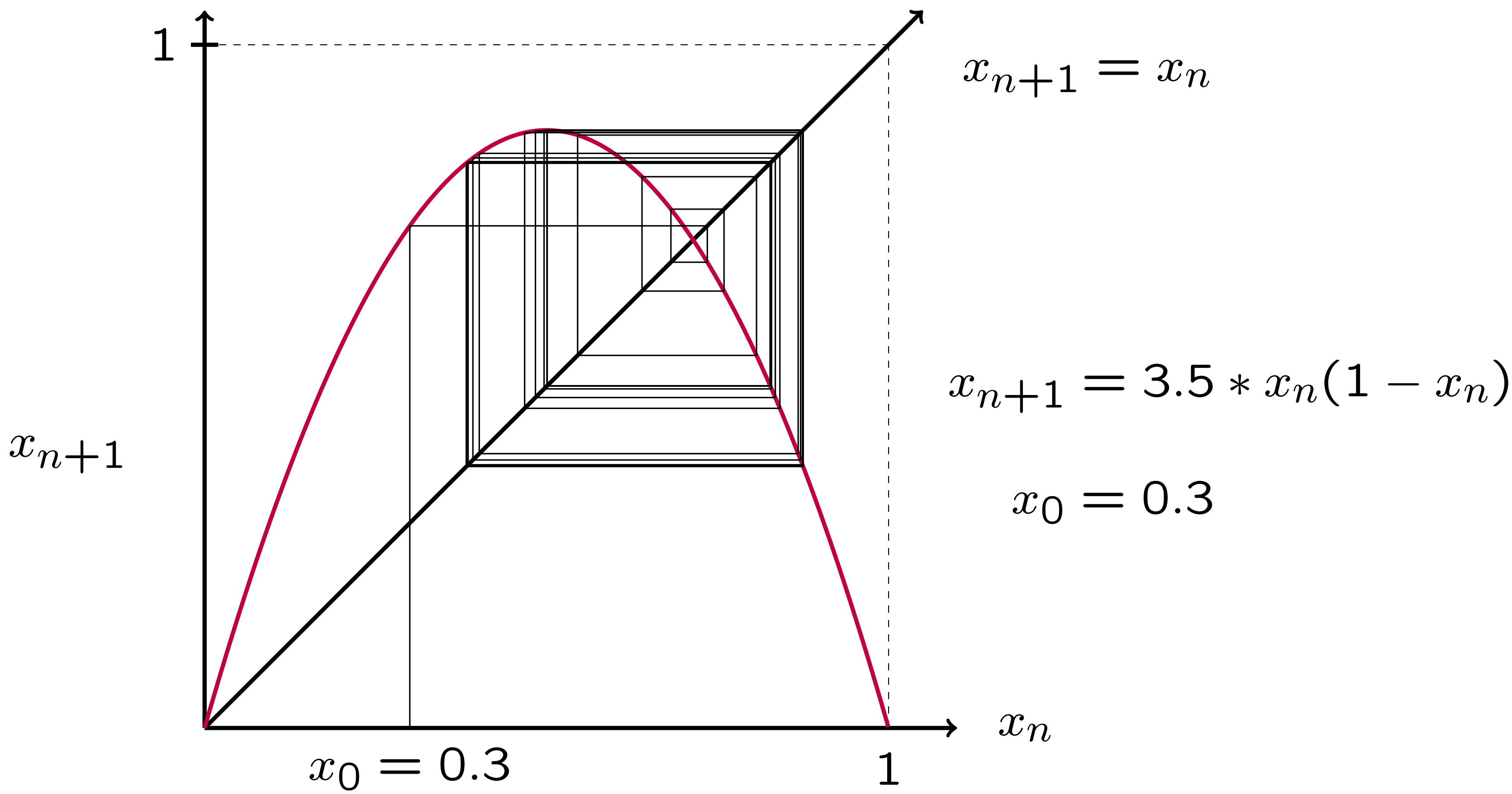
We rotate the graph, and show values of x_n at $R = 3.5$:



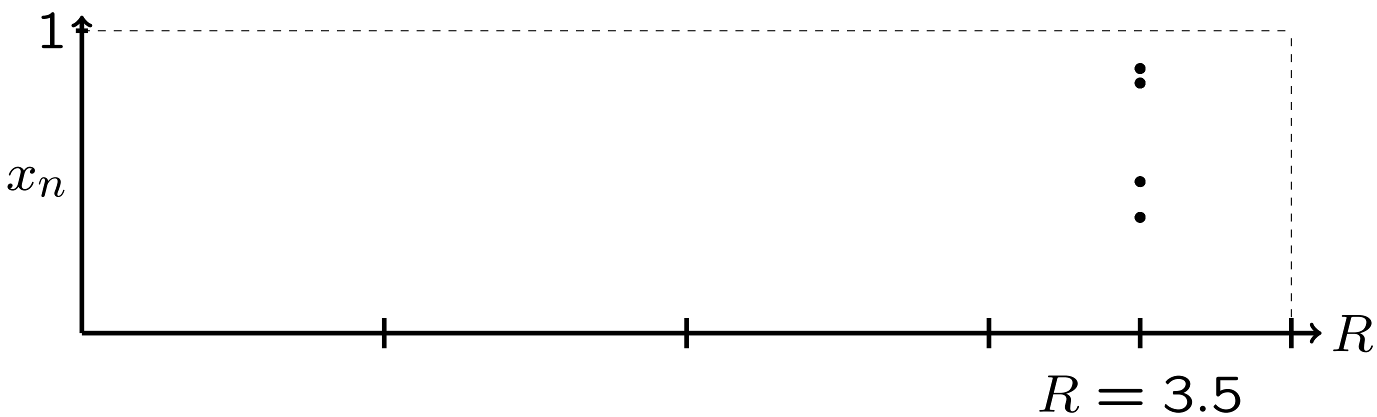
$R = 3.5$

This image is a bit of a smear, and it is hard to see structure in the dynamics. We are actually more interested in the long term behavior of the system. For any particular starting value of x_0, the system is likely to wander around for a while before settling down to its stable (limiting) behavior. We can see this for particular values of R.

Here we are looking at $R = 3.5$. The system wanders for a while, but is settling down to an orbit of period 4:



What we will do is to throw away early values of x_n (often called the *transients*), and show the long term, or *limiting* trajectory. In this picture, we have thrown away the first 80 values of x_n, and plotted the next 40:



In this picture, we can see the period 4 orbit, which is an attracting closed periodic orbit. As n goes to ∞, trajectories approach this limiting orbit.

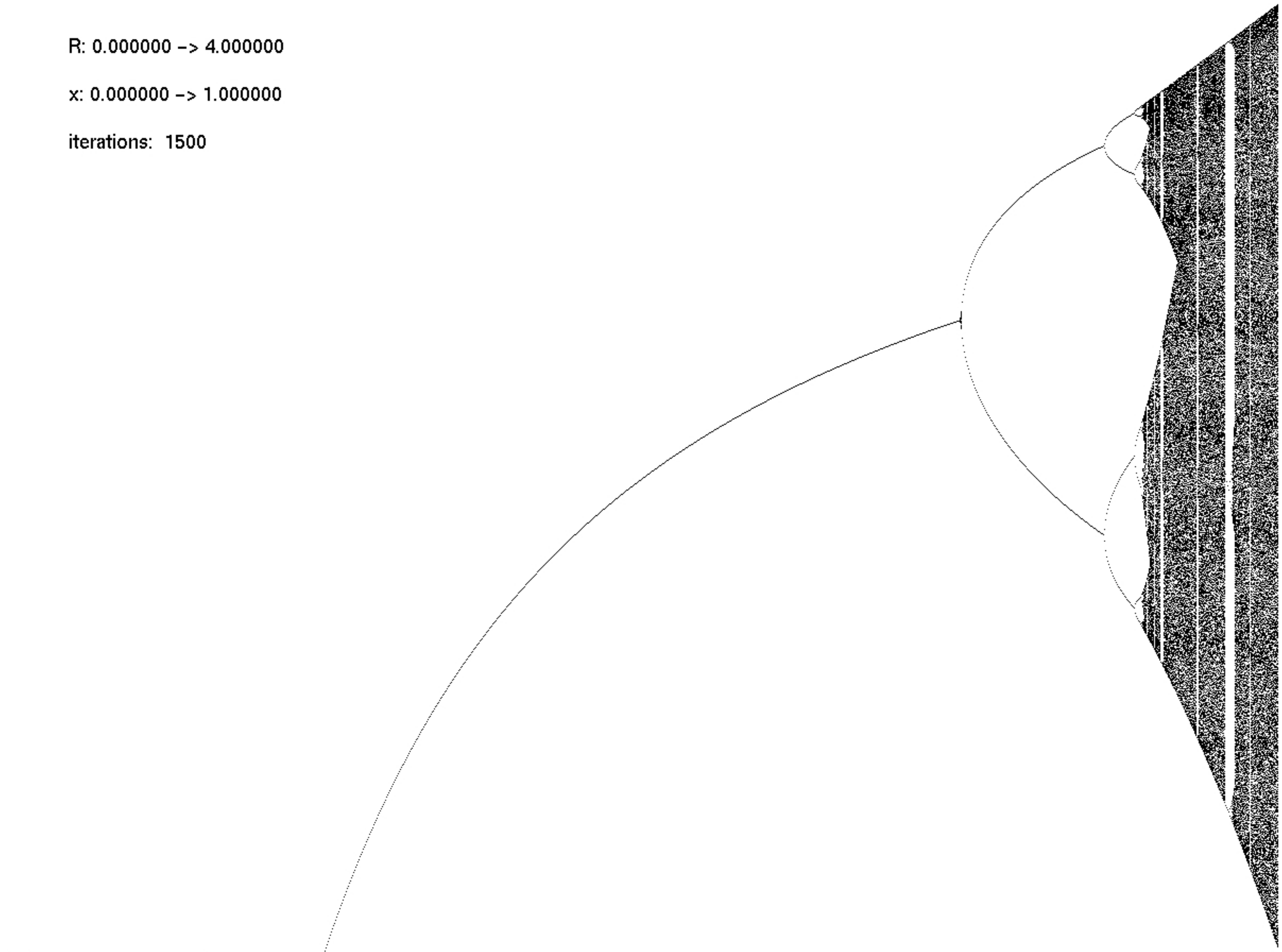
In the following images, we can see "bifurcation diagrams" for the logistic equation. In the first couple of images, things are monochrome, with dots for values of x_n. In the later pictures, there are various colors, reflecting the amount of time the system spends in those regions (closer to white means longer time spent).

The first diagram shows $0 \le R \le 4$, and $0 \le x_n \le 1$. Later we can talk about why we keep $R \le 4$. These images all have $x_0 = 0.3$, and discard the first 80 iterations.

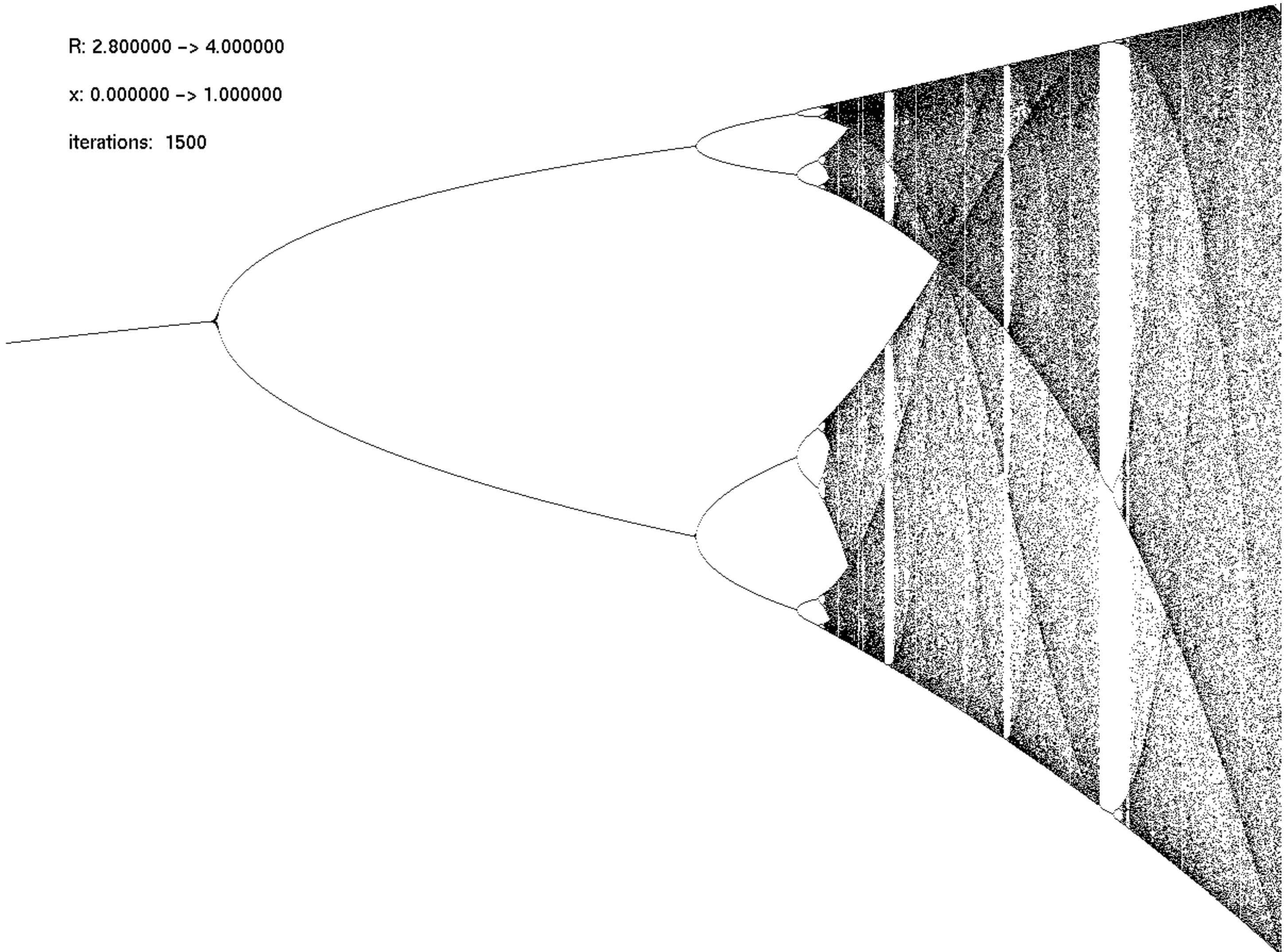
In the first image, you can see the transition from a stable (attracting) fixed point at 0 for $R < 1$. At $R = 1$, that fixed point becomes unstable, and there is a new stable (attracting) fixed point at $\frac{R-1}{R}$ for $1 < R < 3$. At $R = 3$ that fixed point becomes unstable, and a new attractor of period 2 arises. As we increase R, that period 2 orbit becomes unstable, and we get a new orbit of period 4, then period 8, etc. As R grows, we pass through orbits of all periods of the form 2^n. This is often referred to as a *period doubling bifurcation cascade*. This period doubling cascade reaches its limit before R goes to ∞, and we enter a chaotic realm.

Other diagrams show some of what we will see if we zoom in. The captions under the images tell the range of R values, the range of values for x_n being shown, and the number of iterations. In these images, there are 900 possible values for x_n. In generating the images, the logistics equation is iterated for a given value of R, and at each step one of 900 bins for x_n is incremented. These values are then used to select a color.

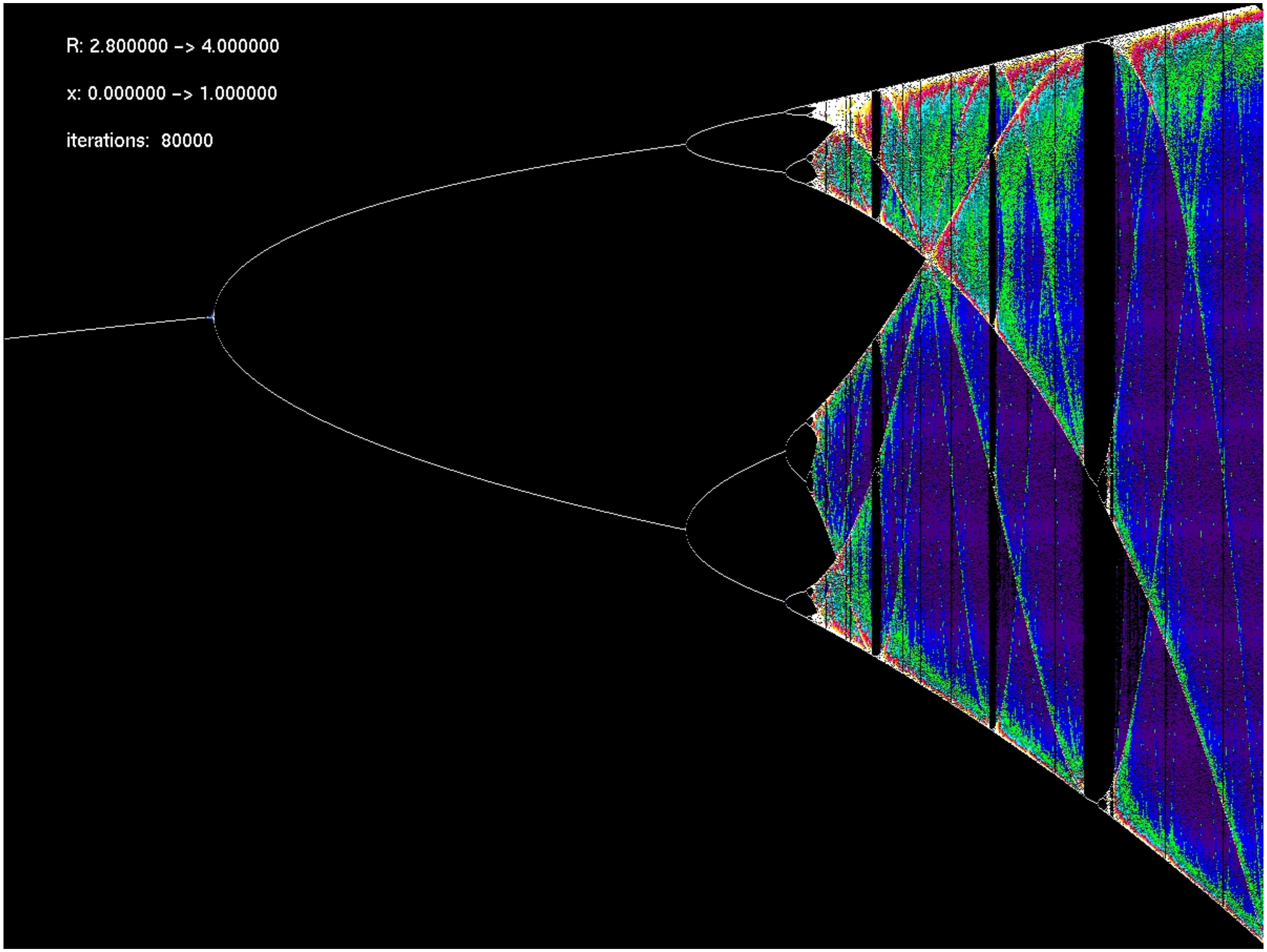
It is worth remembering that these images, as we zoom in, also reveal various artifacts having to do with machine resolution, etc. . . .



