Discrete dynamical systems and the logistic map

an easy introduction

Daan van den Berg
Hogeschool van Amsterdam
Riken brain science institute
東京

Contents

Foreword

This text is intended for non-mathematicians who would like to explore and understand the mathematical side of dynamical systems, the logistic map, and chaos. In particular, it serves to support the comprehension of those participating or interested in Professor C. van Leeuwens research in models such as the coupled logistic maps for a descriptional model of visual information processing in the human visual cortex.
As reading the average mathematical literature on these subjects is harder than decyphering the rosetta stone in pitch-dark, it appears that the gates of insight and understanding only open to those belonging to the die-hard incrowd: mathematicians and those with a strong mathematical background. The latter group does not contain the average psychologist. Nor does it the philosopher or the biologist, the economist. Shame, for if science seeks to stride forward in measures of describing what is happening around us it must at a certain moment find bridges for the yawning chasm that lurks between mathematics and everything else. This text hopes to toss a first line and make a modest bit of mathematical theory accesible to the vast world at the other side of the scientific crevasse.
After each piece of mathematical theory there are some exercises for the reader to test his or her newly aquired knowledge. Answers to the exercises are included. The author likes to stress that this is a first draft; despite careful proofreading errors might be present. Feedback on errors, but also on the difficulty or readibility of this manuscript are welcomed.

Daan van den Berg
D.van.den.Berg@ie.hva.nl
Daan@brain.riken.go.jp

Preliminaries

There are a few things we will have to recall from calculus before we get started. Most of them will be explained in the text as well, but nonetheless a handhold will be nice. Most of these issues are related to our understanding of ‘infinity’. Most non-mathematicians seem to have much difficulty with that concept.

Convergent series

If we start of with one and add one infinitely many times, the result is infinity. This is denoted like this:

$$\sum_{n=1}^{\infty} 1 \quad = \quad 1+1+1+1+1+1+1+1+1 \,....... \quad = \quad \infty$$

It says: start by making n one, add the argument (whatever is in front of the sigma), add one to n, add the argument etcetera until n reaches the value over the sigma. So n functions as a counter while the argument in front of sigma is constantly added to the sum. But what happens if the term is getting smaller all the time? For instance, if we take 1/n instead of n. In the long run, 1/n will tend to zero so less and less is added.

$$\sum_{n=1}^{\infty} 1/n \quad = \quad 1/1 + 1/2 + 1/3 + 1/4 + 1/5 + 1/6+.... \quad = ?$$

This summation in fact goes to infinity. This is because every so many terms are bigger than one. But it is kind of a boundary case. If we take a smaller term

$$\sum_{n=1}^{\infty} 1/10^n = 0.1 + 0.01 + 0.001 + 0.0001 + 0.00001 \,..... \quad = \quad 0.11111 \,...$$

it is easy to see that the result will never exceed 1. There is a critical boundary:

$$\sum_{n=1}^{\infty} 1/a^n \quad = \quad 1/a + 1/a^2 + 1/a^3 + 1/a^4 + 1/a^5 + 1/a^6 +....$$

is convergent for all absolute a greater than one. Convergent means: becomes a certain value and does not go to infinity. This is a very important property for understanding the bifurcation diagram.

Fractions and square roots

Fractions are numbers like ¼ , ½ etcetera. Being either positive or negative, they have an upper and a lower part, both natural numbers. Square roots are not fractions and this is a very important thing to understand. Historically too, it is a much-stirred concept. Hepasos lost hist head because of the proof that a square root is not a fraction. We will show the same thing here without losing any body parts. The thing about square roots is that their binairy expansion have non-repeating tails whereas fractions do. For instance, and ever-repeating of 0.545454.... is a fraction and it is easily shown:

a = 0.54545454.... so 100a = 54.54545454....

100a - a = 99a =54 (because their endlessly repeating tails fall off)

a = 54/99 = 6/11

Open and closed sets

Consider the real line R. These are all square roots and fractions (remember that values like 2 and -4 can also be written as a fraction). So basically, all numbers we usually use. Now we can take a certain piece of this line, an *interval*. We denote such an interval by giving it's boundaries. As such, [2,4] is the interval of all values between 2 and 4, including 2 and 4 themselves. This is a closed set. We could also denote (2,4), the same interval with the boundaries excluded. There is an important difference between these two and that is that the latter does not have a smallest value. The smallest value of [2,4] is obviously zero, but of (2,4) what is the smallest value? It is not 1.99 because 1.999 is smaller. But it is also not 1.9999999..... because that is equal to two! Surprising? By previously mentioned trick:

a= 1.99999..... 100a = 199.99... 99 a = 198 a=2.

Remember that open sets do not have a smallest or largest value.

Density

To be dense for a certain set in another basically means: they are everywhere and infinitely many. In [0,1], the fractions are dense. They are eveywhere since you can find a new fraction between every two fractions you have. Between 0.1 and 0.2, there is 0.15. Between 0.15 and 0.2 there is 0.175 etcetera etcetera.

1. Discrete Dynamical systems

1.1 Orbit and iteration

In common everyday nature, there are many processes that are defined over a time span. A rock falling down, the propagation of a bee colony, an airplane landing. Such processes can be described by a discrete dynamical system. A discrete dynamical system is a formula which describes a certain value (for instance: the height of an airplane) through time. After one second the airplane is at 80 meters height, after four seconds at 41 meters height. Let's expand a little on this example. If we know that an airplane loses 20% of it's height each second, we could describe the height at a certain second by looking at the previous second. If at one moment the airplane is at 100 meters, then the next second it is at 80 metres. Simply take off 20%. This principle, looking at the previous second to determine the next is to be captured in a formula. If we call the height of the plane "x" and time "t" we would, for the airplane example, get the following formula:

$x_{t+1}= 0.8*x_t$

This is a discrete dynamical system. It says: "the value of x at time t+1 is 0.8 times the value at time t". Or, in other words: "if we have the height at a certain second, the height at the next second will be only 80% of that". If we now know that the plane starts its descend from 100 metres, we can quite accurately determine its landing process. The next second will be 80 metres, the second after that will be 64. In such a way we can produce a time series of values, often called the orbit of a certain value. Below is shown the orbit for value x_0 = 100 (which means: the value of x after 0 seconds equals 100, or, after zero seconds the plane is still at 100 meters height) under iteration of our above mentioned formula the following orbit is produced. Note that some values have been rounded off for the sake of convenience.

Time	X_0	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	...	X_{19}	X_{20}	x_{21}
Height	100	80	64	51.2	41	32.8	26.2	21	16.8	13.4	10.7	...	1.4	1.2	0.9

The orbit is generated by constantly imposing our formula over a value and such determining the next. Such a process is called *iteration*. Thus, by iterating a function over a certain starting value, we generate its orbit.
A discrete dynamical system differs from a continuous dynamical system in that we can only take discrete time values. The height of the plane after one, two, four or seven seconds. But not after 2¼ seconds, or after 4¾ seconds. Continuous dynamical systems are given by differential equations. Discrete dynamical systems, the ones we will use, are given by difference equations, though we will hardly use that word.
Finally, a point worth noting is that though x_{t+1} = $0.8*x_t$ is the most formal definition of a discrete dynamical system, it is also often written like f(x)=0.8x stating that this is an iterative function. This is due to practical reasons and mathematical laziness in the first place, but also because when analyzing these systems it is sometimes useful to use merely the graph of the first iteration of the initial values, which is exactly equal to the f(x) notation. Apart from the notation, there is no functional difference.

1.2 fixed point, periodic point, repellor, attrator

From a scientific point of view, it is interesting to understand the behaviour of a dynamical system. For instance, if we look at our previous example,

$x_{t+1}= 0.8*x_t$

we can imagine what happens to any arbitrary value. Any initial value we take gets reduced by 20% in the next time step, and than again reduced by 20%. If we continue to iterate, the initial value will be more and more reduced. This means that an initial value like 100 or 430 ultimately gets smaller and smaller and in the long run will tend to zero. This also applies to negative values, but since we were talking about the height of a plane this fact must be neglected.

So the gross behaviour of the system is a pretty strong element in understanding it. We know from this system that the orbit of any arbitrary value will tend to zero as time proceeds Except for zero itself, of course, which remains in place. The zero value is therefore called an *attractor*: under iteration, it 'attracts' all other values. We can discern local and global attractors, dividing them by the criterion of wether they attract all or merely some points. This one is global: it attracts all other values in the system. The opposite is called a *repellor*. If a system has a repellor, you can pick a value arbitrarily close to the repellor and it will nonetheless tend away from it under iteration. Repellors are often called unstable whereas attractors are sometimes called stable. We'll see examples shortly.

Because the dynamics of this system are too easy to be interesting, we now proceed to our next example. Consider the following discrete dynamical system:

$X_{t+1} = -(X_t)^3$

It is not difficult to see that, like in our previous example, zero remains in place under iteration. After all, if we would iterate zero, we see $-0^3 = 0$ and thus zero projects to zero and doesn't change under iteration. Such points, points invariant under iteration are called *fixed points*. It is possible for a system to have one, multiple or even no fixed points at all.

But this still fairly easy system shows more interesting facts. Lets look at what happens if the starting value 1 is taken, thus $x_0 =1$. If we raise this value to power three and then multiply by -1 as our dynamical system prescribes, the value becomes -1. This is the next value in the orbit. But if we iterate -1 in this formula, the result is again 1. After all, $-(-1)^3 = 1$. Such it comes that -1 evolves to 1, whereas 1 evolves to -1. This said, we can say that choosing either as an inital value, the orbit will only contain 1 and -1 and thus is a finite set consisting of only two elements.

Points like these are called *periodic points*. This because their orbit shows an undeniable periodic structure. The number of iterations in which it cycles through one phase is called the period (more accurately: *prime period*). Our system $x_{t+1} = -(x_t)^3$ has two periodic points: 1 and -1. Their orbits consist of values -1 and 1 only.

Now can we say anything about the behaviour of the system? Most certainly we can. If we take an irrational number between -1 and 1, we see that under iteration, it will tend to zero. This is because an irrational number smaller than one will only shrink. It does this by alternating signs because of the -1 multiplication, but this is a mere detail: all values between 1 and -1 will tend to zero under iteration. Except for zero itself of course, which is fixed. Thus we could call zero a local attractor. It attracts all values between -1 and 1 and this set of numbers is called its *stable set*. A stable set is sometimes referred to as the *bassin* of an attractor.

X0	X1	X2	X3	X4	X5	X6
1.2	-1.7	5.2	-137	2592274	$<<10^{-19}$	$>>10^{57}$
1	-1	1	-1	1	-1	1
0.8	-0.5	0.13	-0.002	$1.4*10^{-8}$	Towards 0	Towards 0
3	-27	19683	$>10^{12}$	$<<10^{-38}$	$>>10^{114}$	$<<10^{-342}$
0	0	0	0	0	0	0
3	-27	19683	$>10^{12}$	$<<10^{-38}$	$>>10^{114}$	$<<10^{-342}$

The behaviour of various initial values under iteration of $x_{t+1} = -(x_t)^3$.

1.3 exercises

1. Consider the discrete dynamical system x_{t+1},= $4x_t$
   a) Determine whether this system has fixed points.
   b) Determine whether this system has periodic points.
   c) What happens to the value -2 under iteration of this system? What happens to 4.1 ?
   d) Determine of all fixed or periodic points wether they are attractors or repellors.

