university of groningen

faculty of science and engineering

Chaos in Newton's method

Master Project Mathematics

July 2021

Student: M.J. van Harten

First supervisor: dr. A.E. Sterk

Second supervisor: prof. dr. H. Waalkens

Abstract
This thesis focusses on the map $T(x) = \frac{1}{2}(x - \frac{1}{x})$. It originates from Newton's method of finding the zeros of $z^2 + 1$ on the real line. The dynamics of the dynamical system are investigated. Different definitions of chaos are considered and we prove that T is chaotic according to some of these definitions. Furthermore, using ergodic theory, we find the probability density function for the invariant measure. Using Birkhoff's ergodic theorem we determine the Lyapunov exponent.

Contents

1 Properties of the map **7**
1.1 Topological Conjugacy . . . 7
1.2 Periodic points . . . 9
1.2.1 Using the conjugacy . . . 10
1.3 Set of points that map to zero K_T . . . 12
1.3.1 K_T using the conjugacy. . . . 13
1.4 Relation between T and the Logistic map . . . 14

2 Chaos **15**
2.1 Li and Yorke Chaos . . . 16
2.2 Marotto . . . 18
2.2.1 Snap-back repeller of the logistic map . . . 18
2.2.2 Snap-back repeller of T . . . 19
2.3 Devaney Chaos . . . 19
2.4 Dispersion Exponent . . . 22
2.4.1 The Doubling map . . . 23
2.4.2 The Logistic map . . . 24
2.4.3 Map T. . . . 25

3 Ergodic theory **27**
3.1 Invariant measure . . . 27
3.2 Birkhoff's ergodic theorem . . . 29
3.3 Lyapunov exponent . . . 30

4 Family N_c **31**

A Appendix **34**
A.1 Inverse functions . . . 34
A.1.1 Inverse T^{-1} . . . 34
A.1.2 Inverse cotangent . . . 35
A.1.3 T and the logistic map . . . 35
A.2 Chaos . . . 36
A.2.1 Li and Yorke . . . 36
A.2.2 Dispersion Exponent Matlab . . . 37
A.3 Ergodic Theory . . . 38
A.3.1 Density functions and invariant measure T . . . 38
A.3.2 Density function and invariant measure F . . . 39
A.4 Lyapunov Exponent . . . 40
A.4.1 Logistic map F . . . 40
A.4.2 Family N_c . . . 41

Introduction

This thesis investigates the dynamics of the map $T(x) = \frac{1}{2}(x - \frac{1}{x})$. Let us begin at the start, where does it come from?

It is often useful to know the zeros of a function $f(z)$, i.e. the values of z for which $f(z) = 0$. For polynomial functions of degree 1 or 2, they are easy to compute. Even for degrees 3 and 4, there have been methods developed to find the zeros. For degrees ≥ 5, there is no method to find the zeros. Luckily, in 1669, Newton found a method to approximate the zeros for any differentiable (complex) function [19]. For this method, we need an initial value z_0 and recursive iterations:

$$z_{n+1} = z_n - \frac{f(z_n)}{f'(z_n)}.$$

This sequence z_n converges to a limit z, which is a zero of f. Functions can have multiple zeros and different initial values lead to different zeros of $f(z)$. In that case, we can define the *basin of attraction* of each zero. The basin of attraction is the set of initial values converging to that specific zero. Now let us consider:

$$f_c(z) = z^2 + c$$

We can find the zeros of f_c by hand:

$$f_c(z) = 0 \Rightarrow z = \begin{cases} \pm\sqrt{-c} & \text{if } c < 0 \\ 0 & \text{if } c = 0 \\ \pm i\sqrt{c} & \text{if } c > 0 \end{cases}$$

Let us also consider Newton's method for this example. For $c < 0$ the basins of attraction of $\pm\sqrt{-c}$ are respectively $\{z \in \mathbb{R} : z > 0\}$ and $\{z \in \mathbb{R} : z < 0\}$. For $c = 0$, the basin of attraction is all of $\mathbb{R}$, all points converge to 0 by Newton's method. For $c > 0$, we have imaginary zeros, and the basins of attraction lie in the complex plane. The basin of attraction of $i\sqrt{c}$ is the upper half complex plane $\{z \in \mathbb{C} : \mathrm{Im}(z) > 0\}$ and the basin of attraction of $-i\sqrt{c}$ is the lower half complex plane: $\{z \in \mathbb{C} : \mathrm{Im}(z) < 0\}$. We can for example look at the convergence of Newton's method for $f_1(z) = z^2 + 1$. Below are the iterations for initial value $z_0 = 1 + 0.5i$, which lies in the basin of attraction of $+i$. Indeed, after 4 iterations, we are already getting close to i:

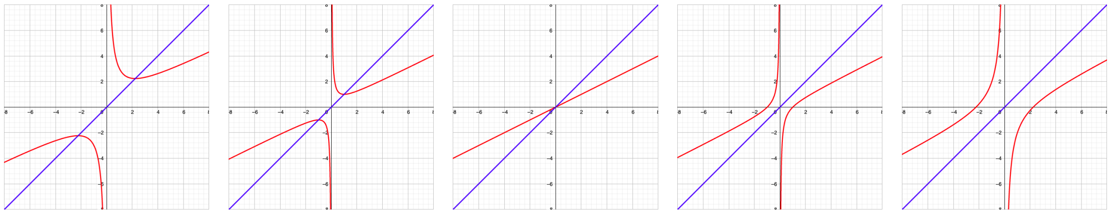
$$\begin{aligned} z_0 &= 1 + 0.5i \\ z_1 &= 0.1 + 0.45i \\ z_2 &= -0.1853 + 1.2838i \\ z_3 &= -0.0376 + 1.0234i \\ z_4 &= -0.0009 + 0.9996i \\ &\vdots \\ z_n &\approx i \end{aligned}$$

A computer can determine basins of attraction by calculating the convergence for all initial values. In the case of f_c, the basins of attraction are simple, but for some functions they are really complicated, and give spectacular images if we colour each basin in a different colour. These images are called fractals and they are not the focus of this paper. More information about this can be found in, for example, these papers: [16, 2].

Going back to our function f_c, we can ask the question of what happens if we apply Newton's method with an initial value that is not in either basin of attraction, for example $z_0 \in \mathbb{R}$ for $f_{\{c>0\}}$? In that case, the iterations do not converge, but they jump around chaotically on the real line. To investigate this behaviour, we will consider the discrete dynamical system that arises from this problem. From now on, we will focus on the real line, so we will change the notation from z to x.

A *discrete dynamical system* is an iterative procedure of a continuous function $f : X \to X$ with initial values in X. We will denote this initial value by x_0 and the sequence of iterates will look like:

$$x_0, f(x_0), f(f(x_0)), \ldots, f^n(x), \quad \text{where } x_n = f^n(x_0)$$



(a) $c = 5$. (b) $c = -1$. (c) $c = 0$. (d) $c = 1$. (e) $c = 5$.

Figure 1: $N_c(x) = \frac{1}{2}(x - \frac{c}{x})$ for different values of c in red and diagonal $y = x$ in blue.

Such a sequence of iterations is called the *orbit* of x_0. The goal of dynamical systems is to understand what happens after a number of iterations. There are some preliminaries that we need to mention. We say x is a *periodic point* x of period n if $f^n(x) = x$. A periodic point of period 1 is called a *fixed point*. A periodic point x of period n is said to be *repelling* if $|\frac{\mathrm{d}}{\mathrm{d}x} f^n(x)| > 1$ and *attracting* if $|\frac{\mathrm{d}}{\mathrm{d}x} f^n(x)| < 1$. Note that if x is a periodic point of period n, then $f(x), f^2(x), \ldots, f^{n-1}(x)$ are also periodic points of period n. This set of points $\{x, f(x), \ldots, f^{n-1}(x)\}$ is called the *period n obit*. We often draw $f(x)$ together with a diagonal line $y = x$ such that the intersection points show where the fixed points are. Note that intersections between the $f^n(x)$ and the diagonal give the periodic points of period n. This diagonal is also a tool to visualize iterations. There is an example in figure 2, the vertical lines show $(x_i, x_i) \to (x_i, f(x_i)) = (x_i, x_{i+1})$ and the horizontal lines show $(x_i, x_{i+1}) \to (x_i, f(x_i)) = (x_{i+1}, x_{i+1})$.

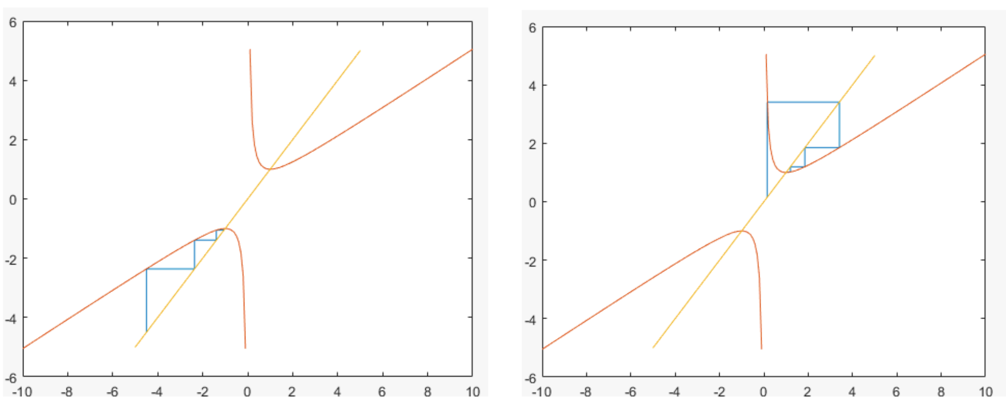
Newton method forms a dynamical system in the following way. We can write the iterations as the so-called *Newton function* [20]:

$$N(x) = x - \frac{f(x)}{f'(x)}, \quad \text{where } x_n = N^n(x_0).$$

In the case of $f_c(x) = x^2 + c$, the Newton function is given by:

$$\begin{aligned} N_c(x) &= x - \frac{x^2 + c}{2x} \\ &= \frac{2x^2}{2x} - \frac{x^2 + c}{2x} \\ &= \frac{1}{2}\left(x - \frac{c}{x}\right) \end{aligned}$$

Now we have our dynamical system, given by $N_c : \mathbb{R} \to \mathbb{R}$. Note that iterations of N_c converge to attracting fixed points of N_c. These fixed points are equal to the zeros of f_c. We have seen that the zeros of f_c are different, depending on c. As a direct result, the dynamics of N_c also differ for



(a) initial condition -4.5. (b) initial condition 0.15.

