0.1 The Painlevé equations

0.2 Spontaneous singularities: The Painlevé’s equation P₁

Let us analyze local singularities of the Painlevé equation P₁,

\[ y'' = y^2 + x \] (1)

In a neighborhood of a point where \( y \) is large, keeping only the largest terms in the equation (dominant balance) we get \( y'' = y^2 \) which can be integrated explicitly in terms of elliptic functions and its solutions have double poles. Alternatively, we may search for a power-like behavior

\[ y \sim A(x - x₀)^p \]

where \( p < 0 \) obtaining, to leading order, the equation \( A(p - 1)x^{p-2} = A^2(x - x₀)^2 \) which gives \( p = -2 \) and \( A = 6 \) (the solution \( A = 0 \) is inconsistent with our assumption). Let’s look for a power series solution, starting with \( 6(x - x₀)^{-2} \):

\[ y = 6(x - x₀)^{-2} + c₁(x - x₀)^{-1} + c₂ + \cdots \]

We get: \( c₁ = 0, c₀ = 0, c₁ = 0, c₂ = -x₀/10, c₃ = -1/6 \) and \( c₄ \) is undetermined, thus free. Choosing a \( c₄ \), all others are uniquely determined. To show that there indeed is a convergent such power series solution we substitute \( y(x) = 6(x - x₀)^{-2} + \delta(x) \) where for consistency we should have \( \delta(x) = o((x - x₀)^{-2}) \) and taking \( x = x₀ + z \) we get the equation

\[ \delta'' = \frac{12}{z^2} \delta + z + x₀ + \delta^2 \] (2)

Note now that our assumption \( \delta = o(z^{-2}) \) makes \( \delta^2/(\delta/z^2) = z^2 \delta = o(1) \) and thus the nonlinear term in (2) is relatively small. Thus, to leading order, the new equation is linear. This is a general phenomenon: taking out more and more terms out of the local expansion, the correction becomes less and less important, and the equation is better and better approximately by a linear equation. It is then natural to separate out the large terms from the small terms and write a fixed point equation for the solution based on this separation. We write (2) in the form

\[ \delta'' - \frac{12}{z^2} \delta = z + x₀ + \delta^2 \] (3)

and integrate as if the right side was known. This leads to an equivalent integral equation. Since all unknown terms on the right side are chosen to be relatively smaller, by construction this integral equation is expected to be contractive.

[Click here for Maple file of the formal calculation \( (y'' = y^2 + x) \)]

The indicial equation for the Euler equation corresponding to the left side of (3) is \( r^2 - r - 12 = 0 \) with solutions 4, -3. By the method of variation of parameters we thus get

\[ \delta = \frac{D}{z^3} - \frac{1}{10} x₀ z^2 - \frac{1}{6} z^3 + C z^4 - \frac{1}{7 z^3} \int_0^z s^4 \delta^2(s)ds + \frac{z^4}{7} \int_0^z s^{-3} \delta^2(s)ds \]

\[ = -\frac{1}{10} x₀ z^2 - \frac{1}{6} z^3 + C z^4 + J(\delta) \] (4)
the assumption that $\delta = o(z^{-2})$ forces $D = 0$; $C$ is arbitrary. To find $\delta$ formally, we would simply iterate \((4)\) in the following way: We take $r := \delta^2 = 0$ first and obtain $\delta_0 = -\frac{1}{10}x_0z^2 - \frac{1}{6}z^3 + Cz^4$. Then we take $r = \delta_0^2$ and compute $\delta_1$ from \((4)\) and so on. This yields:

$$
\delta = -\frac{1}{10}x_0z^2 - \frac{1}{6}z^3 + Cz^4 + \frac{x_0^2}{1800}z^6 + \frac{x_0}{900}z^7 + ... \quad (5)
$$

This series is actually convergent. To see that, we scale out the leading power of $z$ in $\delta$, $z^2$ and write $\delta = z^2u$. The equation for $u$ is

$$
u = -\frac{x_0}{10} - \frac{z}{6} + Cz^2 - \frac{z^{-5}}{7} \int_0^z s^8u^2(s)ds + \frac{z^2}{7} \int_0^z su^2(s)ds
\quad = -\frac{x_0}{10} - \frac{z}{6} + Cz^2 + J(u) \quad (6)
$$

It is straightforward to check that, given $C_1$ large enough (compared to $x_0/10$ etc.) there is an $\epsilon$ such that this is a contractive equation for $u$ in the ball $\|u\|_{\infty} < C_1$ in the space of analytic functions in the disk $|z| < \epsilon$. We conclude that $\delta$ is analytic and that $y$ is meromorphic near $x = x_0$.

Figure 1: The six Painlevé equations, all equations of the form $y'' = R(y', y, x)$, with $R$ rational, having the PP.

The equation $y'' = y^2 + x^2$ does not have the Painlevé property.

Click here for Maple file of the formal calculation, for $y'' = y^2 + x^2$

1 Discrete dynamical systems

The study of the Poincaré map leads naturally to the study of discrete dynamics. In this case we have closed trajectory, $x_0$ a point on it, $S$ a section through $x_0$ and we take a point $x_1$ near $x_0$, on the section. If $x_1$ is sufficiently close to $x$, it must cross again the section, at $x_1'$, still close to $x_1$, after the return time which is then close to the period of the orbit. The application $x_1 \to x_1'$ defines the Poincaré map, which is smooth on the manifold near $x_0$.

The study of the behavior of differential systems is near closed orbits is often more easily understood by looking at the properties of the Poincaré map.

In one dimension first, we are dealing with a smooth function $f$, where the iterates of $f$ are what we want to understand.
We write $f^n(x) = f(f(...(f(x))))$ $n$ times. The orbit of a point $x_0$ is the sequence $\{f^n(x_0)\}_{n \in \mathbb{N}}$, assuming that $f^n(x_0)$ is defined for all $n$. In particular, we may assume that $f : J \to J$, where $J \subset \mathbb{R}$ is an interval, possibly the whole line.

The effects of the iteration are often easy to see on the graph of the iteration, in which we use the bisector $y = x$ to conveniently determine the new point. We have $(x_0, 0) \to (x_0, f(x_0)) \to (f(x_0), f(x_0)) \to (f(x_0), f(f(x_0)))$, where the two-dimensionality and the “intermediate” step helps in fact drawing the iteration faster: we go from $x_0$ up to the graph, horizontally to the bisector, vertically back to the graph, and repeat this sequence.