2. Consider the discrete dynamical system $x_{t+1} = 2x_t(1-x_t)$
   a) Determine the first piece of the orbit of the intial elements 0, 0.5, 0.1 and 1.
   b) Try to determine wether this system has fixed or periodic points and wether they attract or repel.

3 Consider the discrete dynamical system $x_{t+1} = ¼x$,
   a) Determine wether this system has fixed points.
   b) Determine wether this system has periodic points.
   c) What happens to the value -2 under iteration of this system? What happens to 4.1?
   d) Determine of all fixed or periodic points wether they are attractors or repellors.

4. Consider the discrete dynamical system $x_{t+1} = 3.\ 1x_t - 3.1(x_t)^2$

a) Determine wether this system has fixed points.
b) Determine wether this system has periodic points.
c) What happens to the value -2 under iteration of this system? What happens to 0.9?
d) Determine of all fixed or periodic points wether they are attractors or repellors.

5. A colony of kangaroos inhabits southern Australia. This population doubles each year.

a) Give the formula of the discrete dynamical system modelling the population size over years.
b) If the formula is right, and I know the number of kangaroos within 5% correct, will I, using the correct formula be able to determine the kangaroo population size in five years within that same error margin?

1.4 Hyperbolicity, the meaning of the derivative.

1.4.1. Graphical analysis.

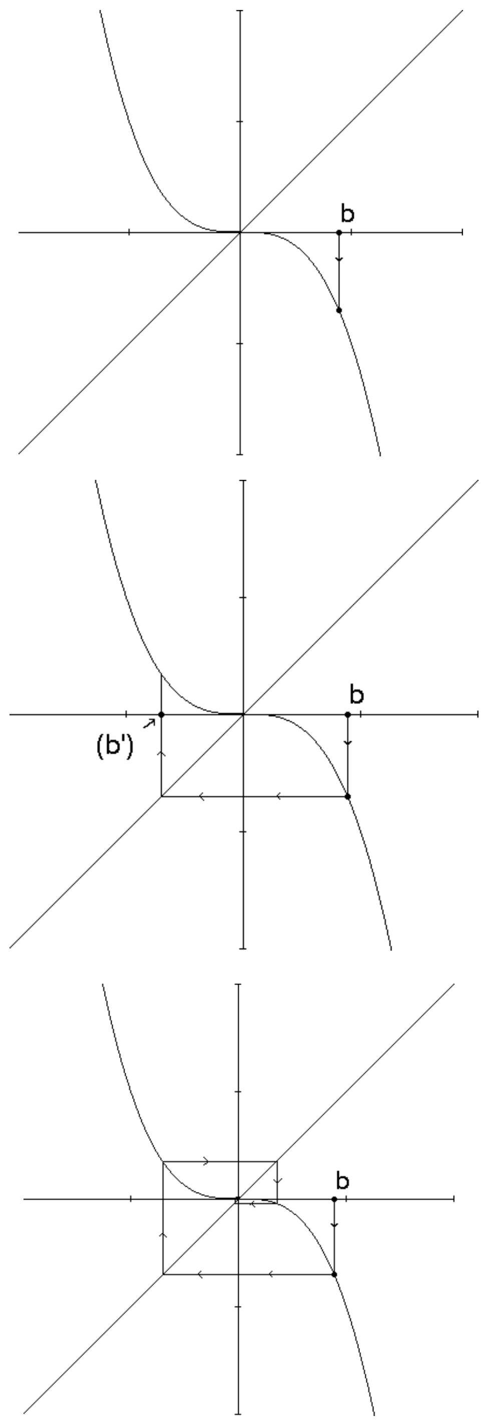
Let's go back to our previous example, the discrete dynamical system $x_{t+1}=-(x_t)^3$. To analyse the behaviour of the dynamical system, we are now going to introduce 'Graphical Analysis', a method which is rather straightforward and visual. Through this we hope to understand more about why certain fixed points are attractors and others are repellors. If we would draw the graph of this function, that is, the graph of $f(x)=-(x)^3$, we could equivalently say we had drawn the first-iteration graph of our dynamical system mentioned above. After all, to each initial value set out on the x-axis, the related point after one iteration is put on the y-axis.

We can now easily find where the fixed points in the dynamical system are. We draw a 45° angle line through the origin. Easy to see, all the points of the form (a,a) lie on this line, such as (0,0), (-2,-2), (¾,¾) and (7,7). If our graph intersects this line, it means that at this intersection lies a fixed point. For instance our graph, intersects line in (0,0). This is hardly surprising. The point 0 under iteration remains in place, it is its fixed point.

But this line has another great advantage: it can also show us the behaviour of other points under iteration. How to work? We take same initial value b, lets say b=0.9 and we wish to know what happens to this point under iteration of our dynamical system. The first iteration is relatively simple. Go vertically to the graph and you will find the corresponding value (for b=0.9 this value is -0.73, let's call it b'). Now to proceed, we go from this point horizontally back to the diagonal line. As soon as we hit the diagonal line we are in point (-0.73, -0.73), because only points of the form (a,a) lie on the diagonal line. Now we can find the first iteration of value -0.73 by vertically returning to the graph. Note that the first iteration of -0.73 is the same as the second iteration of 0.9. Lets say this value is 0.39. Now we can proceed by again horizontally striking the diagonal line and vertically the graph and if we just hold our pen to the paper in this way we would get a spiraling picture of the orbit of our initial value 0.9 falling into the attractor of zero.

Equivalently, we might try the initial value of -1. From -1 a line going straight up strikes the graph at 1. Going horizontally to the diagonal strikes at (1,1). Returning vertically to the graph we come to -1 and if we proceed to go the diagonal we come to (1,-1) forming a perfect square. Proceeding with this process would only lead to repetition of the route we've already been through. This is understandable, since 1 is a periodic point. Thus, in graphical analysis the periodic orbits show through closed cycles like the square we got for our example.

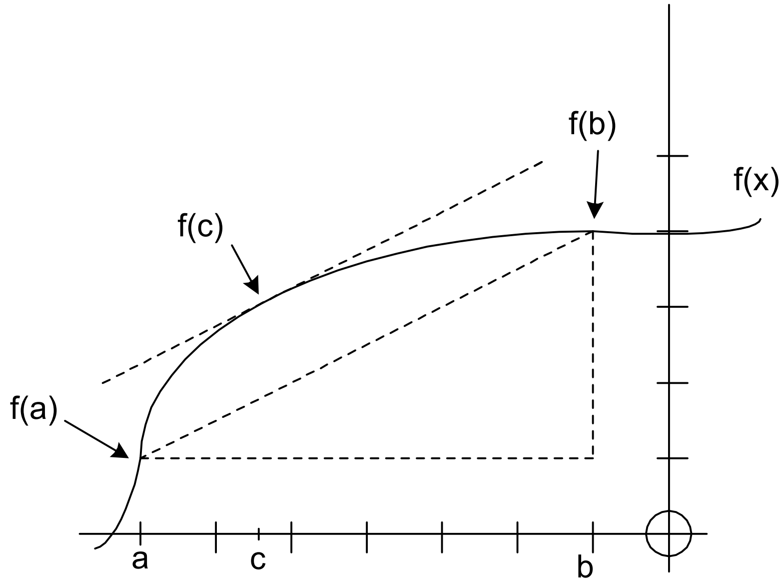
The reader in encouraged to graphically analyze various initial values of this dynamical system with pen and a piece of paper.



T. t. b.: first, second and fifth step of the graphical analysis of $f(x)=-x^3$

1.4.2 The mean value theorem

The mean value theorem is an important theorem underlying many other concepts. For the need of understanding hyperbolicity, or, the behaviour of attractors and repellors, it is explained here. Imagine a function f(x). Just to recap: a function is a prescription that links all values from a certain domain to a certain range. For instance, the function $f(x)=x^2$ links all values of x to a certain other value. It links 2 to 4 and -5 to 25. We might write f(-5)=25 or f(4)=16.



Now imagine a bumpy function f(x). It need not necesserily be bumpy for this theorem, but it makes things a bit more visual. From this function, we cut a certain piece. We take a piece from a to b on the x-axis and all corresponding function values. On this piece of x-axis, the function has a certain way of behaving. As we look from left to right on our piece from a to b, our function might be ascending, descending, or possibly both. If we follow the graph from a to b, the function must arrive either higher, lower or on the same value it left from a. Lets say our function went up from a to b. In that case, we can measure the amount it has gone up. If we divide this amount by the interval [a,b], we get the average climbing rate. Its comparable to a car and its travelled distance. If a car has driven 180 kilometers in 2 hours, the average speed was 90 kilometers an hour. It could have gone faster or slower on certain moments in the2 hours, but the average speed was 90. But there's another important consequence. If the average speed was 90 kilometers an hour, there has been a moment on which the exact speed was 90 kilometers an hour. Consider this: when we started to measure the speed of the car; it was either exactly 90 kilometers an hour or not. If it was, we are done: there is a moment at which the car travels exactly 90 kilometers an hour. If it was not travelling 90 kilometers an hour exactly, its speed was higher or lower. Say it is was higher. Then it means the speed must go down somewhere in the coming two hours. Otherwise the average cannot be 90 kilometers an hour. So at some moment the speed must drop below 90 kilometers an hour, and since it drops from above 90 to below 90, 90 itself has to be somewhere in between.
This is an informal way of proving the mean value theorem. It says that if the average derivative over a certain trajectory is d, at some point in this trajectory the derivative must exactly d. more formally put:

Mean value theorem

If a function f $[a,b]\rightarrow R$ is continuous on [a,b] and differentiable on (a,b) then there is a point $c\in(a,b)$ such that

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

In discrete dynamical system theory, we will use this theorem to investigate the hyperbolicity of fixed points.