Figure 2: Iterations of $N_{-1}(x) = \frac{1}{2}(x + \frac{1}{x})$.

different values of c. Figure 1 shows N_c for different values of c together with the diagonal. The fixed points are found by solving:

$$x = \frac{1}{2}(x - \frac{c}{x})$$
$$\Longrightarrow x = \pm\sqrt{-c}$$

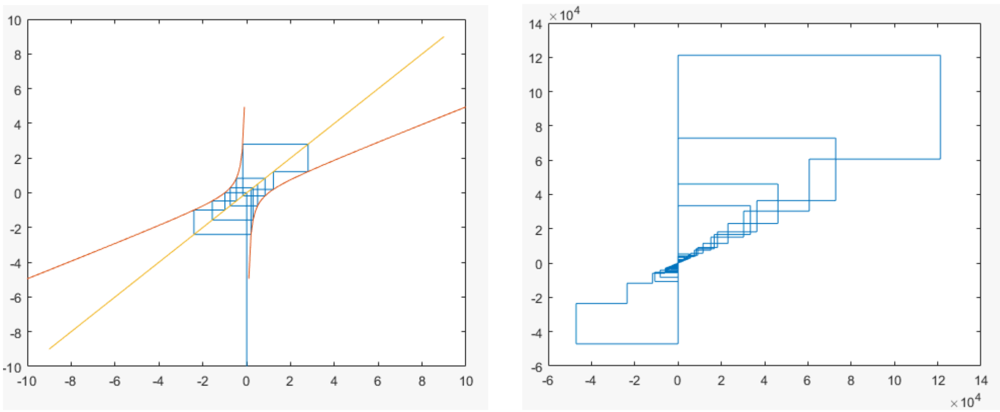
For $c \leq 0$, the dynamics are simple, there are two fixed points at $\pm\sqrt{-c}$. We can check that they are attracting using the derivative:

$$N_c'(x) = \frac{1}{2}(1 + \frac{c}{x^2}).$$

Indeed, $|N_c'(\pm\sqrt{-c})| = 0 < 1$. Figure 2 shows some iterations for $N_{-1}(x)$. The iterations converge to the fixed points $\pm\sqrt{-c}$, as expected. For $c = 0$, $N_c(x)$ reduces to $N_0(x) = \frac{x}{2}$, which has one attracting fixed point at $x = 0$. Indeed, $N_0'(x) = \frac{1}{2} < 1$. For $c > 0$, the fixed points of N_c are imaginary. Figure 1d and 1e show that there are indeed no real fixed points. In both cases, the iterations cannot converge to anything and the dynamics look very interesting. In figure 3b, we see some iterations going from very large numbers back to small values and back to large numbers again. In the rest of this thesis, we will focus on:

$$N_1(x) = \frac{1}{2}(x - \frac{1}{x}) = T(x).$$

We will investigate the dynamics of T. We will start of with some basic properties. We will see a topological conjugacy with the doubling map, and use this to find the periodic points and some other properties of T. Then we will consider different definitions of chaos and we will try to prove that T is chaotic, and finally we will look at T from an ergodic point of view.



(a) 27 iterations. (b) 219000 iterations.

Figure 3: Iterations of $N_1(x) = \frac{1}{2}(x - \frac{1}{x})$ with initial value $\frac{1}{2}$.

1 Properties of the map

We consider the following map, shown in figure 4:

$$T(x) = \frac{1}{2}\left(x - \frac{1}{x}\right). \tag{1}$$

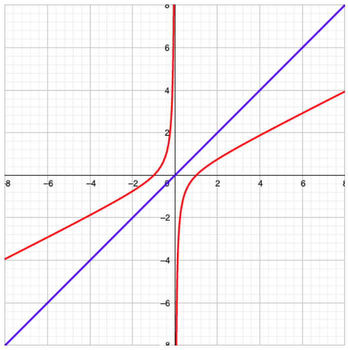


Figure 4: The map T with a diagonal.

Note that T is anti-symmetric, since $T(-x) = \frac{1}{2}(-x + \frac{1}{x}) = -T(x)$. There is an asymptote at $x = 0$, showing a discontinuity at $x = 0$. The map is continuous in the rest of $\mathbb{R}$.

Corollary 1. The map $T : \mathbb{R}\backslash\{0\} \to \mathbb{R}\backslash\{0\}$ is continuous.

Proof. A map f is continuous at a if $\lim_{x\to a} f(x) = f(a)$. Clearly, T is not continuous at $x = 0$, as $\lim_{x\to 0+} T(x) = -\infty$ and $\lim_{x\to 0^-} T(x) = \infty$ and $T(0)$ is undefined. For all other a $\lim_{x\to a} T(x) = \frac{1}{2}(a - \frac{1}{a}) = T(a)$ So T is continuous at $\mathbb{R}\backslash\{0\}$. $\square$

In fact, we need to exclude not only $\{0\}$ from the domain, but all points that eventually map to 0. When looking at orbits of T, we do not want to find any discontinuities. These points form a countable set:

$$K_T = \{x \in \mathbb{R} : T^n(x) = 0 \text{ for some } n \in \mathbb{N}\}.$$

Taking this into account, we find that $T^n : \mathbb{R}\backslash K_T \to \mathbb{R}\backslash K_T$ is also continuous and we have a dynamical system:

$$x_n = T^n(x), \qquad x_0 \in \mathbb{R}\backslash K_T, \tag{2}$$

In this section, we will investigate the dynamics of this system. We will start with the topological conjugacy between T and the well-known doubling map. We will then use it to find the periodic points of the dynamical system and to investigate the set K_T.

1.1 Topological Conjugacy

We want to find a way to investigate what happens to T after n iterations. If we look at the second iteration, we already find a long, complicated formula:

$$x_2 = T^2(x) = \frac{1}{2}\left(\frac{1}{2}\left(x - \frac{1}{x}\right) - \frac{1}{\frac{1}{2}\left(x - \frac{1}{x}\right)}\right).$$

Imagine what the n^{th} iteration would look like! Luckily, there is another way to express x_n [20]. Consider $h : (0, \pi) \to \mathbb{R}$, shown in figure 5, given by:

$$h(\theta) = \cot\theta \tag{3}$$

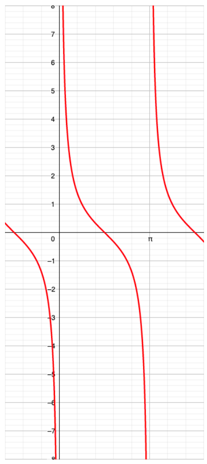


Figure 5: Cotangens $h(\theta) = \cot\theta$.

Note that the cotangens is periodic with period π, we will focus only on $(0, \pi)$. Furthermore, note that h maps to all of $\mathbb{R}$. So for all $x_0 \in \mathbb{R}$, there is $\theta \in (0, \pi)$ such that $x_0 = \cot\theta$. We can use the simple trigonometry double-angle formula: $\cot 2\theta = \frac{1}{2}(\cot\theta - \frac{1}{\cot\theta})$ to find a new expression for $T^n(x)$:

$$\begin{aligned} x_1 &= \frac{1}{2}\left(x_0 - \frac{1}{x_0}\right) \\ &= \frac{1}{2}\left(\cot\theta - \frac{1}{\cot\theta}\right) = \cot 2\theta \\ x_n &= \cot 2^n\theta \quad \forall n \in \mathbb{Z}_{\geq 0} \end{aligned}$$

If $x_i = h(\theta)$, then the angle θ has doubled in the next iteration, namely $x_{i+1} = h(2\theta)$. Since the next iteration can also be written as $x_{i+1} = T(x_i)$, we find a relation between T and the doubling map $g(\theta) = 2\theta \mod \pi$:

$$
\begin{aligned}
T(x_i) &= x_{i+1} \\
T(h(\theta)) &= h(2\theta) \\
&= h(g(\theta))
\end{aligned}
$$

Recall that we excluded the set of points K_T from the domain of T. As a result, the points $\theta \in (0, \pi)$ such that $x_0 = \cot\theta \in K_T$ can be excluded from $(0, \pi)$. We will call this set of points K_g. Note that $T^n(x) = \cot 2^n\theta = 0$ is equivalent to $2^n\theta = \operatorname{arccot} 0 = \frac{\pi}{2}$. The set K_g is given by:

$$
\begin{aligned}
K_g &= \{\theta \in (0, \pi) : x_n = \cot 2^n\theta = 0 \text{ for some } n \in \mathbb{N}\} \\
&= \{\theta \in (0, \pi) : g^n\theta = \frac{\pi}{2} \text{ for some } n \in \mathbb{N}\}.
\end{aligned}
$$

This relationship between T and the doubling map is called a topological conjugacy. It is shown in figure 6.

$$
\begin{tikzcd}
(0,\pi)\backslash K_g \arrow[r, "g"'] \arrow[d, "h"'] & (0,\pi)\backslash K_g \arrow[d, "h"] \\
\mathbb{R}\backslash K_T \arrow[r, "T"'] & \mathbb{R}\backslash K_T
\end{tikzcd}
$$

Figure 6: Conjugacy between the doubling map g to T.

Definition 1. Let $f : X \to X$ and $g : Y \to Y$ be dynamical systems. We say f and g are *topologically conjugate* if there exists a homeomorphism $h : X \to Y$ such that $h \circ f = g \circ h$.