There are simple iterations, for which the result is simple to understand globally, such as

$$f(x) = x^2$$

where it is clear that $x = 1$ is a fixed point, if $|x_0| < 1$ the iteration goes to zero, and it goes to infinity if $|x_0| > 1$.

Local behavior near a fixed point is also, usually, not difficult to understand, analytically and geometrically.

**Theorem 1.** (a) Assume $f$ is smooth, $f(x_0) = x_0$ and $|f'(x_0)| < 1$. Then $x_0$
is a sink, that is, for \( x_1 \) in a neighborhood of \( x_0 \) we have \( f^n(x_1) \to x_0 \).

(b) If instead we have \(|f'(x_0)| > 1\), then \( x_0 \) is a source, that is, for \( x_1 \) in a small neighborhood \( \mathcal{O} \) of \( x_0 \) we have \( f^n(x_1) \notin \mathcal{O} \) for some \( n \) (this does not mean that \( f^n(x_1) \) cannot return “later” to \( \mathcal{O} \), it just means that points very nearby are repelled, in the short run.)

Proof. We show (a), (b) being very similar. Without loss of generality, we take \( x_0 = 0 \). There is a \( \lambda < 1 \) and \( \epsilon \) small enough so that \( |f'(x)| < \lambda \) for \( |x| < \epsilon \). If we take \( x_1 \) with \( |x_1| < \epsilon \), we have \( |f(x_1)| = |f'(c)||x_1| < \lambda|x_1|(< \epsilon) \), so the inequality remains true for \( f(x_1) : |f(f(x_1))| < \lambda|f(x_1)| < \lambda^2|x_1| \) and in general \( f^n(x_1) = O(\lambda^n) \to 0 \) as \( n \to \infty \).

In fact, it is not hard to show that, for smooth \( f \), the evolution is essentially geometric decay.

When the derivative is one, in absolute value, the fixed point is called neutral or indifferent. It does not mean that it can’t still be a sink or a source, just that we cannot resort to an argument based on the derivative, as above.

**Example 1.** We can examine the following three cases:

(a) \( f(x) = x + x^3 \).

(b) \( f(x) = x - x^3 \).

(c) \( f(x) = x + x^2 \).

It is clear that in the first case, any positive initial condition is driven to \(+\infty\). Indeed, the sequence \( f^n(x_1) \) is increasing, and it either goes to infinity or else it has a limit. But the latter case cannot happen, because the limit should satisfy \( l = l + l^3 \), that is \( l = 0 \), whereas the sequence was increasing.

The other cases are analyzed similarly: in (a), if \( x_0 < 0 \) then the sequence still diverges. Case (c) is more interesting, since the sequence converges to zero if \( x_1 < 0 \) is small enough and to \( \infty \) for all \( x_1 > 0 \). We leave the details to the reader.

It is useful to see what the behavior of such sequences is, in more detail. Let’s take the case (c), where \( x_1 < 0 \). We have

\[
x_{n+1} = x_n + x_n^2
\]

where we expect the evolution to be slow, since change is vanishingly small, even in a relative sense (that is, relative to \( x_n \)). We then approximate the true evolution by a differential equation, pretty much as it is often done in models (in biology, economics and so on),

\[
(d/dn)x = x^2
\]

giving

\[
x_n = (C - n)^{-1}
\]
We can show rigorously that this is the behavior, by taking \( x_n = -1/(n+c_0) + \delta, \) \( \delta_{n_0} = 0 \) and we get

\[
\delta_{n+1} - \delta_n = \frac{1}{(n+c)^2(n+c+1)} - \frac{2}{n+c} \delta_n + \delta_n^2
\]  

(7)

and thus

\[
\delta_n = \sum_{j=n_0}^{n} \left( \frac{1}{(j+c)^2(j+c+1)} - \frac{2}{j+c} \delta_j + \delta_j^2 \right)
\]  

(8)

**Exercise 1.** Show that (8) defines a contraction in a ball of radius \( C \) in the space of sequences with the norm \( \sup_{n \geq n_0} |n^2 \delta_n| \), where you choose \( C \) carefully.

**Exercise 2.** Find the behavior for small positive \( x_1 \) in (b), and then prove rigorously what you found.

### 1.1 Bifurcations

The local number of fixed points can only change when \( f'(x_0) = 1 \). As before, we can assume without loss of generality that \( x_0 = 0 \).

We have

**Theorem 2.** Assume \( f(x, \lambda) \) is a smooth family of maps, that \( f(0,0) = 0 \) and that \( f_x(0,0) \neq 1 \). Then, for small enough \( \lambda \) there exists a smooth function \( \varphi(\lambda) \), also small, so that \( f(\varphi(\lambda), \lambda) = \varphi(\lambda) \), and the character of the fixed point (source or sink) is the same as that for \( \lambda = 0 \).

**Exercise 3.** Prove the theorem, using the implicit function theorem.

The only other type of change of local behavior near one fixed point is the change in nature (from sink to source).

**Exercise 4.** Show that a sink (source) remains a sink (source, resp.) under small perturbations. (Use of the the implicit function theorem is again a simple way.)

A typical bifurcation near \( f' = -1 \) is a period doubling bifurcation. Say, the fixed point is \( x^* = 0 \). For small \( x \) we have

\[
f(x; \lambda) = f(0, \lambda) + f'(0; \lambda)x + f''(c; \lambda)x^2/2 \approx \mu x + f''(0; \lambda)x^2 = \mu x + ax^2
\]

where \( \mu \) is close to \( -1 \), \( \mu = -1 + \epsilon \) where \( \epsilon \) is small and we assume that \( a = a(\epsilon) \) is far enough from zero.

We can calculate

\[
f(f(x)) - x = x(x^2a^2 + x\epsilon + \epsilon)(-2 + \epsilon + xa)
\]  

(9)

We see that, in this setting, \( x = 0 \) remains a fixed point; this is by normalization, since we choose to shift \( x^* \), which in general depends on \( \lambda \), to zero.
The multiplier at zero \( f'(f(0))f'(0) = f'(0)^2 = -1 + \epsilon(-2 + \epsilon) \). This changes from less than one to bigger than one, meaning that the two-period switches character between stable and unstable.

There is a second periodic orbit, when \( ax - 2 - \epsilon \) and that one (typically) preserves its character (check). But we see something else happening: the discriminant

\[ \Delta(x^2a^2 + xa\epsilon + \epsilon) = -4\epsilon + \epsilon^2 \]

changes sign when \( \epsilon \) goes through zero. This means the birth of two more periodic orbits.