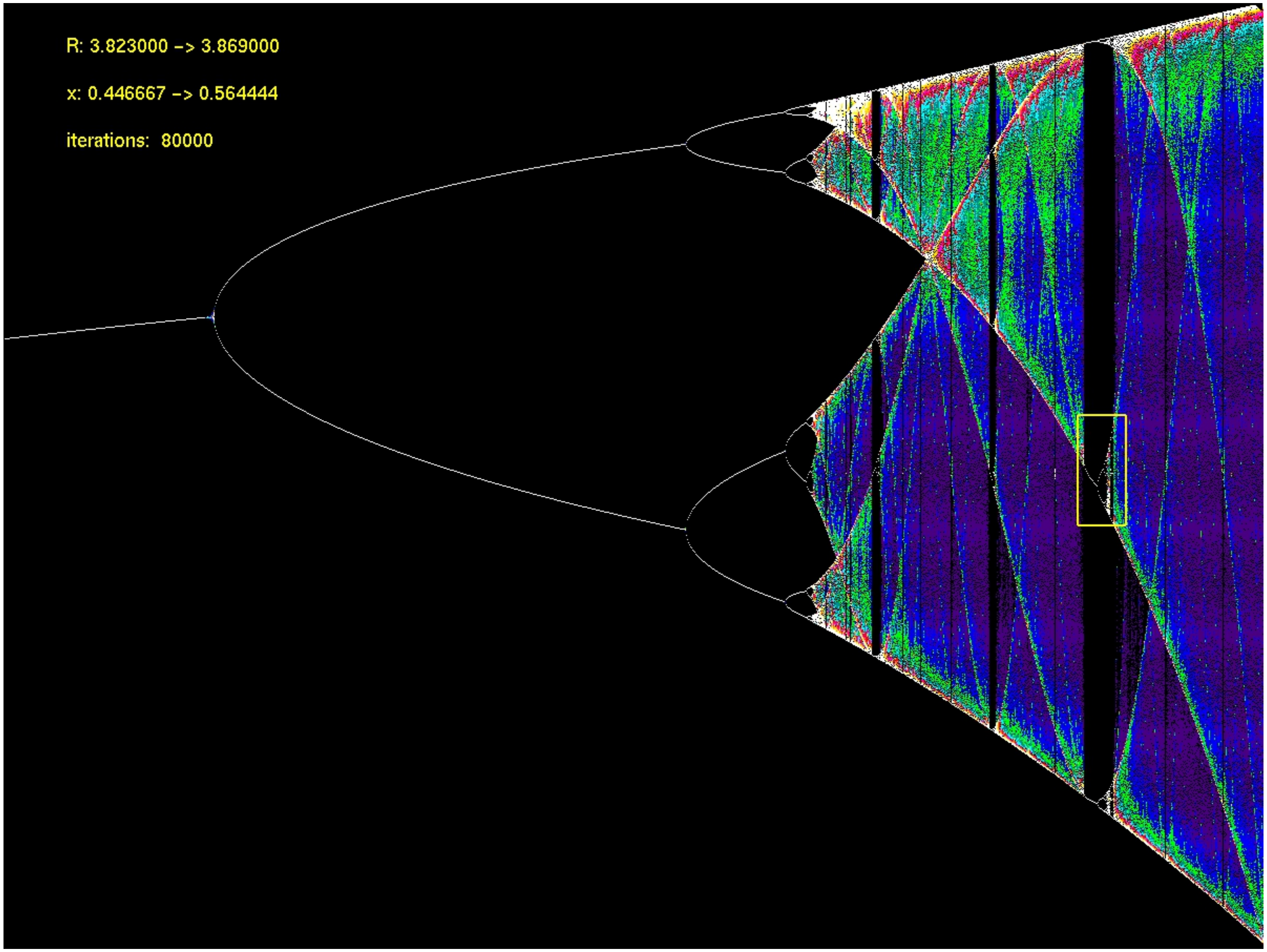
R=0.000-R=4.000-x=0.000-x=1.000-iter=1500



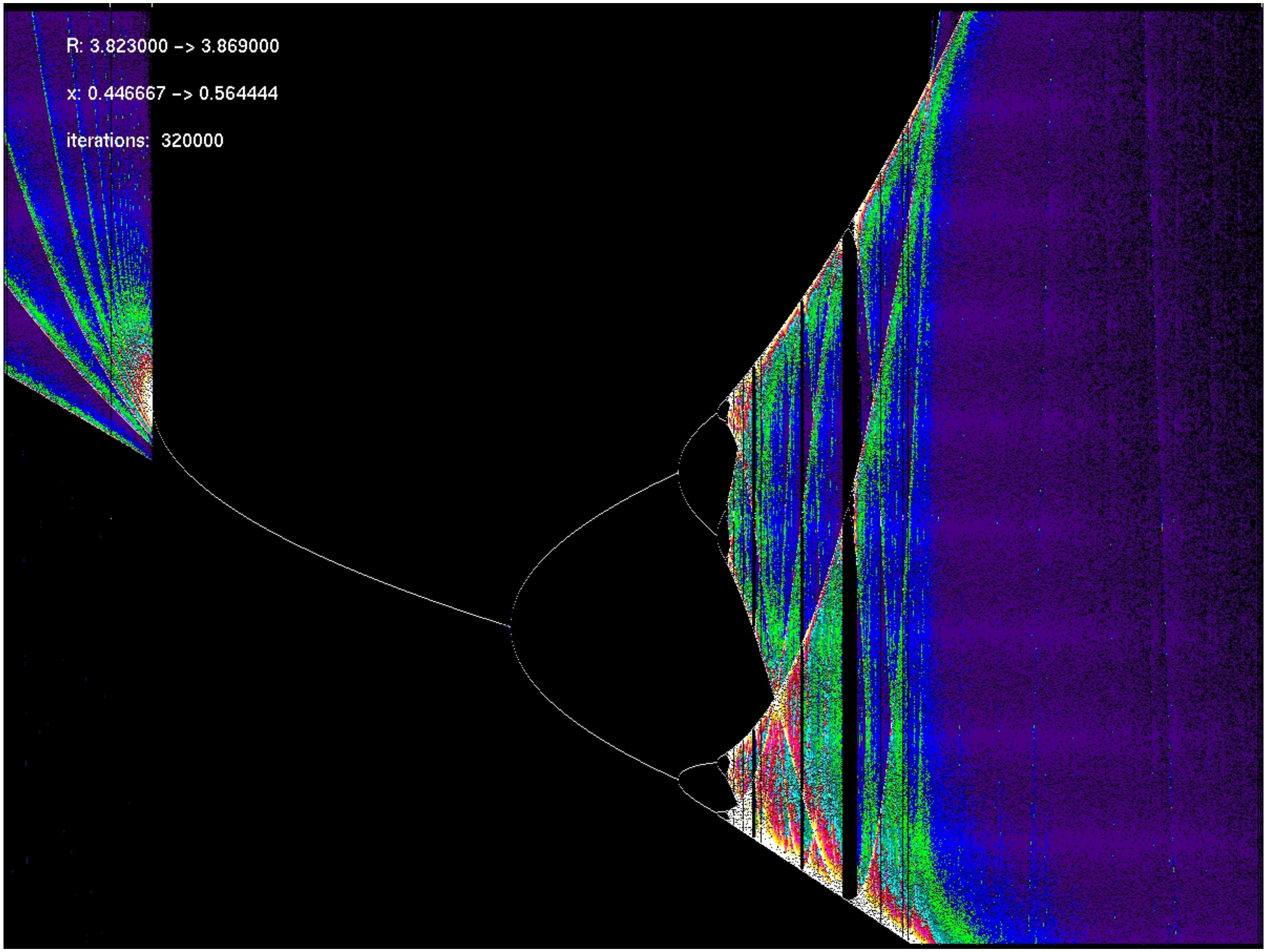
R=2.800-R=4.000-x=0.000-x=1.000-iter=1500



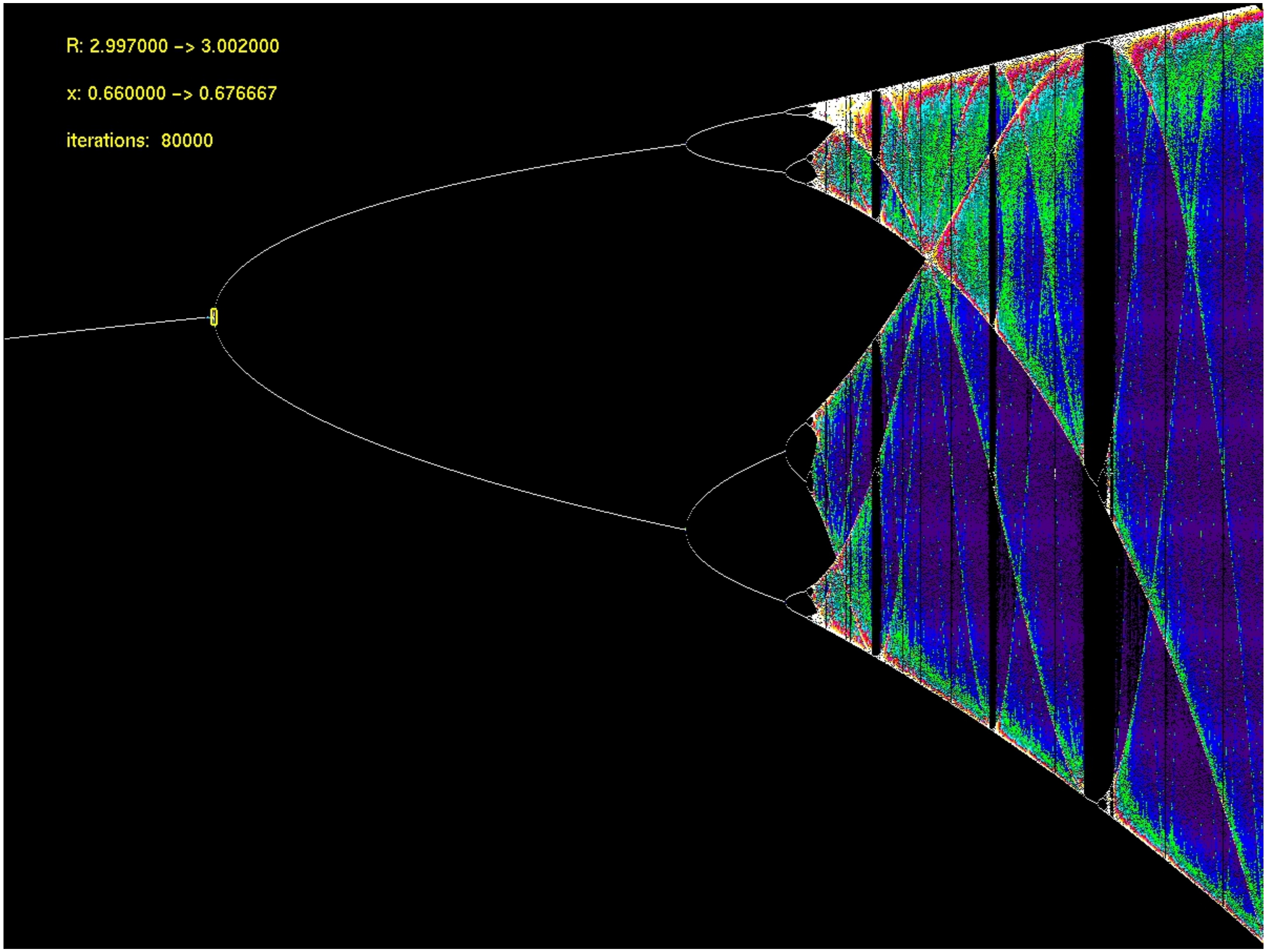
R=2.800-R=4.000-x=0.000-x=1.000-iter=80000



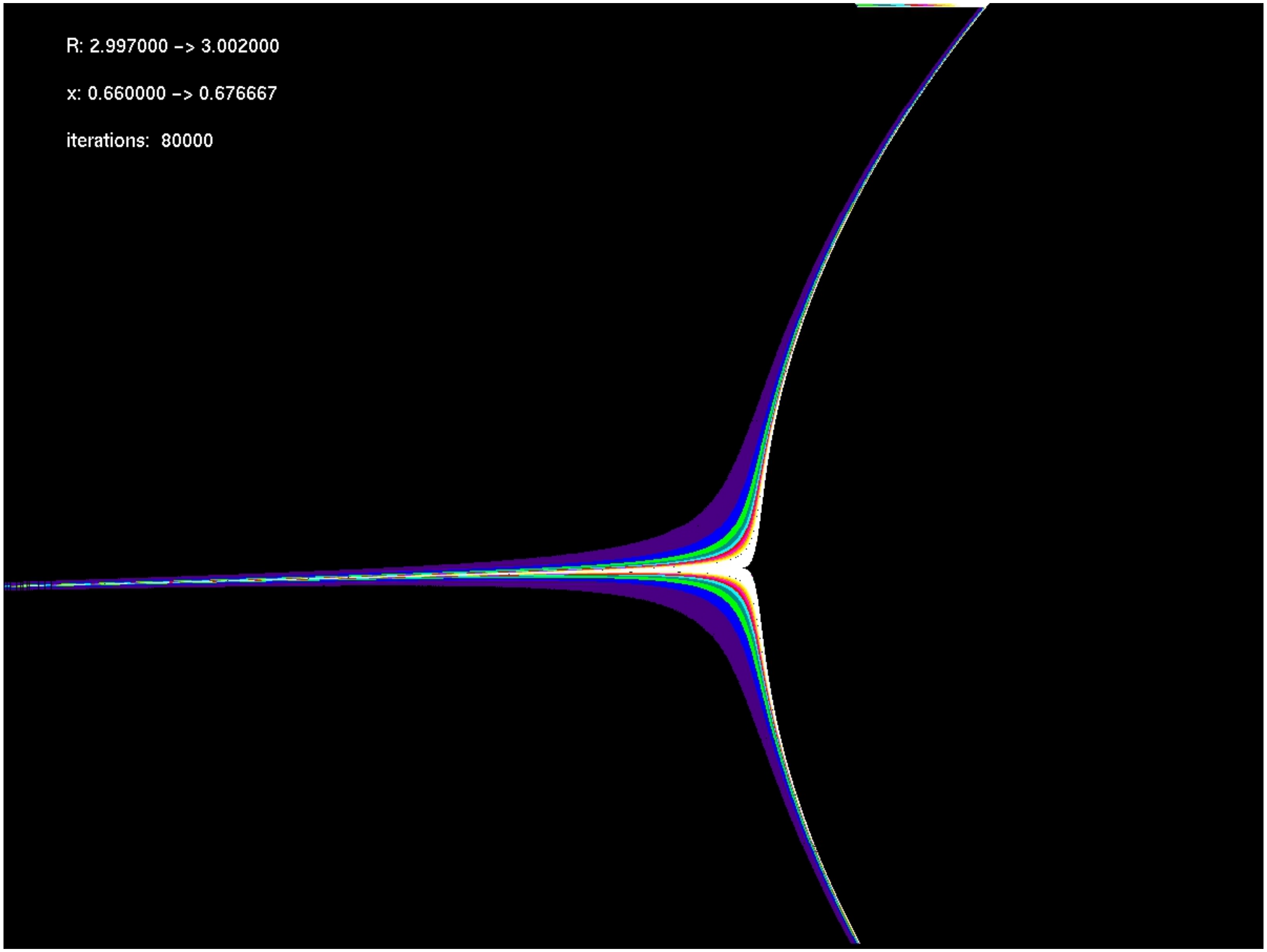
R=2.8000-R=4.000-x=0.000-x=1.000-iter=80000-z



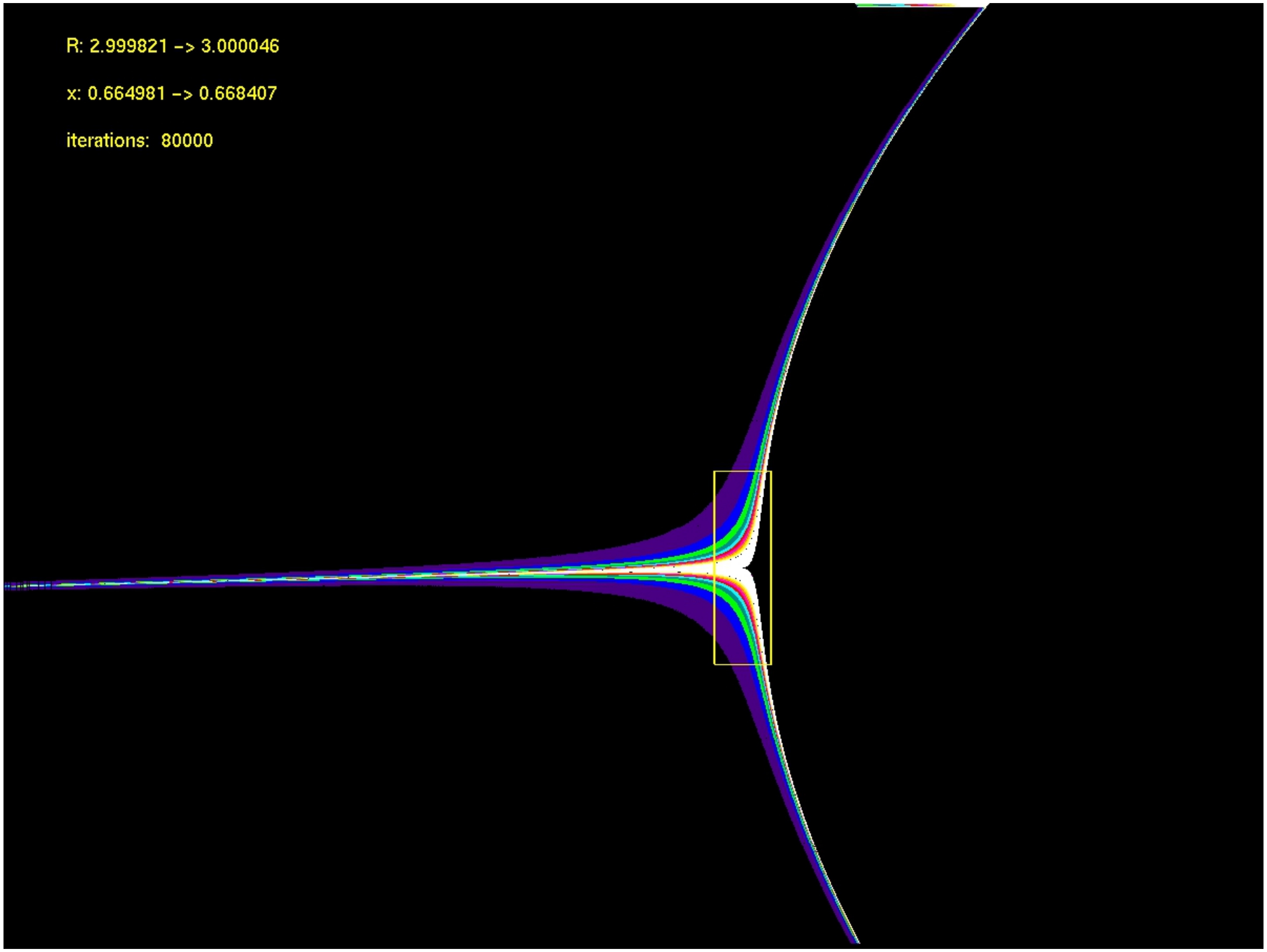
R=3.823-R=3.869-x=0.447-x=0.564-iter=320000



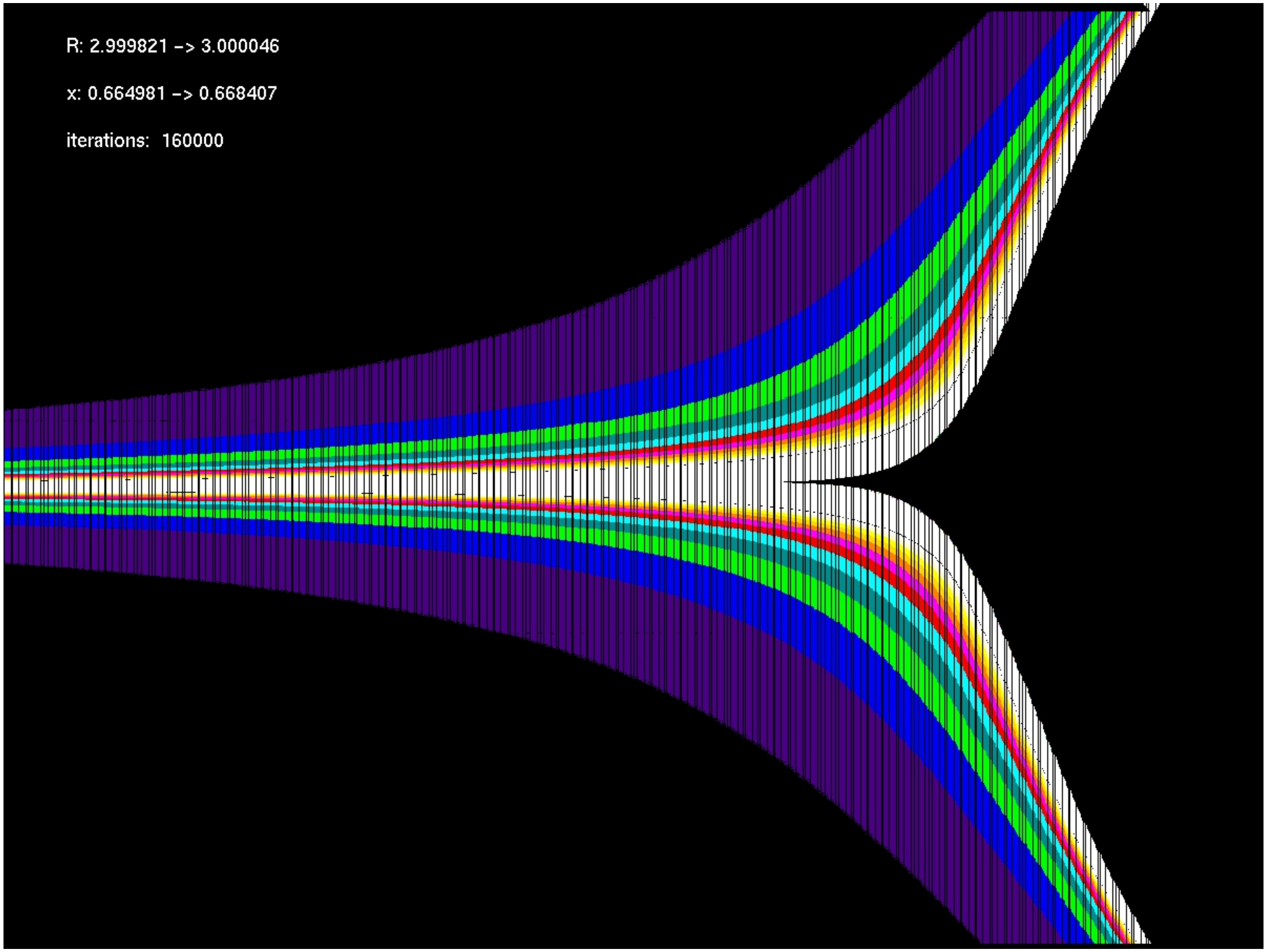
R=2.8000-R=4.000-x=0.000-x=1.000-iter=80000



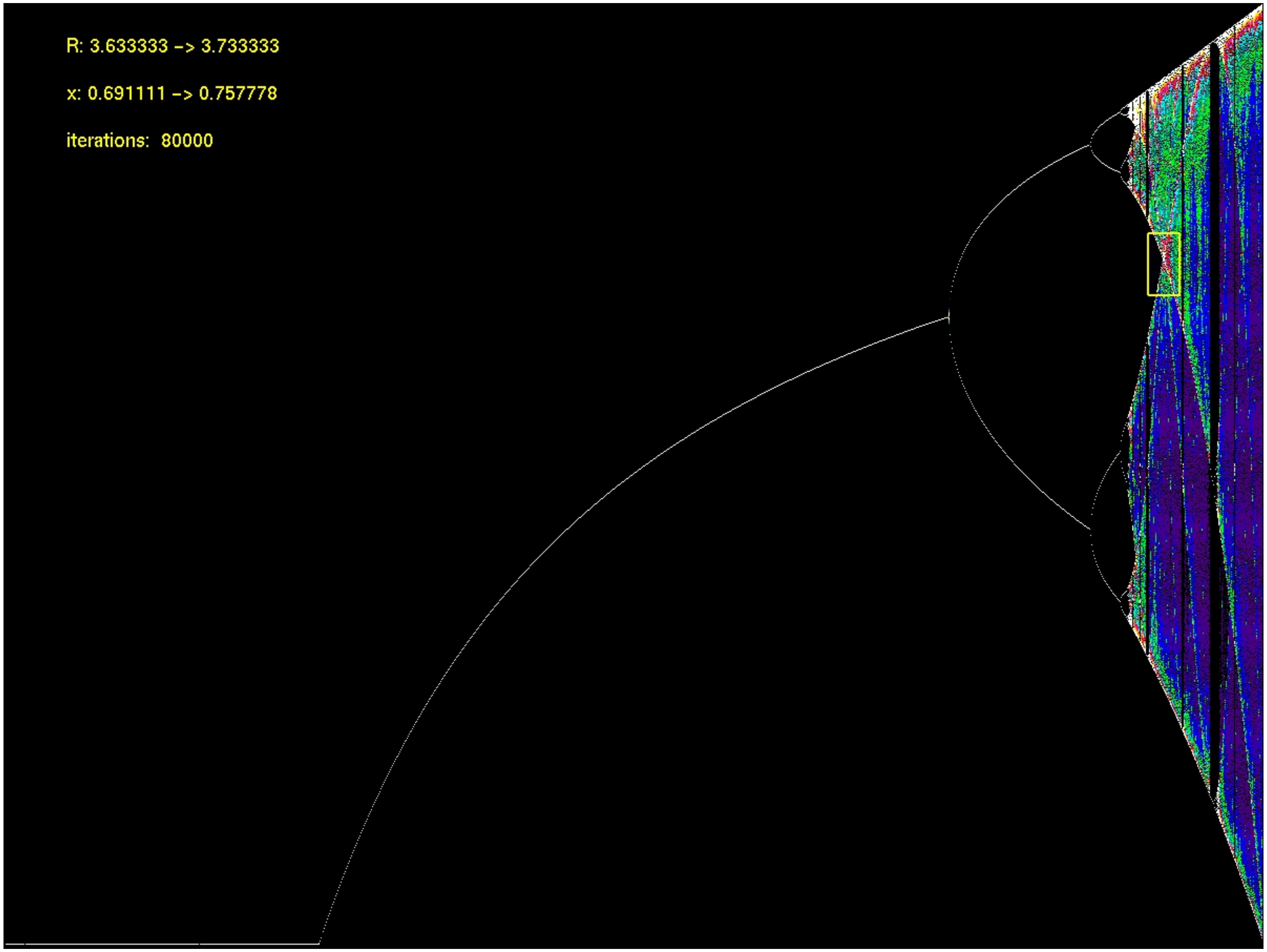
R=2.997-R=3.002-x=0.660-x=0.677-iter=80000