Proposition 1. *The map $h : (0, \pi)\backslash K_g \to \mathbb{R}\backslash K_T$ given by $h(\theta) = \cot\theta$ is a homeomorphism, i.e. it is continuous, bijective and has a continuous inverse.*

Proof. We can easily see that h is bijective. For continuity consider the topological space $(\mathbb{R}, \mathcal{T}_d)$ where $\mathcal{T}_d$ is the usual topology on $\mathbb{R}$, namely the family of all d-open subsets of $\mathbb{R}$. We use the metric $d(x, y) = |x - y|$. Consider also the topological space $((0, \pi), \mathcal{T}_d)$. Then $\tilde{h} : (0, \pi) \to \mathbb{R}$ and $\tilde{h}^{-1} : \mathbb{R} \to (0, \pi)$ are continuous. See also appendix A.1.2. Then the subset $(0, \pi)\backslash K_g \subset (0, \pi)$ gives the subspace topology

$$\mathcal{T}_{(0,\pi)\backslash K_g} = \{(0, \pi)\backslash K_g \cap U : U \in \mathcal{T}_d\}.$$

The inclusion map $i : (0, \pi)\backslash K_g \to (0, \pi)$ is defined by $i(a) = a \quad \forall a \in (0, \pi)\backslash K_g$. The inclusion map is continuous[21]. The map $h : (0, \pi)\backslash K_g \to \mathbb{R}$ is a composition $h = \tilde{h} \circ i$ and since $\tilde{h}$ and i are continuous, h must also be continuous. The proof of $h^{-1}(x) = \operatorname{arccot} x$ being continuous in $(0, \pi)\backslash K_g$ is similar. So h is bijective, continuous and has a continuous inverse, which makes it a homeomorphism. $\square$

Proposition 2. *There is a conjugacy between the doubling map $g(\theta) = 2\theta \mod \pi$ and T.*

Proof. Following definition 1, we have two dynamical systems $(\mathbb{R}\backslash K_T, T)$ and $((0, \pi)\backslash K_g, g)$, and a homeomorphism $h(\theta) : (0, \pi) \to \mathbb{R}\backslash K_T$ given by $h(\theta) = \cot\theta$. We need to check the following quality:

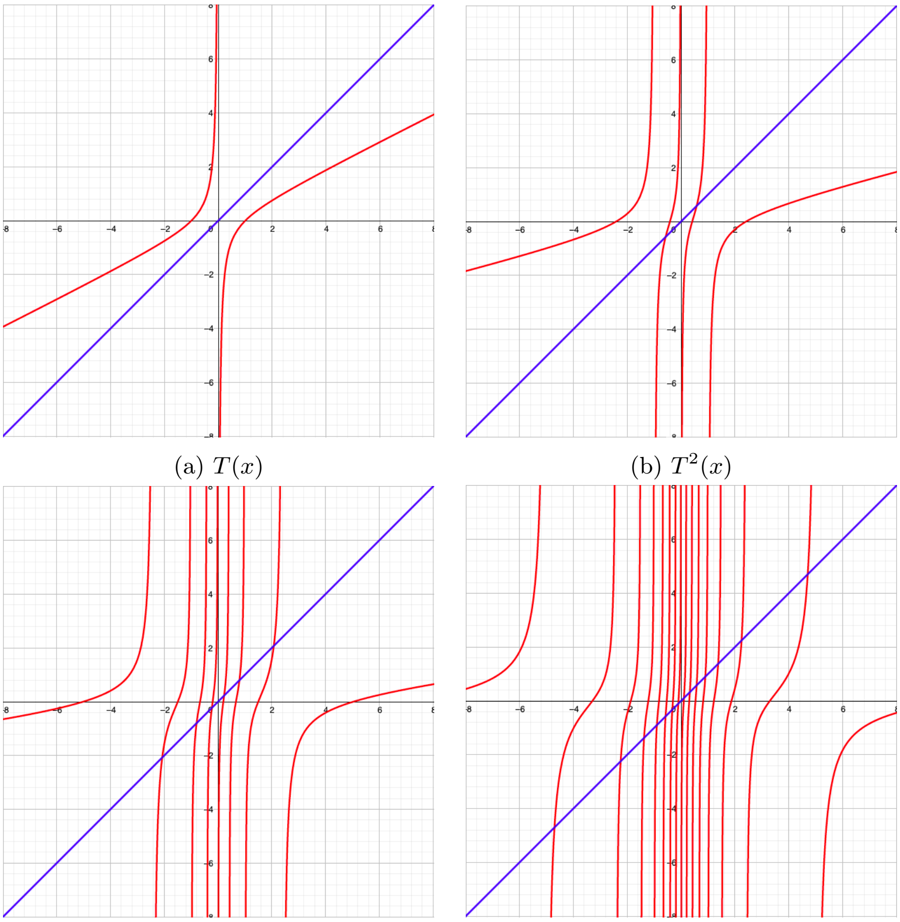
$$
\begin{aligned}
T \circ h(\theta) &= h \circ g(\theta) \\
T(\cot\theta) &= h(2\theta \mod \pi) \\
T(x_1) &= \cot 2\theta \\
x_2 &= x_2,
\end{aligned}
$$

so, indeed T and the doubling map g are topologically conjugate. $\square$

Note that we can write $f = h \circ g \circ h^{-1}$. Then $f^n = h \circ g \circ h^{-1} \circ \cdots \circ h \circ g \circ h^{-1} = h \circ g^n \circ h^{-1}$ holds for all n. As we will see, a topological conjugacy preserves many dynamical properties. In the next section, we will use the conjugacy to find the periodic points.

1.2 Periodic points

Recall that periodic points of period n can be found where T^n intersects the diagonal line $y = x$. In figure 7, we see such graphs showing the periodic points of period 1,2,3,4. Figure 7a shows that there are no fixed points,



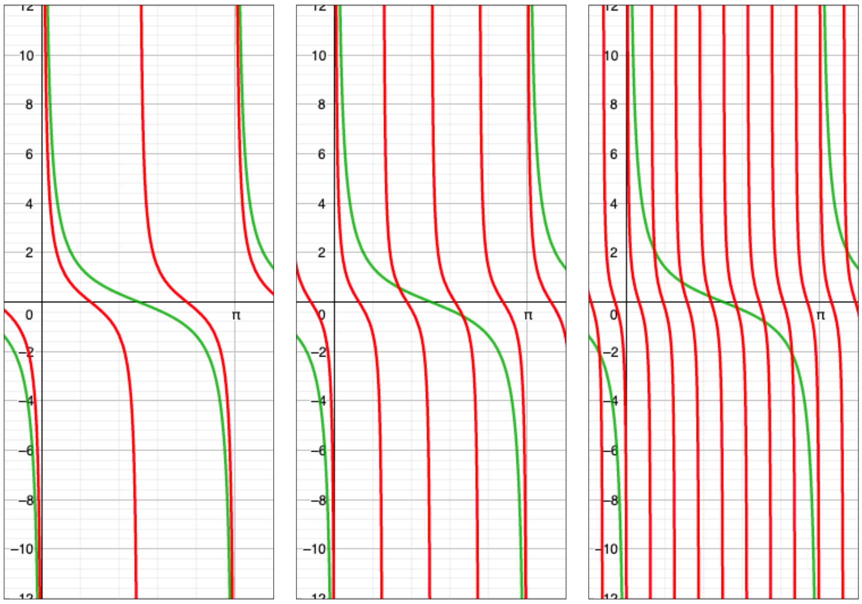
(a) $T(x)$ (b) $T^2(x)$ (c) $T^3(x)$ (d) $T^4(x)$

Figure 7: Graphs of iterations of $T(x)$ with a diagonal $y = x$.

Figure 7b shows that there are two periodic points of period 2. We can calculate them analytically:

$$\begin{aligned} T^2(x) = \frac{1}{2}\left(\frac{1}{2}\left(x - \frac{1}{x}\right) - \frac{1}{\frac{1}{2}\left(x - \frac{1}{x}\right)}\right) &= x \\ \frac{1}{2}\left(x - \frac{1}{x}\right)^2 - 2 &= 2x\left(x - \frac{1}{x}\right) \\ 3x^4 + 2x^2 - 1 &= 0 \\ &\implies x^2 = \frac{1}{3} \text{ or } x^2 = -1 \\ &\implies x = \pm\sqrt{\frac{1}{3}} \end{aligned}$$

The two periodic points at $\pm\sqrt{\frac{1}{3}}$ are both repelling since $\frac{\mathrm{d}}{\mathrm{d}x}T^2\left(\pm\sqrt{\frac{1}{3}}\right) = 4 > 1$. For periodic points of higher order, it becomes increasingly difficult to compute them by hand. Figure 7c shows



(a) $x_0 = \cot\theta$ with $x_1 = \cot 2\theta$

(b) $x_0 = \cot\theta$ with $x_2 = \cot 4\theta$

(c) $x_0 = \cot\theta$ with $x_3 = \cot 8\theta$

Figure 8: Graphs x_0 in green and x_1, x_2, x_3 in red.

6 periodic points of period 3, and figure 7d shows 14 periodic points of period 4. This trend suggests that T has $2^n - 2$ periodic points of period n. Furthermore, the graphs in figure 7 suggest that all periodic points are repelling. A periodic point x is repelling if $|\frac{\mathrm{d}}{\mathrm{d}x}T^n(x)| > 1$, i.e. the tangent line at x has a slope larger than 1. The tangent line of all periodic points is steeper than the diagonal with slope 1, indicating that the periodic points are repelling. With the topological conjugacy, we can prove these conjectures.

1.2.1 Using the conjugacy

We can use the conjugacy to find the periodic points. Suppose we have a periodic point of g such that $g^n(\theta) = \theta$. Then $h(\theta)$ is a periodic point of T:

$$
\begin{aligned}
T^n \circ h(\theta) &= h \circ g^n(\theta) \\
T^n(h(\theta)) &= h(\theta)
\end{aligned}
$$

So if θ is a periodic point of g then $h(\theta) = \cot(\theta)$ is a periodic point of T of the same period. For example, $\theta = \frac{\pi}{3}$ and $\theta = \frac{2\pi}{3}$ are the periodic points of period 2 of g. Indeed, $\cot\frac{\pi}{3} = \sqrt{\frac{1}{3}}$ and $\cot\frac{2\pi}{3} = -\sqrt{\frac{1}{3}}$ are exactly the periodic points of T.
Determining the periodic points of T was quite difficult. Fortunately, the periodic points of $g(\theta) = 2\theta \mod \pi$ are easier to find. Note that if $\frac{\theta}{\pi}$ is rational, then θ is (eventually) periodic [20]. If $\frac{\theta}{\pi}$ is irrational, we have a non-periodic orbit. Non-periodic orbits can be chaotic, this will be investigated later on.