**Exercise 5.** Make the conclusions and analysis above rigorous.

See next section for an example.

### 1.2 The logistic map

This is one of the most studied iteration maps. It is only in small part due to its interpretation as a population model, more realistic than the continuum one, \( x' = ax(1 - bx) \). One of the more important reasons is that it is in some sense generic, the simplest interesting iteration, in which we take two Taylor coefficients, and also that it captures much of the complexity of map iterations. (Just one term would lead to a triviality). It is the quadratic logistic map

\[ x_{n+1} = \lambda x_n(1 - x_n) \quad (10) \]

Surprisingly perhaps, understanding the details of this iteration is highly non-trivial and it took many years until it has been completely (or nearly so) understood in the last few years.

There is a different representation often used. With the transformation \( x = -\lambda^{-1}X + 1/2 \) we get

\[ X_{n+1} = X_n^2 + c; \quad c = \frac{\lambda}{4}(2 - l) \quad (11) \]

The transformation is trivial, but it turns out that some features are much more easily understood if looking at both representations.

We consider (10) on the interval \([0, 1]\), which is invariant under the transformation, until \( \lambda = 4 \). Check this.

In the representation (10) it is clear that the map has two fixed points, \( x = 0 \) and \( x = (\lambda - 1)/\lambda \). Check that the first one is attracting for \( \lambda < 1 \) and repelling for \( \lambda > 1 \), and also that the second one, outside our interval if \( \lambda < 1 \), is attracting for \( \lambda \in (1, 3) \).

At \( \lambda = 3 \) the character of the nonzero fixed point changes. We can look at either representation, for instance at (11) where \( \lambda = 3 \) corresponds to \( c = -3/4 \). The fixed point of interest is \( X = -1/2 \). There, we have \( f' = 2X = -1 \), when \( c = -3/4 \). The fixed point goes from stable to unstable. At the same time, a period doubling bifurcation occurs. Let us draw the pictures of \( f(x) - x \) and \( f(f(x)) - x \), near \( c = -3/4 \).
Figure 3: \( x^2 + c - x, c = -3/4 - 0.1, c = -3/4, c = -3/4 + 0.1. \)

**Exercise 6.** Prove rigorously that there is a period doubling at \( c = -3/4 \) and that one of the two new periods is stable and the other one is unstable.
Figure 4: \((x^2 + c)^2 + c - x, c = -3/4 - 0.1, c = -3/4, c = -3/4 + 0.1\). The “discontinuity” is due to intentional limitation of the window of the plot.
Figure 5: A few iterations of the logistic map $\lambda x(1 - x)$ for different values of $\lambda: -0.5, 1, 1.5, 2, 3.2, 3.5, 3.7, 4$ respectively, from left to right and top to bottom. The very last figure is for $\lambda = 4$, and $f^5$ only.

Exercise 7 (+). Prove rigorously, as much as you can, of the structure of the iterations, for $\lambda \leq 3.2$.

We see that at $\lambda = 4$, the picture looks quite complicated and no convergence is apparent on the figure. The point is that now the map $f$ is two-to-one, and onto. This means that $[0, 1/2]$ is mapped onto $[0, 1]$, and the same is true for $[1/2, 1]$. Then, $f^2$ will map, by symmetry, $[0, 1/4]$ onto $[0, 1]$ and similarly for $[1/4, 1/2]$, $[1/2, 3/4]$, and $[3/4, 1]$. In general $f^n$ will map any interval of the form $[2^{-n-1}j, 2^{-n-1}(j + 1)]$ onto the whole of $[0, 1]$. In general then, there are about $2^{n+1}$ periodic points, of period $n$. The evolution of an “arbitrary” point is quite random: it is arbitrarily close to one of the true periods, and thus it will return to almost the same value, infinitely often. But if it returns to the almost same value, it will return to almost all points it has already covered, and that suggests the trajectory is dense, perhaps in the whole interval.

1.3 Predictable behavior versus chaos

As we have seen, for $\lambda = 4$ the orbit of one, typical, seed is very complicated. We will return to this in a moment. The value of the parameter is crucial in this. For $\lambda$ less than 3.5 or so the behavior of the solutions of the logistic recurrence can be quite well predicted, at least when the seed is in $[0, 1]$. We will also look later into what happens if the seed is complex, since that is important too. Things get quite a bit more complicated there.

For now let us take $f(x) = \lambda x(1 - x)$, $\lambda \in [0, 1]$, and, of course, $x_0 \in [0, 1]$ too. Since $f([0, 1]) \subset [0, 1]$ the orbit of $x_0$ is well defined and it is confined in the interval $[0, 1]$. We can determine, with good precision the fate of every seed in $[0, 1]$. 

9
Indeed, the sequence
\[ x_{n+1} = \lambda x_n (1 - x_n) \]  \hspace{1cm} (12)
is monotonic decreasing thus convergent, and since it decreases the limit is zero. Thus, at every step, when \( x \) becomes very small, \( x_{n+1}/x_n \approx \lambda \). Thus we expect \( x_n \) to be of order \( \lambda^n \). We therefore write \( x_n = \lambda^n u_n \), and get
\[ u_{n+1} = u_n (1 - \lambda^n u_n) \]  \hspace{1cm} (13)
Since the \( x_n \) were all positive, so will be the \( u_n \). But then \( u_n \) is also monotonic, and will converge to a constant \( C = C(x_0) \) (which we cannot determine easily from (13), but there are methods to do that.)

We conclude that, for a constant \( C \) we have
\[ \lim_{n \to \infty} \frac{x_n}{C \lambda^n} = 1 \]  \hspace{1cm} (14)
We can relatively easily calculate corrections to this expansion. Indeed we have, since \( u_n \to C \),
\[ u_n - C = -\lambda^{n+1} u_{n-1} + (u_{n-1} - C) \]  \hspace{1cm} (15)
and thus \( |u_n - C| \leq 2C \lambda^{n+1} \), for large \( n \). Then,
\[ x_n = C(x_0) \lambda^n + O(\lambda^{2n}) \]  \hspace{1cm} (16)
We can proceed further and get as many corrections as we want. As mentioned, there are good methods to calculate \( C(x_0) \), but for now we could calculate an approximate \( C \) by taking \( n \) fairly large and using (16) to determine \( C \).

Let us take the middle range, \( \lambda = 2 \) this turns out to be one of the (few) exactly solvable cases.