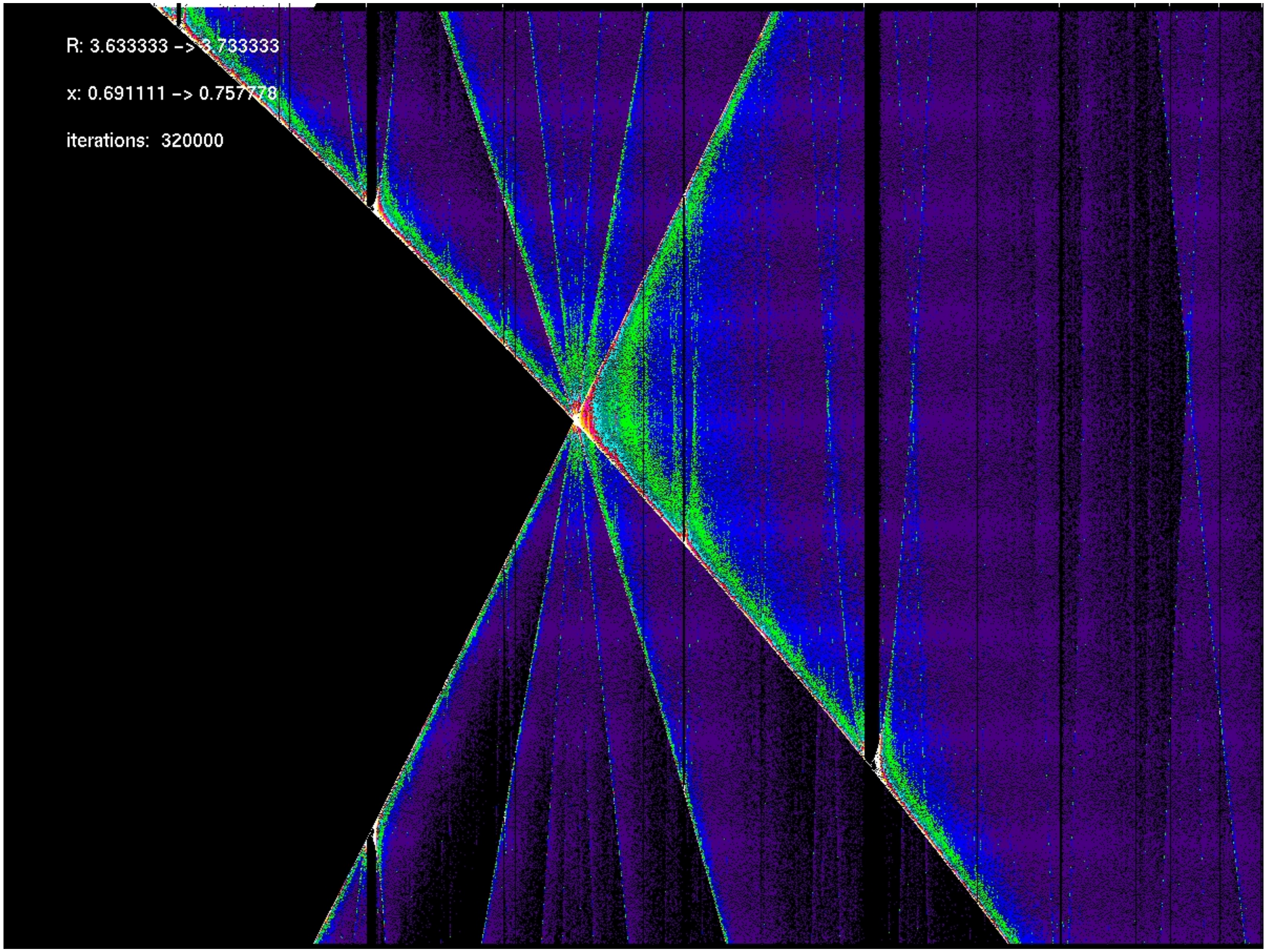
R=3.000-R=3.000-x=0.665-x=0.668-iter=80000-z



R=3.000-R=3.000-x=0.665-x=0.668-iter=160000



R=2.800-R=4.000-x=0.000-x=1.000-iter=80000 z



R=3.633-R=3.733-x=0.691-x=0.758-iter=320000

- There are various tools for exploring such bifurcation diagrams, such as:

  http://csustan.csustan.edu/~
  tom/SFI-CSSS/nonlinear/Chaos.html

Universality ←

- An important result in the study of nonlinear (and chaotic) systems is a *universality* property discovered by Mitchell Feigenbaum. As we increase R from 0, we see a sequence of bifurcations. In particular, starting at $R = 3$, we see a sequence of period doubling bifurcations. by the time we get to $R = 3.57$, we have had infinitely many period doublings, and we reach a chaotic realm.

  Feigenbaum observed that the distance between successive period doubling bifurcations in the logistics map exhibited an interesting property. If we let R_i be the location of the ith period doubling bifurcation after $R = 1$, and let $d_i = R_{i+1} - R_i$, then we can look at the sequence $\frac{d_i}{d_{i+1}}$. Feigenbaum showed that this sequence has a limit:

$$\delta = \lim_{i \to \infty} \frac{d_i}{d_{i+1}}$$

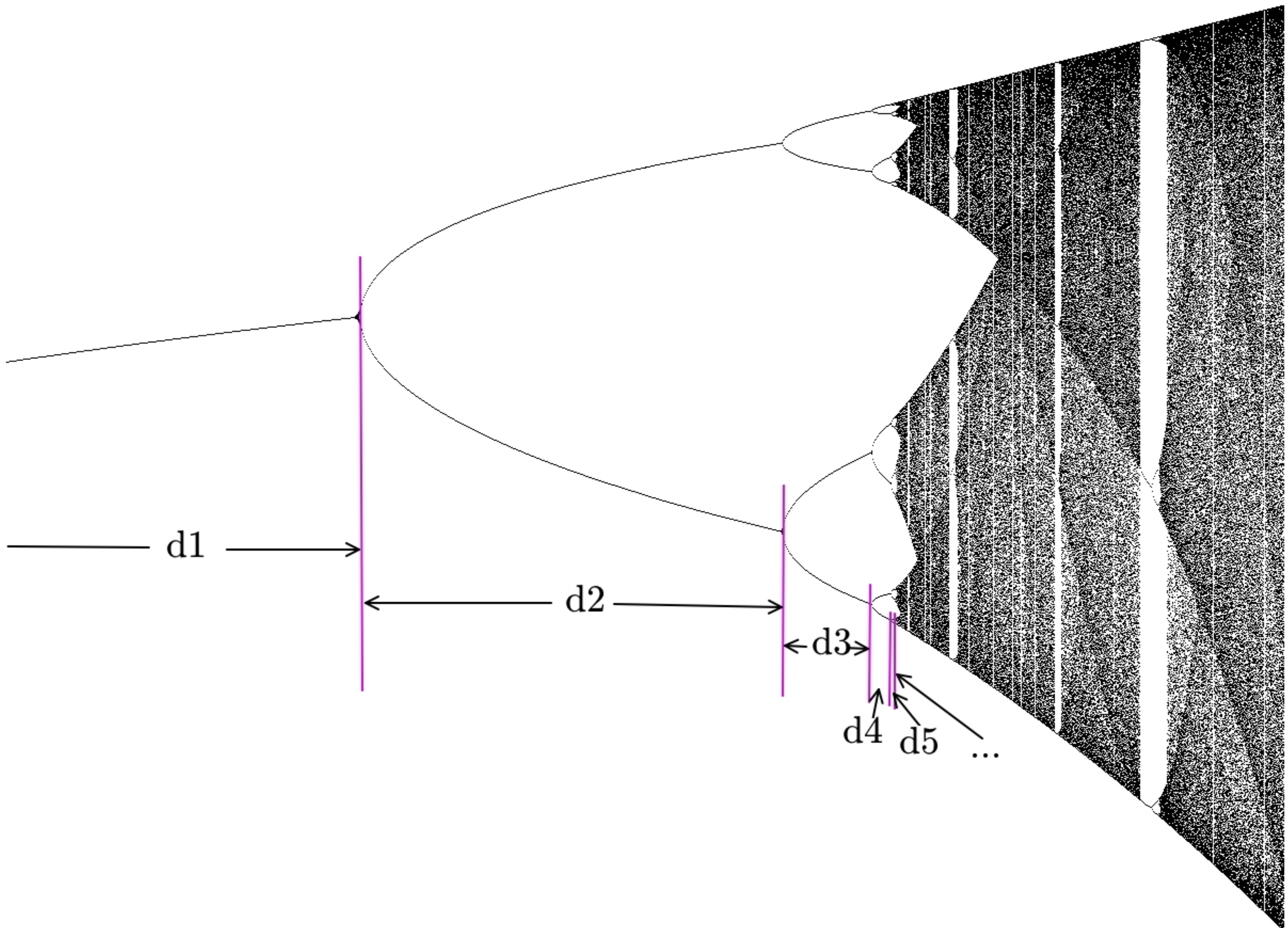
The value of δ has been calculated (to many decimal places):

$$\delta = 4.669201609102990671853203 82\ldots$$

This value is approached as R approaches the accumulation point where the bifurcation cascade ends, and a chaotic region begins:

$$R_\infty = 3.569934669\ldots$$

Feigenbaum also made a more striking and important observation. This process (period doubling bifurcation cascade route to chaos) occurs in may systems, and the value of δ is *universal* for a broad range of systems. This universality property opened up the possibility of finding *quantitative* properties, rather than just *qualitative* properties, for a variety of nonlinear systems.



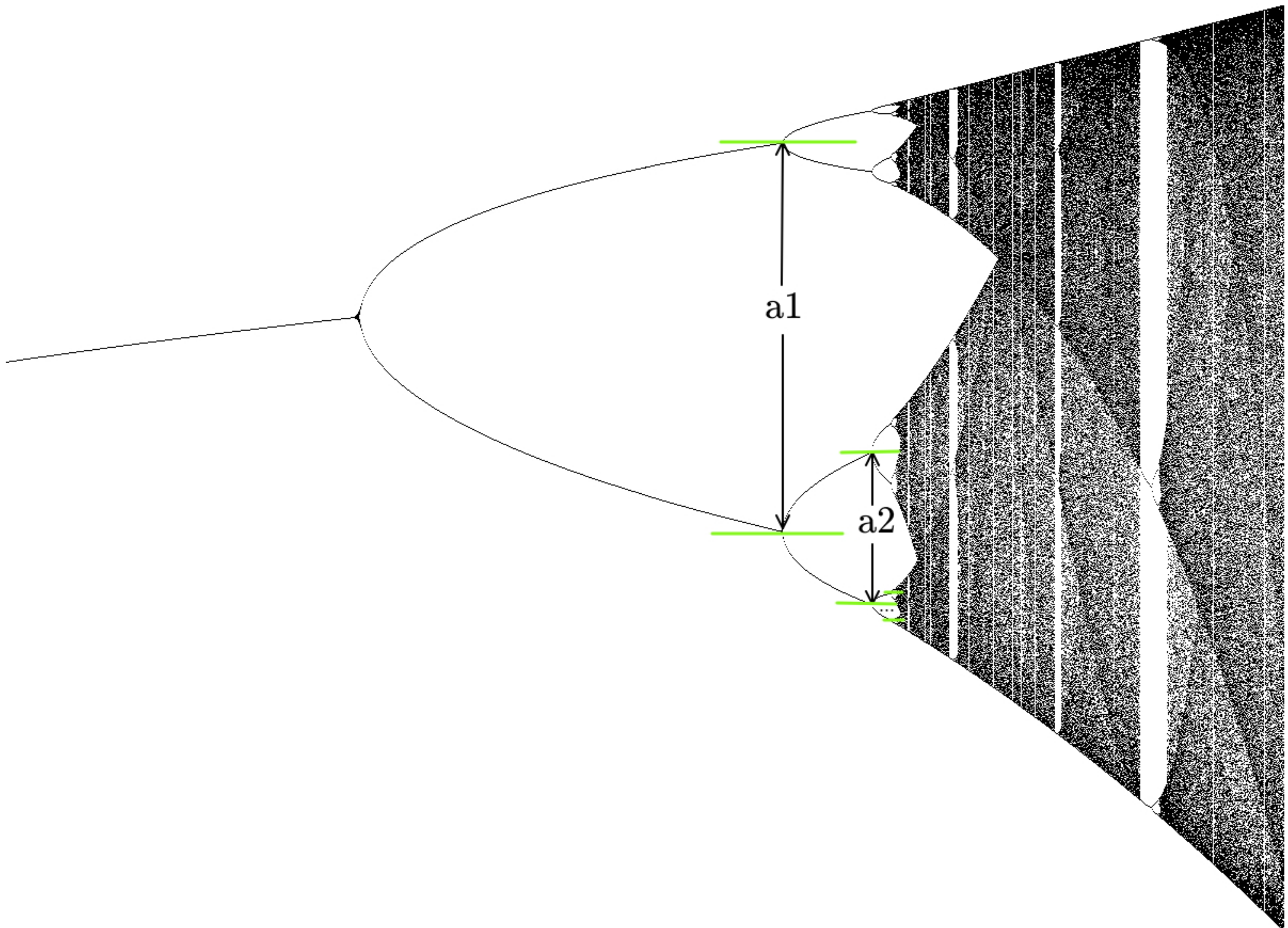
R=2.6-R=4.0-x=0.0-x=1.0-iter=1500 $\mathbf{d_i}$

Feigenbaum also discovered another *universality* principle concerning the width of the tines in the pitchfork bifurcation diagrams. In this case, we look at the successive widths of the tines, and find that the limiting value of the ratios of these widths is

$$\alpha = 2.5029078750958928222839028732182\ldots$$

On the next page is a brief pictorial overview.

For more discussion of these ideas, see, for example, Fiegenbaum [19] and Fillion [20].



R=2.6-R=4.0-x=0.0-x=1.0-iter=1500 $\mathbf{a_i}$

An example of the range of universality of these properties can be seen from the following conditions, which are sufficient to give δ as the limiting value for the bifurcation cascade:

1. $f : [0,1] \to \mathbb{R}$ is continuous, with a unique differentiable maximum $\bar{x}$;

2. $f(0) = f(1) = 0$, $f(x) > 0$ for $x \in (0,1)$, f is strictly increasing on $(0, \bar{x})$, and strictly decreasing on $(\bar{x}, 1)$;

3. For some parameter value, f has two fixed points which are both unstable; and,

4. In the interval N containing $\bar{x}$ such that $|f'(x)| < 1$, f is concave downward.

The Шарковский Theorem ←

- The Шарковский Theorem tells us about *periodic points* of continuous functions. (Note: Шарковский is a Ukrainian mathematician. His name (here written in Cyrillic) is transliterated as "Sarkovskii" or "Sharkovsky" or "Sharkovski" or . . . Since I can, I've gone ahead and used the Cyrillic :-)

  Suppose$f(x)$ is a function. Then a point x is a *periodic point of period* n for f if

$$x = f^n(x) = f(f(\ldots(f(x))\ldots)),$$

  and

$$x \neq f^k(x) \text{ for } 1 \leq k < n.$$

The theorem looks at the special case of a continuous function $f : \mathbb{R} \to \mathbb{R}$.

Consider the following ordering of the positive natural numbers:

$$\begin{array}{cccccccc}
3 & 5 & 7 & 9 & 11 & \dots & (2n+1)\cdot 2^0 & \dots \\
3\cdot 2 & 5\cdot 2 & 7\cdot 2 & 9\cdot 2 & 11\cdot 2 & \dots & (2n+1)\cdot 2^1 & \dots \\
3\cdot 2^2 & 5\cdot 2^2 & 7\cdot 2^2 & 9\cdot 2^2 & 11\cdot 2^2 & \dots & (2n+1)\cdot 2^2 & \dots \\
3\cdot 2^3 & 5\cdot 2^3 & 7\cdot 2^3 & 9\cdot 2^3 & 11\cdot 2^3 & \dots & (2n+1)\cdot 2^3 & \dots \\
 & \vdots & & & & & & \\
\dots & 2^n & \dots & 2^4 & 2^3 & 2^2 & 2 & 1
\end{array}$$

(note that every positive natural number occurs exactly once in this list).

The Шарковский Theorem states that if $f : \mathbb{R} \to \mathbb{R}$ is continuous, and has a periodic point x_n of period n for some value n on this list, then it also has points y_m of period m for all m later in the list than n. In particular, if f has a point x_3 of period 3, then it has periodic points of all possible periods.

This leads to the famous saying that "period 3 implies chaos" (although this may need to be taken with a grain of salt . . .). There is a classic paper about this: *Period Three Implies Chaos* by Li and Yorke [24].

A Hint of Topics (likely :-) to Come: $\leftarrow$

- . . .

The Cantor Set ←

- An important example (of a fractal, among other things) and tool for studying nonlinear dynamical systems is the *Cantor set*. This set was developed by Georg Cantor in the 1880's, and has led a long and fruitful life.

  We can develop the Cantor set by starting with the unit interval, and iterating steps. At each step, we will remove the middle thirds of each (remaining) piece:

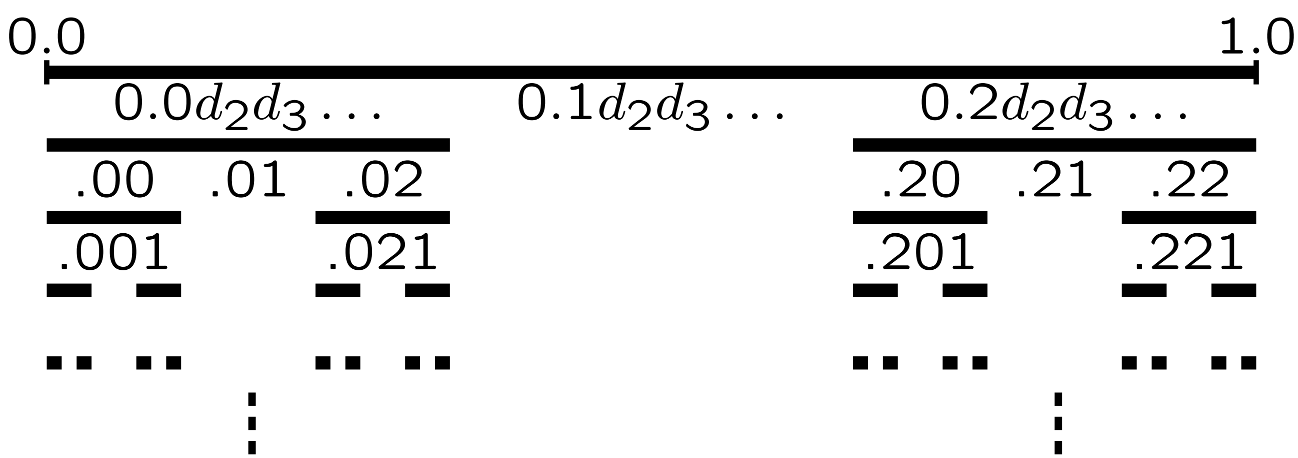
The Cantor set is what is left after we iterate this process countably many times.

We can calculate how much "length" is left after the process:

$$\begin{aligned}
length &= 1 - \left(\frac{1}{3} + \frac{2}{9} + \frac{4}{27} + \dots\right) \\
&= 1 - \left(\frac{2^0}{3^1} + \frac{2^1}{3^2} + \frac{2^2}{3^3} + \dots\right) \\
&= 1 - \frac{1}{3}\sum_{n=0}^{\infty}\left(\frac{2}{3}\right)^n \\
&= 1 - \frac{1}{3}\left(\frac{1}{1-\frac{2}{3}}\right) \\
&= 1 - 1 \\
&= 0.
\end{aligned}$$

It may appear from this that there is nothing left in the set, but that is far from the case.

Let's look at the set again, but labeling things with ternary (base 3) numbers. Each point in the interval $[0, 1]$ can be written as a ternary number $0.d_1d_2d_3\ldots$ where each of the digits d_i is 0, 1, or 2:



In this form, we can see that in the first step, we remove all points of the form $.1d_2d_3d_4\ldots$. In the second step, we remove all points of the form $.01d_3d_4\ldots$ and $.21d_3d_4\ldots$. In the nth step, we remove all points that have their first (ternary) digit 1 in their nth place. We will be left with (only) all points that have no 1 in their ternary expansion.

Note: We do have the (potential) problem that some points, such as $0.222\ldots = 1.0$ have two representations. We will use the representations with the repeated 2s. Thus, for example, the point $0.0222\ldots = 0.1000\ldots = \frac{1}{3}$ will not be removed.