Recall that x_n can be written as $x_n = \cot 2^n\theta$. A point is periodic if $x_0 = x_n$, so a way to find periodic points of period n is to look for intersections in the graphs of $x_0 = \cot\theta$ and $x_n = \cot 2^n\theta$. In figure 8 we see such graphs. In figure 8a, we see that x_0 and x_1 have no intersection points, as expected. Figure 8b shows that $x_0 = \cot\theta$ and $x_2 = \cot 4\theta$ intersect in two points. These are the two periodic points at $\frac{\pi}{3}, \frac{2\pi}{3}$. Figure 8c shows 6 periodic points of period 3 at $\frac{\pi}{7}, \frac{2\pi}{7}, \ldots, \frac{6\pi}{7}$. These correspond to the 6 periodic points in figure 7c. This method of comparing $\cot\theta$ with $\cot 2^n\theta$ can be used to find all the periodic points of T:

Proposition 3. *The map $T = \frac{1}{2}(x - \frac{1}{x})$ has $2^n - 2$ points of period n. For $n = 2, 3, \ldots$ they are* *given by*

$$\cot\left(\frac{\pi}{2^n-1}\right), \cot\left(\frac{2\pi}{2^n-1}\right), \cot\left(\frac{3\pi}{2^n-1}\right), \dots, \cot\left(\frac{(2^n-2)\pi}{2^n-1}\right).$$

Proof. A periodic point of period n is found when $x_0 = x_n$. Let $m := 2^n$, then:

$$\begin{aligned} x_0 &= x_n \\ h_0(\theta) &= h_n(\theta) \\ \cot\theta &= \cot m\theta \\ \cos\theta \sin m\theta &= \sin\theta \cos m\theta. \end{aligned}$$

By the angle sum and difference identities:

$$\begin{aligned} \sin(m\theta \pm \theta) &= \sin m\theta \cos\theta \pm \cos m\theta \sin\theta \\ \sin(m\theta + \theta) + \sin(m\theta - \theta) &= 2\sin m\theta \cos\theta \\ \sin(m\theta + \theta) - \sin(m\theta - \theta) &= 2\cos m\theta \sin\theta. \end{aligned}$$

Now we can solve:

$$\begin{aligned} \sin(m\theta + \theta) - \sin(m\theta - \theta) &= \sin(m\theta + \theta) + \sin(m\theta - \theta) \\ \sin{(m-1)\theta} &= 0 \\ \Rightarrow \theta &= \frac{\pi}{m-1}, \frac{2\pi}{m-1}, \dots, \frac{(m-2)\pi}{m-1} \end{aligned}$$

So indeed, for all $n = 2, 3, \dots$, there are $m - 2 = 2^n - 2$ values of θ such that $x_0 = x_n$. Finally, $x = \cot\theta$ gives the periodic points of T, finishing the proof. □

We can write the set of all periodic points of T as follows:

$$\mathrm{Per}(T) = \bigcup_{n=2}^{\infty} \bigcup_{k=1}^{2^n-2} \cot\frac{k\pi}{2^n-1}$$

The periodic points of g can easily be found with the conjugacy. This set can be written as:

$$\mathrm{Per}(g) = \bigcup_{n=2}^{\infty} \bigcup_{k=1}^{2^n-2} \frac{k\pi}{2^n-1}$$

Note that the periodic points of g are evenly distributed in $(0, \pi)$. We can prove that set of periodic points is dense. First let us recall the definition:

Definition 2. A subset $A \subset X$ is dense in X if for any point $x \in X$, any neighbourhood of x contains at least one point of A.

Proposition 4. *Per*(g) *is dense in* $(0, \pi)\backslash K_g$.

Proof. Consider a point $\theta \in (0, \pi)\backslash K_g$. Let x, y be two periodic points of period n such that they are of consecutive k. The distance between them is:

$$d_n(x, y) = \left|\frac{k\pi}{2^n-1} - \frac{(k+1)\pi}{2^n-1}\right| = \frac{\pi}{2^n-1}.$$

Note that $d_n(x, y) \to 0$ as $n \to \infty$. Let $B_\epsilon(\theta)$ be an open ball around θ of radius $\epsilon > 0$. For all $\epsilon > 0$ there exists a value of n such that $d_n(x, y) < \varepsilon$. Therefore, there exists a periodic point x of period n, such that $d_n(x, \theta) \leq d_n(x, y) < \epsilon$. Hence, there must be at least one periodic point inside the open ball, making $\mathrm{Per}(g)$ a dense subset of $(0, \pi)\backslash K_g$. □

The periodic points of T are not evenly distributed in its domain. As we saw in figure 7 the periodic points are mostly centered around 0. As n increases, the periodic points move away from 0, but the periodic points around 0 stay closer together than those further away. This is because of the nature of $h(\theta) = \cot(\theta)$ and the fact that $\mathbb{R}$ is not bounded. We can see in figure 5 that h maps values $\theta \in (\frac{\pi}{4}, \frac{3\pi}{4})$ to $(-1, 1)$, values of $\theta \in (0, \frac{\pi}{4})$ are mapped to $(1, \infty)$, and $\theta \in (\frac{3\pi}{4}, \pi)$ are mapped to $(-\infty, -1)$. So if we have a certain distance $d(\theta, \tilde{\theta})$ between two periodic points θ and $\tilde{\theta}$ of g, then $d(h(\theta), h(\tilde{\theta}))$ gets larger as θ is closer to 0 or π.

Theorem 1. All periodic points of g and T are repelling.

Proof. Consider the doubling map g. A periodic point of period n is repelling if $|\frac{\mathrm{d}}{\mathrm{d}\theta}g^n(\theta)| > 1$. This is the case since $|2^n| > 1$ for $n \geq 2$. The topological conjugacy preserves this quality of the periodic points, as we can see below:

$$\frac{\mathrm{d}}{\mathrm{d}\theta}T^n \circ h(\theta) = \frac{\mathrm{d}}{\mathrm{d}\theta}h \circ g^n(\theta)$$
$$\frac{\mathrm{d}}{\mathrm{d}h(\theta)}T^n(h) \cdot \frac{\mathrm{d}h(\theta)}{\mathrm{d}\theta} = \frac{\mathrm{d}}{\mathrm{d}g^n(\theta)}h(g^n(\theta)) \cdot \frac{\mathrm{d}g^n(\theta)}{\mathrm{d}\theta}(\theta)$$

If we consider this at a periodic point $\theta_0 = \theta_n = g^n(\theta_0)$, we find:

$$\frac{\mathrm{d}}{\mathrm{d}h(\theta_0)}T^n(h(\theta_0)) \cdot h'(\theta_0) = h'(\theta_0) \cdot 2^n$$
$$|\frac{\mathrm{d}}{\mathrm{d}x}T^n(x)| = 2^n \text{ for } x = h(\theta_0).$$

Indeed, periodic points of g and T are periodic. □

1.3 Set of points that map to zero K_T

Recall that the domain of T is $\mathbb{R}\backslash K_T$ where $K_T = \{x \in \mathbb{R} : T^n(x) = 0 \text{ for some } n\}$. In this section we will investigate this set K_T. For any $x \in K_T$ we have $T^n(x) = 0$ for some n. Such a point x is mapped to $\pm\infty$ by the next iteration, and then again to $\pm\infty$ for all future iterations:

$$T(0) = \lim_{x\to 0^\pm} \frac{1}{2}(x - \frac{1}{x}) = \mp\infty,$$
$$T(\pm\infty) = \lim_{x\to\pm\infty} \frac{1}{2}(x - \frac{1}{x}) = \lim_{x\to\pm\infty} \frac{x}{2} = \pm\infty.$$

Let us take a look at the zeros of T^n for $n = 1, 2, 3, 4$ in figure 7. The points where $T^n(x) = 0$, are of course found on the intersections with the $x-$axis. Note that if $T^i(x) = 0$, then there is an asymptote at x for iteration T^n for all $n > i$.

The first point in K_T is $x = 0$. For $n = 0$, we find $T^0(x) = x$, which is of course 0 at $x = 0$. Figure 7a shows that $T(x)$ has an asymptote at $x = 0$ and two new zeros at ± 1. These new zeros can be computed with the inverse of T (see also appendix A.1):

$$T_\pm^{-1}x = x \pm \sqrt{x^2 + 1}$$

Note that this is a double inverse, the pre-image of any points consists of two points. To find the points of K_T, we need to look at the pre-image of 0. Indeed, we find the two zeros $T^{-1}(0) = \pm 1$. We can do this again to find the four zeros of T^2:

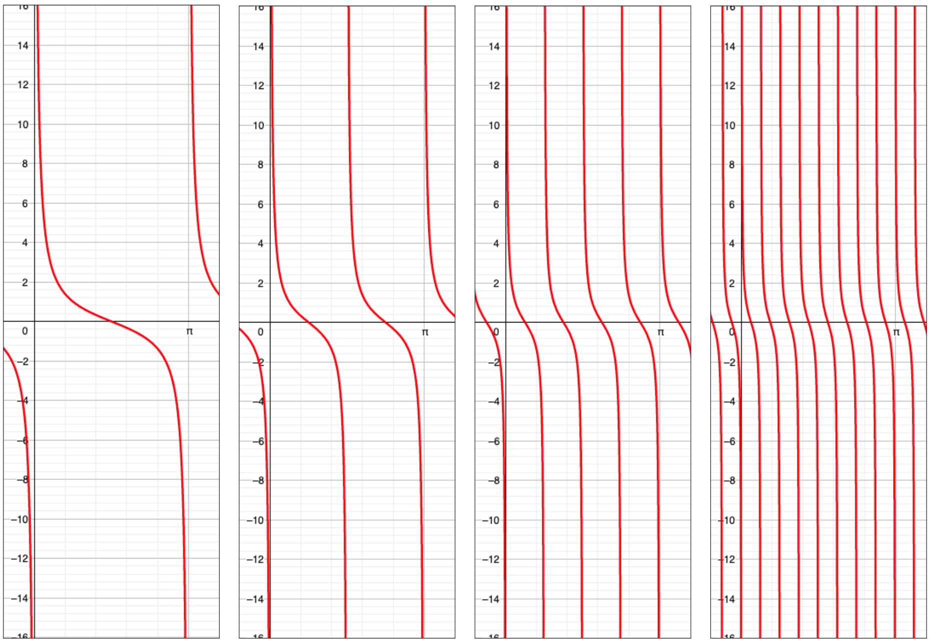
$$T_\pm^{-1}(\pm 1) = \pm 1 \pm \sqrt{2}.$$

Indeed, figure 7b shows four zeros and 3 asymptotes at $-1, 0, 1$. Using the same method, we can calculate that the zeros of T^3 are $\pm(1 + \sqrt{2}) \pm \sqrt{4 + 2\sqrt{2}}$ and $\pm(1 - \sqrt{2}) \pm \sqrt{4 - 2\sqrt{2}}$. Indeed, figure 7c shows 8 zeros and 7 asymptotes. Lastly, figure 7d shows the 16 zeros and 15 asymptotes of T^4. In theory, we can find all the zeros using the inverse, but it would take too much time. We can however conclude how much zeros and asymptotes there are.