Take \( x_n = 1/2 - z_n/2 \). The recurrence becomes
\[ z_{n+1} = \frac{z_n}{2^n} \]  \hspace{1cm} (17)
with initial condition \( z_0 = 1 - 2x_0 \). We see that the maps (12) and (17) with \( z_0 \in [-1, 1] \) are analytically conjugated (in fact linearly). Of course, the solution to (17) is \( z_0^{2^n} \) and thus
\[ x_n = 1/2 - (1 - 2x_0)^{2^n} \]  \hspace{1cm} (18)
and the fixed point is (super)attracting, and the basin of attraction of the fixed point 1/2 (we’ll talk some more about that later) includes (−1, 1).

The behavior is very deterministic. The precision in knowing the asymptotic fate of a given seed is extremely good.

Note that \( 1 - 2x_0 \) maps \([0, 1]\) (linearly) onto \([-1, 1]\). Thus, \( |z_0| < 1 \) if \( x_0 \in (0, 1) \). The solution of the iteration is

Let us return to “the other extreme”, \( \lambda = 4 \). This is another value for which the recurrence can be solved in closed form. If we take \( x_n = \sin^2 2\pi \theta_n \) we get
\[ \sin^2 2\pi \theta_{n+1} = 4 \sin^2 2\pi \theta_n (1 - \sin^2 2\pi \theta_n) = \sin(4\pi \theta_n)^2 \]  \hspace{1cm} (19)
Thus, if we consider the auxiliary sequence
\[
\theta_{n+1} = 2\theta_n \mod 1 \tag{20}
\]
or
\[
\theta_{n+1} = \begin{cases} 
2\theta_n & \text{if } \theta_n \leq 1/2 \\
2\theta_n - 1 & \text{otherwise} \end{cases} \tag{21}
\]
then \(x_n = \sin^2 2\pi \theta_n\) solves the original equation. This is called the doubling map, perhaps the simplest one which exhibits chaos. We will look at this map in detail. To see however the nature of the orbits, note that the recurrence is nicely viewed if we represent the numbers in \([0,1]\) in base 2: It then amounts to left shifting the binary representation and removing the integer part:
\[
x_n = 0.a_1a_2a_3...a_n,... \rightarrow x_{n+1} = 0.a_2a_3a_4...a_n,... \tag{22}
\]
Suppose all this was in base 10. (This makes no qualitative difference). Take \(x_0 = \pi\). The recurrence takes us through the digits of \(\pi\), one by one. Of course, this is not random behavior, since there is an underlying function behind it, but there is of course no way to find the orbit of \(x_n\) short of calculating all digits of \(\pi\). Returning to base 2, suppose we started with the seed \(\gamma\), the Euler constant, and we asked the “simple” question whether the trajectory is (eventually) one of the periodic ones. This is of course equivalent to the question whether \(\gamma\) is rational. This is not known yet, and it is suspected to be a very difficult question (Hilbert believed it is “unapproachable”). So the behavior of an individual point even quite explicit, becomes a daunting problem.

**Note 2.**

1. The transformation
\[
x_n = \frac{1}{2} - \frac{1}{4}(z_n + z_n^{-1}) \tag{23}
\]
transforms again into \(z_{n+1} = z_n^2\). But now, the initial condition becomes \(z_0 = (1-2x_0)\pm 2i\sqrt{x_0 - x_0^2}\). We check immediately that \(|z_0| = 1\) if \(x \in [0,1]\). The map \(23\) is now 2 to 1, from \([0,1]\) to the unit circle.

2. We choose one of the two solutions and get \(\exp(2\pi ia_{n+1}) = \exp(2\pi i2a_n)\) and again we see that a solution is provided if \(a_n\) solves the doubling iteration, \(a_{n+1} = 2a_n \mod 1\).

3. It turns out that much more generally, quadratic maps are conjugated (generically very non-explicitly so), in certain regions of \(\mathbb{C}\), to the canonical map \(17\); the question is whether the initial condition falls inside the unit disk (in which case \(z_n \to 0\)), outside \((z_n \to \infty)\) or exactly on the unit circle, when \(z_n\) is chaotic.

4. Note also that even for \(\lambda = 2\), if \(x_0 = 1/2(1 - e^{2\pi i\theta})\) the trajectory is, once more, chaotic. Everything depends on the initial condition.

But what is chaos? We explore this issue in the next section.
2 Chaos

It is not easy to define chaos in a way that at the same time captures our intuition about “complete disorder”, is mathematically rigorous, and is testable in sufficiently many cases.

So we take a reasonable approximation of our intuition and make it into a definition.

**Definition 1.** *The iteration of a map* $f : J = [a, b] \rightarrow [a, b]$ *is said to chaotic if*

1. Periodic points are dense in $J$.
2. $f$ is transitive on $J$, that is given arbitrary subintervals $U_1$ and $U_2$ there is a seed $x_0 \in U_1$ and an $n \in \mathbb{N}$ such that $f^n x_0 \in U_2$. (The image of any interval under all iterations is the whole of $J$.)
3. $f$ has sensitive dependence in $J$: there exists a constant $\beta$, the “sensitivity” constant such that, for any $x_0 \in J$ and any open interval $U$ around $x_0$, there is some seed $x_1 \in U$ so that

$$|f^n(x_0) - f^n(x_1)| > \beta$$

(24)

The definition is, as mentioned, debatable and in fact debated; for instance there are systems which are considered chaotic and which do not have periodic orbits.

Transitivity is equivalent to the existence of a dense orbit. In one direction, the result is immediate (check!). In the opposite direction, the proof relies on the Baire category theorem, and works more generally, for $T : X \rightarrow X$ a continuous transformation on a complete separable metric space without isolated points (see [4]). It goes as follows. Choose countable a dense set $x_j$ in $J$ (say, the rationals). Consider a sequence of smaller and smaller balls $B_m(x_j)$ (say of radius $1/m$) around $x_j$. We denote, as usual, by $T^{-1}(A)$ the preimage of $A$. We note that the condition that the trajectory of $x$ is dense can be written as

$$\forall (m,j) \exists n : T^n(x) \in B_m(x_j)$$

(25)

or

$$\forall (m,j) \exists n : x \in (T^n)^{-1} B_m(x_j)$$

(26)

Note that the sets $(T^n)^{-1} B_m(x_j)$ are open, and the union over $n$ is dense in $X$, by transitivity.

Now, (26) is equivalent to

$$x \in S = \bigcap_{m,j} \bigcup_n (T^n)^{-1} B_m(x_j)$$

(27)

and $S$ is a countable intersection of open, dense sets, and thus dense in $X$.