1.4.3 Hyperbolicity

In previous sections we have seen that fixed points are important factors in describing the behaviour of a dynamical system. Some are attractors, drawing all values in the neighbourhood towards it. Others are repellors. Initial values close to a repellor will tend away from it under iteration. In this section we will examine a little analysis on the attractiveness of fixed points.
The behaviour of a system around a fixed point generally depends on wether the derivative of the system in the fixed point is greater or smaller than one. Since also the negative values must be included, we are better off by saying that the *absolute value of the derivative* determines wether a fixed point is a repellor or an attractor. If the absolute derivative in that point is smaller than one, the point is an attractor. If it is greater than one, it is a repellor. We'll start with the analytical side of statement and subsequently add some graphical analysis to make it a little more visual, a little more intuitive.
Let f(x) be a discrete dynamical system and let p be a fixed point of f(x), with $|f(p)|<1$. Thus, p is a fixed point of the system in which the derivative is lower than 1. Then p will attract all nearby values under iteration. Nearby means a neighbourhood around p in which the absolute derivative of f(x) is smaller than 1 everywhere. Because we only examine continuous functions with continuous derivatives, such a neigbourhood always exists. It might be very small, but it exists. We might say this is the stable set or the bassin of the attractor. Now if we take any x out this bassin, it will under iteration unstoppable approach the attractor. Why?
If we would measure the difference quotient between f(x) and f(p) it will be smaller than one. After all, we decided to call the bassin a stretch on which the derivative is lower than one and therefore, the average can never be higher than one. The mean value theorem then garantees us that at some point between x and p, the derivative of the function is exactly equal to the difference quotient mentioned above. We will call this value A. Thus,

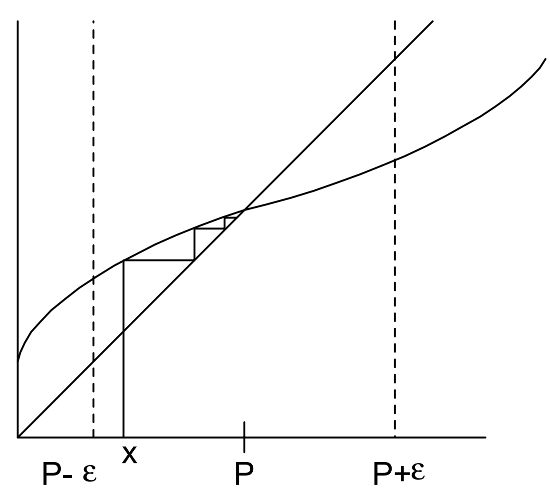
$$\frac{|f(x) - f(p)|}{|x-p|} = A$$

and now by multiplying both sides of the equation by Ix - p|, we get

$$|f(x) - f(p)| = A|x-p|$$

But look to what the formula has resulted now. It says, the difference between f(x) and f(p) equals A times the difference between x and p. And since A is smaller than one we can conclude: the difference between f(x) and f(p) is smaller than the difference between x and p. Or, iterating brings x and p closer and since p is a fixed point and does not move, it is x moving towards p under iteration.

The 'attractiveness' of the fixed point also shows by performing graphical analysis. As we have seen earlier, graphical analysis is a way of visualizing the orbit of an initial value. If we would perform this trick on our value x, the value from the bassin, the results become intuitively clearer. From iterating x, striking the graph, we can only go rightwards, towards the attractor. There is no other way to draw, you can only go towards the attractor.
Repellors and attractors share the common lable *hyperbolic*. Any point or orbit is called hyperbolic if its derivative is unequal to one. To formalize: a hyperbolic fixed point is a fixed point of which the derivative is unequal to one. If the derivative of the point is exactly one the point is labeled 'nonhyperbolic'.



1.5 Exercises

1. Consider the discrete dynamical system $x_{t+1} = 2x_t$ on R→R
   a) Decide wether this system has fixed points.
   b) Decide of all fixed points whether they are stable or unstable

2. Consider the discrete dynamical system $x_{t+1} = X_t^2$ on R→R
   a) Decide whether this system has fixed points.
   b) Decide of all fixed points wether they are stable or unstable

3. Consider the discrete dynamical system x_{t+1},= $2+\sqrt{x_t}$ on [0, ∞) →[0, ∞)
   a) Decide whether this system has fixed points.
   b) Decide of all fixed points wether they are stable or unstable

2. A particular kind of discrete dynamical system: the logistic map

2.1 Origin of the logistic map

Imagine a certain tropical island. This island is inhabited by two lizards. Of lizards it is known that they repoduce exponentially. That is, each month the number of lizards is doubled in respect to the previous month. This is a non realistic model. As time proceeds, the island will get overpopulated and many lizards will die. Therefore, if the number of lizards is very high, the population will drop rapidly. If the population is small however, it will grow through the abundant availability of food.
This principle of competing forces, proliferation on the one hand and food shortage on the other, was modeled (originally by the Belgian mathematician Pierre Verhulst) to a discrete dynamcal system called the Verhulst-function, or more commonly, the logistic map:

$x_{t+1} = 2x_t(1-x_t)$, or equivalently: $f(x) = 2x(1-x)$.

This model describes the population size (x) relative to the time (t). A few formal remarks are to be made. First of all, the function is a normalized parabola (it intersects the x-axis at 0 and 1). This is by convention and makes the mathematics a little easier.
As long as A remains smaller than 4, all points from [0,1] map back into [0,1] under iteration This means that if we take an initial value from between 0 and 1 and iterate it, the result will also be a value somewhere between 0 and 1. And as all values between 0 and 1 result in a value between 0 and 1, any initial value of between 0 and 1 will under iteration produce an orbit which in entirely confined within the [0,1] region. All values outside this interval, that is, values smaller than 0 or greater than 1 are in the stable set of minus infinity.
This does mean, however, that all values on the orbit of some initial value lie between 0 and 1. Note therefore, that we are (if talking about population magnitudes) talking about thousands, tenthousands or some other multiplicative factor of the original values. Thus if an orbit rends the values 0.22, 0.84 and 0.11 we would be speaking of populations of 220, 840 and 110 lizards in the previous example.
The number 2 in the above example represents the growth rate of the lizards. If there was no constraining factor like food shortage, the population would double each year. We will in this chapter examine the behaviour of the logistic map for various growth rates, and therefore call it A. Parameter A can be any number between 0 and 4 for the behaviour we will study. Thus, the general notation will be:

$f(x) = Ax(1-x)$

A brief survey will ensure us the logistic map has at least two fixed points. First of all 0 because the parabola is normalized and furthermore the intersection with the function y=x. Recall from 1.4.1 that fixed points lie on the diagonal of 45°, which is equivalent to the function y=x. Thus, solving the equation $Ax(1-x) = x$ results in finding the fixed points 0 and $1-1/A$. For most parameter values, the second fixed point will lie in our area of investigation. Only for $A<1$ it does not.
The logistic map is a nice function to model population growth. But it has also found applications in chemistry, physics, economy and even the science of psychology. This is due to the rich dynamics of the map we will describe shortly.

2.2 Behaviour of to logistic map for various parameters of A

2.2.1 A first investigation: 0<A<-1

If A is smaller than one, this means disaster for the lizards. Since A represented their growth rate, the population will drop and the species will become extinct. This is a nice intuitive guess of what happens to orbits under iteration of this dynamical system, but since this text is intended explanatory mathematical, we will have to formalize as well.
If A is smaller than one, the value of the second fixed point, 1-1/A is smaller than zero and not in our area of investigation. Therefore we ignore it and focus on the other fixed point, 0. If we derive the function f(x) = Ax(1-x) we would get f'(x) = A-2Ax. If we now investigate the fixed point for hyperbolicity, thus determining the derivative in that point, we get f (0) = A - 2*A*0. This is equal to f (0) = A and since A is smaller than one the fixed points derivative is smaller than one and it is an attractor. This matches our intuition perfectly. If the reproduction rate is smaller than one, the population will drop to zero.
Now left to be made is a remark about what happens to A=1. In other words, when the system is non-hyperbolic. Is this case it is weakly attracting. Non hyperbolic points can also be weakly repelling. We will not pay too much attention to non-hyperbolicity, it is uninteresting in respect to the rest of the dynamics of the logistic map.

2.2.2 Investigation for 1<A<3; bifurcation

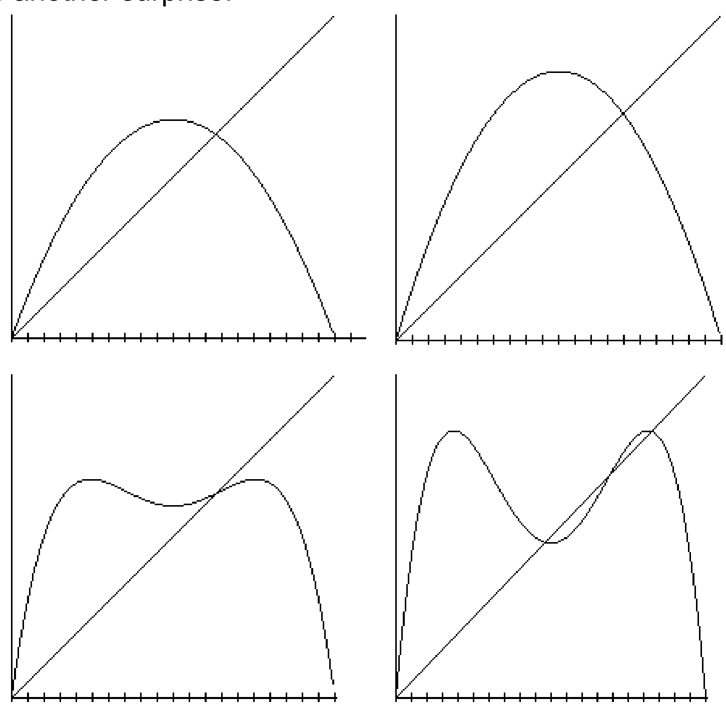
As mentioned earlier, for A<1 the logistic map has no fixed points apart from 0 in the interval [0,1]. This interval, all points between 0 and 1, is our area of investigation, because this is the place where all the action happens. As soon as A becomes greater than one, two things happen. First of all, the derivative of the fixed point 0 passes the critical 'greater than 1 boundary'. As we have seen, a fixed point with a derivative smaller than one is an attractor, and with greater than one is a repellor. So if A grows from smaller than one to greater than one, the stability of the 0 fixed point changes radically. But something else happens. From the moment A is greater than one, our equation 1-1/A results in a value between 0 and 1. In other words, our function gains a fixed point within our area of investigation.
Thus, as A grows greater than 1 the dynamics of the system changes drastically. Such a sudden change in the system's dynamics when a parameter is varied is called a *bifurcation*. The parameter change for which the bifurcation occurs is called the *bifurcation point*. A bifurcation of this type is called a *transcritical bifurcation*. Other types of bifurcations are for instance the *saddle node bifurcation* and the *period doubling bifurcation*. The latter is also present in the logistic map. We will investigate it shortly.

2.2.3 Investigation for 3<A<3.57; period-doubling-bifurcation, Feigenbaum's constant.

But the previous bifurcation is not the only one in the parameter range 0<A<4. There is another bifurcation point at A=3. Thus as A changes from smaller than three to greater than three, there is another dramatic change in the dynamics of the system.

As A grows, the height of the parabola changes. It gets higher as A gets higher. But another thing happens as A grows: the derivative of the fixed point other than zero changes. The fixed point is given by 1-1/A and the derivative is given by A-2Ax. As we are investigating the fixed point's derivative, we might rewrite A-2Ax to a form in which x is the fixed point. Thus, x=1-1/A. Substitution rends A-2A(1-1/A) = A - (2A - 2A/A) = A-(2A -2) = -A+2 = 2-A. So the derivative of the fixed point relates to the A-parameter as 2-A. If A=1½, the derivative of the fixed point is ½, thus it is stable. If A=2.1, the fixed point is stable, for its derivative is -0.1, in absolute value smaller than one. If A=3 its derivative is exactly one and we are at a bifurcation point because it goes from being stable to being unstable. But there's another surprise.