Proposition 5. *The set K_T is countable and consists of $\cup_{n=0}^\infty 2^n$ points. Each graph of T^n has $2^n - 1$ asymptotes.*

Proof. T^n has 2^n zeros for any n, given by $T^{-n}(0)$. The zeros turn into asymptotes in the next iteration, so the number of asymptotes is just the sum of all previous zeros: $\sum_{i=0}^{n-1} 2^i = 1 + 2^1 + 2^2 + 2^3 + \cdots + 2^{n-1}$. We can prove by induction that this sum equals $2^n - 1$ For $n = 1$ both are equal to 1. Next, assume this holds for $n = k$: $\sum_{i=0}^{k-1} 2^i = 2^k - 1$ and prove it for $n = k + 1$:

$$\sum_{i=0}^{k} 2^i = 2^k + \sum_{i=0}^{k-1} 2^i$$
$$= 2^k + 2^k - 1 = 2 \cdot 2^k - 1 = 2^{k+1} - 1.$$



(a) $\cot\theta$ (b) $\cot 2\theta$ (c) $\cot 4\theta$ (d) $\cot 8\theta$

Figure 9: Graphs of $\cot 2^n\theta$ for $n = 0, 1, 2, 3$

So indeed, K_T consists of a countable number of points, a collection of 2^n zeros in each iteration of T^n and each graph of T^n has $2^n - 1$ asymptotes. $\square$

Figure 7 also shows that most zeros are centered around 0, but as n increases, they slowly move further away from 0. We will consider the outer asymptotes and see how far they can move away from 0. Since T is anti-symmetric we can focus on only the most right asymptote, which we find by iterating the inverse with the + sign: $T_+^{-1}x = x + \sqrt{x^2+1}$. Let a_n be the most right asymptote for some n, and consider the sequence $\{a_n\}$ defined by $a_{n+1} = T^{-1}(a_n)$. Consider the ratio:

$$\begin{aligned}\left|\frac{a_{n+1}}{a_n}\right| &= \left|\frac{a_n + \sqrt{1+a_n^2}}{a_n}\right| \\ &= \left|1 + \sqrt{\frac{1}{a_n^2} + 1}\right| > 1 \quad \forall n.\end{aligned}$$

According to the ratio test, this means that this sequence is divergent. This shows that as $n \to \infty$, the outer asymptotes will move to $\pm\infty$. The set of asymptotes has no bound and neither does the set of periodic points, since all periodic points lie between the outer asymptotes. This is something we already knew, since we computed all of the periodic points in the previous section.

1.3.1 K_T using the conjugacy.

It turned out to be very time-consuming to calculate all points in K_T, but again, we can use the conjugacy to find them. Recall that:

$$\begin{aligned}K_g &= \{\theta \in (0,\pi) : x_n = \cot 2^n\theta = 0 \text{ for some } n \in \mathbb{N}\} \\ &= \{\theta \in (0,\pi) : g^n\theta = \frac{\pi}{2} \text{ for some } n \in \mathbb{N}\}.\end{aligned}$$

Figure 9 shows the graphs of $\cot 2^n\theta$ for $n = 0, 1, 2, 3$. Figure 9a shows no asymptotes in $(0,\pi)\backslash K_g$, but one zero at $\frac{\pi}{2}$. Note that this corresponds to the zero we found for $n = 0$ as $x = \cot\frac{\pi}{2} = 0$. Figure 9b shows one asymptote at $\frac{\pi}{2}$ and two zeros at $\frac{\pi}{4}, \frac{3\pi}{4}$. Indeed, $\cot\frac{\pi}{4} = 1$ and $\cot\frac{3\pi}{4} = -1$ Figure 9c shows 3 asymptotes and 4 zeros and figure 9d shows 7 asymptotes and 8 zeros. This corresponds exactly to what we found for T.

More generally, we can find all points for which $g^n(\theta) = \frac{\pi}{2}$ for some n. This is equivalent to saying $h(g^n(\theta)) = \cot 2^n\theta = h(\frac{\pi}{2}) = 0$:

$$\begin{aligned} 0 &= \cot 2^n\theta \\ 0 &= \frac{\cos 2^n\theta}{\sin 2^n\theta} \\ 0 &= \cos 2^n\theta \\ \Rightarrow 2^n\theta &= \frac{\pi}{2} + k\pi \text{ for } k \in \mathbb{N} \\ \theta &= \frac{\pi}{2^{n+1}} + \frac{k\pi}{2^n} = \frac{(1+2k)\pi}{2^{n+1}}. \end{aligned}$$

Now we know that:

$$K_g = \bigcup_{n=0}^{\infty} \bigcup_{k=0}^{2^n-1} \frac{(1+2k)\pi}{2^{n+1}},$$

and:

$$K_T = \bigcup_{n=0}^{\infty} \bigcup_{k=0}^{2^n-1} \cot \frac{(1+2k)\pi}{2^{n+1}}.$$

1.4 Relation between T and the Logistic map

There is a link between T and the logistic map $F_\mu(x) = \mu x(1-x)$ for $\mu = 4$. Recall that T comes from the Newton's method when trying to find the zeros for x^2+1 on the real line. The iterations $y_n = x_n^2 + 1$ show how Newton's method moves over the vertical axis [20]. We substitute $x_{n+1} = \frac{1}{2}(x_n - \frac{1}{x_n})$ to find that:

$$\begin{aligned} y_{n+1} &= x_{n+1}^2 + 1 \\ &= \frac{1}{4}\frac{y_n^2}{y_n - 1} \end{aligned}$$

And then a second substitution $z = \frac{1}{y}$ gives us the quadratic equation:

$$\begin{aligned} y_{n+1} &= \frac{1}{4}\frac{y_n^2}{y_{n-1}} \\ z_{n+1} &= 4z_n(1-z_n) \end{aligned}$$

The details of this computation and more information about the link between T and F_4 can be found in Appendix A.1.3.

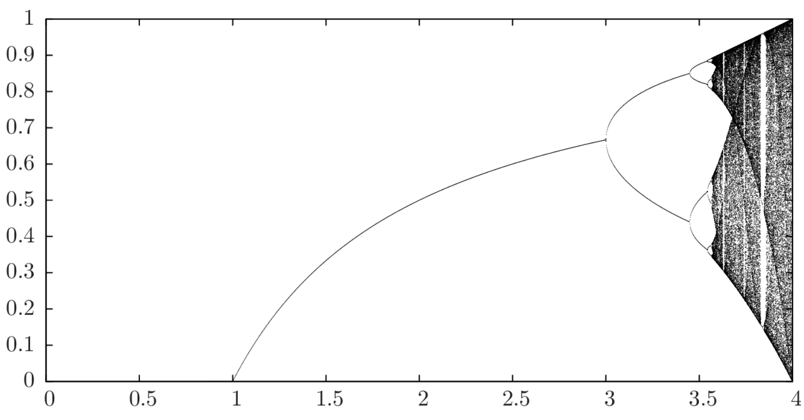


Figure 10: Feigenbaum diagram

Note that:

$$z_n = \frac{1}{x_n^2 + 1} \tag{4}$$

is not a homeomorphism, so we do not have a topological conjugacy between the logistic map and T, but the chaotic properties of the logistic map are extensively studied and might give us an idea of the chaotic dynamics of T.

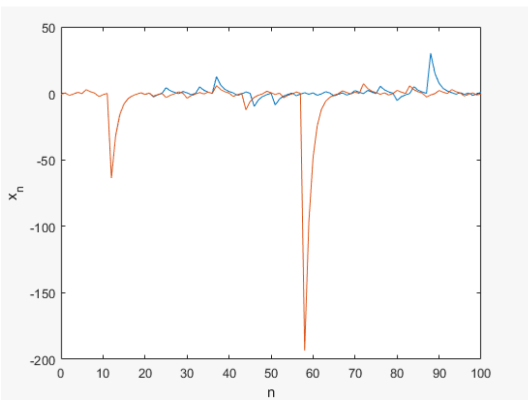
The logistic map has different dynamics for different values of μ. In figure 10, we find the Feigenbaum diagram, which is the bifurcation diagram for the logistic map. On the horizontal axis we find values for μ and on the vertical axis values of $x \in [0, 1]$. The bifurcation diagram shows all attracting periodic points for all values of $0 < \mu < 4$. We can see that for $0 < \mu < 1$, there is only one attracting periodic point at $x = 0$. At $\mu = 1$, there is a bifurcation and for $1 < \mu < 3$ there are two periodic points. The one at $x = 0$ is now repelling and the other one is attracting. At $\mu = 3$, there is another bifurcation point and for $\mu > 3$, the situation becomes more complicated and eventually become chaotic around $\mu \approx 3.6$. For $\mu \geq 4$, the logistic map is also chaotic, but it will have a different kind of chaos.

In the next section we will investigate different definitions of chaos and try to prove that T and the logistic map are chaotic.

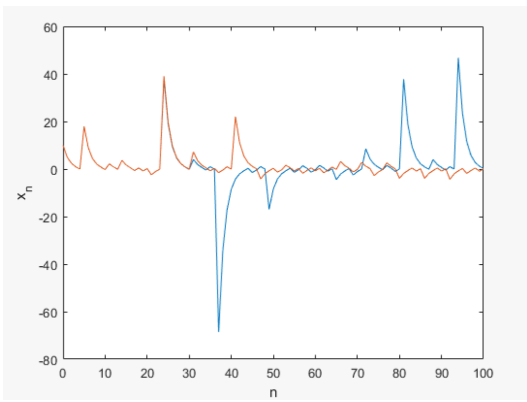
2 Chaos

In this section we will try to answer the question of whether $T(x) = \frac{1}{2}(x - \frac{1}{x})$ is chaotic. As mentioned before, there are different definitions for chaos. We will start at the beginning. Edward Lorenz was one of the pioneers of chaos theory. In a book written by him, "The Essence of Chaos", he describes chaos as something looking random but not being random. In fact, in a deterministic dynamical system there is no randomness [13]. Every next iteration of an evolution is precisely defined by the dynamical system. When Lorenz worked on weather predictions, he once noticed that a slight difference in initial conditions led to a completely different outcome of his calculations. It is interesting how a deterministic system can have behave this way. This phenomenon is known as 'sensitive dependence on initial conditions' and it is an important characteristic of chaos. Even though there exist different definitions of chaos, they all agree on this.