**Exercise 1.** *Show that the conclusion remains true, and the proof is similar, if $T$ is piecewise continuous on an interval $J$.**
**Definition 2** (Conjugacy, semiconjugacy). Consider two maps \( f_{1,2} : J_{1,2} \to J_{1,2} \). Then

1. The maps \( f_1 \) and \( f_2 \) are conjugate if there is a homeomorphism \( h : J_1 \to J_2 \) such that \( h \circ f_1 = f_2 \circ h \).

2. The maps \( f_1 \) and \( f_2 \) are semiconjugate if there is a continuous map \( h : J_1 \to J_2 \) and which is onto nd at most \( n \) to one such that \( h \circ f_1 = f_2 \circ h \).

**Note 3.** We see that \( f_1 \) and \( f_2 \) are semiconjugate if there is a continuous map \( h : J_1 \to J_2 \) and which is onto nd at most \( n \) to one such that \( h \circ f_1 = f_2 \circ h \).

**Exercise 2.** Prove 3. in the case \( h \) is continuous and at most \( n \) to one. One possibility is to show first that, under this assumption, for every point there is some \( x_1 \in J_1 \) such that \( h(x_1) = x_2 \) and an open set \( U \) around \( x_1 \) so that \( h \) is a homeomorphism between \( U \) and \( h(U) \).

Thus we see that chaoticity of the doubling map immediately establishes chaoticity of \( f_4 \).
2.1 The doubling map

If we consider the base 2 shift $S$, we can see that periodic points are dense. All rational numbers are periodic wrt $S$, and they are indeed dense. Rational numbers are the only periodic points, of course. A number has a dense trajectory iff its representation contains all finite strings of the form $a_1 a_2 \cdots a_n$, $a_i \in \{0, 1\}$ and $n$ arbitrary. Check that this is a set of measure one, in particular dense and uncountable. (However, note that density of the orbit of many specific numbers, such as $\pi, \gamma$ is highly nontrivial. In some sense, this is part of the fact that the behavior is chaotic.) Finally, sensitivity is immediate. Thus $f_4$ is chaotic too.

2.2 The tent map

This is the map

$$f_t = \begin{cases} 2x & \text{if} \ x < 1/2 \\ 2 - 2x & \text{if} \ x \geq 1/2 \end{cases} \tag{28}$$

Note that, if $f_d$ is the doubling map, we have

$$f_t(f_t(x)) = f_t(f_d(x)) \tag{29}$$

so that $h = f_t$ itself provides a semiconjugation map with $f_d$. Thus, again, chaoticity of $f_t$ is inherited from that of $f_d$. 

Figure 6: The first $10^4$ iterates of $f_4$, for the seed $x_0 = 0.33$. 

3 Symbolic dynamics

Chaotic dynamical systems of sufficiently simple systems can be often brought to a canonical form, the shift map, very similar to the doubling map.

We start by illustrating how this can be done for the logistic map, with \( \lambda > 4 \), since the reduction process is typical.

Note that for \( \lambda > 4 \), \( f_\lambda([0, 1]) \supseteq [0, 1] \) and some points exit the interval. We note that if \( f(x_0) \notin [0, 1] \), then necessarily \( f(x_0) > 1 \) and then \( f(f(x_0)) < 0 \).

From that point on, \( f^n(x_0) \) stays negative; if \( x = -s \) then \( f(s) = \lambda s(1 + s) \).

Since \( \lambda > 1 \) we then have \( f^n(x_0) \to -\infty \). The behavior of these points is easy to understand.

Do all points eventually exit \([0, 1]\)? Certainly not. For instance \( \{0, 1\} \mapsto \{0\} \) and 0 is a fixed point. But there are (uncountably) many more points which are confined to \([0, 1]\) for all \( n \).

Let \( A_0 \) the open interval where \( f > 1 \). Note that to exit \( J \) in one step, \( x_0 \) must be in \( A_0 \). Note also that the endpoints of \( A_0 \) are fixed.

Let \( A_1 = f^{-1}(A_0) \). We note that \( A_1 \) is disjoint from \( A_0 \) since \( f(A_0) \cap [0, 1] = \emptyset \). This \( A_1 \) is then the set of points which exit \( J \) after exactly two iterations, or enter \( A_0 \) after one. \( A_1 \) consists in two disjoint intervals, as it is easy to see. We note that the preimage of each of the two intervals composing \( A_1 \) consists in two disjoint intervals, and that \( A_2 = f^{-1}(A_1) \) consists in 4 disjoint open intervals, and in general, if we let we let \( A_n = f^{-1}(A_{n-1}) \), \( A_n \) will consist in \( 2^n \) disjoint open intervals. We also see that the points in \( A_n \) enter \( A_0 \) in exactly \( n \) steps, and exit the interval in exactly \( n + 1 \). In particular we see that \( A_n \supset A_{n+1} \). We also see that any point in \( J \) which is not in \( A_n \) will remain in \( J \) after at least \( n + 1 \) iterations. Any point leaving \( J \) eventually does so in \( n + 1 \) steps for some \( n \) and thus belongs to \( A_n \).

Therefore

\[
\Lambda = J \setminus \bigcup_{j=0}^{\infty} A_j = J \cap \bigcap_{j=0}^{\infty} A_j^C = \bigcap_{j=0}^{\infty} A_j^C
\]

consists in points which never leave \( J \); here the superscript \( C \) denotes the complement with respect to \( J \). \( \Lambda \) is forward invariant, and the map \( f \) is well defined from \( \Lambda \) to \( \Lambda \). For, \( f(x_0) \in \Lambda \) if \( x_0 \in \Lambda \), or else \( x_0 \) would exit \( J \).

So, \( f \) induces a self-consistent iteration, not of an interval, but of the set \( \Lambda \).

We will see shortly that \( \Lambda \) is an uncountable set.

We let \( I_0 \) and \( I_1 \) be the two components (note that these are closed intervals) of \( A_0^C \). Clearly, the entire orbit of a point in \( \Lambda \) is in \( I_0 \cup I_1 \). We take \( x_0 \in \Lambda \) and simply follow the sequence of intervals its successive iterations belong to. Say \( x_0 \in I_1 \), \( f(x_0) \in I_1 \), \( f^2(x_0) \in I_0 \), etc. We simply associate to \( x_0 \) the sequence \( I_1, I_1, I_0, ..., \) or, more simply the sequence \( 1, 1, 0, ... \). We will see shortly that there is a one-to-one correspondence between the space of sequences

\[
\Sigma = \{(s_0, s_1, s_2, ..., s_n, ...): s_i \in \{0, 1\}\}
\]

and \( \Lambda \). For a point \( x_0 \in \Lambda \), its associated sequence in \( \Sigma \) is called “the itinerary of \( x_0 \)".
For now, let us introduce a (relatively natural) metric on $\Sigma$. It is closely related to the usual distance on $[0, 1]$ where we interpret an element of $\Sigma$ as a number written in base 2.