In this form, we can see that there is a $1-1$ onto function from the Cantor set to the interval $[0,1]$ – we simply replace every 2 in the ternary expansion with a 1. From this, we can see that the Cantor set is an uncountable set. In other words, even though we have removed "length" 1, we nonetheless have uncountably many points still in the set.

Definition of chaos ←

- At some point it becomes worthwhile to work with a specific definition of "chaos," think some about to what extent it might be a useful definition, and to see the definition in action.

  The definition we will work with is (generally) attributed to Robert Devaney [17]. Our version is not exactly Devaney's, but is fairly close. It contains three pieces, which do a reasonable job of capturing intuitive ideas of deterministic chaos. We'll start by stating the definition, with some brief discussion of some aspects of the definition (e.g., redundancy among parts of the definition), and then apply the definition to some specific cases.

Definition. *Suppose X is a metric space, $f : X \to X$ is continuous, and $S \subseteq X$ is an infinite subset with the property that $f(S) = S$. Then we say that f is* chaotic *on S if*

1. f has sensitive dependence on initial conditions *on S.*

2. Periodic points of f are dense *in S, and*

3. f exhibits topological transitivity *on S.*

We need to clarify some terms in this definition.

- A *metric space* is a set of points (a *space*) X for which we have a *metric function* $d : X \times X \to \mathbb{R}$ such that, for all $x_1, x_2, x_3 \in X$,

  1. $d(x_1, x_2) \geq 0$,

  2. $d(x_1, x_2) = 0$ iff $x_1 = x_2$,

  3. $d(x_1, x_2) = d(x_2, x_1)$, and

  4. $d(x_1, x_3) \leq d(x_1, x_2) + d(x_2, x_3)$.

  For such a space, we will use the topology induced by the metric (on both X and S), so that, for example, continuity of a function is relative to the metric topology. The metric topology is the topology generated by ϵ-neighborhoods (for any $\epsilon > 0$):

$$B(x_0, \epsilon) = \{x \in X | d(x_0, x) < \epsilon\}.$$

- For *sensitive dependence on initial conditions* (SDIC), we will use the property that, for any $s_0 \in S$, and for any ϵ-neighborhood $B(s_0, \epsilon)$ of s_0, there exist $s_1 \in B(s_0, \epsilon)$, $\delta > 0$, and $\lambda > 0$ such that

$$d(f^n(s_0), f^n(s_1)) \approx e^{\lambda n}\delta$$

(at least for a while ...). Here $f^n(s)$ is the nth iteration of f on s (i.e., $f^n(s) = f(f(\ldots f(s)) \ldots)) = f \circ f \circ \ldots \circ f \circ f(s))$.

At this point, this is not a "mathematical" statement – we would need to clarify what we mean by "$\approx$" and "for a while".

In fact, for at least some of the examples we will work with, we can use the more specific (and in some sense stronger) condition that there exists a $\lambda > 0$ such that, for any $s_0 \in S$, for any ϵ-neighborhood $B(s_0, \epsilon)$ of s_0, for any $n > 0$, there exist $\delta > 0$ and $s_1 \in B(s_0, \epsilon)$ such that

$$d(f^k(s_0), f^k(s_1)) \geq e^{\lambda k}\delta \text{ for } k \leq n.$$

We use this sort of notation (and definition) to correspond with the concept of the ляпуно́в *(lyapunov) exponents* of continuous dynamical systems.

This condition says that no matter how closely we specify initial conditions (s_0, but with "error" of a small ϵ), our prediction error can grow exponentially . . .

There is ongoing discussion among researchers about this part of the definition (see, for example, the discussion in Bishop [4]). In particular, Devaney originally actually only required *linear growth* rather than *exponential growth*.

- The condition that *periodic points are dense* means that for any ϵ-neighborhood $B(s_0, \epsilon)$, there exist an $n > 0$ and an $s \in B(s_0, \epsilon)$ such that $f^n(s) = s$.

  One way to think about this is that no matter where we look in a *chaotic* system, we will still see regularities (periodic points) – in particular, our system is not just *pure noise*, but has structure within it.

  This is another area of ongoing discussion (again, see Bishop [4]).

- The condition of topological transitivity means that for any two points $s_0, s_1 \in S$, and for any ϵ-neighborhoods $B(s_0, \epsilon_0)$ and $B(s_1, \epsilon_1)$ of these two points, there is a point $s \in B(s_0, \epsilon_0)$ with $f^n(s) \in B(s_1, \epsilon_1)$ for some n.

  This condition is often called *mixing*. Since the point s_1 can be anywhere in S, this means that points in any neighborhood of a given point s_0 will, over time, spread throughout all of S, and this is the case for any $s_0 \in S$. Thus, over time, the system gets thoroughly mixed (despite the densely distributed periodic points).

  Some workers in the field take this property to be the most important property of *chaos* (i.e., the system looks "noisy," in the sense "most points get sent almost anywhere over time" . . .).

  This also is an area of active discussion among workers in the field . . .

- It turns out that this definition (well, more precisely, Devaney's original definition) is not minimal. For example, Banks, et al.[1] showed that condition 1. (sensitive dependence on initial conditions) follows from 2. and 3. (periodic points are dense, and topological transitivity). We could thus leave condition 1. out of the definition, and not change the collection of systems we would call *chaotic*. This also says that despite the fact that condition 1. (sensitive dependence) looks like a *metric* property, in fact *chaos* (as so defined) is really a topological property (mixing, with periodicity).

- We can also observe that one way to show that a system is *chaotic* is to show that it is homeomorphic to a system we have already shown to be *chaotic*. In general form, if we have a commutative diagram

$$
\begin{tikzcd}
X \arrow[r, "f"] \arrow[d, "h"'] & X \\
Y \arrow[r, "g"] & Y \arrow[u, "h^{-1}"']
\end{tikzcd}
$$

where $h : X \to Y$ is a homeomorphism, and $g : Y \to Y$ is known to be chaotic on Y, then we can conclude that f is also chaotic on X.

The general idea here is known as *topological conjugacy*.

- It is also worth noting that this definition is far from universally accepted among mathematicians and/or physicists. For discussion of some of the issues, see Bishop[4] and Knudsen[23]. But, to a reasonable extent, this definition captures a reasonable amount of what a variety of workers in the field seem to have in mind when they speak of *chaos*.

  In any case, this is a useful definition. As we will see, there are systems that are very easy to describe that are *chaotic* under this definition, and, conveniently, the proofs that they are chaotic are relatively straightforward. Of course, it shouldn't be any great surprise that the proofs are comparatively easy in at least some specific cases, because part of the motivation for choosing a specific "definition" is to make particular proofs relatively easy . . .

A Chaotic Map $\leftarrow$

- Let's look at a specific map. Consider $f : [0,1] \to [0,1]$ given by

$$f(x) \;= \begin{cases} 2x & \text{if } 0 \leq x < \frac{1}{2} \\ 2x - 1 & \text{if } \frac{1}{2} \leq x \leq 1. \end{cases}$$

  A convenient way to look at the map (iterations) of the function is to use the binary representation of the unit interval. Each point in the unit interval has a representation as $x = 0.d_1d_2d_3d_4\ldots$ where each binary digit $d_i \in \{0,1\}$. Our function f is then realized as

$$f(x) = f(0.d_1d_2d_3d_4\ldots) = 0.d_2d_3d_4\ldots$$

  In other words, our function is just a *binary shift* – it shifts all the binary digits one place to the left (and throws away the leading digit if it was 1).

Note: as we did with the Cantor set, we will make the binary representations unique by using repeating 1's, so, for example, we will use 0.111... instead of 1.000...

We will now claim that this function f is *chaotic* of the entire unit interval. Let's check against our 3-part definition:

1. *Sensitive Dependence on Initial Conditions*: Suppose we are given $x_0 \in [0,1]$, $\epsilon > 0$, and $n > 0$.

   Let $\lambda = \ln(2)$ so that $e^{\lambda k} = 2^k$. Let $m \geq 0$ be such that $2^{-(m+n)} < \epsilon$. Let $\delta = 2^{-(m+n+1)}$. Now, if $x_0 = 0.d_1 d_2 d_3 d_4 \ldots$, we let $x_1 = 0.d_1 d_2 d_3 d_4 \ldots d_{m+n}(1 - d_{m+n}) d_{m+n+1} d_{m+n+2} \ldots$. In other words, we change the $m+n$ digit of x_0 (from 0 to 1, or 1 to 0) to get x_1.

We then will have that $d(x_0, x_1) = 2^{-(m+n)} < \epsilon$, and $d(x_0, x_1) > 2^{-(m+n+1)} = \delta$. Now, as we iterate, we get

$$d(f^k(x_0), f^k(x_1)) = 2^k * 2^{-(m+n)} > 2^k * \delta = e^{\lambda k}\delta$$

for $k \leq n$, as desired. In other words, we get local exponential spreading of iterates of (at least some) nearby points.

2. *Periodic Points Are Dense*: This is easy . . . Every rational number in $[0, 1]$ has periodic binary expansion, and therefore is periodic under the shift map.

3. *Topological Transitivity* (mixing): Suppose we are given $x_0, x_1 \in [0, 1]$, $\epsilon_0 > 0$, and $\epsilon_1 > 0$. Then, we let n be such that $2^{-n} < 2^{-2}\epsilon_0$. Now if $x_0 = 0.d_1d_2d_3d_4\ldots$ and $x_1 = 0.e_1e_2e_3e_4\ldots$, we let $x = 0.d_1d_2d_3d_4\ldots d_ne_1e_2e_3e_4\ldots$. In other words, x has

digits from x_0 for the first n positions, and then digits of x_1 concatenated after that. We will then have that $x \in B(x_0, \epsilon_0)$, and $f^n(x) = x_1 \in B(x_1, \epsilon_1)$, as desired.

From this, we see that the binary shift map is *chaotic* on the unit interval $[0, 1]$ (under the definition we have outlined above . . .).

Note: Full disclosure – the binary shift map isn't continuous on $[0, 1]$ – in particular, there is a discontinuity at $1/2$. With some care, we can handle that, so let's just move forward :-)

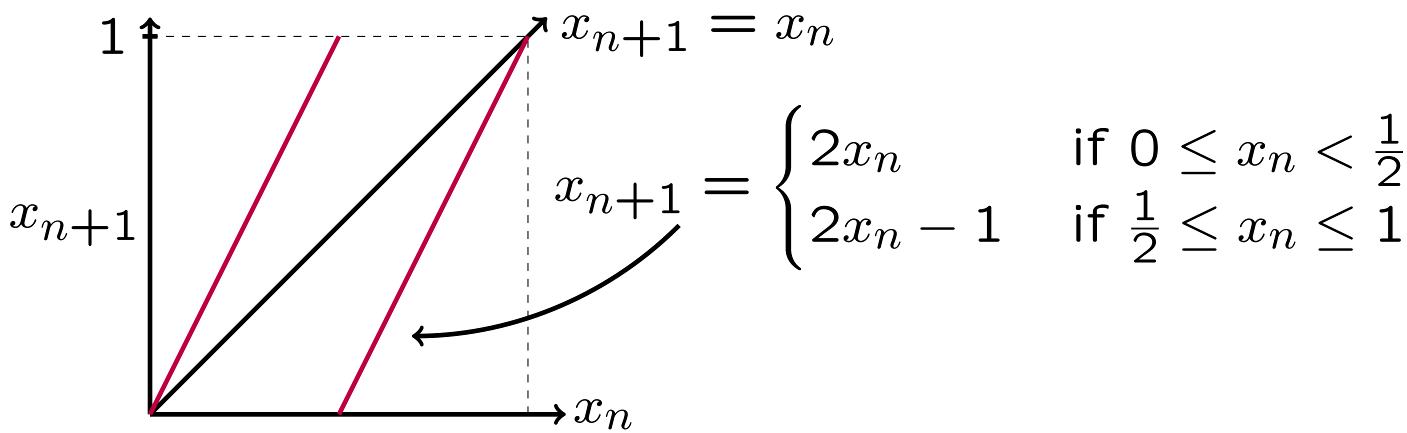
Some other chaotic maps

- Now that we have an example of a chaotic map, we can use the "topological conjugacy" idea to show that some other related maps are chaotic.

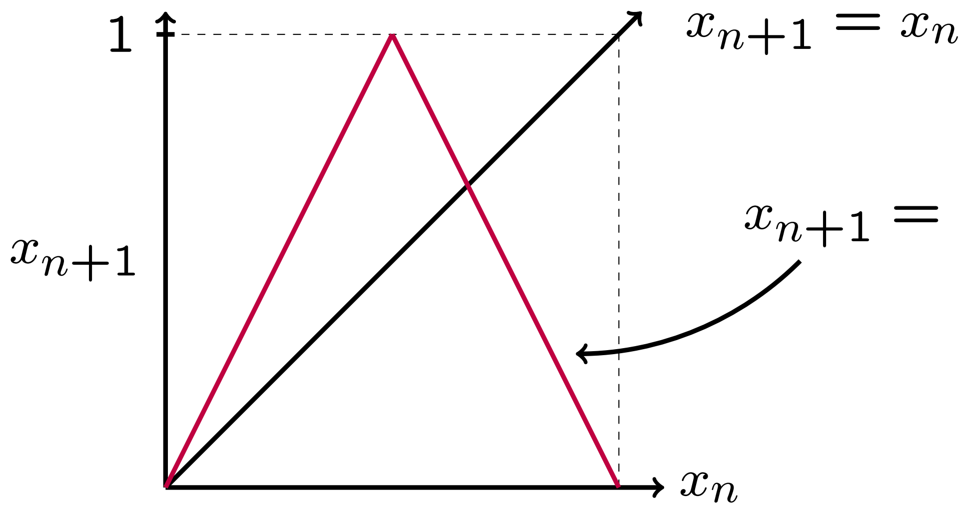
  Let's start with a picture of the binary shift map:

$$f(x) \;=\begin{cases} 2x & \text{if } 0 \le x < \frac{1}{2} \\ 2x-1 & \text{if } \frac{1}{2} \le x \le 1 \end{cases}$$

  looks like this:



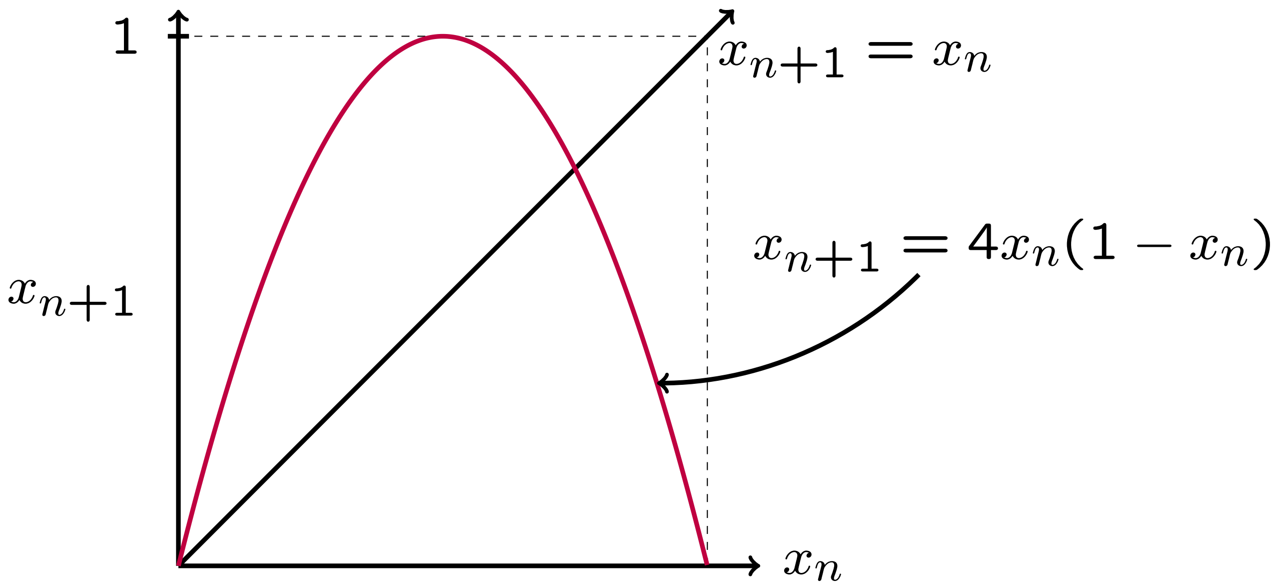
- If we flip the second half, we get the *tent map*, which looks like this:



$$x_{n+1} = \begin{cases} 2x_n & \text{if } 0 \leq x_n < \frac{1}{2} \\ 2 * (1 - x) & \text{if } \frac{1}{2} \leq x \leq 1 \end{cases}$$

A relatively straightforward topological conjugacy argument shows that this map also is chaotic.

- Now if we slightly modify the tent map (make is "smooth" at the top), we get the logistic map with $R = 4$, which looks like:

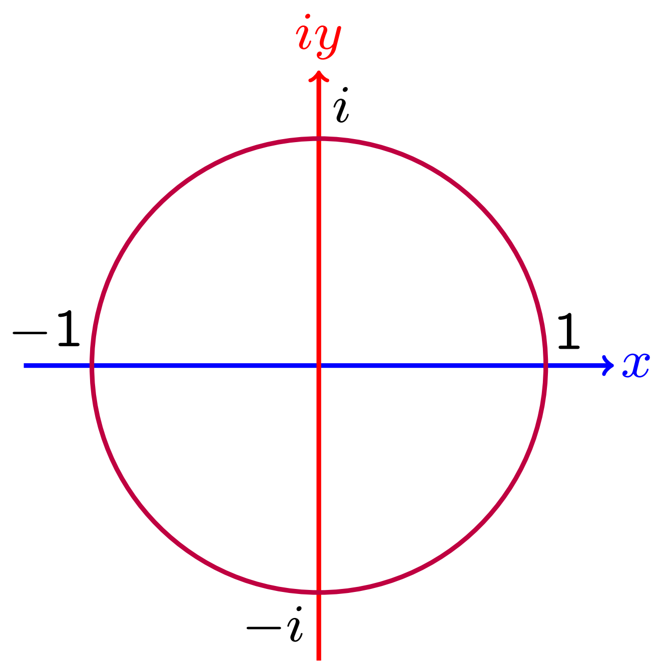


Again, by a relatively straightforward topological conjugacy argument, we can see that the logistics map for $R = 4$ is chaotic on $[0, 1]$.

- One more example: we'll now move into the realm of the complex numbers, $\mathbb{C} = \{x + iy | x, y \in \mathbb{R}, i^2 = -1\}$.

  Consider the unit circle in the complex plane $C_1 = \{z \in \mathbb{C} | \ |z| = 1\}$, where $|z| = |x + iy| = (x^2 + y^2)^{\frac{1}{2}}$.

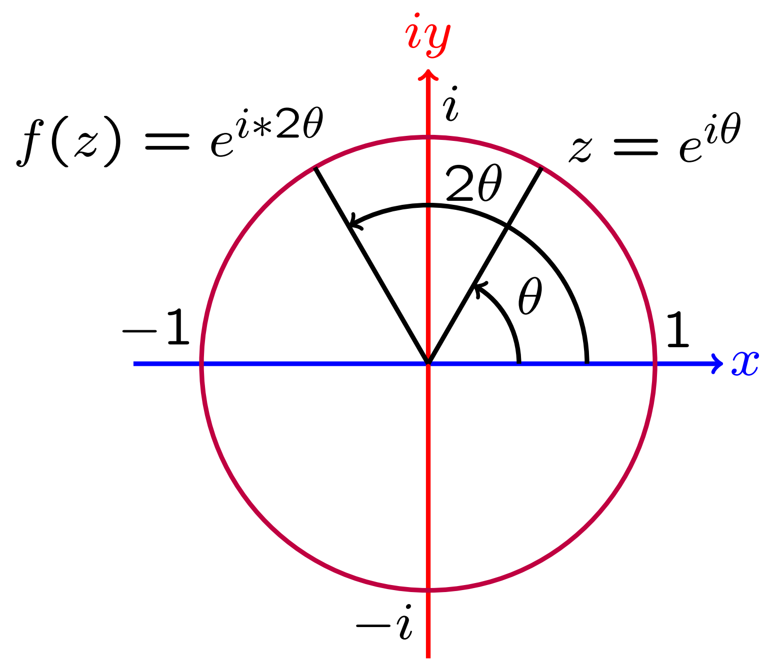
  C_1 looks like:



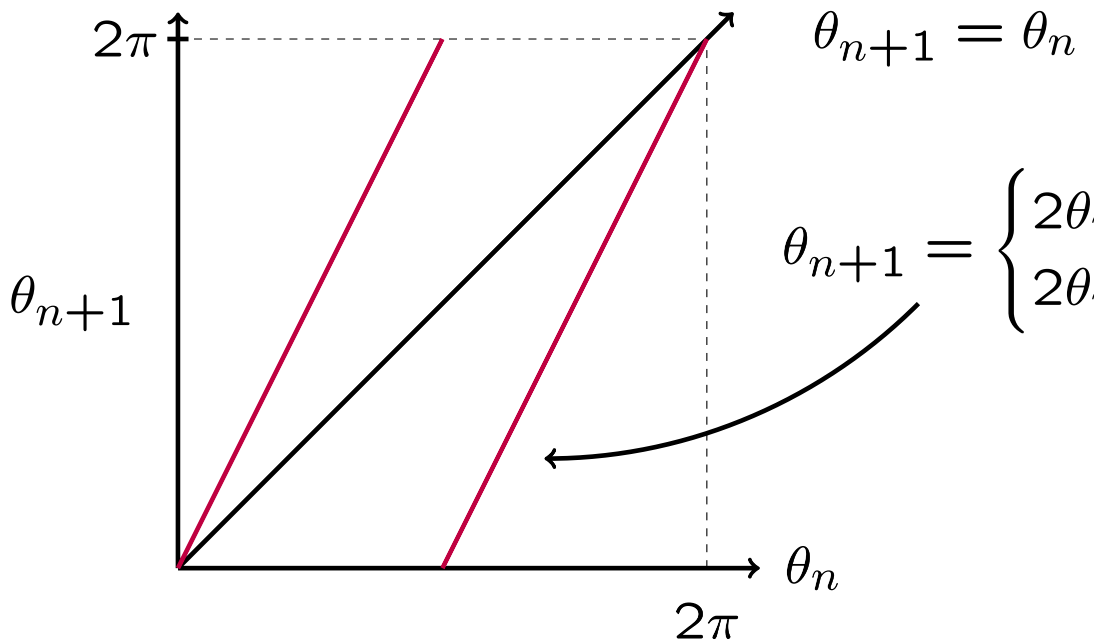
Now consider $f : C_1 \to C_1$ given by $f(z) = z^2$. We can observe that any point in C_1 can be written as $z = x + iy = \cos(\theta) + i\sin(\theta)$ for some θ, and then remember (from above) that $\cos(\theta) + i\sin(\theta) = e^{i\theta}$. We then see that our function $f(z)$ is given by

$$f(z) = f(e^{i\theta}) = (e^{i\theta})^2 = e^{i*2\theta}.$$

In pictorial form:



Another way to look at this (just keeping track of θ) is:



$$\theta_{n+1} = \begin{cases} 2\theta_n & \text{if } 0 \leq \theta_n < \pi \\ 2\theta_n - 2\pi & \text{if } \pi \leq \theta_n \leq 2\pi \end{cases}$$

Once again, it is easy to see that this map is chaotic on C_1.