For the map T, we can plot iterations $T^n x$ vs n to see what happens if we slightly change the initial condition. This is shown in figure 11a for initial conditions $x_0 = 0.5$ and $x_0 = 0.50000001$. The iterations are similar for the first 20 iterations, but after that they start to differ and around the 58th iteration, there is a spike in the orange line but not in the blue line. If we look at the same perturbation for initial condition $x_0 = 20$, we see that it takes longer for the iterations to differ. The iterations are similar for the first 30 iterations and again we see that their behaviours be-



(a) I.c. 0.5 (orange) and i.c. 0.50000001 (blue).



(b) I.c. 20 (orange) and i.c. 20.00000001 (blue).

Figure 11: Iterations for different initial conditions (i.c.)

come completely different. We will investigate different definitions of chaos and we will investigate whether these definitions are satisfied by the map T.

2.1 Li and Yorke Chaos

In 1975, Li and Yorke published the well-known paper "Period Three Implies Chaos". They investigated a non-periodic orbit of a continuous map f on an interval. They proved that if f has a periodic point of period 3, then there exist periodic points of all periods. Furthermore, there will exist an uncountable set which contains no asymptotically periodic points. If this is the case, we say f has Li-Yorke chaos:

Theorem 2. Let J be an interval and let $f : J \to J$ be continuous. Assume there is a point $a \in J$ for which the points $b = f(a)$, $c = f^2(a)$ and $d = f^3(a)$, satisfy

$$d \leq a < b < c \text{ (or } d \geq a > b > c\text{)}.$$

Then

1. for every $k = 1, 2, \ldots$ there is a periodic point in J having period k.
2. there is an uncountable set $S \subset J$ (containing no periodic points), which satisfies the following conditions:

   a) For every $p, q \in S$ with $p \neq q$,

   $$\limsup_{n\to\infty} |f^n(p) - f^n(q)| > 0$$

   and

   $$\liminf_{n\to\infty} |f^n(p) - f^n(q)| = 0$$

   b) For every $q \in S$ and periodic point $p \in J$,

   $$\limsup_{n\to\infty} |f^n(p) - f^n(q)| > 0.$$

Proof. Li and Yorke [12]. □

Note that the existence of a periodic point of period 3 satisfies the condition for this theorem. The first consequence of having periodic points of all periods is actually a special case of Sarkovskii's theorem. We will not prove it here, the proof can be found in [7]. Below is Sarkovskii's theorem written, note that if we take $k = 3$, then $3 \triangleright l$ for all l and indeed, there are periodic points of all periods l.

Theorem 3. Consider the Sarkovskii ordering of natural numbers, first list all the odd numbers except 1, then 2 times all odd numbers except 1 etc, then left are all powers of 2, which are listed at the end of the ordering in decreasing order:

$$3 \triangleright 5 \triangleright 7 \triangleright \cdots \triangleright 2\cdot 3 \triangleright 2\cdot 5 \triangleright \cdots \triangleright 2^2\cdot 3 \triangleright 2^2\cdot 5 \triangleright \ldots \cdots \triangleright 2^3 \triangleright 2^2 \triangleright 2 \triangleright 1$$

Suppose $f : \mathbb{R} \to \mathbb{R}$ is continuous. Suppose f has a periodic point of period k. If $k \triangleright l$ in the Sarkovskii ordering, then f also has a periodic point of period l.

The second consequence of Li and Yorke's theorem is the existence of an uncountable set $S \subset J$ of points that are not asymptotically periodic. This set S is also known as the *scrambled set.* Li and Yorke define an asymptotically periodic point in the following way:

Definition 3. A point $x \in J$ is asymptotically periodic if there is a periodic point p for which

$$|f^n(x) - f^n(p)| \to 0 \text{ as } n \to \infty.$$

This is equivalent to the better known definition of asymptotically periodicity:

Definition 4. We say that $x \in J$ is asymptotically periodic if there exists a periodic point $p \in J$, of period k, such that

$$d(f^n(x), \{p, f(p), \ldots, f^{k-1}(p)\}) \to 0 \quad \text{as} \quad n \to \infty. \tag{5}$$

Recall that the distance between a point $x \in \mathbb{R}$ and a set $A \subset \mathbb{R}$ is defined as

$$d(x, A) = \inf\{|x - a| \,:\, a \in A\}.$$

In particular, we have that

$$d(f^n(x), \{p, f(p), \ldots, f^{k-1}(p)\}) = \min\{|f^n(x) - f^i(p)| \,:\, i = 0, \ldots, k-1\}.$$

Note that taking a minimum instead of an infimum is justified since we are dealing with a finite set. The proof is found in appendix A.2.1.

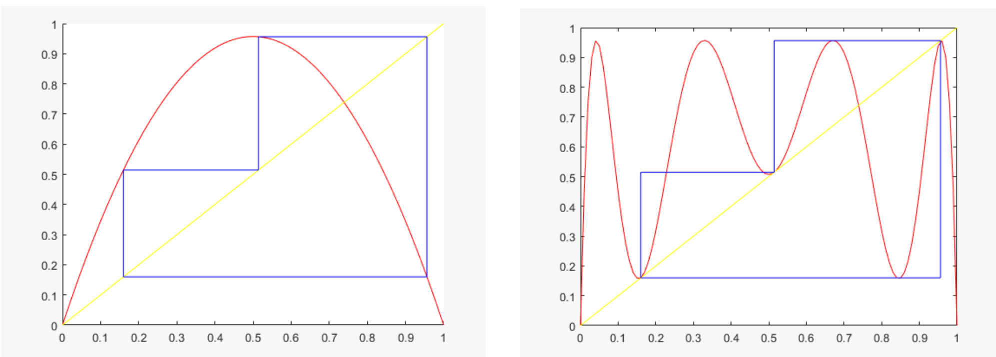
Property 2b of Li and Yorke's theorem states that no point q in S is asymptotically periodic. The distance between $f^n(p)$ and $f^n(q)$ cannot converge to 0, for any periodic point p, since the supremum of this distance is larger than 0. Property 2a states that for any two points p, q in S, their orbits will never converge. However, at some point in the orbit, $f^n(p)$ and $f^n(q)$ might be close together, since the infimum of this same distance is zero.

We know that T has a period 3 orbit, namely $\cot(\frac{\pi}{7}) \to \cot(\frac{2\pi}{7}) \to \cot(\frac{4\pi}{7}) \to \cot(\frac{8\pi}{7}) \equiv \cot(\frac{\pi}{7})$. This would imply Li-Yorke chaos, but unfortunately, T is only defined on $\mathbb{R} \backslash K_T$, which is not an interval, and if we would consider T on all of $\mathbb{R}$, it is not continuous. Therefore, we cannot use this theorem to prove chaos.

An example of a map that we can apply this theorem to is the logistic map $F_\mu(x) = \mu x(1-x)$ on $[0, 1]$, which was also investigated in Li and Yorke's paper. The requirements of the theorem are met, as we have a continuous map on an interval. For certain values of μ, the logistic map has Li-Yorke chaos. When writing their paper, Li and Yorke did not know exactly for which μ there are points that are not asymptotically periodic. They show that at least up till $\mu = 1 + \sqrt{6} \approx 3.449$, each point in $[0, 1]$ is asymptotically periodic [12]. Later, it was found that the onset of chaos is at $\mu_c \approx 3.56995...$ [9]. Recall that in the Feigenbaum diagram in figure 10, the dynamics of F_μ seem to become very chaotic around μ_c.

The first orbit of period 3 occurs at $\mu = 1 + 2\sqrt{2} \approx 3.8284...$ [8]. Figure 12 shows this orbit of period 3, starting at $x_0 = 0.159929$. Figure 12a shows the orbit together with F_μ, it shows clearly that $x_3 = x_0$. Next to that, in figure 12b we see the orbit together with F_μ^3, such that we can see that the period 3 orbit intersects with the diagonal exactly where F_μ^3 also intersects the diagonal.

According to Li and Yorke's theorem, we need a periodic orbit of period 3 in order to have Li-Yorke chaos, so how can we have chaos before the first period 3 orbit? There are values before $\mu = 1 + 2\sqrt{2}$ for which there is an orbit of period 6, for example at $\mu \approx 3.627$. In this case the theorem can be applied to F_μ^2. If there exist points for F_μ^2 that are not asymptotically periodic, then the same is true for F_μ. The same reasoning with $F_\mu^3, F_\mu^4, \ldots$, leads to the onset of chaos at μ_c.



(a) $F_\mu(x)$. (b) $F_\mu^3(x)$

Figure 12: Three orbit of F_μ, with $\mu = 1 + 2\sqrt{2}$.

For $\mu_c < \mu \leq 4$, the logistic map has Li-Yorke chaos. Note that for $\mu > 4$, certain orbits start to leave $[0, 1]$, so we cannot talk about $F_\mu : J \to J$ anymore. For $\mu_c < \mu \leq 4$, there are periodic points of all periods and drawing all these periodic points gives the very dark section in the Feigenbaum diagram. Furthermore, there exists an uncountable scrambled set of points that are not asymptotically periodic. It is difficult to pinpoint this scrambled set. Of course the (eventually) periodic points are not in the scrambled set. The union of the n^{th} pre-images of all periodic points form the set of eventually periodic points. Excluding this countable set from the domain gives an uncountable set of points that could very well be the scrambled set.

2.2 Marotto

In 1978, Marotto expanded on Li and Yorke's theorem. He defined a so-called snap-back repeller for multi-dimensional maps and proves that if a dynamical system contains such a snap-back repeller, then it is Li-Yorke chaotic.
Below is the exact definition of a snap-back repeller, [15]:

Definition 5. Suppose $\tilde{x}$ is a fixed point of differentiable map $f : \mathbb{R}^n \to \mathbb{R}^n$ with all eigenvalues of $Df(\tilde{x})$ exceeding 1 in magnitude, and suppose there exists a point $x_0 \neq \tilde{x}$ in a repelling neighbourhood of $\tilde{x}$, such that $x_M = \tilde{x}$ and $\det(Df(x_k)) \neq 0$ for $1 \leq k \leq M$, where $x_k = f^k(x_0)$. Then $\tilde{x}$ is called a *snap-back repeller* of f.