**Definition 4.** We endow $\Sigma$ with the distance

$$d(s, t) = \sum_{i=0}^{\infty} \frac{|s_i - t_i|}{2^i}$$

(32)

**Exercise 1.** Check that (32) is indeed a distance function on $\Sigma$.

**Proposition 5.**

1. If $s_i$ and $t_i$ coincide for $i \leq n$, then $d(s, t) \leq 2^{-n}$.

2. Conversely, if $d(s, t) < 2^{-n}$, then $s_i = t_i$ for all $i \leq n$.

**Proof.**

1. Since $|s_i - t_i| \leq 1$, we see that the series is majorized by

$$\sum_{j=n+1}^{\infty} 2^{-j} = 2^{-n+1} \sum_{j=0}^{\infty} 2^{-j} = 2^{-n}$$

(33)

2. Very similar, left as an exercise.

We now study the mere correspondence between $\Lambda$ and $\Sigma$, as metric spaces.

**Theorem 3.** $\Lambda$ and $\Sigma$ are homeomorphic, and the itinerary map $S$ is the homeomorphism.

**Proof.** We only show this in the simpler case $\lambda > 2 + \sqrt{5} \approx 4.23$ when we have $\inf_{I_0 \cup I_1} |f' | > K > 1$.

We first show that $S$ is a bijection.

1. Injectivity: suppose the itineraries of $x$ and $y$ coincide. Then, of course, $f^n(x_0)$ and $f^n(x_1)$ belong to the same interval for all $n$. Say for a given $n$, $f^n(x_0) < f^n(x_1) < 1/2$ Then $f$ is increasing $[f^n(x_0), f^n(x_1)]$, all points in $[f^n(x_0), f^n(x_1)]$ will be mapped into points between $[f^{n+1}(x_0), f^{n+1}(x_1)]$. So the points travel along the same itinerary together with all points in-between, and we are dealing with the trajectory of a whole interval, $[x_0, x_1]$, which remains an interval at every iteration. But

$$\text{diam}[f^{n+1}(x_0), f^{n+1}(x_1)] > K \text{diam}[f^n(x_0), f^n(x_1)] > K^n \text{diam}[x_0, x_1]$$

Since $K^n \to \infty$, this is only possible if $\text{diam}[x_0, x_1] = 0$.

2. Surjectivity. We must now show that, given an itinerary, there is always one point (we know now that it’s unique) which corresponds to this trajectory.
We let $S = (s_0, s_1, \ldots)$ and denote the corresponding intervals by $J_0, J_1, \ldots$. We have $x_0 \in J_0$, $f(x_0) \in J_1$, $f^2(x_0) \in J_2$ and so on. Then, we take an $x_0$ in $I_{s_0, s_1, s_2} = J_0 \cap f^{-1}(J_1) \cap f^{-2}(J_2)$. We show by induction that this is a sequence of closed nonempty intervals. Of course $J_0$ is nonempty. We assume that the result is true for any combination of intervals $J_0, \ldots, J_l$, and any $l \leq n - 1$.

**Note.** For any closed nonempty interval $[a, b]$ we have $f^{-1}([a, b])$ is the union of two closed nonempty intervals, one in $I_0$ and the other one in $I_1$.

By induction we know that the set belonging to step $n - 1$

$$A = \{ x : x \in J_1, f(x) \in J_2, \ldots, f^{n-1}(x) \in J_n \} \quad (34)$$

is a closed nonempty interval.

We take $f^{-1}(A)$. By the note above, it consists in exactly two nonempty closed intervals, one in $I_0$ and the other one in $I_1$. We take the one in $J_0$. By construction then, for all $x \in A \cap J_0$ we have

$$x \in J_0, f(x) \in J_1, \ldots, f^n(x) \in J_n \quad (35)$$

and the induction step is proved.

The intervals are evidently nested, since at every step an extra restriction is imposed on $x$. The intersection of all these intervals, nonempty too, is the desired point.

3. Continuity. Take any $\epsilon > 0$ and let $\epsilon \in (2^{-L-1}, 2^{-L})$. Given $x$, we need to find a $\delta$ so that the trajectory of $x$ is the same as the trajectory of $y$ for the first $L + 1$ steps, the trajectories of $x$ and $y$ coincide. Assume first $x \in \text{int} (I_0 \cup I_1), \ldots, f^{L+1}(x) \in \text{int} (I_0 \cup I_1)$ (the boundary is left as an easy exercise). Let $d_k = \text{dist} (f^k(x), \partial(I_0 \cup I_1))$ and $\Delta = \min\{d_1, \ldots, d_{L+1}\}$. The vector valued function $(f(x), f^2(x), \ldots, f^{L+1}(x))$ is continuous and there is a For any $L$, there is a $\delta$ so that we have $|f^l(x) - f^l(y)| < \Delta/2$ for all $l \leq L + 1$. Then, clearly $x$ and $y$ have the same trajectory up to $L + 1$, and thus $d(S(x), S(y)) \leq 2^{-L-1}$.

☐

3.1 A digression: continued fractions and approximations of irrational numbers

Let us look at a map which is in some ways related to the doubling map: rotation by a given angle on the circle. The application is $x \mapsto x + \alpha \mod 2\pi$. To normalize things, we can take a circle of circumference one, measure the angle in terms of the arclength $s$ and then the map becomes $x_{n+1} = x_n + s \mod 2\pi$. We note that in this form, as a map from $[0, 1]$ to $[0, 1]$, with fixed $a$ this is not chaotic according to our definition, since, for instance, there are no periodic points if $a$ is irrational. Also, it is easy to see that the map is not
sensitive, the points on the circle move in sync. Nonetheless, if we see it as
a map on the arclengths \( a, a \mapsto na \) mod 1 then we have all the hallmarks of
chaoticity, since a subset of the orbit is the doubling map.