Julia Sets ←

- Now we are about ready to dive in to the *Mandelbrot set*. We'll start with Julia sets.

  We observed that $f(z) = z^2$ is chaotic on the unit circle in $\mathbb{C}$. In particular, the function keeps elements of the unit circle on the unit circle. In other words, if $|z| = 1$, then $|z^2| = 1$ also.

  On the other hand, if $|z| < 1$, then iterates of z go to 0, which is an attracting fixed point of the map. For any point z with $|z| < 1$, we have $\lim_{n\to\infty} f^n(z) = 0$. On the other hand, if $|z| > 1$, then $\lim_{n\to\infty} f^n(z) = \infty$.

The complex plane $\mathbb{C}$ thus breaks up into 4 pieces:

1. An *attracting fixed point* at 0,

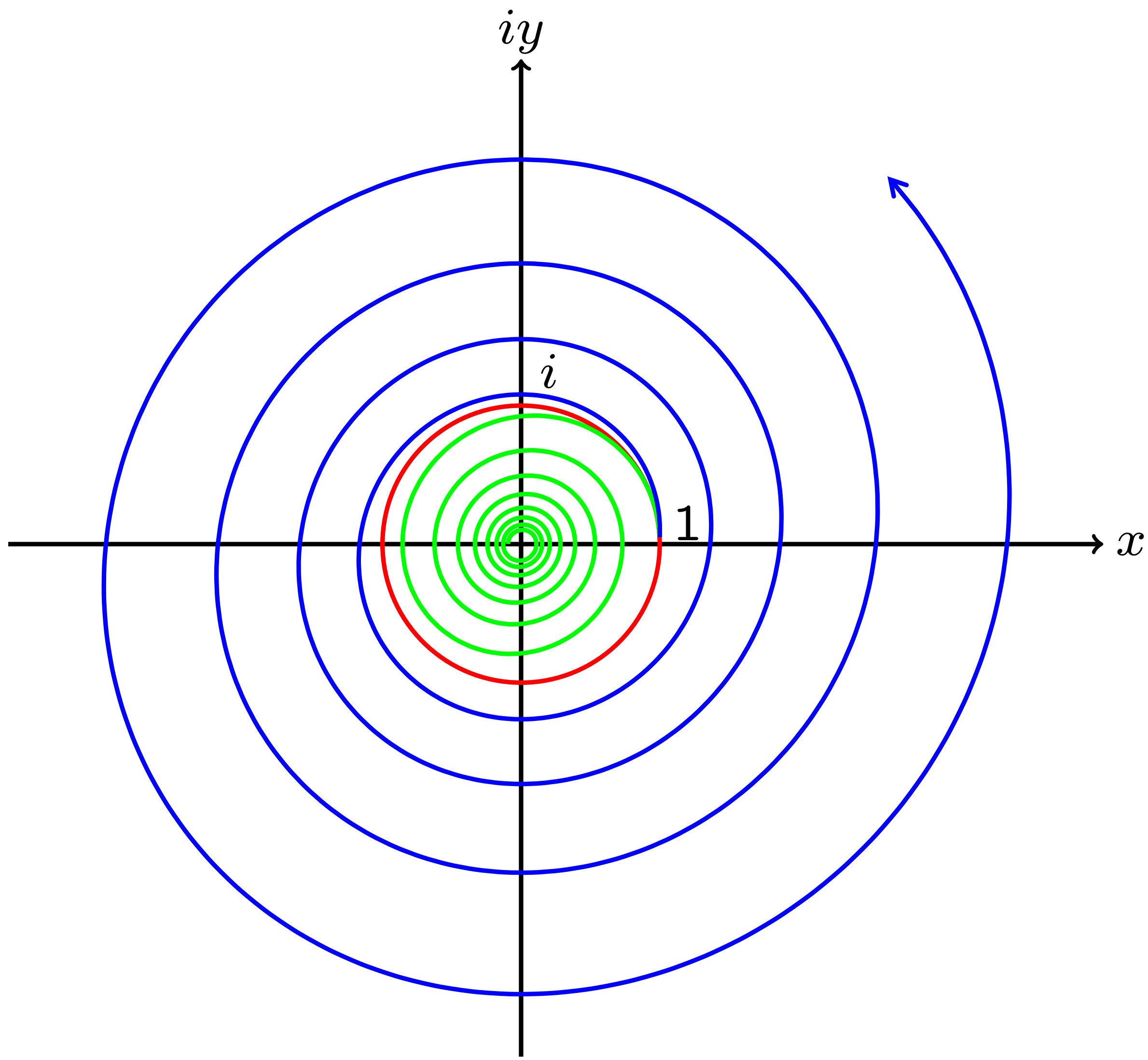
2. A *basin of attraction* of the attracting fixed point – in other words, a set of points whose trajectories go toward the attracting fixed point

3. A *basin of attraction* of the attracting point at infinity (more on this later . . .), and

4. A *boundary* between these two basins – in this case, the unit circle.

In the following picture, regions 2, 3, and 4 are indicated in green, blue, and red.

This is a view of the regions associated with $f(z) = z^2$:



- The set of points that form the boundary of the basin of attraction of ∞ are particularly interesting. In the case of the function we have been looking at ($f(z) = z^2$), the boundary is just the unit circle ($\{z \mid |z| = 1\}$).

  Now let's generalize slightly, and consider the set of functions $f_c(z) = z^2 + c$, for $c \in \mathbb{C}$. We have been looking at $f_0(z)$.

  For these functions, we can define the *Julia Set* of the function by:

  $$J(f_c(z)) = \text{the boundary of the basin of attraction of } \infty$$

  where the *basin of attraction of* ∞ is the set

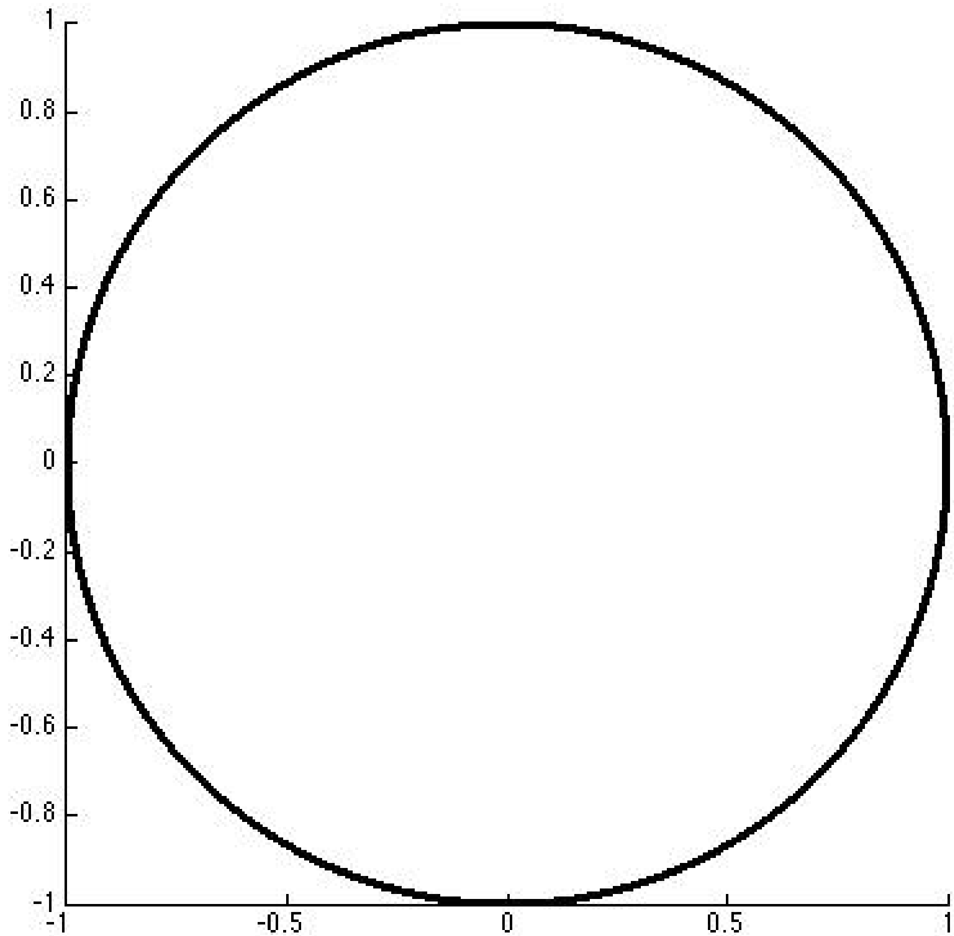
  $$B(\infty) = \{z \mid \lim_{n\to\infty} f_c^n(z) = \infty\}.$$

On the following pages are some examples of Julia sets for various values of c. These are actually calculated by inverse iteration of the relevant function:

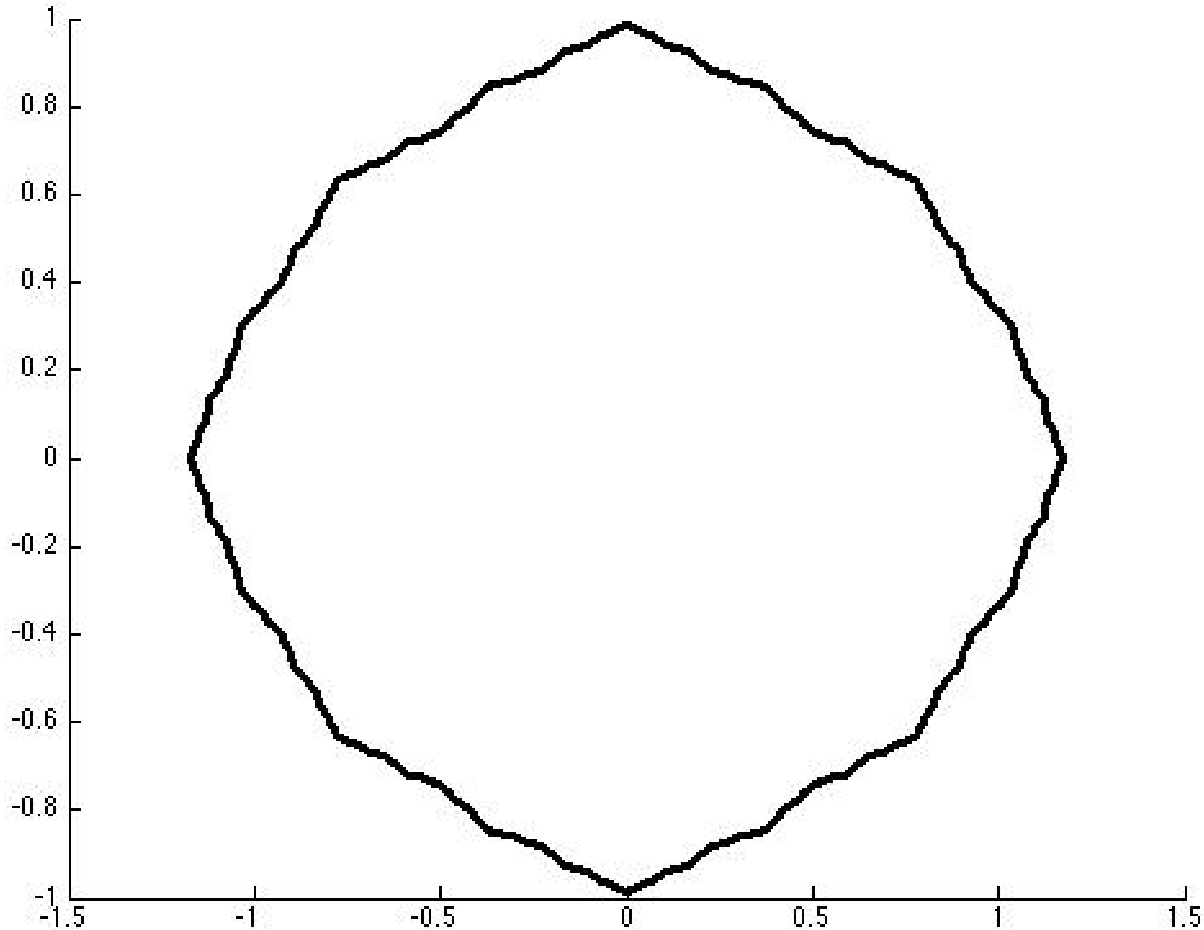
$$f_c^{-1}(z) = sqrt(z - c)$$

However, since there are two square roots, we randomly choose one of the square roots. Similarly to what we did in the bifurcation diagrams, we discard the first few iterations, and then plot from there on.

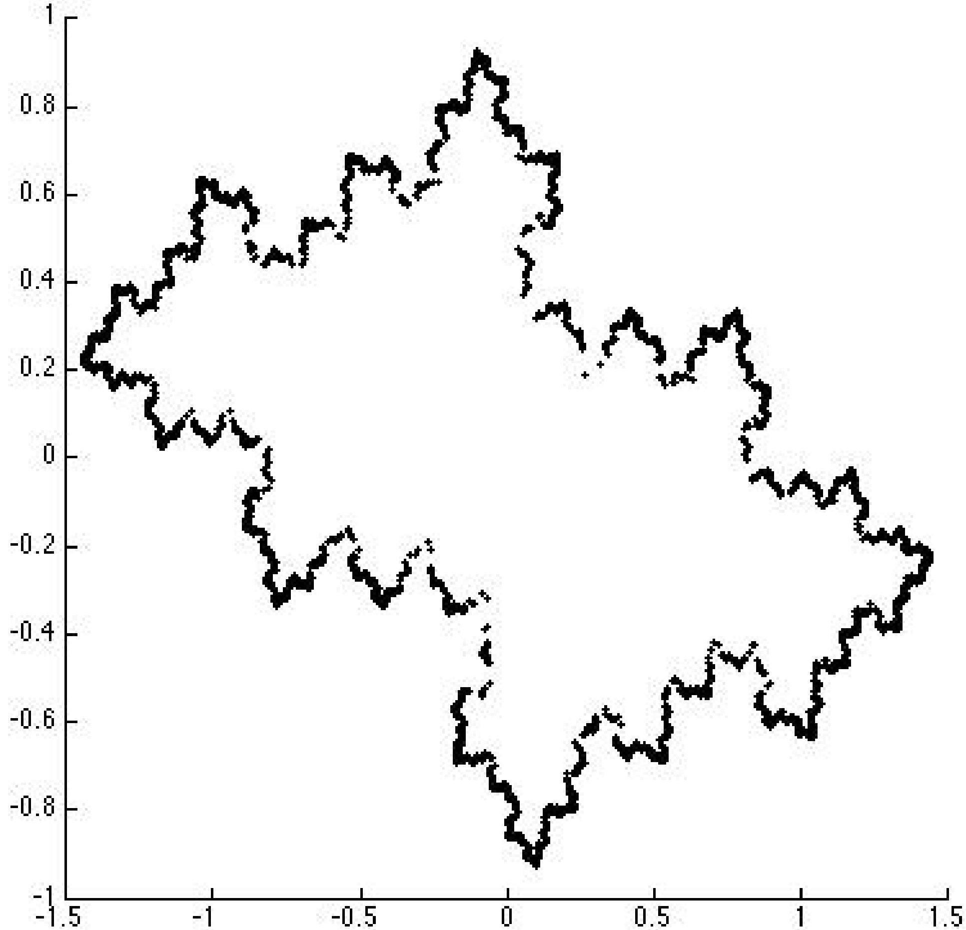
MatLab source code follows the Julia sets.