Lets have a look at the second-iteration function for a change. To visualize, check the illustrations on this page. Top-left is the firstiteration-graph of the logistic map, top-right is the second-iteration map. As the first-iteration graph is given by Ax(1-x), the second is given by the same formula iterated twice. This is A(Ax(1-x))(1-(Ax(1-x)) which equals $A^2x-A^2x^2-A^3x^2+2A^3x^3-A^3x^4$. The exact formula matters not. What is important is to realize that a twice iterated logistic map equals a once iterated fourth-degree polynomial. This second-iteration-map has a graph with two humps instead of one. Now if the A-parameter gets higher and higher, the graph gets more eccentric. In case of A=2.9 (top left and right) the fixed point shows to be stable in the left picture. The right picture, the second-iteration-graph, also shows one stable fixed point and if we look closely, we see it is the same point as in the leftmost graph. This doesn't surprise us, because what is fixed after one iteration is also fixed after two iterations. But what happens as the A-parameter gets bigger? As we have seen, the top of the first-iterationgraph gets higher, but also the second-iteration-graph gets more eccentric. The maxima grow higher and the minimum deepens and as soon as the A -parameter passes the value 3, the minimum suddenly crosses the 45° line. This means it suddenly own two more periodic points of period 2 (after all, these are fixed points in the second-iteration-graph). These periodic points are stable and thus we can draw a remarkable conclusion:



Top-left: first iteration graph for A<3 Top-right: same for A>3
Bottom-left:2nd-iteration graph for A<3 Bottom-right: same for A>3

As A passes the value 3 and the fixed point changes from being stable to being unstable, the function gains a stable period 2 periodic point.

To transfer to our lizard-island concept: if the rate of lizard reproduction grows beyond 3 (for instance, because of a fertility drug) the population doesn't grow towards a singular equilibrium. Rather it keeps alternating each year. One year 760 lizards, the other year 240, then 760 again. Remarkable but true, this periodicity seems to be a property of the system itself rather than induced from an external source.
Because the function behaves extremer as the A-parameter increases there is a point for A where in the first-iteration graph the fixed point becomes unstable and the second iteration graph gains two fixed points which manifest themselves in the first iteration function as a stable orbit of period two. This transition from stable to unstable is an issue in bifurcation theory. Notice how when A was smaller than one, the zero fixed point was stable and as A increased became unstable. With it came a new stable fixed point. The stability of this fixed point became questionable as well as A approached 3. After A passed 3, this point became unstable as well and made way for a stable period 2 orbit. We could see this in the second-iteration graph.
Now, wonders the curious spirit: "as A slowly increases from 3 to 4, and the graph of first- and second-iteration behave more extreme, could the fixed points in the second iteration graph also make a transition from stable to unstable?". And indeed this happens. At a value real close to A=3.45, another period doubling bifurcation occurs. This because, as one might expect, the fourth iteration graph is a polynomial of degree 16 with even more humps (maxima and minima) than the seconditeration graph. Result is that at some point some of these minima cross the 45° line and a stable period four occurs. Very similar to the period two. And as A increases, so does the period of the stable orbit.
After the first period doubling bifurcation (for A=3) the stable attractor changes from period one to period two. At the second bifurcation (for A = almost 3.45) it increases to period four and as A increases towards 3.55, another period doubling bifurcation occurs and the stable attractor's period increases to eight.
An interesting fact is that these period-douling bifurcation occur at a fixed ratio. As the parameter distance between the first and second period doubling bifurcation is 3.45-3 = 0.45 and the distance between the second and the third is 3.55-3.45 = 0. l. The second distance is therefore about 4.669 times shorter than the first distance. The distance between the third and the fourth is again 4.669 times shorter and for all subsequent period-doubling bifurcations the distance gets shorter by the factor 4.669. This result was aquired by Mitchell Feigenbaum. This distance reduction factor has ever since been known as Feigenbaum's constant:

$\delta = 4.669\ 20...$

This constant is universal. This means that this factor goes for all unimodal functions. Technically spoken, these are all continuous functions with one zero-deravtive, 'single hump-functions' one might say. The fact that all these kinds of iterated functions are subject to Feigenbaum's constant is known as the principle of universality.
But, as the curious mind continues, "if distance between the bifurcation points gets about five times shorter at each bifurcation point, this sequence of bifurcations converges to a certain value; this sequence is not infinitely long". Indeed it is not, and the series converges to a certain point called the accumulation point. The accumulation point, written A_{acc} is about 3.57, thus $A_{acc} \approx 3.57$. "But what happens BEYOND this accumulation point?" The answer is... chaos, to which we will dedicate the next paragraph.

2.2.4 Beyond the accumulation point: Chaos

It is a good thing for us to know what happens for parameter values in the range 0 to 3.57. All initial values are either a point of the repelling orbit, a point of the attracting orbit or a point in its bassin. But what happens beyond this accumulation point?
On the one hand, since the Feigenbaum constant is a convergent series, we know that the all periodic points of period 2' have become unstable. Therefore we can expect all periodic orbits of 2' when A is greater the the accumulation point. Thus there are infintely many periodic points. But on the other hand we must consider the fact they are all repelling. Repelling towards what? The answer is that for values of A greater than the accumulation point, the function Ax(1-x) behaves in a chaotical manner. We'll take A=4 as an example.
What is a chaotical function anyway? There are several definitions of chaos, non of them universally accepted. We will adopt the very common definition by Robert L Devaney.

Definition (Devaney).
A function f. V→V is chaotic if
- periodic points are dense in V
- f is topologically transitive
- f has sensitive dependence on initial conditions

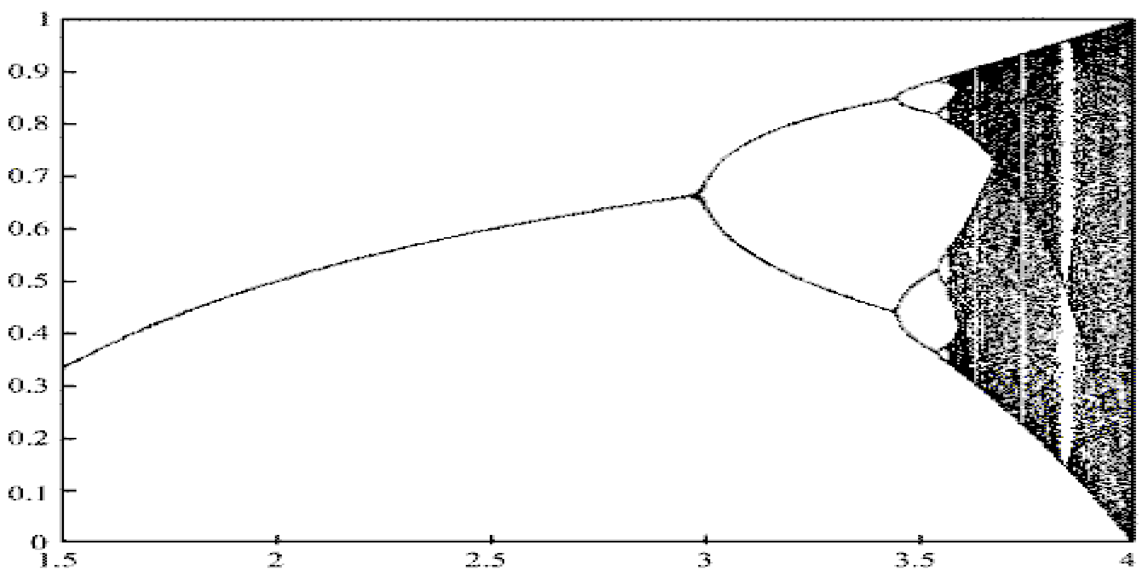
Consider our function f(x) = 4x(1-x). This function maps [0,1] to [0,1] and matches all above criteria. A set being dense means that between all two numbers, there is another number. For instance, between 2 and 2.1 there is 2.05. But between 2 and 2.05 is 2.005. Between 2 and 2.005 again, there is again a number. It follows therefore that a dense set has infinitely many points. Now if we look at the way periodic points are spread, we see that there are infinitely many and between any two lies another third. We can- imagine this if we recall how the stable attractor was transformed during the perioddoubling bifurcation cascade. After one bifurcation, when the stable period two emerged, there was an unstable fixed point. After two bifurcations, the period of the stable attractor is four, the unstable period 1 still exists but an unstable period two is added. Thus two more periodic points. In the next bifurcation, more unstable periodic points are added. And this all the way to infinity. In other words: in [0,1] under iteration of 4x(1-x) there are infinitely many periodic points, and they are so densely packed that between every two, we can find another third. Periodic points are said to be dense in [0,1]
Toplogical transitivity is different matter. If we pick an open interval of any size in [0,1], there is a point in this interval that under iteration will go to any other interval. Thus, is we pick any small neighbourhood in [0,1], it has a point in it that under iteration goes anywhere else in the domain. We might say this function 4x(1 -x) well-mixes the domain [0,1]. Like a blender in a bowl of brown and white sugar, we can intuitively feel why chaotic means well-mixed. If the brown and the white sugar remain side by side we would say it is quite orderly. Whereas if it is well-mixed, we would say it is a chaotic disorganization of sugar crystals. The well-mixedness is an intuitive obligation for chaos.
But the most intuitive hallmark is the sensitive dependence on initial conditions. This means that if we take two initial values very close to each oter, the will under iteration become more and more separate. We can compare this to a river. Put two pieces of cork very close to each other in the water. If the river is nice and calm, both of them will flow to the middle, towards the mainstream of the river but if the river is wild and turbulent, the cork pieces soon move their own way, which is not structural and fully unpredictable. The same goes for the chaotic logistic map. Take some initial value and on forehand there is no predicting on where it is going to be.

2.3 The bifurcation diagram: chaos and order.

What more can we say for the logistic when $A>A_{acc}$, apart from the fact that for A=4, it is chaotic. To investigate, we shall make use of the very important theorem of Fatou:

Fatou's theorem: If a quadratic function has a stable periodic orbit, then the critical point is in the stable set of a point from this orbit.

So what this theorem says is: "If you iterate the critical point, it will fall into the attractor (given it exists). The critical point is the point where the derivative is zero. In case of the logistic map, therefore, the critical value is always 0.5. This is a fact we can make very good use of, for we are now able to draw a so-called bifurcation diagram. This goes as follows. On the the x-axis, we put the range of parameter A, form 0 to 4 that is. On the y-axis, we will place the value of the stable point. Note therefore that for A<1 there is no value drawn. The stable point is 0 and therefore when drawn falls exactly on the x-axis. Between 1 and 3, the function has one stable point. After 3, the function has two stable fixed points: a stable periodic orbit, one of the points somewhere between 0 and 0.5, the other one somewhere between 0.5 and 1. After another bifurcation it has a stable period four. Note that distance between bifurcationpoints gets shorter and shorter. The period eight attractor is still wellvisible but one needs to look real close for the period sixteen and the period 32 seems to disappear into the black mass behind it.



This black mess is where chaos enters the system. Just beyond the accumulation point, iterating the critical point doesn't get it into a nice stable attractor anymore. It moves around like a bumblebee, without showing any particular pattern so it seems. This is the region of the chaotic orbits. Chaotic orbits are non-periodic (they never repeat) nor are they in a bassin of a periodic attractor. But the most significant fact is that they are sensitive to initial conditions. This means that an orbit nearby, even an orbit very nearby, will eventually move away from it and go it's own path under iteration.
A striking fact are the white bars which run cross the chaotic region. If we look closely, we see a periodic orbit. And this is true. All of a sudden, at varous parameters of A, stability occurs. A stable periodic orbit resides amidst the chaos. If you look closely at the bifurcation diagram depicted on this page you might find more of them. In fact, there is a stable periodic attractor arbitrarily close to any chaotic attractor.