**Note.** The orbit is periodic iff \( a \) is rational. Indeed, clearly, if \( a = p/q \)
is rational, then the orbit of \( a \) is periodic, since \( qna = pn \) mod 1 for all \( n \).
Conversely, if the orbit is periodic, then \( qna = pma \) mod 1 or \( qna = pma + N \)
for some \( N \in \mathbb{Z} \) and thus \( a \) is rational.

We will look at this map geometrically, as it gives an interesting perspective
on the relation between rational and irrational numbers. We will see that for
any irrational number, the unit circle is densely covered by the trajectory.

![Figure 7: \( a \mapsto na \) mod 1 (the circle has circumference one) for \( a = 1/9, 3/7, \sqrt{5} \). The last three correspond to \( a = \sqrt{5} \), and the number of iterations is 10, 50, 150 resp.](image)

### 3.2 Approximation by rational numbers

If \( a \not\in \mathbb{Q} \) and we need to find the best approximants by rational numbers, we
can use this circle dynamics to find them. First, what do we mean by best
approximant? We will understand by a *best approximant of the first kind* any
ratio \( p/q \) (we assume the fraction irreducible) with the property that \(|a - p/q| < \)
\[|a - p'/q'| \text{ for any } q' < q. \] Note that there are finitely many \(p\) for a given \(q\) so that \(|a - p/q| < 1\). The candidates are among numbers of loops for which \(f^n a\) returns close to the origin. There is another notion of proximity, that inherited directly from the circle rotation, namely \(p/q\) is a best approximant if for all \(q' < q\) and all \(p'\) we have

\[|qa - p| < |q'a - p'| \quad (36)\]

This is called a best approximant of the second kind.

**Exercise 2.** Show that a best approximant of the second kind is always also a best approximant of the first kind.

We can proceed as follows. Let’s start with some \(a\) irrational, and we may assume without loss of generality that \(a < 1\) (if \(a > 1\) we write \(a = [a] + a'\), \(a' < 1\) and the rotation by integers is identified with the zero rotation.

We join the successive points, to obtain a never closing polygonal line.

We rotate by \(a\) until we go one full loop, and we let \(k + 1\) be the first number such that \((k + 1)a > 1\). Clearly, we must have \(1 - ka < a\). Otherwise either \(1 - (k + 1)a < 0\), or \((k + 1)a = 1\) which would make \(a\) rational.

We thus have

\[ka = 1 - a'_1; \quad a'_1 < a \quad (37)\]

Let us assume for now that \(a'_1\) is less than a fraction of \(a\), as this is a simpler case to visualize. It is clear that after another \(k\) rotations, the point will return at \(-2a_1\), and so son, until the side of the initial polygon is covered by the now slower-sliding corners of the first polygon. This requires \([a_1/a_2]\) rotations. Then the point will return near the origin, evidently even closer than before. And so on. Let’s see what we get this far, and then once we know what we need to do we’ll proceed analytically; this will now be far simpler.

Click to get the animated gif of the rotation

Note that we have

\[a = \frac{1}{k_0} - \frac{a'_1}{k_0} \text{ where } a'_1 > 0; \quad \frac{a'_1}{k_0} \leq \frac{1}{k_0} \quad (38)\]

since \(a < 1/k_0\).

Therefore the first best approximant is just \([1/a]\). What is the value of \(a_1\), which guides us to the next best approximant?

We have

\[a_1 \approx \frac{1}{k_1}; \quad k_1 = \left\lfloor \frac{a}{1 - ka} \right\rfloor \quad (39)\]

or

\[k_1 = \left\lfloor \frac{1}{a_1} \right\rfloor; \quad a_1 = \{1/a_1\}; \quad (40)\]

The process applied to \(a_1\) is the same as the one we applied to \(a\). Furthermore, we see from (39) that in the next approximation we have

\[\frac{1 - ka}{a} \approx \frac{1}{k_1}; \quad \frac{1}{a} \approx k + 1/k_1 \text{ or } a \approx \frac{1}{k + 1/k_1} \quad (41)\]
So, algebraically, the process to get here is the following: We write
\[ a = \frac{1}{n_0 + a_1} \] (42)
with \( a_1 \in (0, 1) \) and in this form we know that \( 1/n \) is a best approximant. Now, as we saw, we take \( a_1 = 1/(n_1 + a_2) \) where, by the same argument we used for \( a \), we know that \( |a_1 - 1/n_1| \leq n_1^{-2} \). So we continue writing
\[ a_1 = \frac{1}{n_1 + a_2}; \quad a_2 = \frac{1}{n_2 + a_3} \] (43)
etc. and in the same way, inductively, we get
\[ a_0 = a = \frac{1}{n_0 + \frac{1}{n_1 + \frac{1}{n_2 + \cdots + \frac{1}{n_k + \epsilon_k}}}} \] (44)
where
\[ a_k \text{ and } \epsilon_k \text{ are in } (0, 1) \] (45)
Now, we also see that
\[ a_j = \frac{1}{n_j + \frac{1}{n_{j+1} + \cdots + \frac{1}{n_k + \epsilon_k}} \frac{\epsilon_k}{n_k}} \] (46)
To estimate the error in approximating \( a \) by the \( k-1 \)th partial quotient or \( k-1 \)th convergent
\[ T_{k-1} = \frac{1}{n_0 + \frac{1}{n_1 + \frac{1}{n_2 + \cdots + \frac{1}{n_k-1}}}} \] (47)
we use the mean value theorem: for some \( c \in (0, 1) \) we have
\[ a - T_{k-1} = \left( \frac{da_1}{da_k} \right) \bigg|_{\epsilon_k = c} \epsilon_k = \left( \prod \frac{da_j}{da_{j+1}} \right) \bigg|_{\epsilon_k = c} \epsilon_k \] (48)
which gives
\[ |a - T_{k-1}| \leq \frac{1}{n_0^2 n_1^2 \cdots n_{k-1}^2 n_k} \] (49)
Thus, in particular, \( \lim_{k \to \infty} T_k = a \), and we write
\[ a = \frac{1}{n_0 + \frac{1}{n_1 + \frac{1}{n_2 + \cdots + \frac{1}{n_k + \cdots}}}} \] (50)
We note that the denominator \( q_{k-1} \) of \( T_k \) (after brought to an irreducible fraction) satisfies (for all \( k \))
\[
q_{k-1} \leq n_0 n_1 \cdots n_{k-1} \tag{51}
\]
and we get
\[
|a - T_{k-1}| \leq \frac{1}{q_{k-1}q_k} \tag{52}
\]

**Remark 5.** The notation, if \( a \) is as in \((50)\), is
\[
a = [0, n_0, n_1, n_2, \ldots, n_k, \ldots]; \quad N + a = [N, n_0, n_1, n_2, \ldots, n_k, \ldots], \quad N \in \mathbb{N} \tag{53}
\]
The \( n_i \) are called *quotients*.