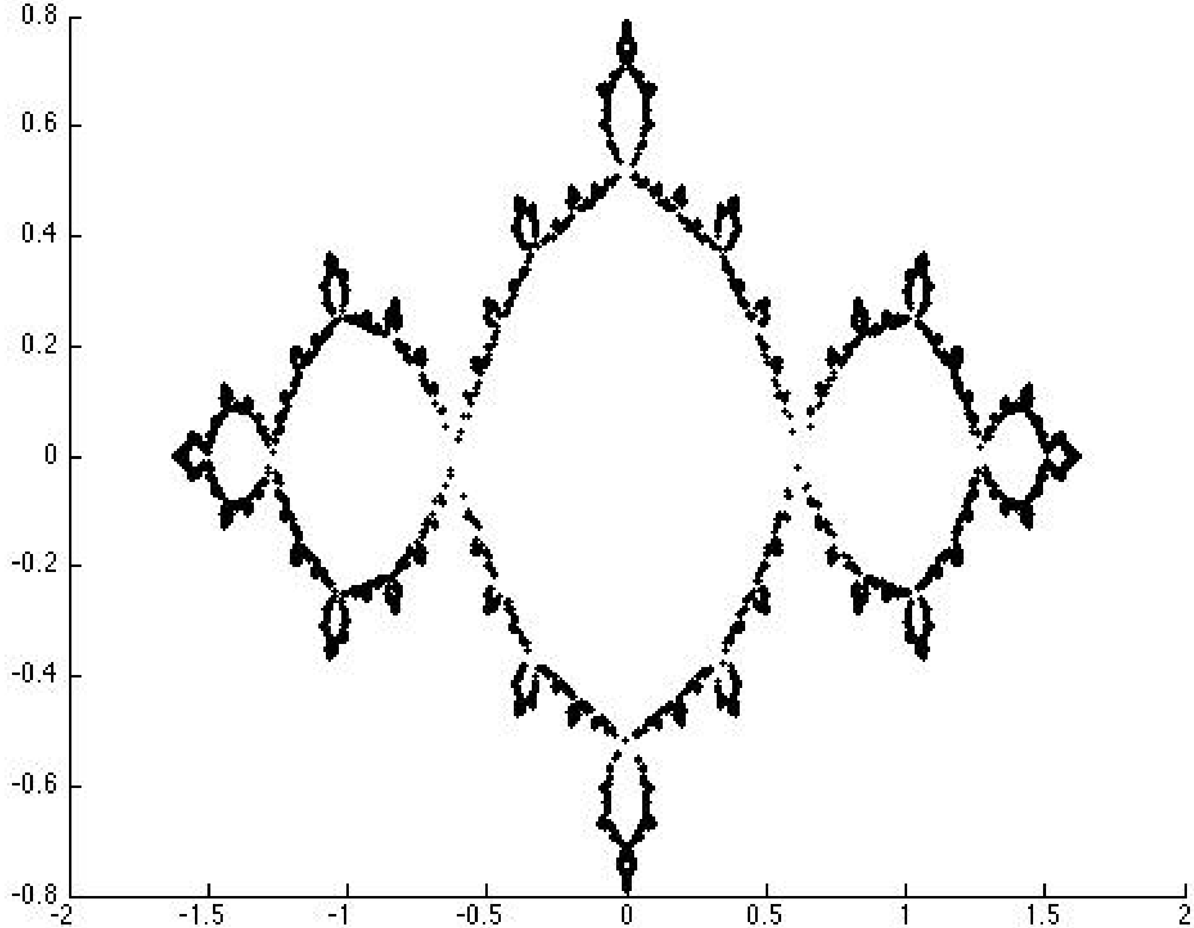
$c = 0.0 + 0.0i$



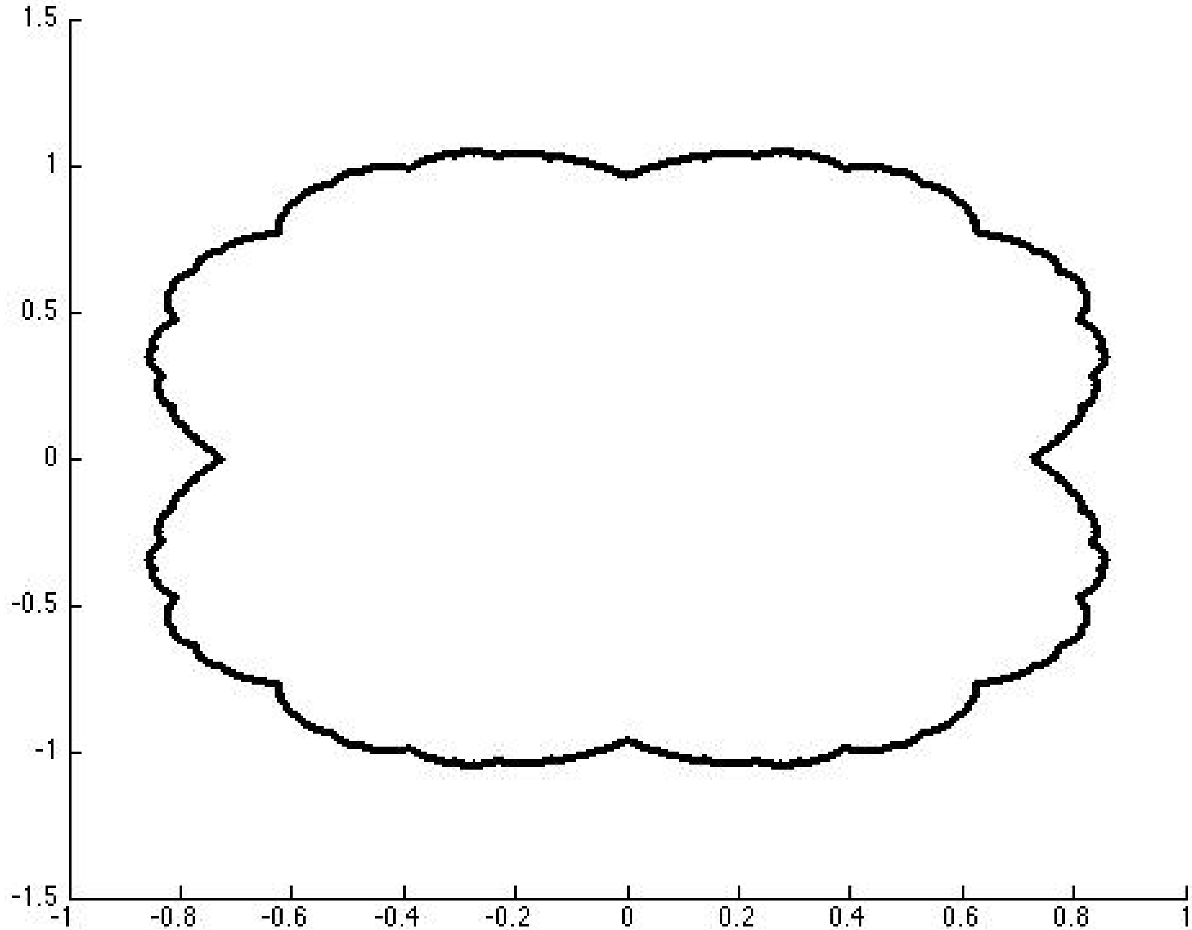
c = -0.2 + 0.0i



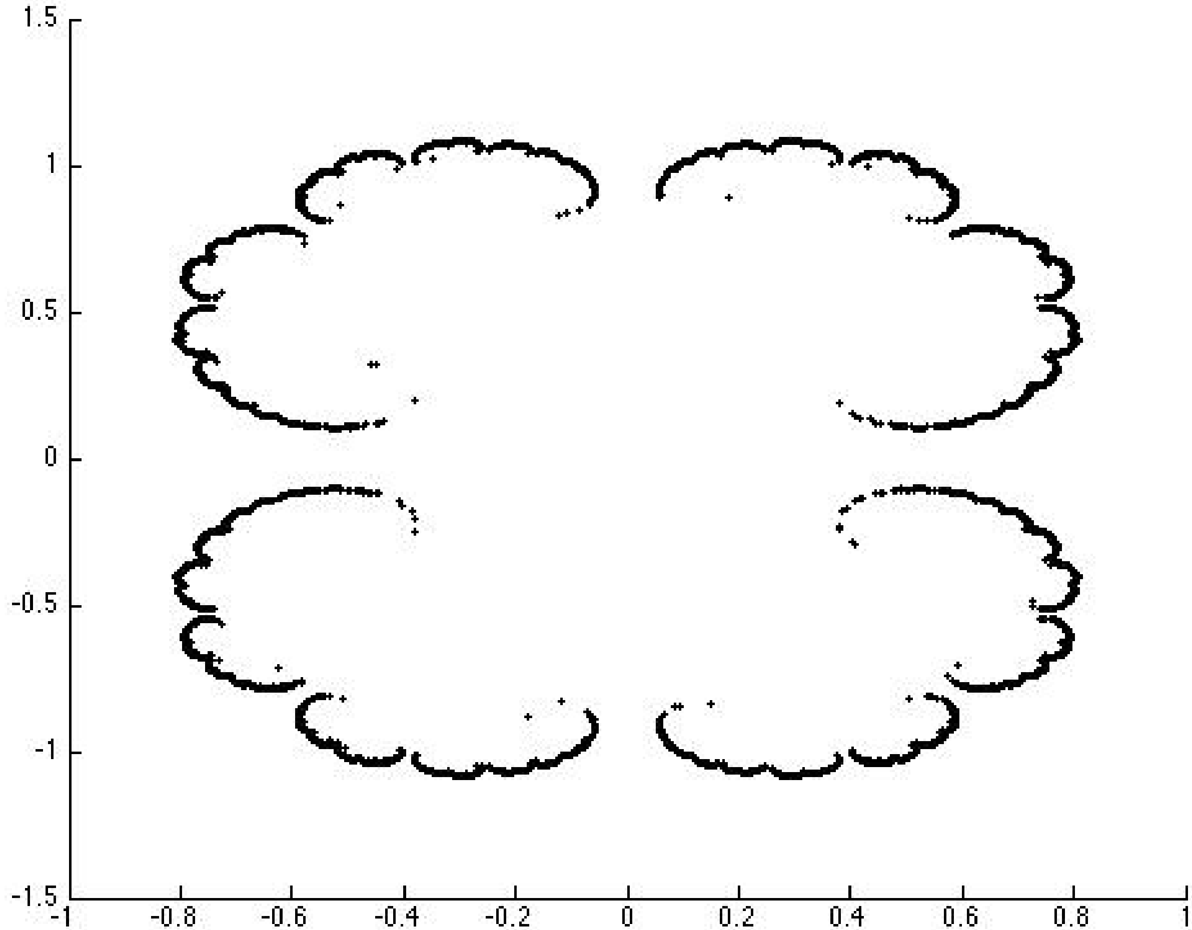
$c = -0.6 + 0.4i$



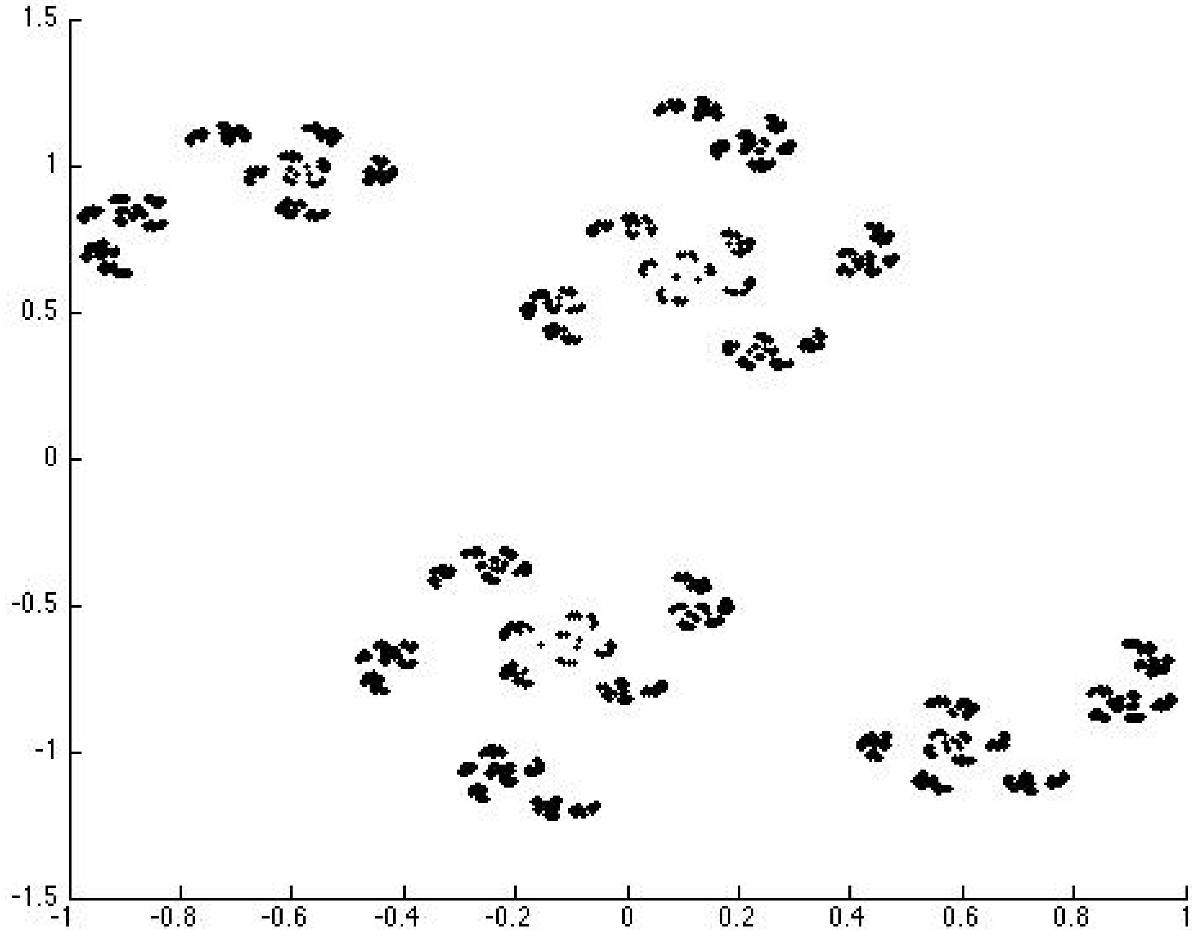
$c = -1.0 + 0.0i$



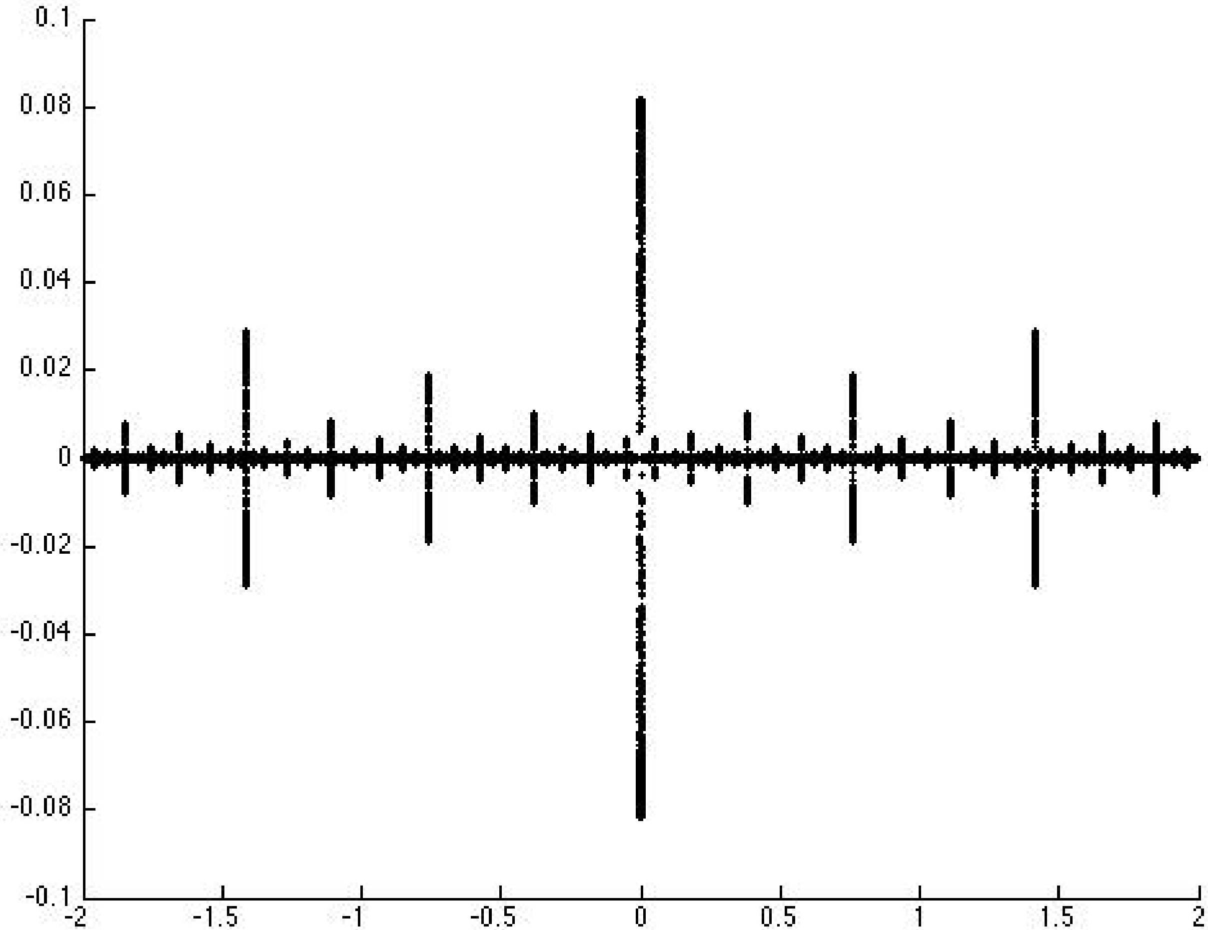
$c = 0.2 + 0.0i$



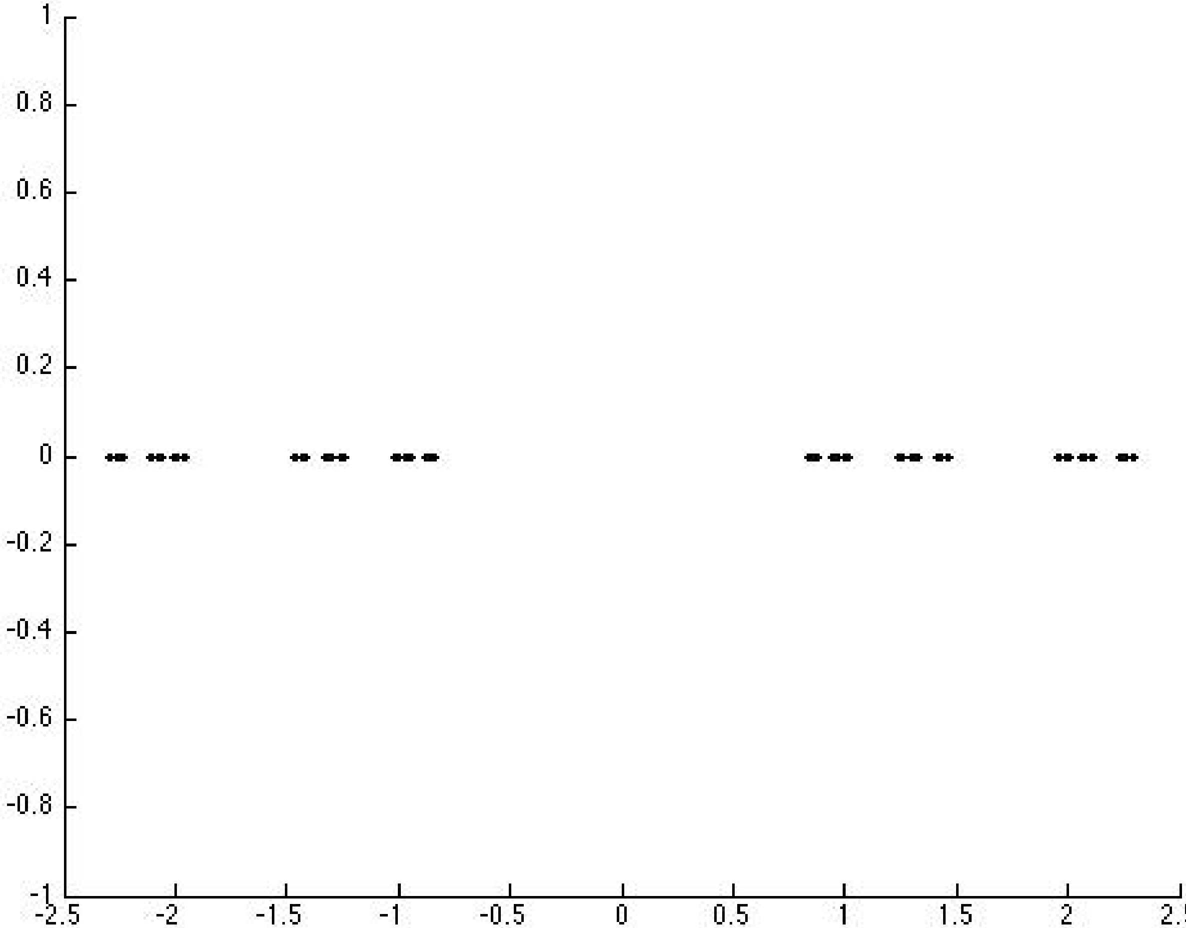
c = 0.3 + 0.0i



c = 0.5 + 0.5i



c = -1.99 + 0.0i



c = -3.0 + 0.0i

```
% Matlab source code:  myjulia.m
%
% calculating Julia sets by inverse iteration,
%    and random choice of positive or negative
%    square root
%
hold off; clf; clear
z = 0.00001+0.0i;
c = -0.6+0.4i
for n=1:20
    if (random('unid',2) == 1)
        z = sqrt(z - c);
    else
        z = (-1)*sqrt(z - c);
    end
end
```

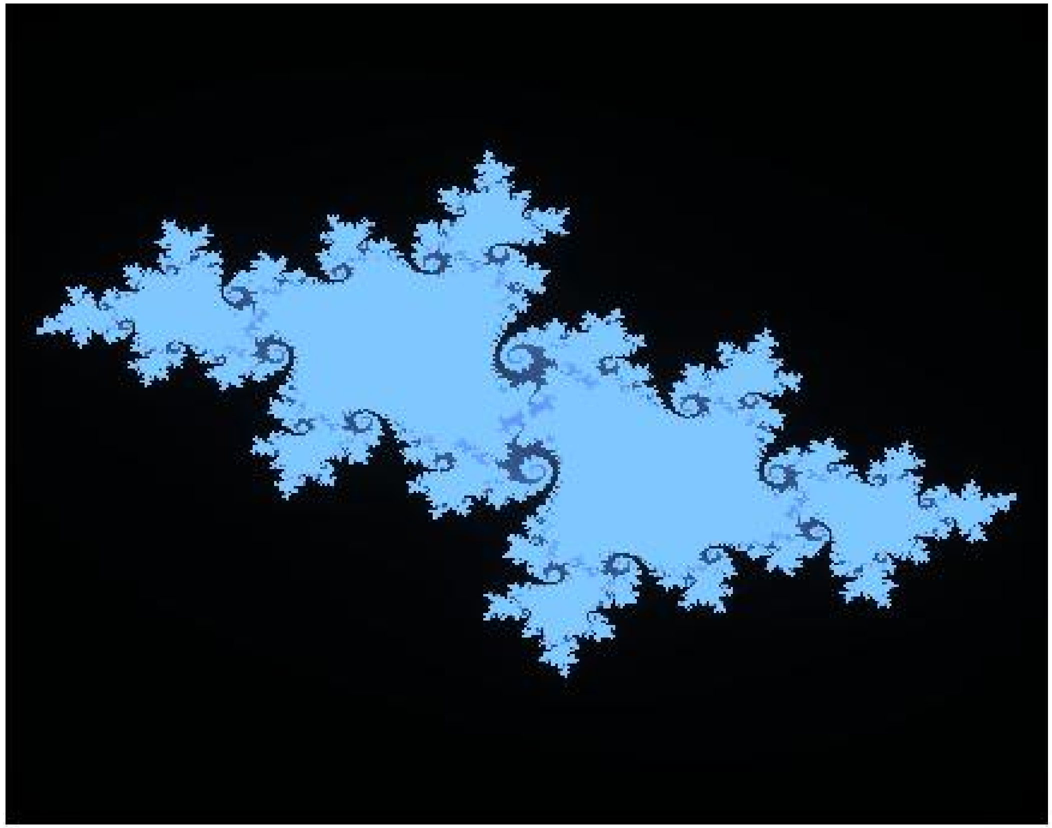
```
y=zeros(1,40000);
y(1) = z;
for n=2:40000
    if (random('unid',2) == 1)
        y(n) = sqrt(y(n - 1) - c);
    else
        y(n) = (-1)*sqrt(y(n - 1) - c);
    end
end

%pbaspect([1 1 1]);
hold on;
scatter(real(y),imag(y),4,[0 0 0],'filled');
```

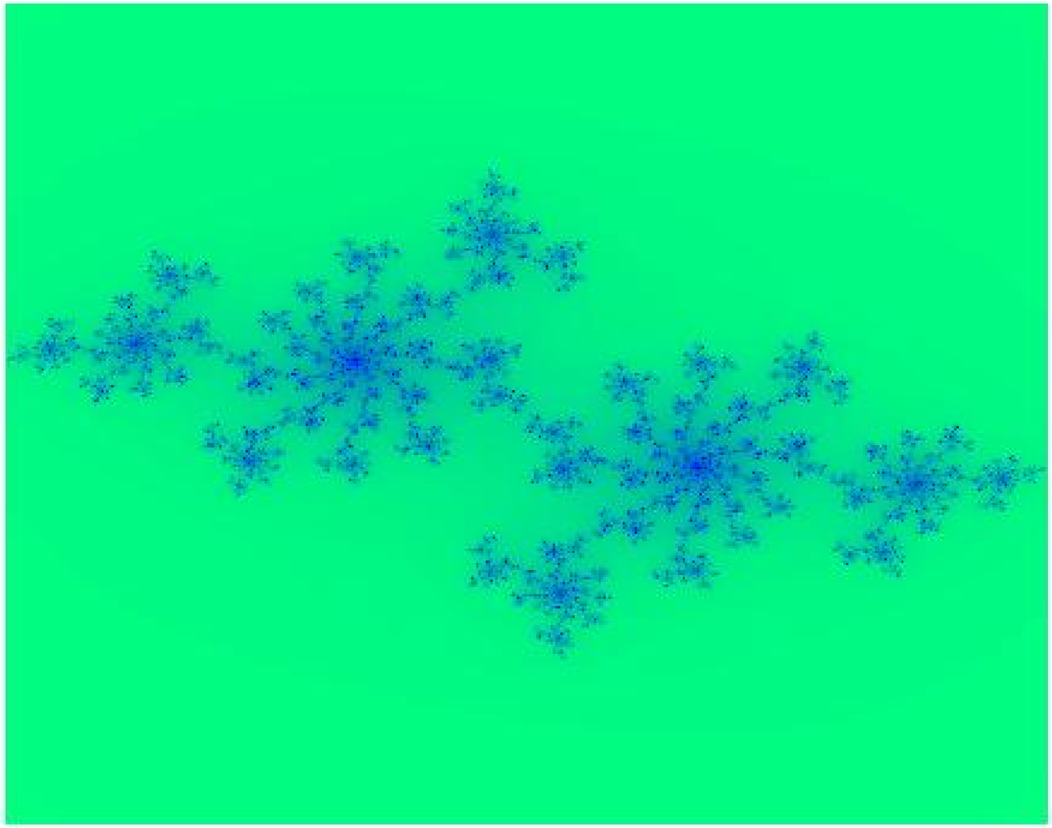
- There is another way to visualize Julia sets. Instead of doing inverse iteration, we look at the behavior of many different points in a region. In particular, we iterate the function $f_C(z)$ for each value of z in a rectangular region containing the Julia set. Most of the points will run off to ∞, at varying rates of speed.