2.4 Nearly period three, intermittency

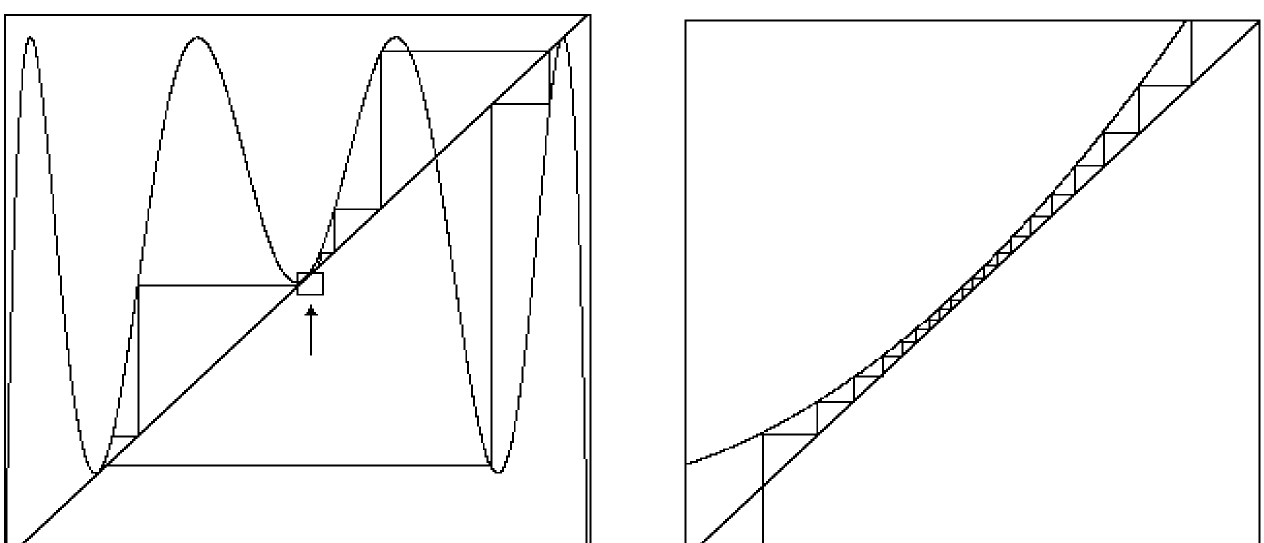
Chaos is stunning. It is surprising to see how a relatively simple formula like the logistic map shows such complex behaviour in the chaotic region. Another strange thing is the periodic windows inside the chaotic region; sudden areas of order which seem to emerge from nowhere. Let's have a closer look at the period three window, because more interesting behaviour is showing there.

If we want to find the values for which a period three exists, we have to solve f(t(f(x))) = x in which f(x) stands for Ax(1-x), the logistic map. That is, we look for values of A for which the thirditeration-graph of f(x) strikes the 45° diagonal. Two of these points are obvious; the fixed points 0 and 1-1/A. These have prime period one, so they also have period three (what is fixed after one iteration remains fixed after three as well).

But at A = 3.828... another period three orbit exists, and this one, remarkably enough, is. stable. But an even more interesting phenomenon displays in the regio just before the period three window. If we look closely at the graph depicted on- the. page we can. see that it nearly touches the diagonal.- And the keyword is nearly. Graphical analysis should make things a bit clearer. If we describe the orbit of a certain initial value with graphical analysis and we get near the point where the stable period three is about to appear, the graphical analysis squeezes through the small space left between the graph and the diagonal.

This is a very interesting phenomenon. Because the distance between the graph and the diagonal is so small, the steps taken in our graphical analysis are also very small. This means that two subsequent steps in the analysis are very close to each other in numerical value. And since we are analyzing the third-iteration graph, we might conclude that all third values of the orbit are very close together. We could call this quasi periodicity. Because every third value is alike, so are the second and the first in some way, because these are projections of the third and since the logistic map is not too bumpy, the projections of the points close together are also rather close together.

This quasiperiodicity is more commonly known as intermittency or sometimes called the laminar phase. Intermittency is a period of quasi periodicity followed by a chaotic burst. The latter follows when the orbit escapes from the narrow channel again. In physical systems, intermittency is a well-known route to chaos.



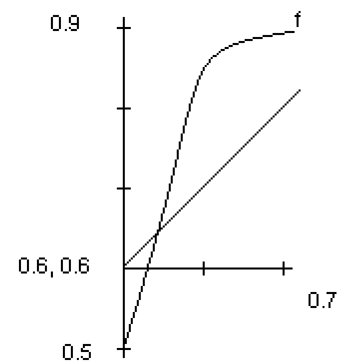
(above) Graphical analysis of the third-iteration function of the logistic map in the intermittent phase (A=3.828). Note the little box in the left picture is enlarged in the right picture.

2.5 Period three, Sarkovskii's theorem

The intermittent regime just before the period three attractor shows surprising behaviour. At the moment the period three attractor is established though, another strange phenomenon occurs. Behold the Sarkovskii theorem:

Theorem (Sarkovskii):
Let f be a continuous function. If f has a periodic point of period three, it has periodic points of all other periods as well.

It looks like we have quite a strong statement on our hands here. The proof is coming up, again in an informal setting since this text is intended for non-mathematicians. It is supported by two pillars, the first them being the *intermediate value theorem* (not to be confused with the *mean value theorem*).
Imagine two intervals, I and J such that $I \subset J$. A function f is defined such that $J \subset f(I)$. Then f has at least one fixed point in I. First, $I \subset J$ means J is a subset of I. Second, f(I) are all the points in interval I after one iteration. For the sake of the example, we'll take the interval of J to range from 0.6 to 0.8 and the interval I to range from 0.55 to 0.85. We'll take the range of our function even a little bigger than that, from 0.5 to 0.9, also for the sake of example. Since f is continuous (i.e.: it has no sudden breaks, we can draw it without taking our pencil off the paper) it has to cross the diagonal y = x at least once, because its range is encompassing its domain (see illustration). This guarantees us that if the condition $I \subset J \subset f(I)$ is true, the function has at least one fixed point in f.



The second pillar is the following: suppose we have two intervals called A_0 and A_1. We have chosen these intervals such that $A_1 \subset f(A_0)$. That is, if we would iterate all points in A_0 we would get the area of A_1 and a little more, we say $f(A_0)$ *covers* A_1. This 'little more' is important. We could take a somewhat smaller piece of A_0, such that its points would under iteration *exactly* make up A_1 and call this A_0-piece J_0. Now imagine we take another A called A_2 such that $A_2 \subset f(A_1)$. Since J_0 was the exact precursor of A_1 and $f(A_1)$ is a larger interval than A_2, there must be a piece of J_0 which exactly makes up A_2. We will call this piece J_1. We could do this again, Adding an A_3 and a J_2 and continue in this fashion until A_n and J_{n-1}, for as long as we wish.
Now look at the following. We have three periodic points, because our function has period three. We will call them a, b and c such that a is smaller than b and b is smaller than c. In mathematical terms: $a<b<c$ and $f(a)=b$, $f(b)=c$ and $f(c)=a$. The interval between a and b will be given the name I_0 and the interval between b en c will be called I_1. Since the borders of I_0 project to the borders of I_1 under iteration of f, it follows that $f(I_0)$ covers I_1. But for the similar reason, $f(I_1)$ covers $I_0 \cup I_1$. If you're not too sure of this, note the projection of b and c onto c and a (i.e. I_0 and I_1 and maybe even a little more).
This last fact is utmost important for the theorem. Since $f(I_1)$ also covers I_1, there must be part of I_1 which exactly makes up I_1. We'll call this A_0. As such, $f(A_0) = I_1$. But moreover, we can find an interval A_1 such that $f(A_1)$ covers A_0, or better, such that $f(A_1) = A_0$. See the inductiveness: we can find an interval A_2 which exactly maps to A_1. We can find an A_3 for A_2 and an A_4 for that one. In fact, we can go on an find an A_n as large as we wish. From this it follows that in A_n there lies a period n periodic point. But how do we make sure this is its *prime period*? After all, a fixed point is also a point of period 3,4 and 6 because it returns after each iteration. We would falsely be claiming we found a point of period 3, 4 or 6 whereas it is actually a period 1 point returning, returning and returning.

But here comes the trick. Let's have a look at our period 2 point. Since f(I1) covers I0 and f(I0) covers I1, it is possible to find a period two point which lies in I1 but who's first iterate lies in I0. Result: it must be period 2, and cannot be period one.
But this also works a point with period 5. Let's choose the following intervals: A_0, A_1, A_2, A_3, A_4. Each interval covers the previous one under iteration of f. Thus $A_3 \subset f(A_4)$ and $A_2 \subset f(A_3)$ and so on. Second $A_0 = I_1$. So this is a sequence of nested subintervals, all in I_1. But I_1 covers I_0 and I_0 again covers I_1. As a consequence, $A_0 = I_0 \subset f^5(A_4)$. In words: there exists an interval such that it's fifth iteration covers I_0. But the last iteration of this interval is actually We choose our A_0 such, that it exactly covers I_0 instead of I_1. If we now go looking for subsequent A_1, A_2 A_3 etcetera and find their periodic points, all their iterations will be in I_1 except for the last. Result: it cannot have a period smaller than the prime period we were looking for, otherwise it would totally be contained within I_1.
Convince yourself that this also works for any other period. Just choose any period n and there will be intervals $A_0 ... A_{n-1}$ such that A_{n-1} contains a point p which under iteration resides the first n-2 iterations result in a I_1-point and its n-1th iteration results in a I_0-point. Iteration number n has it returning to its starting point, since it is periodic.

2.6 A=4: the how and why of chaos

In this section we will try to uncover the deeper mathematical meaning of the chaotic behaviour of the logistic map when A=4. We will do this by proving the tent-map is chaotic and subsequentially showing that its behaviour is in some essential way equivalent to the behaviour of the logistic map. The tent map is much simpler and given by the formula $f(x) = 1 - 2|x - ½|$

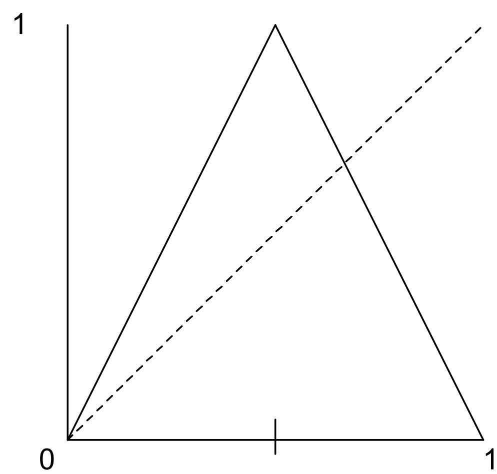
2.6.1 The tent map is chaotic

The chaotic behaviour of the logistic map is a profoundly studied phenomenon. There are several definitions for it and we will stick to a very famous one, the one as stated by Robert L. Devaney.

Defnition: a discrete dynamical system h(x) on (D→D) is chaotic if:

- the system is topologically transitive on D.
- the system is sensitive to initial conditions.
- periodic points are dense in D, the domain of the system.