This is a fascinating result as you would not expect points repelled away from a fixed point to return to it. This result makes it possible to prove chaos for a differentiable map using a repelling fixed point. Note that f cannot be invertible. If there exists an inverse map f^{-1} such that $f^{-1}(f(x)) = x$, then for a fixed point, we would also find that $f^{-1}(x) = x$, and the only point in the pre-image of repelling fixed point x is x itself. If f is not injective, the pre-image consists of multiple values, and we might be able to find a snap-back repeller. In that case, the following propostition states:

Proposition 6. *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be differentiable. If f possesses a snap-back-repeller then f is chaotic in the sense of Li and Yorke.*

Proof. [14] □

Note, that although Marotto expands on Li and Yorke's theorem, there is a slight difference in the function requirements. Li and Yorke require a continuous function on an interval, while Marotto requires a differentiable function on $\mathbb{R}^n$. It is not a problem to consider a subset of $\mathbb{R}^n$, but it is unclear if this subset needs to be an interval. In the following subsections, we will find snap-back repellers for F_4 and T.

2.2.1 Snap-back repeller of the logistic map

In this example, we will show that there exists a snap-back repeller of the logsitic map for $\mu = 4$. $F_4(x) = 4x(1 - x)$ has fixed points $x = 0$ and $x = \frac{3}{4}$, and we will focus on $x = \frac{3}{4}$. Note that the derivative $|F_4'(\frac{3}{4})| = 2 > 1$ shows that this fixed point is repelling. The inverse of F_4 is:

$$F_{4,\pm}^{-1}(x) = \frac{1}{2} \pm \frac{1}{2}\sqrt{1 - x}.$$

Note that F is not injective, as this is a double inverse, $F_{4,+}^{-1}(\frac{3}{4}) = \frac{3}{4}$ and $F_{4,-}^{-1}(\frac{3}{4}) = \frac{1}{4}$. With the inverse function we can find a sequence of points $\{x_{-1}, x_{-2}, \dots\}$. This sequence is called the *backward orbit of* x_0. Note that there are many possible backward orbits $\{x_{-n}\}$, by using $F_{4,-}^{-1}$ and $F_{4,+}^{-1}$ in different orders. Note that $T^i(x_{-n}) = x_0$ for all $i > n$. To find a backward orbit of the fixed point $\frac{3}{4}$, we first have to use $F_{4,-}^{-1}$ to find its first value x_{-1}. Then we use the inverse to find more points in the backward orbit until for some M, x_{-M} is close to $\frac{3}{4}$. In this case, the iterations of X_{-M} first repell away from $\frac{3}{4}$, but eventually, after M iterations, the orbit reaches the fixed point $x = \frac{3}{4}$, and so, the repelling fixed point $\frac{3}{4}$ is a snap-back repeller.
Figure 13 shows such an orbit. The first dotted line is at the fixed point $\frac{3}{4}$ and the second dotted line is at the value at 0.761.... Note that the iterations are first clearly repelled away from $\frac{3}{4}$ and then are snapped back to $\frac{3}{4}$ exactly.

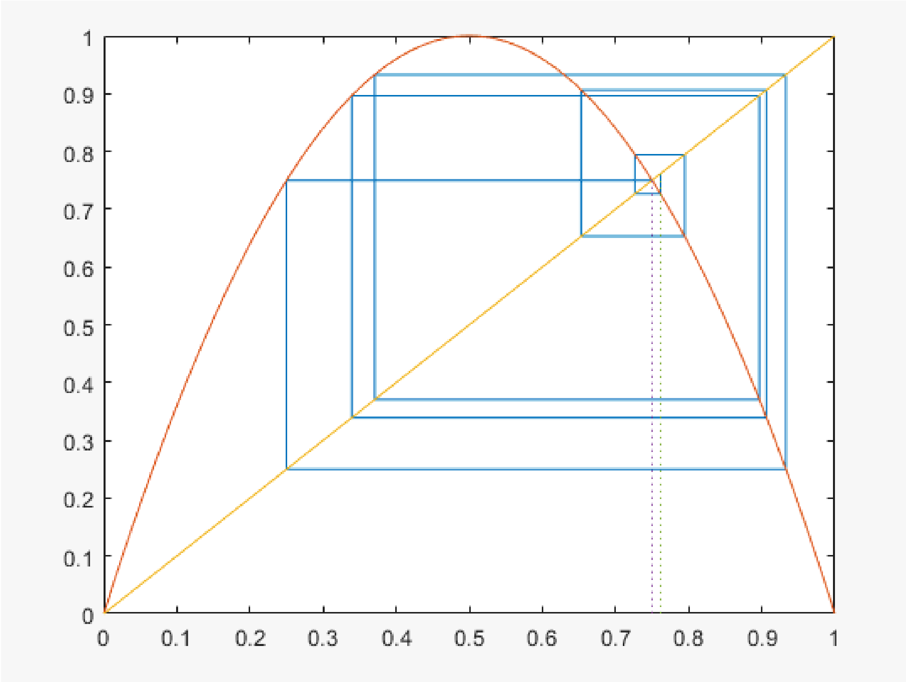


Figure 13: Snap-back repeller for F_4.

There exist snap-back repellers for more values of μ. Recall from the previous section that the logsitic map F_μ has Li-Yorke chaos for $\mu > \mu_c \approx 3.5699$. Marotto shows that F_μ is chaotic by finding snap-back repellers for $\mu > 3.680$. He also applies this method to F_μ^2 and finds chaos for $\mu > 3.595$. It is likely that if we continue this process for $F_\mu^4, F_\mu^8, \dots$ we will find the same chaos onset at μ_c.

2.2.2 Snap-back repeller of T

Recall that T has no fixed points, so we will apply this theorem to T^2. The fixed points of T^2 are found at $\pm\frac{1}{\sqrt{3}}$. We already know that all periodic points of T are repelling, and we know the inverse:

$$T^{-1}x = x \pm \sqrt{x^2+1}$$

Figure 14 shows that $\frac{1}{\sqrt{3}}$ is a snap-back repeller. The first dotted line is at the fixed point $\frac{1}{\sqrt{3}} \approx 0.577$. The second dotted line is at 0.594... and we see again that the iterations are first repelled away from the fixed point and then snapped back to the fixed point.
We need to keep in mind that T is not defined on all of $\mathbb{R}$, which is something that Marotto assumes. So although we have found a snap-back repeller, we cannot say for sure that T has Li-Yorke chaos, but we can secretly suspect it.

2.3 Devaney Chaos

In 1989, Robert L. Devaney also gave a definition for chaos. He stated that a chaotic map has the following three properties [7]:

Definition 6. Let X be a metric space. A continuous map $f : X \to X$ is said to be *chaotic* on X if

1. f has sensitive dependence on initial conditions
2. f is topologically transitive
3. periodic points of f are dense in X

Devaney says that a chaotic system has three ingredients: unpredictability, undecomposability and an element of regularity. Sensitive dependence on initial conditions gives a sense of unpredictability.

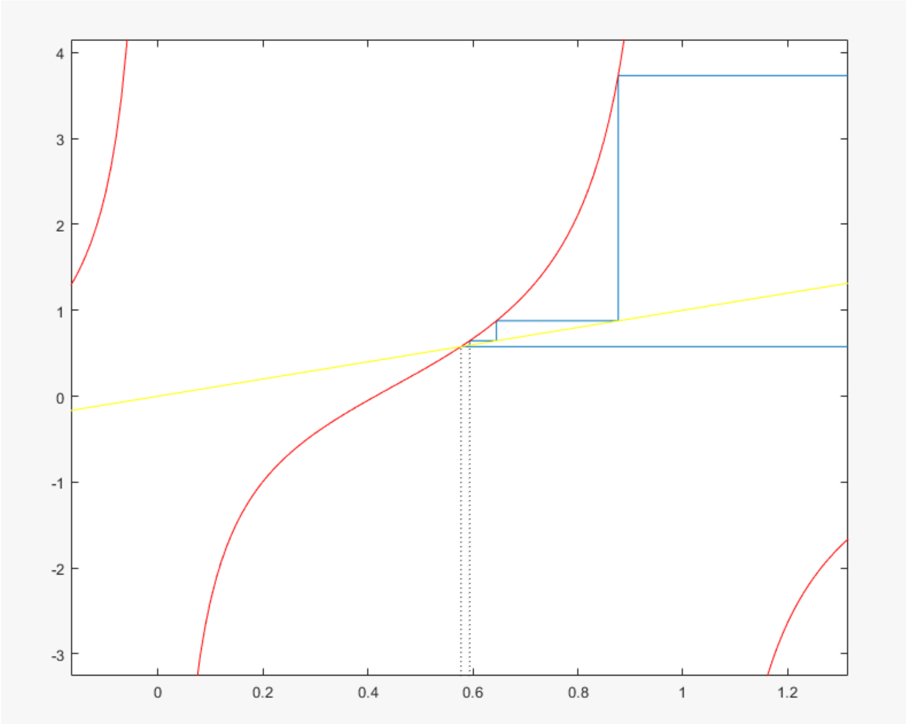


Figure 14: Snap back repeller at $\frac{1}{\sqrt{3}}$ for T^2x

Topologically transitivity makes f undecomposable, which means that it cannot be broken down into two invariant open subsets which do not interact under f. Lastly, the dense periodic points give an element of regularity.

Definition 7. A map $f : X \to X$ has *sensitive dependence on initial conditions* if there exists $\delta > 0$ such that, for any $x \in X$ and any neighborhood $B(x)$ of x, there exists $y \in B(x)$ and $n \geq 0$ such that $|f^n(x) - f^n(y)| > \delta$.

Definition 8. A map $f : X \to X$ is said to be *topologically transitive* if for any pair of open sets $U, V \subset X$ there exist $k > 0$ such that $f^k(U) \cap V \neq 0$.