  Hence, what we will do is assign a color to each pixel in the rectangle according to the "escape time" of the corresponding z value. In particular, for each point z, we will see how many iterations of the function it takes for the value to have magnitude $>$ 2. We will also have a cutoff for total iterations, because points in (or inside of) the Julia set will never run off to ∞. We will then assign a color according to the number of iterations.

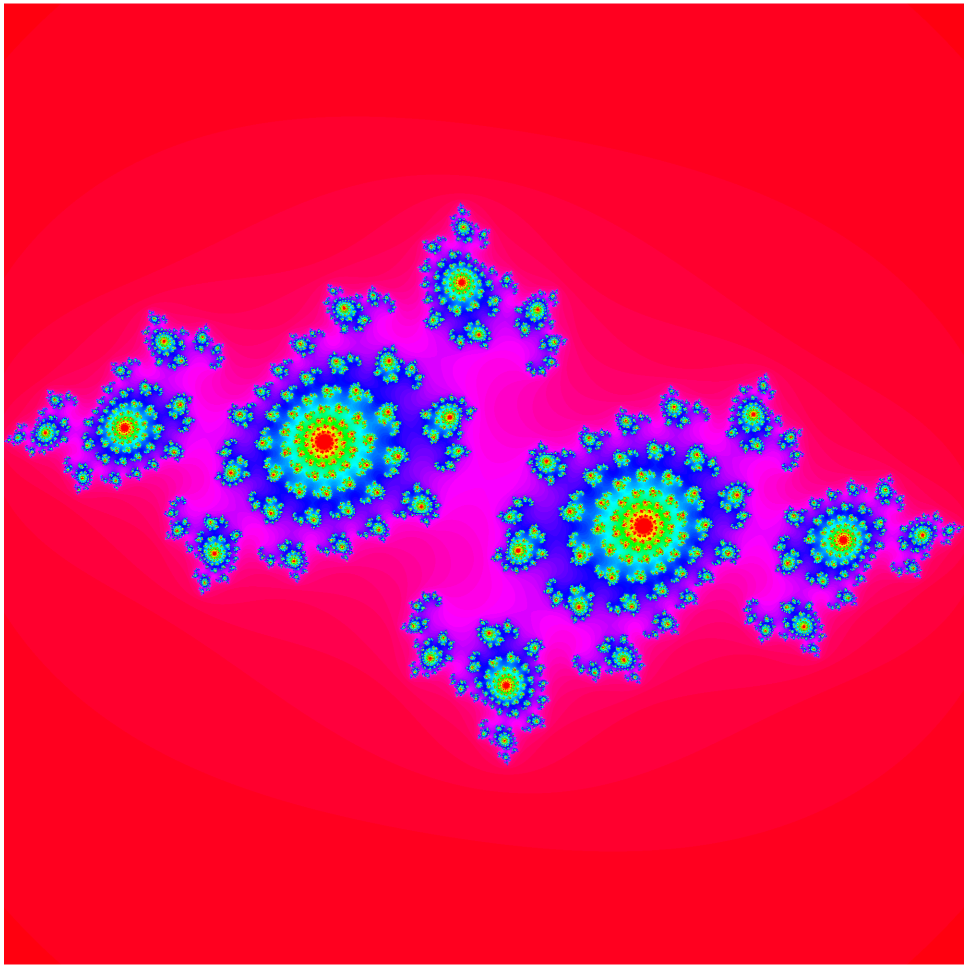
  Following are some examples of this:



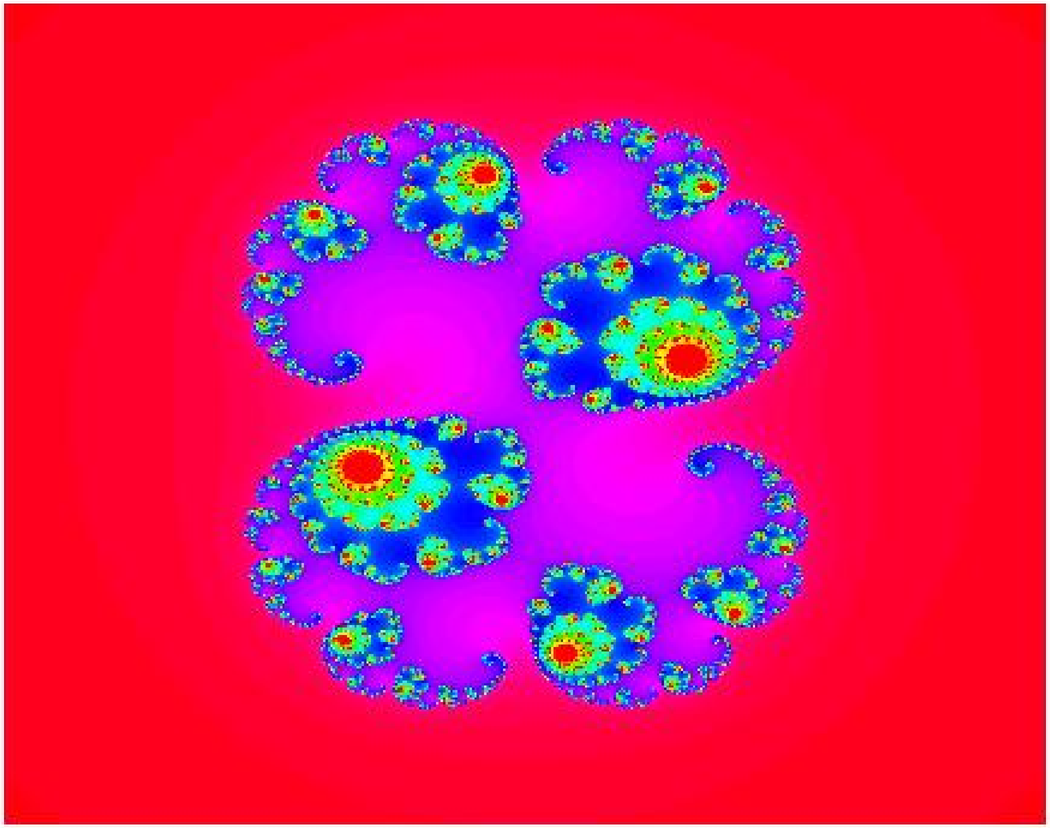
c = -0.4942+0.5229ii



$c = -0.7017 + 0.3842i$



c = 0.7322 + 0.2628i



$c = 0.2850 + 0.0100i$

The Mandelbrot Set $\leftarrow$

- The next step is to look at the Mandelbrot set. Consider the functions we were looking at for Julia sets:

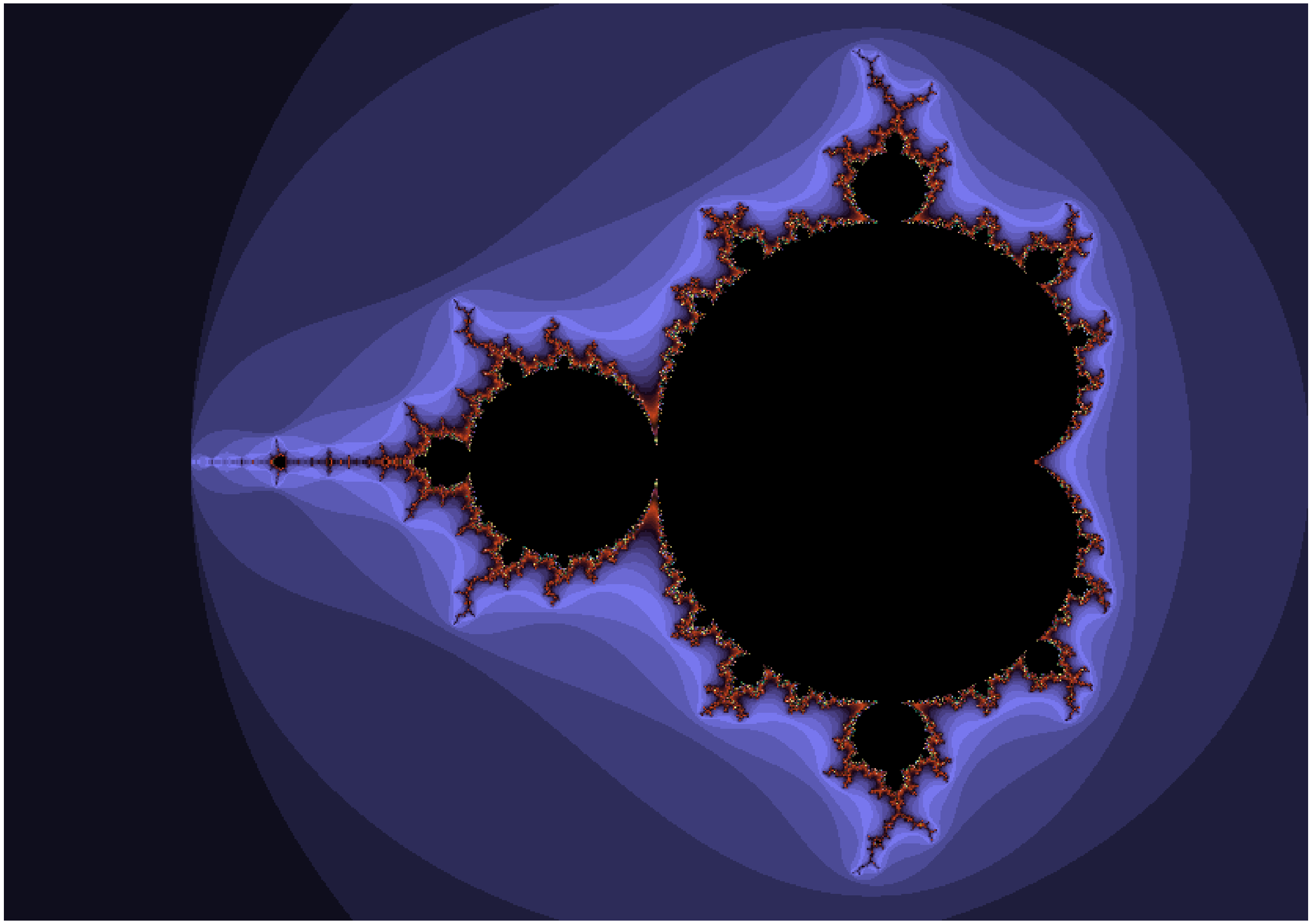
  $$f_c(z) = z^2 + c, \text{where } c \in \mathbb{C}.$$

  We now ask, for $c \in \mathbb{C}$, what happens to 0 under iterations of f_c – i.e., what is the behavior of $f_c^n(0)$ as $n \to \infty$? There are two possibilities: either $f_c^n(0)$ stays bounded for all n, or $f_c^n(0) \to \infty$ as $n \to \infty$. We can now define the *Mandelbrot Set*:

  $$\textit{Mandelbrot} = \{c \in \mathbb{C} : f_c^n(0) \text{ stays bounded for all n}\}$$

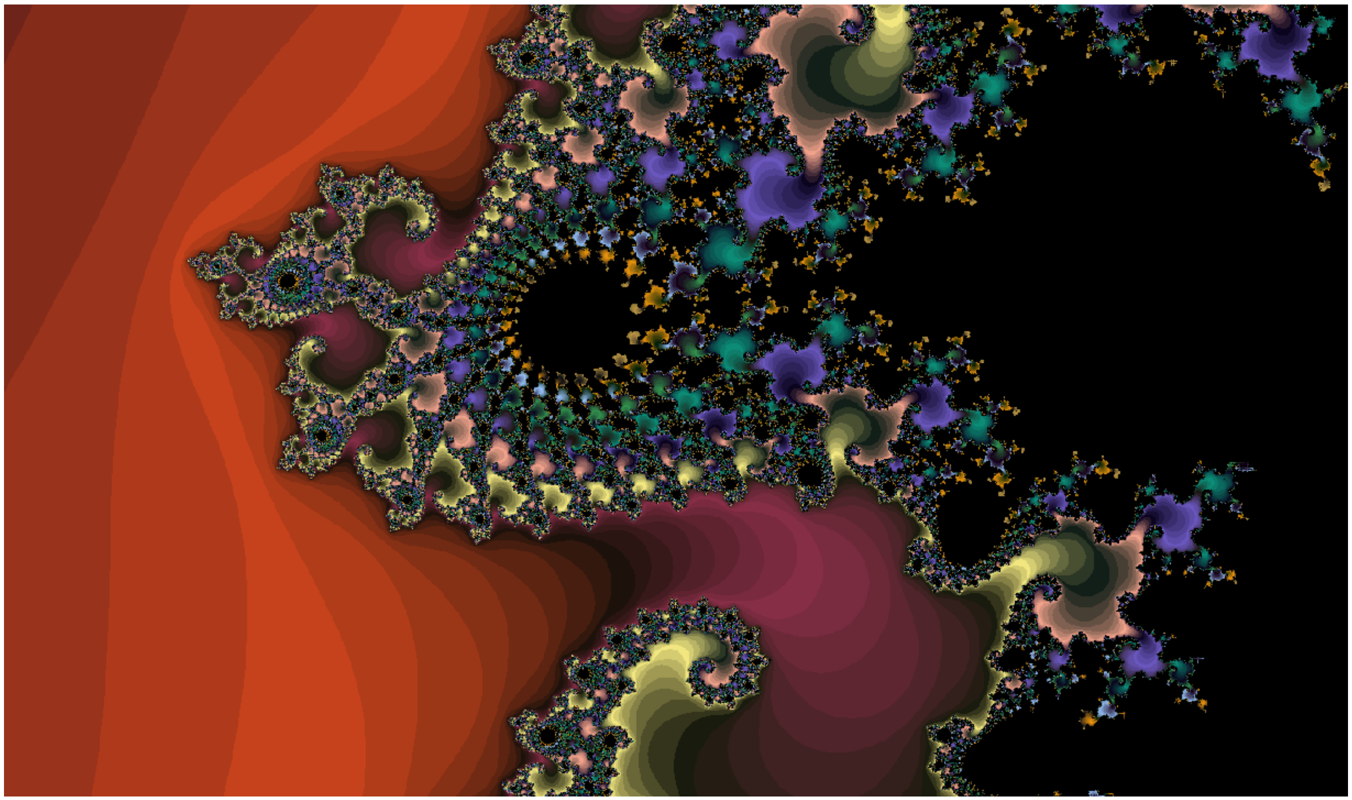
  There is a property of the *Mandelbrot set* that is similar to the Julia sets: if $|f_c^n(0)| > 2$ for some $n \geq 0$, then $c \notin$ *Mandelbrot*. We can use this to make an *escape time* picture of *Mandelbrot*.

Escape time picture (black central area is *Mandelbrot*):



Mandelbrot Set

In all the fancy colorful pictures you have seen of the *Mandelbrot set*, only the black parts are actually in the set – the rest of the points are running away to ∞ at various rates:



Mandelbrot Set around $-0.742 - 0.147i$

- Another way to view this is to think about whether 0 is "trapped" inside the Julia set. Another definition of the *Mandelbrot set* is this:

  $$\textit{Mandelbrot} = \{c \in \mathbb{C} : \text{the Julia set of } f_c(z) \text{is a connected set}\}$$

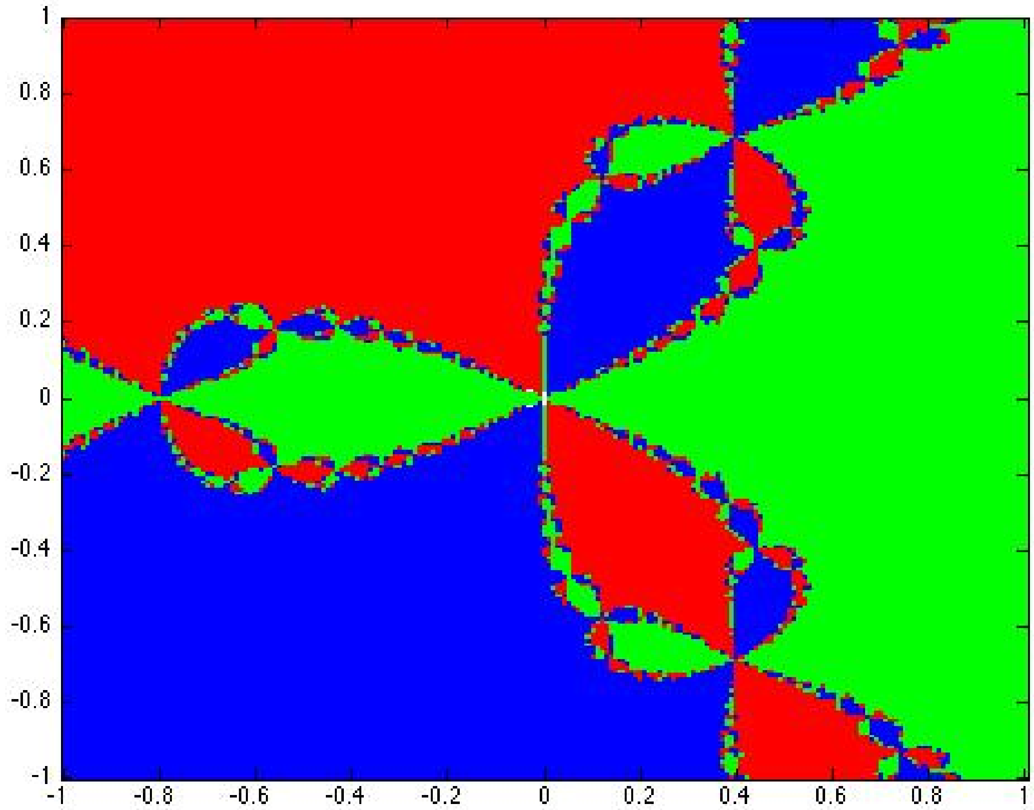
  For points in the *Mandelbrot set*, the Julia set is one connected piece. For points outside the *Mandelbrot set*, the Julia set breaks up into a "dust" of isolated points. "Dusts" like these are sometimes called "Cantor dusts," because they have strong similarities to (and properties like) the *Cantor set*.

Fixed points, limit sets, stable sets, attractors $\leftarrow$

- ...

Basins of attraction ←

- . . .



Newton's method, roots of $z^3 - 1$

Statistical measures ←

- Suppose we have a data set. A thing we might want to do is characterize the data set using a relatively small collection of numbers. Or, at least, we might wish to use a relatively small collection of numbers to distinguish one data set from another. Each of these (hopefully characteristic) numbers is called a *statistic*. An example of a *statistic* is the *mean*. Other examples are the *median*, the *mode*, the *variance*, the *standard deviation*, etc.

  In some special cases, things can be particularly nice. A *normal distribution* (*Gaussian distribution*)

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

  is completely characterized by its mean (μ) and standard deviation (σ).

This becomes more significant in view of the (traditional) *central limit theorem*:

If $\{X_i\}$ is a sequence of i.i.d. (independent, identically distributed) random variables, with mean $E[X_i] = \mu$ and variance $Var[X_i] = \sigma^2$ (in particular, finite mean and variance), then we have the convergence:

$$\sqrt{n}\left(\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) - \mu\right) \xrightarrow{d} \mathcal{N}(0,\ \sigma^2),$$

where $\mathcal{N}(0,\ \sigma^2)$ is the normal (Gaussian) distribution with mean 0 and variance σ, and $\xrightarrow{d}$ is *convergence in distribution*.

Convergence in distribution is the property, for a sequence of random variables $\{X_n\}$, that

$$\lim_{n\to\infty} F_n(x) = F(x)$$

for all $x \in \mathbb{R}$ at which $F(x)$ is continuous, where F_n is the *cumulative distribution function* of X_n, and F is the cumulative distribution function of the limit distribution.

This leads to a paradigm concerning noisy data. If we assume that *noise* (or *error*) is primarily the result of an accumulation of independent random external influences, then it (often) will make sense to assume that any noisiness in our data will be normally distributed, and hence that we can profitably assume that the basis for statistical analysis will rest on studying the means and variances of our data sets.

One could more or less identify this paradigm as a Gaussian / Normal Distribution world, with mean and variance as *sufficient statistics* for analysis of error / noise.

If we have a data set, and we have good reason (e.g., a *central limit theorem*) to believe that a particular parametrized statistical model (e.g., a Gaussian distribution, with mean μ and variance σ^2 as parameters) is appropriate (correct?), then a *sufficient statistic* (or set of statistics) for the model is a statistic (or set) such that no other statistic that could be calculated from our data set could give additional information about the parameters of the model for that data set. In the case of the Gaussian / normal distribution, the sufficient statistics would be the *sample mean* and the *sample variance*. In particular, there would be no reason to calculate any higher *moments* of the data set.

- If we have reason to believe our data set does not come from a Gaussian distribution, or, more generally that Gaussian models are inappropriate or insufficient, we may find it necessary to calculate additional or alternative statistics.

  One example of additional statistics is higher moments. For a distribution X, for any $k \geq 1$, we have the kth *central moment*:

$$\mu_k = E\left[(X-\mu)^k\right] = \left\langle \; (X-\mu)^k \; \right\rangle,$$

  where $E[\;]$ is the *expected value*, and μ is the mean of the distribution.

  The second central moment μ_2 is the variance of the distribution. The third central moment μ_3 is zero for symmetric distributions (such as the Gaussian), and in general is a measure of the asymmetry of the distribution.

Standardized versions of the 3rd and 4th central moments are called the *skew* and *kurtosis* or the distribution:

$$\text{skew}(X) = \gamma_1 = E\left[\left(\frac{X-\mu}{\sigma}\right)^3\right]$$

$$\text{kurtosis}(X) = \gamma_2 = E\left[\left(\frac{X-\mu}{\sigma}\right)^4\right] - 3$$

where σ is the standard deviation. The adjustment of -3 is such that the *kurtosis* of the Gaussian is zero.

We can also *standardize* the higher central moments in the same general way, as $\frac{\mu_k}{\sigma^k}$.

One issue that needs to be kept in mind is that the moments may not exist – the relevant sums (or integrals) may be unbounded.