Proposition: the tent map is chaotic



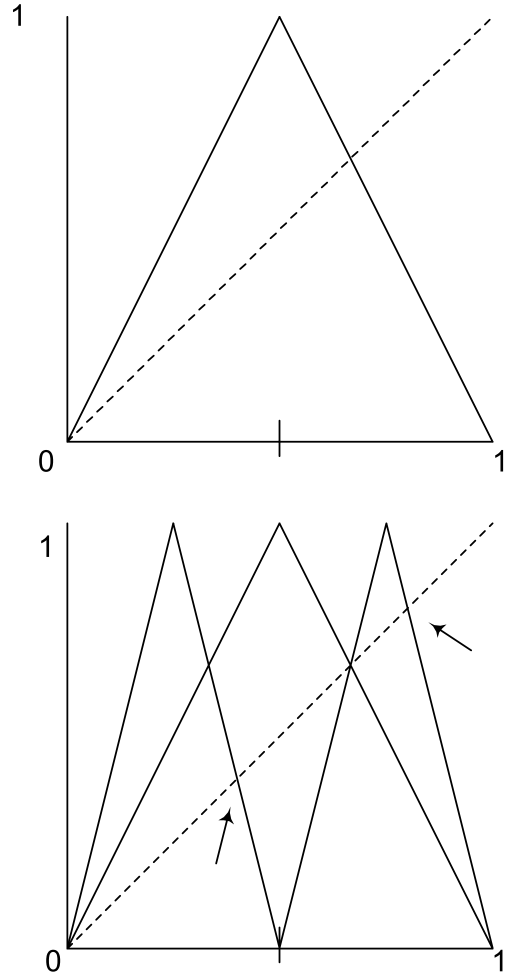
The tent map.

Well, that gives us something to chew on. The definition of a chaotic system. How can we actually prove that the tent map is chaotic? That sensitive dependence was one of the prerequisites was clear to us, but what about the other two? Imagine the domain of the function as an interval. For the logistic map, this is the interval [0,1]. If we take any point from this line and it turns out not to be a periodic point, and therefore a point of a chaotic orbit it should travel trough the entire domain under iteration of the function. After all, if its travel permit was restricted to a certain piece of the domain it wouldn't correspond to our intuition of chaos. "Why should it not come in this particular area?" Formally spoken the definition is as follows: if a system is topologically transitive, we can choose any open interval. This interval then contains a point that under iteration becomes an element of every other interval of the domain. The last requirement is one of regularity. Amidst the chaos there are periodic points as well, and even infinitely many and everywhere in the domain.

Propostition: periodic points are dense in [0,1] under iteration of the tent map

Lets recap on density. We call a set ‘dense’ if there’s a point to be found between every other two points. Think of the set of fractions. Now to address the question why periodic points are dense in [0,1] under iteration of the tent map.

Considering this context our question means that between any two periodic points on the interval [0,1], a third one exists We know of the existence of two fixed points, namely zero and two-third in the first-iteration-graph. Now if we proceed to constructing the second-iteration-graph, the interval on which it projects from 0 to 1 gains a fixed point. This interval is [0,½], but also [½,0]. After all, if some area projects to [0,1] after one iteration, then the second iteration is again a tent map. It follows that in the second-iteration graph two ‘small tents’ be visible. These small tents also go from zero to one, but on half the domain. You can imagine that for the third iteration graph, two more tents become visible. And these tents, whose number increase as the iteration graph increases, all go from zero to one on a certain domain and thus cross the y=x line.



The first and second-iteration graph of the tent map.

Subproposition: The set of (eventually) periodic points is equal to the set of fractions.

Proof: each fraction can be written as a/b with $a,b \in Z$. But the number of fractions with b in the lower part is limited. In this sense only five fractions of four exist in [0,1], 0/4, 1/4, 2/4, 3/4, 4/4. And since iteration only incorporates doubling with or without a substraction of either 2 or ½, and in this sense remains a fraction or at most is turned into a fraction in which a and b are multiplied by two (for the ½ compatibility). Considering that the is only a limited number of a fraction with a certain b it follows that if we iterate a fraction it must repeat itself under iteration. In other words: all fractions are periodic points. Is the converse also true? That is, are all periodic points fractions, or are there also periodic points that are not fractions?

The truth reveals itself by looking at the n^{th}-iteration graph of the tent map. As we have seen, each iterations more tents emerge, the same height but on ever smaller domain pieces. This means the derivative of each linepiece after n iterations is $\pm 2^n$. Since each of these linepieces goes from 0 to 1 they each cross the y=x line (guaranteed by the intermediate value theorem). So solving all these line formulas to y=x gives all the fixed points. Because of their form $y = \pm(2^n x) \pm c$ (with $c \in N$) these values are all fractions. Concluding we may confidently say that all fractions are periodic points and all periodic points are fractions. The sets are equal. Since fractions are dense in D, so are periodic points.

Proposition: the tent map is topologically transitive

Definition: a function f: D→D is topologically transitive if for two open intervals U and V in D there is a point z in U such that $f^n(z) \in V$ for some n.

So a function from and to the same interval (in our case [0,1]) is topologically transitive if from any open interval you pick a point it travels to any other open interval under iteration of f. We could say such a function 'well- mixes the domain'. You might wonder: "why not say: pick any point and under iteration it goes everywhere in the domain?". This would be a reasonable question but wrong in the details. We cannot pick just any point, for if we pick a fraction it does not go everywhere since it's periodic. Therefore, we need the open interval.

We have seen that the set of fractions is equal to the set of periodic points. From this it follows that any point which is no fraction is no periodic point. Lucky us, for there's the square roots. Square roots are not fractions as we've seen in the preliminaries and second: they're dense. So if we pick any open interval, however small, there's a square root in it. So if we can prove that every arbitrarily chosen square root travels around in the entire interval [0,1] we have proven that the tent map is topologically transitive. Here goes.

Take a square root w. For instance w = √(1/6). This square root has two precursors, one on the left and one on the right of ½. Remind yourself that this is true because the graph of the tent map goes from 0 to 1 on the interval [0, ½] but also on the interval [½,1]. So on both these intervals there must be a value (let's call them w_1 and w_2) which after iteration takes on the value √(1/6).

And here's the recursive trick again. These values w_1 and w_2 of course *also* have two precursors, both on one side of ½. Let's give these precursors of w_1 the name w_{1a} and w_{1b} and similarly w_{2a} and w_{2b} for w_2. Because they are precursors $f(w_{2a})=w_2$ and $f(w_{2b}) = w_2$. And since iteration of w_2 yields w, it follows that $f^2(w_{2a}) = w$. The further we're looking at precursors the more we find. There are two points for one iteration (which we called w_1 and w_2), there are four for two iterations (which we labeled w_{1a}, w_{1b}, w_{2a} and w_{2b}) and similarly eight points for third-iteration-precursors of w.

Since the two points of a precursors are located on either side of ½, that is, one in the first half of the domain and one in the second half of the domain, we can imagine that second-iteration-precursors all lie in their respective quarters of the domain. This is not strange to imagine because the second-iteration-graph of the tent map assumes all values, including w, four times (check out the earlier shown pictures if you're not too sure).

Continuing is this fashion, there are eight points in [0,1] which project to w after three iterations and 16 points for four iterations. The further we look the more we find. For 30 iterations we find 1073741824 ($=2^{30}$) different points that eventually yield √(1/6), and all nice and equally distributed on [0,1]. In other words, precursors of √(1/6) are *dense* in [0,1]. This again means that for any open interval we choose, there is a precursor of √(1/6) in it. You might get the point right now, but for the sake of completeness we'll have a look at the proposition again.

Proposition (restated): The tent map is topologically transitive if for any two open intervals U and V in [0,1] there is a point in U that under iteration goes to V.

Take some open interval V. Take some square root value. This is possible because square roots are dense in [0,1]. Since it's precursors are also dense there is a precursor in U and therefore there is a point in U which under iteration reaches V (eventually). Ergo: the tent map is topologically transitive.

Proposition: the tent map is sensitive to initial conditions

This last requirement might just be the most characteristic for a chaotic dynamical system: sensitivity for initial conditions. This is a very easy requirement. Of you take any two values really close together two things can be the case:

Take two initial conditions, x_0 and y_0, which are very at at a very small distance (let's call this distance δ (delta)). Now two situations are possible:

1) x_0 and y_0 are on the same side of ½ (both under the same diagonal)
2) x_0 and y_0 are on different sides of ½,

Note that if 2) is the case they lie very close to ½ just because we chose δ so small. Because of this, they might even be *closer* together after one iteration, but at least then they're on the same half of the domain by which situation 1) is the case. As soon as they are, they each iteration, their difference will get twice as big because they project to a line that, due to its steepness, covers twice the distance vertically as it does horizontally.

2.6.2 Topological conjugacy

If you would ever pick up a dutch bible, the first words you would read would be "in den beginne...". If you would ever read the first words of an english bible it would say "In the beginning...". The exact same idea though in a different language. You could prove this by constructing a mapping from english to dutch {(in,in), (the,den), (beginning,beginne), ... }. Note that this mapping also works the other way around, from dutch to english. The essence, the structure, is the same though the language is different.
We have reached the final stage of our proof. We have just shown that the tent map matches all criteria for being a chaotical function and now we will show, by a mapping called a *topological conjugacy* that the structure of the tent map and the logistic map for A=4 are the same though their appearance is different. This topological conjugacy could be seen as the language mapping for our maps.

Definition: We have two functions on their domains: $f:D\rightarrow D$ and $g:E\rightarrow E$. A topological conjugacy is a homeomorphism $\tau:D\rightarrow E$ such that $\tau(f(x)) = g(\tau(x))$. If such a topological conjugacy exists between f and g we call them topologically conjugated.

This definition ircorporates the word homeomorphism

Definition: Een homeomorphisme is een function that is surjective, injective, continuous and has a continuous iverse.

In an injective function, every projection is unique. The function $f(x)=x^2$ for example, is not injective, because both -2 and 2 project to 4. A surjective function uses the whole codomain. The function $f(x)=\sqrt{x}$ is surjective on $[0,4]\rightarrow[0,2]$ because for every value in codomain [0,2] there is an original in [0,2] that projects to it. For the same reason $[4,16]\rightarrow[0,4]$ is not surjective. If we take the value 1 from [0,4], there is no original in [4,16].

Example homeomorphism: $f(x)=\tan(x)$ is a homeomorphism from $(-\pi/2, \pi/2)$ to R. For all values in $(-\pi/2, \pi/2)$ there is a uniquely corresponding value from R. is er een unieke bijbehorende waarde uit R. Besides, the entire R is used. Tan(x), and arctan(x) are both continuous and therefore, tan(x) is a homeomorphism from $(-\pi/2, \pi/2)$ to R.

Example topological conjugacy: Two functions $f(x) = ½x :[0,2]\to[0,2]$ and $g(x)= ¼x:[0,4]\to[0,4]$ are topologically conjugate. The topological conjugacy used is $\tau(x)=x^2$. Look what happens if we take the highest initial value:

$f(x) = ½x :[0,2]\to[0,2]$	2	1	0.5	0.25	0.125
$g(x) = ¼x:[0,4]\to[0,4]$	4	1	0.25	0.0625	0.015625

Clearly to be seen is that both system's initial values tend towards the 0 attractor. The vertical quadratical relation is clearly visible. For explanatory values, we give a commutative diagram In which the arrows indicate the appliances of $f(x)$,$g(x)$ en $\tau(x)$. Check the diagram to convince yourself, that improves understanding of the concept.