**Remark 6.** We have
\[
T_1 < T_3 < \cdots < a < \cdots < T_2 < T_0 \tag{54}
\]
This can be seen for instance by induction and looking at the signs of the successive approximations and at the estimates, noting that we have both lower and upper bounds on \( \epsilon_k \).

**Proposition 6.** The \( T_k \) provide all best approximations of the second kind of \( a \).

**Proof.** Indeed, if \( p/q \) would be another approximant, by \((54)\) we would have that \( p/q \) is between two successive convergents, \( T_k \) and \( T_{k+2} \). We assume \( k \) to be odd, the other case being similar. We thus have
\[
\frac{1}{qq_k} \leq \left| \frac{p_k}{q_k} - \frac{p}{q} \right| < \left| \frac{p_k}{q_k} - \frac{p_{k+1}}{q_{k+1}} \right| < \frac{1}{q_kq_{k+1}} \tag{55}
\]
The first inequality comes from the fact that \( p/q \) and \( p_k/q_k \) are irreducible and different, and the middle one from the fact that \( p_{k+1}/q_{k+1} > a > p/q \). Thus we must have \( q > q_{k+1} \). We similarly have
\[
\left| a - \frac{p}{q} \right| > \left| \frac{p_{k+2}}{q_{k+2}} - \frac{p}{q} \right| \geq \frac{1}{qq_{k+2}} \tag{56}
\]
so that
\[
|qa - p| > 1/q_{k+2} \tag{57}
\]
whereas we know that
\[
|q_k a - p_k| \leq 1/q_{k+2} \tag{58}
\]
from our general inequalities, and the fact that \( q > q_k \) yields a contradiction. \( \square \)

**Exercise 3.** Show that \( T_k \) has denominator \( n_0 n_1 \cdots n_k \), that is the fractions are from the beginning irreducible.
3.3 First recurrence process

By looking how we obtain \( n_0 \) from \( a \), \( n_1 \) from \( a_1 \) and so on, we see that we simply follow the simple recurrence

\[
 n_k = \left\lfloor \frac{1}{a_k} \right\rfloor; \quad a_{k+1} = \left\{ \frac{1}{a_k} \right\} \tag{59}
\]

3.4 Some special numbers

1. First we see that the continued of rational numbers is finite. The number itself is its best approximant up to that level, and from that point on too. The series of approximants stops there.

2. We also see easily that quadratic numbers (solutions of quadratic equations with integer coefficients) have periodic continued fractions, and in fact they are the only ones with this property. Indeed, a continued fraction is periodic iff

\[
a = \frac{1}{n_0 + \frac{1}{n_1 + \frac{1}{n_2 + \ddots}}} = \frac{q_1 + a}{q_2 + q_3 a} \tag{60}
\]

We see that the right hand side is a sequence of linear fractional applied to \( a \) and thus, after simplification, we have

\[
a = \frac{q_1 + a}{q_2 + q_3 a} \tag{61}
\]

where \( q_i \) are rational. But this happens iff \( a \) is quadratic.

3. some special values:

\[
\sqrt{2} = [1, 2, 2, 2, \ldots]; \quad \sqrt{3} = [1, 1, 2, 1, 2, 1, 2, \ldots] \tag{62}
\]

while

\[
[1, 1, 1, \ldots] = \frac{1}{2} (1 + \sqrt{5}) \tag{63}
\]

which is the golden ratio \( \varphi \). This is systematically the farthest possible from rationals as seen by looking at the left side of \( \sqrt{2} \).

More difficult to show are the following

\[
\tanh 1 = [0, 1, 3, 5, 7, \ldots]; \quad e = [2, 1, 1, 2, 1, 1, 4, 1, 1, 6, \ldots] \tag{64}
\]

while \( \pi \) has no known pattern in its continued fraction

\[
\pi = [3, 7, 15, 1, 292, 1, 1, 1, 2, 1, 3, 1, 14, 2, 1, 1, 2, 1, 2, 1, 1, 2, 1, 84, 2, 1, 1, 15, 3, \ldots] \tag{65}
\]

but we see that if we stop just short of 252, we get a great approximation for \( \pi \), which is the “famous” \( \pi \approx 355/113 \).
Proposition 7 (Typicality). Note that the most numbers in the set of finite decimal expansions of length less than $q_k/2$ cannot be approximated better than $1/q_k^2$ by rational numbers with denominator less than $q_k$.

Proof. We just sketch an idea of a proof here: It is clear that if this was not the case, then we could establish a one-to-one correspondence between the real numbers with exactly $n$ decimal places and a smaller subset of the same (writing, say, the numerator followed by zero followed by the denominator as a string). It is an information-content approach.

Theorem 4 (Loch, 1964). For almost all real numbers in the interval $(0, 1)$, the number of terms $m$ of the number’s continued fraction expansion needed to determine the first $n$ places of the number’s decimal expansion behaves asymptotically as follows:

$$\lim_{n \to \infty} \frac{m}{n} = \frac{6 \ln 2 \ln 10}{\pi^2} \approx 0.97027014$$

Remark 7. The set of real numbers for which the best approximants satisfy

$$|x - p_k/q_k| < q_k^{-2-\epsilon}, \forall k$$

has zero measure.

In particular, This plays an important role in ODEs and many other problems, since in power series representations of linearization maps, we naturally have denominators of the form $k_1a + k_2$, “small denominators”, where $k_1$ and $k_2$ are integers, positive or negative, and the convergence of the series is contingent upon these denominators not being too small. The corollary gives a lower bound, with probability one.

A deep, recent result, by Roth states that the lower bound in the corollary is also an upper bound, for any $\epsilon > 0$, for all algebraic numbers.

3.5 Recurrence relations

We drop the primes now. We have

$$a_k = \frac{1}{n_k + \frac{1}{n_{k+1} + \cdots}}$$ (66)

and then

$$a_k = \frac{1}{n_k + a_{k+1}}; \quad a_{k+1} = \frac{1}{a_k} - n_k$$ (67)

we see that we get $a_{k+1}$ simply as the integer part of $1/a_k$ (since $a_k \in (0, 1)$ and $n_k$ is an integer. At every step we take the reciprocal of the previous number, we set aside the integer part as our $n_k$, keep the fractional part and repeat.
References