The various moments may be *sufficient statistics* to characterize a distribution. More generally, we can use various statistics to distinguish between distributions – in particular, if two distributions have at least one difference in their statistics, then we can have confidence that they are different distributions.

Over time, various measures other than the *moments* have been developed for the analysis of data. An important element of the evolving paradigm of nonlinear systems and chaos is the observation that what may look like random or noisy data may actually be a deterministic chaotic system. An example of this, which was observed by people such as John von Neumann and Stan Ulam, is the logistic equation with $R = 4$ (for appropriate initial values).

In the next section we'll look some at a particular statistic that has seen wide use in nonlinear systems – the *fractal dimension*.

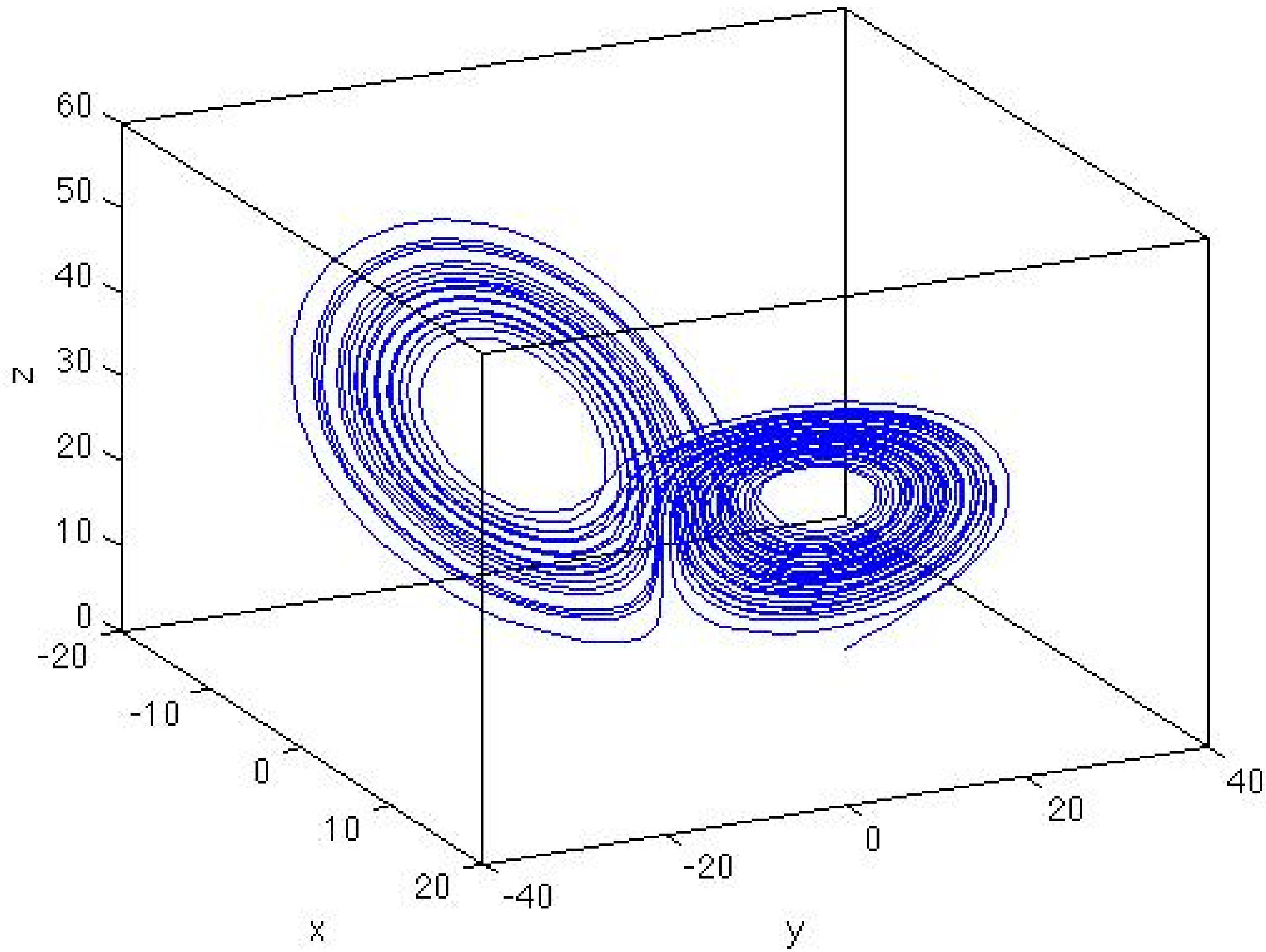
Fractal dimensions, related measures $\leftarrow$

- There are various ways to think about *dimension*. Let's look for a little while at the Lorenz equations:

$$
\begin{aligned}
\dot{x} &= \sigma(y - x) \\
\dot{y} &= x(\rho - z) - y \\
\dot{z} &= xy - \beta z
\end{aligned}
$$

  where σ is the *Prandtl number*, ρ is the *Rayleigh number*, and β is another adjustable parameter. All of σ, ρ, and β are positive, with typical values $\sigma = 10$, $\beta = 8/3$, and $\rho = 28$.

  If we pick particular values for σ, ρ, and β, and initial values x_0, y_0 and z_0, we can use a numerical solver to follow a single trajectory. The picture on the following page uses the MatLab built-in ODE45, which is a variable step Runge-Kutta (4,5) solver:



$\sigma = 10,\ \beta = 8/3,\ \rho = 28,\ x_0 = y_0 = z_0 = 10$

- One way to think about this is that the trajectory is 1-dimensional – as time passes, the system moves along a (curved) line in space. A second way to think of it is that the system lives in 3-dimensional (x,y,z) space. A third way is that there are 3 adjustable parameters (σ, ρ, and β) that control the behavior of the system. A fourth way is that any particular trajectory is determined by six numbers, σ, ρ, and β, and the initial conditions x_0, y_0, and z_0.

  There are a couple of other considerations. One is that there is actually an *attractor* inside the system. For a comparatively large range of initial conditions, the trajectories will approach a limiting stable set. By following a trajectory using a numerical solver, we get an approximation to the attractor. We could explore the dimension(s) of the attractor.

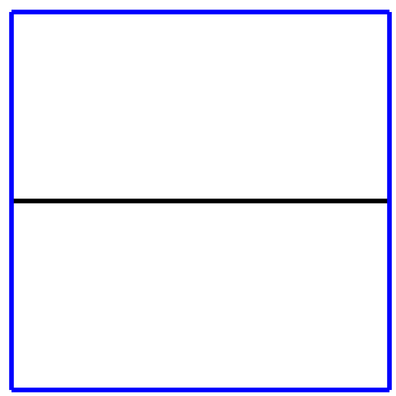
A second consideration is what we might think of as the "real world" problem. For real systems that we might study, we typically don't actually know the equations of motion of the system. We are likely just to have a set of measured data values extracted from the system through some instrumentation.

We'll look at a relatively generic approach to assigning a number to a data set (i.e., calculating a statistic) which corresponds with and generalizes the idea of *dimension* in simple cases. We will be developing (one notion of) the *fractal dimension* of a data set.

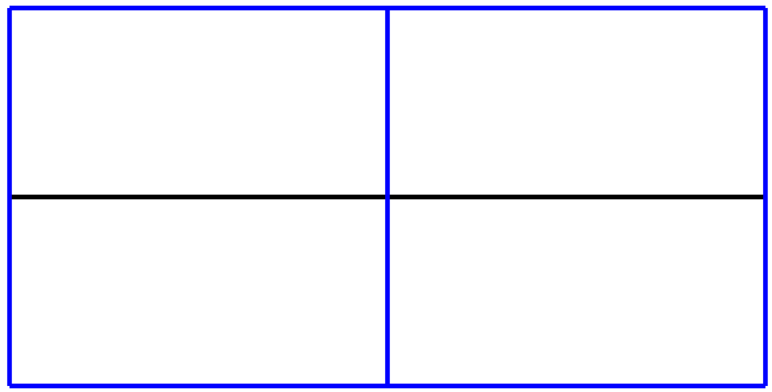
- What we will do is to cover our data set with “boxes” (or “balls” – see Carter [8]), and see how the number of boxes grows as we decrease the sizes of the boxes.

  We begin with a line (segment) of length 1:

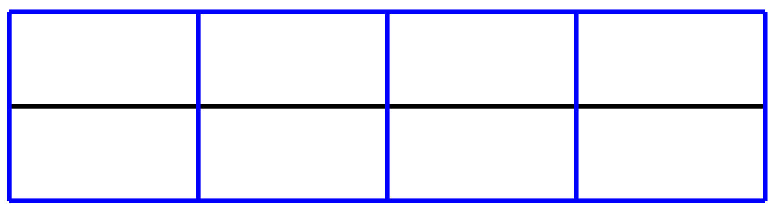
  We can cover this line segment with 1 box of side $1 = 1/2^0$:



We can cover the line segment with 2 boxes of side $1/2 = 1/2^1$:

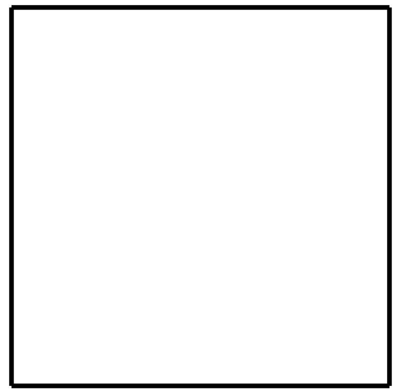


or 4 boxes of side $1/2^2$:

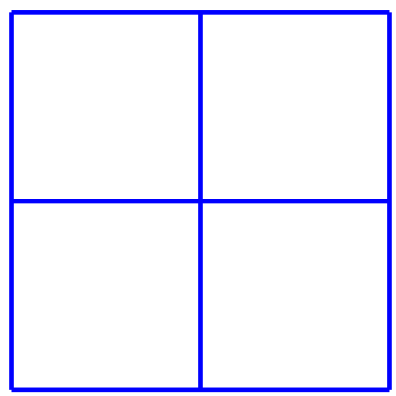


In general, it can be covered with 2^n boxes of side $\frac{1}{2^n}$.

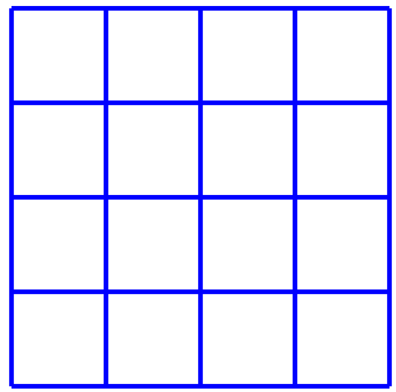
Now consider a square:



We can cover this with 1 box of side 1, or 4 of side $1/2^1$:



or 16 of side $1/2^2$:



In general, we can cover the square with 2^{2n} boxes of side $1/2^n$.

In a similar fashion, we can cover a cube with 2^{3n} boxes of side $1/2^n$.

We can now make a table, showing how the number of boxes needed to cover a “cube” in dimension k grows as the length of the side of the boxes (r) decreases.

r	num_1	num_2	num_3	$\cdots$	num_k	$\cdots$
$\frac{1}{2^0}$	1	1	1	$\cdots$	1	$\cdots$
$\frac{1}{2^1}$	$\frac{1}{2^1}$	$\frac{1}{2^2}$	$\frac{1}{2^3}$	$\cdots$	$\frac{1}{2^k}$	$\cdots$
$\frac{1}{2^2}$	$\frac{1}{2^{2*1}}$	$\frac{1}{2^{2*2}}$	$\frac{1}{2^{2*3}}$	$\cdots$	$\frac{1}{2^{2*k}}$	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\cdots$
$\frac{1}{2^n}$	$\frac{1}{2^{n*1}}$	$\frac{1}{2^{n*2}}$	$\frac{1}{2^{n*3}}$	$\cdots$	$\frac{1}{2^{n*k}}$	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\cdots$

Putting all of this together, we can define a *dimension* by:

$$\dim(S) = \lim_{r \to 0} \frac{\log(num(r))}{\log\left(\frac{1}{r}\right)}$$

where r is the length of the side of a box, and $num(r)$ is the number of boxes of side r needed to cover our set S.

This *dimension* is often called the *fractal dimension* of the set S.

For some (informal) calculations of *fractal dimensions* of some example sets, you can look at Carter [10].

(I need to put more discussion of naming and details here . . .)

Strange attractors $\leftarrow$

- . . .

About fractals $\leftarrow$

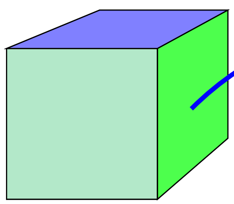
- . . .

Continuous systems and flows $\leftarrow$

- . . .

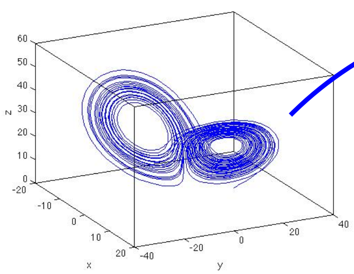
Some history (e.g., Poincaré) $\leftarrow$

- . . .



. . . 0.7, 0.3, 0.8, 0.2, 0.5 . . .

μ σ



. . . 0.7, 0.3, 0.8, 0.2, 0.5 . . .

Some other history (e.g., catastrophe theory) $\leftarrow$

- ...

Poincaré sections $\leftarrow$

- ...

Entropy and information theory $\leftarrow$

- ...

Diagnostics and control of chaotic systems $\leftarrow$

- . . .

Other discrete systems – e.g.: $\leftarrow$

- . . .

Cellular automata $\leftarrow$

- . . .

Iterated function systems ←

- . . .

Complex adaptive systems ←

- . . .

and complex systems more generally ←

- . . .

Various other topics . . . ←

- . . .

Appendix 1 $\leftarrow$

Just to check (and practice our derivatives :-) with the damped harmonic oscillator example . . .

We had:

$$x(t) = e^{-\frac{b}{2}t} * 2 * \cos\left(\frac{\sqrt{4-b^2}}{2}t\right)$$

from which,

$$\begin{aligned} v(t) = \dot{x}(t) &= -\frac{b}{2}e^{-\frac{b}{2}t} * 2\cos\left(\frac{\sqrt{4-b^2}}{2}t\right) \\ &\quad - e^{-\frac{b}{2}t} * \sqrt{4-b^2}\sin\left(\frac{\sqrt{4-b^2}}{2}t\right) \\ &= -e^{-\frac{b}{2}t}\left(b * \cos\left(\frac{\sqrt{4-b^2}}{2}t\right) + \sqrt{4-b^2} * \sin\left(\frac{\sqrt{4-b^2}}{2}t\right)\right) \end{aligned}$$

Combining the terms:

$$
\begin{aligned}
&- x(t) - bv(t) \\
&= -e^{-\frac{b}{2}t} * 2\cos\left(\frac{\sqrt{4-b^2}}{2}t\right) \\
&\quad + be^{-\frac{b}{2}t}\left(b * \cos\left(\frac{\sqrt{4-b^2}}{2}t\right) + \sqrt{4-b^2}\sin\left(\frac{\sqrt{4-b^2}}{2}t\right)\right) \\
&= e^{-\frac{b}{2}t}\left((b^2-2)\cos\left(\frac{\sqrt{4-b^2}}{2}t\right) + b\sqrt{4-b^2}\sin\left(\frac{\sqrt{4-b^2}}{2}t\right)\right)
\end{aligned}
$$

And, sure enough, when we calculate $\dot{v}(t)$, we get the same thing:

$$\begin{aligned}
\dot{v}(t) = {} & \frac{b}{2}e^{-\frac{b}{2}t}\left(b * \cos\left(\frac{\sqrt{4-b^2}}{2}t\right) + \sqrt{4-b^2}\sin\left(\frac{\sqrt{4-b^2}}{2}t\right)\right) \\
& - e^{-\frac{b}{2}t}\left(-b * \frac{\sqrt{4-b^2}}{2}\sin\left(\frac{\sqrt{4-b^2}}{2}t\right) + \frac{4-b^2}{2}\cos\left(\frac{\sqrt{4-b^2}}{2}t\right)\right) \\
= {} & e^{-\frac{b}{2}t}\left(\frac{b^2}{2}\cos\left(\frac{\sqrt{4-b^2}}{2}t\right) + b\frac{\sqrt{4-b^2}}{4}\sin\left(\frac{\sqrt{4-b^2}}{2}t\right)\right) \\
& + e^{-\frac{b}{2}t}\left(b\frac{\sqrt{4-b^2}}{2}\sin\left(\frac{\sqrt{4-b^2}}{2}t\right) - \frac{4-b^2}{2}\cos\left(\frac{\sqrt{4-b^2}}{2}t\right)\right) \\
= {} & e^{-\frac{b}{2}t}\left((b^2-2)\cos\left(\frac{\sqrt{4-b^2}}{2}t\right) + b\sqrt{4-b^2}\sin\left(\frac{\sqrt{4-b^2}}{2}t\right)\right)
\end{aligned}$$

Appendix 2 $\leftarrow$

Something will go here . . .

←

References

[1] Banks, J., Brooks, J., Cairns, G., Davis, G., and Stacey, P., *On Devaney's Definition of Chaos*, American Mathematical Monthly, Vol. 99, No. 4, Apr. 1992, Mathematical Association of America. http://www.jstor.org/pss/2324899

[2] Bar-Yam, Yaneer, *Dynamics of Complex Systems (Studies in Nonlinearity)* , Westview Press, Boulder, 1997.

[3] Bhaumik,Indranil, and Choudhury, Binayak S. *The Shift Map and the Symbolic Dynamics and Application of Topological Conjugacy* Journal of Physical Sciences, Vol. 13, 2009, 149-160 ISSN: 0972-8791. http://www.vidyasagar.ac.in/journal/ http://vidyasagar.ac.in/journal/maths/vol13/art13.pdf

[4] Bishop, Robert,
"Chaos", The Stanford Encyclopedia of Philosophy (Fall 2009 Edition), Edward N. Zalta (ed.).
http://plato.stanford.edu/archives/fall2009/entries/chaos/

[5] Briggs, Keith Martin
Feigenbaum Scaling in Discrete Dynamical Systems
Ph.D. Thesis to University of Melbourne, Australia
http://keithbriggs.info/documents/Keith_Briggs_PhD.pdf

[6] Çambel, Ali Bulent,
Applied Chaos Theory: A Paradigm for Complexity
Academic Press, 1992.

[7] Carter, Thomas,
Nonlinear Systems (...and chaos) a brief introduction
(this document, with LaTeX source, etc.)
http://csustan.csustan.edu/~tom/Lecture-Notes/Nonlinear-Systems/

[8] Carter, Thomas,
What Shape is a Circle?
http://csustan.csustan.edu/~tom/Lecture-Notes/Circles/Circles.pdf

[9] Carter, Thomas,
Entropy, Power Laws, and Economics
http://csustan.csustan.edu/~
tom/Lecture-Notes/Econ101/econ101.pdf

[10] Carter, Thomas,
Some Fractals and Fractal Dimensions
http://csustan.csustan.edu/~
tom/Lecture-Notes/Fractal-Dimension/Fractal-dimension.pdf

[11] Carter, Thomas,
An introduction to information theory and entropy
http://csustan.csustan.edu/~
tom/Lecture-Notes/Information-Theory/info-lec.pdf

[12] Carter, Thomas,
A brief survey of linear algebra
http://csustan.csustan.edu/~
tom/Lecture-Notes/Linear-Algebra/lin-alg.pdf

[13] Carter, Thomas,
The Logistic Flow (Continuous)
http://csustan.csustan.edu/~
tom/Lecture-Notes/Logistic/Logistic-continuous.pdf

[14] Carter, Thomas,
Peripatetic (A random walk in random walks)
http://csustan.csustan.edu/~
tom/Lecture-Notes/Random-Walk/peripatetic.pdf

[15] Carter, Thomas,
Symbolic Dynamics (a brief introduction)
http://csustan.csustan.edu/~
tom/Lecture-Notes/Symbolic-Dynamics/Symbolic-Dynamics.pdf

[16] Davies, Brian,
Chaos Explorer
http://maths.anu.edu.au/~ briand/chaos/software-applet.html

[17] Devaney, Robert,
An Introduction to Chaotic Dynamical Systems, 2nd Edition
Westview Press, Boulder, 2003.

[18] Ekeland, Ivar,
Mathematics and the Unexpected
University of Chicago Press, 1990.

[19] Feigenbaum, Mitchell,
Universal behavior in nonlinear systems
Los Alamos Science, Summer 1980, 1(1):4-27,
Los Alamos National Laboratory Library, 1980.
http://library.lanl.gov/cgi-bin/getfile?01-01.pdf

[20] Fillion, Nicolas,
Order in Chaos: Universal Behavior in Nonlinear Systems
Final project for AM9570B: Introduction to Mathematical Modelling and Simulation, Presented to Professor Pei Yu
Department of Philosophy
The University of Western Ontario, April 21, 2010
http://www.nfillion.com/docs/project.pdf

[21] Girvan, Ray,
The Mandelbrot Monk
Web publication – 1999/04/01
http://classes.yale.edu/fractals/Mandelset/MandelMonk/MandelMonk.html
Caution: Do not stop on the tracks (Furthur - 2011/10/07)

[22] Gleick, James,
Chaos: Making a New Science,
Penguin (Non-Classics); Anv Rep edition, 2008.

[23] Knudsen, Carsten
Chaos Without Nonperiodicity
American Mathematical Monthly, Vol. 101, No. 6, Jun. - Jul. 1994, Mathematical Association of America.
http://www.jstor.org/stable/2975324

[24] Li, Tien-Yien, and Yorke, James A.
Period Three Implies Chaos
The American Mathematical Monthly, Vol. 82, No. 10 (Dec., 1975), pp. 985-992
Mathematical Association of America
http://www.jstor.org/stable/2318254

[25] Schroeder, Manfred,
Fractals, Chaos, Power Laws, Minutes from an Infinite Paradise, W. H. Freeman, New York, 1991.

[26] Strogatz, Steven,
Nonlinear Dynamics And Chaos: With Applications To Physics, Biology, Chemistry, And Engineering (Studies in Nonlinearity)
Westview Press, Boulder, 2001.

[27] Zhan, Mei-Qin,
Course notes, on Systems of First Order Differential Equations
http://www.unf.edu/~ mzhan/chapter4.pdf

To top ←