2	$\to f(x) \to$	1	$\to f(x) \to$	0.5	$\to f(x) \to$	0.25	$\to f(x) \to$	0.125
$\downarrow$		$\downarrow$		$\downarrow$		$\downarrow$		$\downarrow$
$\tau(x) = x^2$		$\tau(x)$		$\tau(x)$		$\tau(x)$		$\tau(x)$
$\downarrow$		$\downarrow$		$\downarrow$		$\downarrow$		$\downarrow$
4	$\to g(x) \to$	1	$\to g(x) \to$	0.25	$\to g(x) \to$	0.0625	$\to g(x) \to$	0.015625

But analytically it also makes sense: The paramount prerequisite for a topological conjugacy is:

$\tau(f(x))=g(\tau(x))$, In our case: $(½x)^2= ¼(x^2)$, written out: $¼x^2 = ¼x^2$

In other words: since this equation is firm like a rock and $\tau(x)$ [0,2] -> [0,4] is a homeomorphism, f(x) and g(x) are topologically conjugate.

The tent and it's logistics. There is a topological conjugacy from the tent map to the logistic map, though it's split up in two parts because of the tent's sharp top. This is a mere detail. The topological conjugacy for $x\leq½$ is $\tau(x)= \sin^2(½\pi\, x)$. It relies on some mathematical formulas and there's no very easy way of explaining them in the proof, so we will just apply them:

1) $\sin^2(\pi x) = 4\sin^2(½\pi x) \cos^2(½\pi x)$
2) $\cos^2(½\pi x) = 1- \sin^2(½\pi x)$

$\tau(T(x)) = h(\tau(x))$, filled out:

$\sin^2(½\pi\, 2x) = 4\sin^2(½\pi\, x)\,(1-\sin^2(½\pi\, x))$ by expanding on 2:

$\sin^2(½\pi\, 2x) = 4 \sin^2(½\pi\, x)\cos^2(½\pi x)$ and then by expanding on 1:

$4 \sin^2(½\pi\, x)\cos^2(½\pi x) = 4 \sin^2(½\pi\, x)\cos^2(½\pi x)$

The function $\sin^2(\frac{1}{2}\pi\ x)$ is continuous, injective, invertible and has a continuous inversion on [0,1]. Therefore it is a homeomorphism and so the tent map and the logistic map are topologically conjugate. They have similar dynamics and since the tent map was proven chaotic, so is the logistic map. By now, you should be aware of the powerful forgery that a homeomorphism is.

2.7 Back to biology

What implications do all these results have for the real world? Just imagine our model for population growth, Ax(1-x) in which A is the fertility rate, is correct. If the biologist measures the number of lizards on the island and wants to make estimates for the future, he could be in serious trouble As we have seen, for low A-parameter values the system simply propagates towards its fixed point. This would mean for the biologist that he could predict very accurately what the population goes to. If the fertility parameter is slightly higher, he must realize that the population size will fluctuate between certain values, or , the population fluctuates over a period of two, four or eight years. But what happens if the fertility rate grows beyond 3.57? All predictability ceases to exist and there is no way of saying what next year's number of lizards will be.

This is quite a fundamental argument. Of certain phenomena of nature, such as gravity exerting its influence on a rock, we can very accurately predict what will happen; we have very well-defined formula's and the only chance incorporated is the small perturbation we cannot measure of choose to neglect. But certain aspects of nature so it seems, are fundamentally unpredictable. One such an example is the logistic map. Another such example might be the weather. Fully governed by a large set of differential equations, it shows some of the same characteristics including unpredicatbility. And the most remarkable is: the unpredictability is not due to minor external fluctuations, it is a property of the system itself.

3. Randomness

3.1 Proces & product

It's a rainy day in Amsterdam. You're sitting in your room, bored because it's raining, flipping a guilder coin through the air and seeing which side it lands on. From the outcome, you decide it's totally random whether the heads or tails comes up. As on all rainy days, you're kind of philosophical and wonder whether the chaos of the universe reflects the unpredictability of the coin flip. After quite a few flips, you decide the coin flip is a real random generator. There's no saying which side comes up. They both have the probability ½. You recall a famous statement by Joseph Ford: "Chaos is merely a synonym for randomness." "Is this true?, *Is* chaos the same as randomness?" you wonder. To this rainy day, we'll devote a chapter.
After getting your umbrella and raincoat, you go out to visit a friend. Bad weather as it is, you want to take a taxi. You walk up to Amsterdam central station where lots of taxis are parked waiting for passengers. You skim the numbers briefly to see if your favourite number 71 is in but as you look, you only see even numbers. This strikes you as extremely odd. Only taxi's with numbers 2,4,6,8,10,12.... Why is this? Are only even numbers going on rainy days? It could be a coincidence but that doesn't seem very likely to you. Still it could be just a coincidence, every combination of taxinumbers is equally probable but you refuse to believe it is. It just can't be.

> "In our minds, we organize all events in classes. We regard as special classes those which are very rare. As such, we regard one hunderd consecutive heads from a coin flip as special, because it has much more *structure* and structured sequences are rare as opposed to *unstructured* sequences. That is, sequences with repeative patterns which we can conceive and understand. " [freely translated after Laplace]

A statement by Pierre-Simon Laplace (1749-1827). What Laplace means is further illustrated once you reach your friend. Since the bad weather, he's just as bored as you and you decide to play some scrabble. You both take seven letters from the box and put them on your stand. H-R-N-I-C-E-K. Not very good. You can make 'hire' or 'neck' but it's a nasty combination. Your friend starts and lays down the word 'chaotic'. All letters gone in one go. He's won in the first turn. How is this possible? First the taxi numbers and now this! He must be cheating! How else can he draw the letter C-H-A-O-T-I-C and you a worthless combination like H-R-N-I-C-E-K?

Still, it just *might* be a coincidence and this exactly emphasizes laplace's point. Such an event you would classify as rare, because it's structured and there's more unstructured than structured sequences. It's all about the structure in your interpretation. If the same game would be played in Hungaria the opposite could be the case. The guy with H-R-N-I-C-E-K on his stand lays it all down in one go because 'hrnicek' happens to be a perfectly constructed hungarian word.

Pierre-Simon Laplace (1749-1827)

The guy with ‘chaotic’ would be perplexed, for he doesn’t speak english. (translated from hungarian)“This can not be a coincidence! How can you get ‘hrnicek’ in one go while I have a rubbish combination like C-H-A-O-T-I-C on my stand? You must be cheating!” Aim of the story is that all combinations of letters have the same probability but although a combination like “CHAOTIC” has the same probability as "HRNICEK", "AKNLWMB" or "ENIARKU", it is very clear that our intuition of randomness of objects (like taxinumbers or scrabble-sequences) is dependend on whether is has interpretable structure.
Back to our more universal, language-independent example. We take a fair coin (assume that probabilities are fifty-fifty). If we now flip a number of times adn recors a ‘1’ each time heads comes up, and a ‘0’ for each tails. The result will be a random string, no matter what it be. It might be 100 zeros or 100 ones but since we assumed that our coin is fair, it is random. In such cases you could wonder if the coin really is fair, but that would trap us in a circle-statement: “We flip our fair coin 100 times, but if the outcome is really structured the coin wasn’t fair after all.” But what if we only get a string presented and then have to decide wether the generation of this string was done by a random process?

And here’s our problem presented on a silver plate. It depends on the *process* wether the *product* is random. But what if we only have the product? Can we say anything about its randomness without knowing how it was formed?

3.2 Infinite strings

Let’s examine a string of which we don’t know whether the process was random or not. An infinte string of zeros and ones. The chance process guarantees us that the distribution of zeros and ones should be about fifty-fifty if the string is random. Imagine we do heads or tails a ‘large number of times’. Then half that ‘large number of times’ should be heads, the other half be tails. But is this enough to decide for its randomness? We flip a number of times and get :

1100110011 0011001100 1100110011 0011001100 1100110011 0011001100 1100110011 00....

Now that doesn’t look very random. If this continues in this fashion, you would be able to predict exactly the 1021st outcome. Just a repeating pattern. If the first is ‘1’, then so is the 21st, the 41st, the 81st and so the 1021st. And isn’t it a hallmark of randomness *not* to be predictable? But yet, the distribution of heads and tails exactly matches the fifty-fifty criterion. Obviously, that requirement is not enough. For instance, the subsequence ‘000’ is to be found nowhere. But also five consecuitive tails do not exist in this sequence.
The mathematician E. Borel called a string of zeros and ones ‘normal’ if every block of length n had frequency 2^{-n} in that string. Would you count the number of blocks of length two (00,01,10,11) than every block would have equal probability and thus account for 25%. Would you count all blocks of length four (0000,000... 1111) then each of these should have 1/16 = 6.25%. That seems a likely measure for a random string but there came Champernowne.
Champernowne has show that numbers exist that are normal in Borel’s sense but still not correspond to our intuition of randomness.:

1234567891011121314151617181920212223... (Champernowne's number)

Intuitively very clear that this is not a random string (be it on base 10 except of 2). Still every block of every length has the same frequence. It follows that being-normal according to Borel is not a sufficient measure for randomness. On the other hand, if a string is random it is also normal, otherwise we would have an unequal disribution of certain subsequences and that would not be the result of a fair coin.

Valid is: "x is random $\Rightarrow$ x is normal"
And also: "x is not normal $\Rightarrow$ x is not random",

But *not*: "x is normaal $\Rightarrow$ x is random" (because of Champernowne's number, which Is normal but still not random)

The randomness of an individual infinite string is quite hard to pinpoint. Even be there a statistical measure, like some advanced-Borel-normality there would be a practical problem. The string would in the long run converge to match this measure. But what is the long run? A thousand coin flips? A million? 10^{600} coin flips? None. Because even after 10^{600} coinflips there could come 10^{700} zeros messing up our statistical measure. As Keynes told us "The long run? In the long run we'll all be dead!" by which he meant you can never trust a sequence to converge up to a certain point. There is always a subsequent-sequence that trows a rock in the pond. By now it might be clear that the determination of randomness of an individual string is no Swiss cottage cheese.

3.3 Finite strings

Let's have a look at finite strings. If the randomness of an infinite string is hard to pinpoint, what to say about finite strings now that even convergence is not something to rely on? What makes a finite string random or not?

Back to our heads or tails game. We decide to flip our coin twenty times and denote the result. We repeat this experiment five times which rends the following data:

1. 11011100101010001010
2. 11111111111111111111
3. 00011011100000101110
4. 01101010001010001010
5. 10101010101010101010

Some of these strings look very random. Some don't. Even if the outcome seems to match a fifty-fifty criterion in heads-tails it still doesn't need to be random. Borel's normality does not apply for infinte strings, but even apart from that, look at string number 5. Look at string number 2. No sensible human being would call either of those a random string. There is a gap between our intuition and science. Probability theory says each string has the same chance though 2 and 5 feel very unrandom. We can clearly see *structure*. Laplace, appear!

What is this structure we observe? In any case, repeating patterns exhibit a 'rule', a generalized way of its apperance. String two is just 20 times "1", string 5 just 10 times "10". The nice thing is that we can formalize this. Repeativeness, structure means we can give an 'easy formula' (20 x "1") that describes the original object (11111111111111111111). Obviously, the randomness we feel to haunt over certain strings have something to do with it's lack of pattern whereas orderly strings seem to be describable in an easy formula. An appeal to intuition: which telephone number is easier to remember:

1) 0031 - 37501 - 683992 -1143893
2) 4444 - 55555 - 666666 -7777777

Obviously the one with the most structure. Obviously the one for which we can, in our minds, design an easy formula to remember. Is there a way of formalizing this intuition? The answer is of course yes, but we will have to rely on computational theory. You will have to know a little it about computer programming for a thorough understanding but references to 'method' and 'algortihm' will also be given. Check strings A and B:

A: 1111111111 1111111111 1111111111 1111111111 1111111111 1111111111

B: 1010100011 1011111101 1011111010 1000110110 0011111011 1010101001

String B is random, String A is not. String A can be written as:

A2: For a=1 to 60; print"1"; next a

As string B can be written as:

B2: print"1010100011 1011111101 1011111010 1000110110 0011111011 1010101001"

And now we have caught our 'structural intuition' in a computer program. String A can be given by a very short description, a repeative algorithm whereas string B cannot. It is exactly this 'being describable in a short formula' which indicates the randomness of a finite string. In computer science we would say: String A is *compressible*, string b is not. The program A2 required to produce A is shorter than A itself. The shorter this description, the less random it is.

We bump our heads into a number of problems at once. As a practical one, this method is obviously not valid for strings of all lengths. If we have a string of only ten bits, it is impossible to find a computer program which is shorter. Computer programs of less then ten bits do not exist.

Second, it obviously depends on the programming language we use. If we use a cumbersome programming language with long instructions the description will obviously be longer.

Third, how do we know which program is actually the shortest we can find. Obviously, we can add a huge amount of useless instructions to any program, so there are many programs that can produce the string we want to compress. It is also possible to device two sets of instructions totally equivalent in output. Which we use is of course of critical importance to the length of our program. Is there any saying if there is an even shorter equivalent set?

But all these problems can easily be avoided by referring to a Turing Machine. Alan Turing, an englishman and pure genius that decyphered the German enigma-code in the second world war was also foremost concerned with computation theory and artificial intelligence. He designed a machine, later called the Turing Machine, which can be seen as a purely mathematical model of a computer. Using this abstraction we can avoid or solve some of the issues. In essence, all computer-processors are able to perform the same operations. These operations are described by the Turing Machine and as such diffuses the question of different programming languages.

We can very easily make some general remarks about structureless strings now. Imagine we have a 100-bit long string. We call such a string 'random' if it is compressible by 20%, that is, if there is a program of 80 bits or less that exactly reproduces our string . This is a very modest choice. If you WINZIP a text file you can reach compression rates of 50% easily. How many of these strings do exist anyways? Binairy calculation shows there are $2^{100} \approx 1.26*10^{30}$ such strings. But there are only $2^{81}-1 \approx 2.42*10^{24}$ strings of length 80 or shorter. So how many strings are incompressible according to our 20% standard? At least $1.26*10^{30} - 2.42*10^{24}$! More than 99% that is. A simple counting argument gives a striking result. With a very modest compression rate on a relatively short string, the number of random strings very far exceeds the non-random strings. And just how neetly does this fit Laplace's intuition: "*...and structured sequences are rare as opposed to unstructured sequences ...*".

The notion of randomness as stated here is the invention of A.N. Kolmogorov and closely related to Kolmogorovcomplexity. His original notion was far more complicated and thorough, but the main idea is that randomness, lack of structure, is conceivable of as incompressibility.

3.4 Chaos and randomness

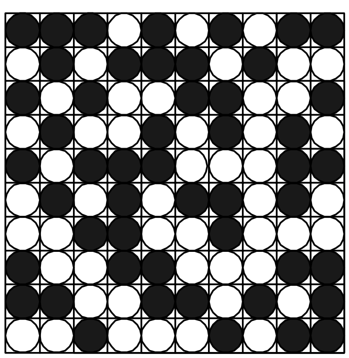
Then the main question. To what extend is chaotic behaviour comparable to randomness? We ask ourselves this question step-by-step, always related to the logistic map.

- As a process: A random process is a process for which a certain probability distribution governs the outcome. The logistic map is different. It is *deterministic*. Each starting value projects only to one subsequent value. Therefore, the logistic can never be regarded of as a random generator.
- As a product. If we take a product, like the binary expanded orbit of initial value ½, is the product it yields by any chance random. No. There is an easy-to-construct computer program that does the iteration for us and therefore the whole string would be compressible.
- So what do they have in common? Answer is unpredictability. For 2000 coinflips there is no saying what the outcome is and so it is for the logistic map. It just follows a certain pattern of probability and so does the logistic map. When A=4, orbits dance around In a negative-exponential manner always showing the same statistical characteristics, never the same pattern.

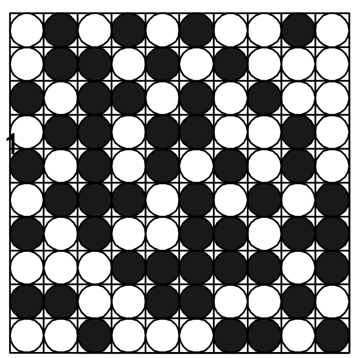
3.5 Exercises

1. Below depictuted are a few samples of the Japanese board game GO. In all of them, fifty squares are covered with white chips and fifty are covered with black chips. Which look the most randomly covered? Why?

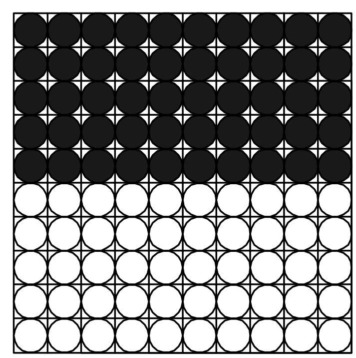
1



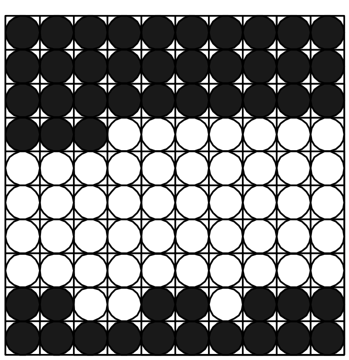
2



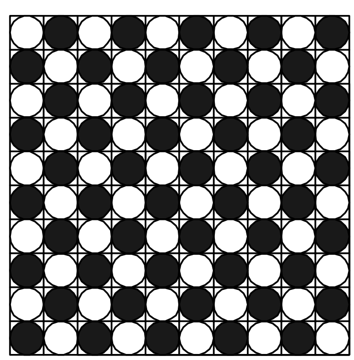
3



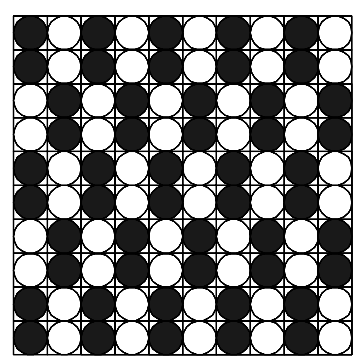
4



5



6



2. Which of the following phonenumbers is easiest to remember? Which is the intuitively most compressible and how is this related?

A: 0041-4572-33819
B: 0011-9999-99999
C: 0057-5757-57162

3. Put these 54-bits strings according to increasing randomness.

A: 111110011010100001110001010000110010101010110001110101
B: 110011001100110011001100110011001100110011001100110011
C: 101001000100001000001000000100000001000000001000000000
D: 001010011101011010010101000011101010001111010111001101
E: 111111111111111111111111111111111111111111111111111111

Solutions to exercises

Paragraph 1.3

1. Consider the discrete dynamical system: $x_{t+1} = 4x_t$
a) The system's only fixed pint is 0.
b) It has no periodic points.
c) The orbit of-2: {-2,-8,-32,-128,-512,...}.
The orbit of 4.1: (4.1, 16.4, 65.6, 262.4, 1049.6,...)
d) 0 is a repellor

2. Consider the discrete dynamical system: $x_{t+1}= 2x_t(1-x_t)$
   a) The orbit of 0: {0}
   The orbit of 0.5: {0.5}
   The orbit of 0.1: {0.1, 0.18, 0.2952, 0.416.... and further towards 0.5)
   The orbit of 1: {1,0}
   b) This system has one repellor at 0 and one attractor at 0.5. In fact, this is a special case of attractor; in this case, the attractor is the critical point and convergence of all orbits towards 0.5 is much faster than in other cases. It is therefore sometimes called *superstable*.

3. Consider the discrete dynamical system: $x_{t+1}= \frac{1}{4} x_t$
   a) It has one fixed point at 0.
   b) It has no periodic points other than zero (which we call fixed, because its period is 1)
   c) They both tend towards zero.
   d) 0 is an attractor.

4. Consider the discrete dynamical system. $x_{t+1}= 3.1x_t - 3.1(x_t)^2$
   a) It has two fixed points: one at 0 andthe other at 21/31
   b) This is best done by graphical analysis
   c) -2 goes to $-\infty$, 0.9 tends to a period two orbit.
   d) Determine of all fixed points are repellors, the periodic orbit is attracting.

5. A colony of kangaroos inhabits southern Australia. This population doubles each year.
   a) $x_{t+1} =2x_t$, in which x is the population size and t the time.
   b) Because your (inaccurate) estimate will also double under iteration the margin remains within 5% correct. In fact, it will stay precisely the same value.

Paragraph 1.5

1. Consider the discrete dynamical system $x_{t+1} = 2x_t$ on R->R
   a) It has one fixed point at 0, found by solving 2x=x
   b) It is a repellor (unstable), since the derivative is 2 which is greater than 1.

2. Consider the discrete dynamical system $x_{t+1} = (x_t)^2$
   a) It has two fixed points: 0 and 1
   b) The derivative of $f(x)=x^2$ is f(x)=2x. We can determine the hyperbolicity by determining f'(0) and f'(1).In this case, f '(0)=0, so this one is stable whereas f'(1)=2 which implies it is unstable.

3. Consider the discrete dynamical system x_{t+1},= $2+\sqrt{x_t}$ on $[0, \infty) \rightarrow [0, \infty)$
   a) It has two fixed points: one in 0 en and one in 4.
b) The derivative of f(x) = 4x is given by f (x) = $1/(2\sqrt{x_t})$. In 0, the derivative is not defined, but any value arbitrarily close to zero is very high and thus we may conclude that it repels. In 4 the derivative is 0.25 and it is therefore an attractor.

Paragraph 3.5

1. The least random are obviously boards 5,3 and 6. They are all comprised of repeating patterns. Intuitively, we can remember them easily "Oh, it's just half-black half-white." Or we could easily construe a computer program for them. The most random are probably boards one and two and board four is somewhere in between. It could be recreated by a computer program which lines 1,2,3 are generated by a repeating loop, 4 is literally printed, 5,6,7,8 are again looped, 9 is again printed and 10 is a loop again.

2. Telephone number B is most easily remembered and most easily reconstructable from an easy algorithm. C has some repetition but a small irregularity on the end and is therefore more random whereas A is the most patternless. Note that this is an appeal to intuition more than to computation theory.

3. E,B,C,A,D