It has been proven in 1992, by Banks, that properties 2 and 3 imply property 1 [1]. Although this means we do not have to prove sensitive dependence on initial conditions, it is such an important characteristic of chaos that we will explore it. A map has sensitive dependence on initial conditions if for every neighbourhood $B(x)$ of every x, there exists $y \in B(x)$ such that for some future iteration, the distance between the two orbits is at least δ. For the doubling map, this can be proven as follows:
Take a point $x \in (0, \pi) \backslash K_g$ and y in an open neighbourhood $B_\varepsilon(x)$ of x, such that $d(x, y) < \epsilon$. The doubling map doubles this distance in each iteration, so there exists n such that $d(f^n(x), f^n(y)) = |f^n(x) - f^n(y)| = 2^n|x - y| > \delta$.

For T, it is a bit harder to prove. This is because x and y can either diverge or converge, depending on their location in the domain. We can write the distance $d(T(x), T(y))$ as follows:

$$\begin{aligned}
|T(x) - T(y)| &= |\frac{1}{2}(x - \frac{1}{x}) - \frac{1}{2}(y - \frac{1}{y})| \\
&= \frac{1}{2}|(x - y) + (\frac{1}{y} - \frac{1}{x})| \\
&= \frac{1}{2}|(x - y) + \frac{x - y}{xy}| \\
&= \frac{1}{2}|x - y| \cdot |1 + \frac{1}{xy}|.
\end{aligned}$$

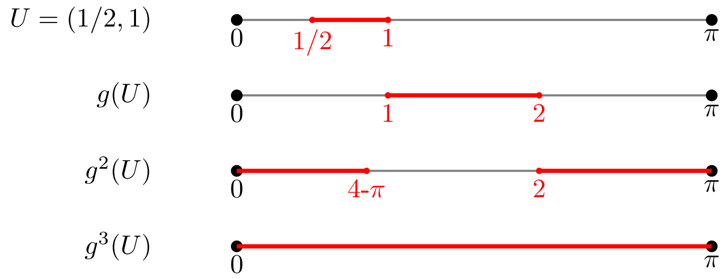


Figure 15: Topological transitivity doubling map.

We need to distinguish between $|x| < 1$ and $|x| > 1$. If $|x| < 1$, there exist $|y| < 1$ in any neighbourhood $B(x)$ such that $|T(x) - T(y)| > |x - y|$, which satisfies the definition. If $|x| > 1$, we will find that $|T(x) - T(y)| < |x - y|$, the distance between the orbits decreases. For any neighbourhood $B_\varepsilon(x)$, there exists a point $y \in B(x)$ such that $|x - y| < \varepsilon$. As long as $T^i(x) > 1$, the next iteration will bring the orbits of x and y closer together, so we still have $|x-y| < \varepsilon$. When $|x| < 1$ at some point, the orbits will diverge in the next iteration, such that $|T^n(x) - T^n(y)| < \delta$ for some $\delta > 0$. The distance between the orbits after n iterations can be expressed by induction:

$$\begin{aligned}
|T^2(x) - T^2(y)| &= \frac{1}{2}\,|T(x) - T(y)| \cdot \left|1 + \frac{1}{T(x)T(y)}\right| \\
&= \frac{1}{2}\left(\frac{1}{2}|x-y| \cdot |1 + \frac{1}{xy}|\right) \cdot \left|1 + \frac{1}{T(x)T(y)}\right| \\
&\vdots \\
|T^n(x) - T^n(y)| &= |x - y| \cdot \prod_{i=0}^{n-1} \frac{1}{2}\left|1 + \frac{1}{T^i(x)T^i(y)}\right|. \qquad (6)
\end{aligned}$$

The distance $|T^n(x) - T^n(y)|$ is dependent on $T^i(x)$ and $T^i(y)$ for all previous $i < n$. From figure 11 we see that the orbit has $|T^i(x)| < 1$ most of the time. Therefore, we expect this product to be larger than 1. It is difficult to make this proof rigorous, but hopefully the reader is convinced of the sensitive dependence on initial conditions. We will see now that T satisfies the other two properties of definition 6 and thereby also proving that T has sensitive dependence on initial conditions.
To prove that T is topologically transitive, we will use the conjugacy with the doubling map. This property is conserved by a conjugacy:

Lemma 1. Assume that the maps $f : X \to X$ and $g : Y \to Y$ are topologically conjugate. If f is topologically transitive, then so is g.

Proof. Maps f and g are topologically conjugate. This means that there exists a homeomorphism $h : X \to Y$ such that $h \circ f = g \circ h$. Assume that $U, V \subset Y$ are open sets. Since h is continuous, the sets $U' = h^{-1}(U)$ and $V' = h^{-1}(V)$ are open in X. Since f is topologically transitive, there exists $n \in \mathbb{N}$ such that $f^n(U') \cap V' \neq \emptyset$. This implies that

$$\emptyset \neq h(f^n(U') \cap V') \subset h(f^n(U')) \cap h(V') = g^n(h(U')) \cap h(V') = g^n(U) \cap V,$$

which shows that g is also topologically transitive. $\square$

To prove that the doubling map g is topologically transitive, take an open set $U \in (0, \pi)\backslash K_g$. Figure 15 shows an example of what happens to U after a number of iterations. Suppose $U = (\frac{1}{2}, 1)$, then $g(U) = (1, 2)$ and $g^2(U) = (0, 4 \mod \pi) \cup (2, \pi)$ and lastly, $g^3(U) = (0, 8 \mod \pi) \cup (4 \mod \pi, \pi)$ which covers the whole interval $(0, \pi)$. If U is smaller it will take more iterations, but there will always, for any open set U, exist k such that $g^k(U)$ covers all of $(0, \pi)\backslash K_g$. In that case, $g^n(U)$ must have a non-empty intersection with any other open set $V \in (0, \pi)\backslash K_g$ and hence g is topologically transitive. By lemma 1, T is also topologically transitive.

The last property of Devaney's definition is dense periodic points. Proposition 4 already proved that the periodic points of the doubling map are dense. This property is also conserved by the topological conjugacy. Recall that a subset $A \subset X$ is called dense in X if for any $x \in X$ and any open neighborhood U of x we have that $A \cap U \neq \emptyset$. The following result states that maps which are both surjective and continuous preserve dense sets. In particular, homeomorphisms preserve dense sets.

Lemma 2. Let X and Y be topological spaces and assume that the map $f : X \to Y$ is both surjective and continuous. If A is dense in X, then $f(A)$ is dense in Y.

Proof. Let $y \in Y$ be arbitrary and let $V \subset Y$ be any open set containing y. We need to show that $f(A) \cap V \neq \emptyset$. Since f is surjective, there exists $x \in X$ such that $f(x) = y$. Since f is continuous, the set $U = f^{-1}(V)$ is open in X. In addition, we have that $x \in U$. Since A is dense in X, we have that $A \cap U \neq \emptyset$, which implies that

$$\emptyset \neq f(A \cap U) \subset f(A) \cap f(U) \subset f(A) \cap V.$$

This completes the proof □

Now we can summarize all these results in the following two propositions:

Proposition 7. *If continuous maps f and g are topologically conjugate, and f is chaotic as defined by Devaney, then so is g.*

Proof. If f is chaotic, then it has sensitive dependence on initial conditions, it is topologically transitive and it has dense periodic points. By the conjugacy, g is topologically transitive and it has dense periodic points. These two properties imply sensitive dependence on initial conditions, so g also satisfies all three properties, and is therefore chaotic as defined by Devaney. □

Proposition 8. *The doubling map $g(\theta) = 2\theta \mod \pi$ and the map $Tx = \frac{1}{2}(x - \frac{1}{x})$ are both chaotic as defined by Devaney.*

Proof. The doubling map satisfies all three properties of definition 6. Because of the topological conjugacy between g and T, as described in proposition 1.1 and proposition 7, we know that both g and T are chaotic as defined by Devaney. □

Note that this is a different kind of chaos than Li-Yorke chaos. Devaney chaos and Li-Yorke chaos are both well known and there are many papers on the relation between the two. We will not go into details, more can be read in for example [18] and [10].

2.4 Dispersion Exponent

Yet another definition of chaos is given by Broer and Takens [4]. They define a dynamical system as chaotic if it has a positive *dispersion exponent*. To explain the motivation of a dispersion exponent, first recall that it is generally agreed upon that sensitive dependence of initial conditions is characteristic of a chaotic system. Sensitive dependence on initial conditions means that we cannot accurately predict the future based on the approximate initial condition. Broer and Takens consider the so-called principle *l'histoire se répète*, which works as follows: Take a segment of an orbit $\{x_1, x_2, \ldots, x_n\}$. To predict future values of this orbit, consider $m \in [0, \frac{1}{2}n]$, such that the distance $d(x_n, x_m)$ is minimal. Then according to this principle, iterations in the future, denoted by $\hat{x}$, can be approximated like this:

$$\hat{x}_{n+s} = x_{m+s}.$$

The idea is that, because x_n and x_m are so close together, we can approximate x_{n+s} by x_{m+s}. Of course, in a chaotic dynamical system, this way of predicting the future will not work. From sensitive dependence on initial conditions, we know that such a small difference does not have to stay a small difference after a number of iterations. The principle *l'histoire se répète* leads to the following definition of the dispersion exponent:

Definition 9. Let (M, f) be a dynamical system. Assume M is a complete metric space with metric d and that f is continuous. Consider an orbit $\{x_n\}$, that is positively compact and not (eventually) periodic.

1. For $s > 0$ and for $\varepsilon > 0$ we define
$$E(s,\varepsilon) = \sup_{0<n<m;\, d(x_n,x_m)<\varepsilon} \frac{d(x_{n+s}, x_{m+s})}{d(x_n, x_m)}$$
2. For $s > 0$ we define
$$E(s) = \lim_{\varepsilon\to 0} E(s,\varepsilon)$$
3. Finally, we define the *dispersion exponent*
$$E = \lim_{s\to\infty} \frac{1}{s} \log E(s).$$

So to determine the dispersion exponent, we consider a single orbit of f and we try to apply the principle *l'histoire se répète*. We find points x_n and x_m on the orbit such that $d(x_n, x_m)$ is small. Then we see what the distance between the points is after s iterations. The more they have diverged, the larger $E(s,\varepsilon)$ will be. Then, by letting ε go to zero and s to infinity, we find the dispersion exponent.

Note that we have to consider an orbit that is not (eventually) periodic. If we do try to calculate $E(s,\varepsilon)$ for a periodic orbit, we have to divide by 0 at some point, as $d(x_n, x_m) = 0$. For a chaotic orbit, we will have that x_{n+s} and x_{m+s} will disperse as s increases. In that case $E(x,\varepsilon) > 1$ and we will find a strictly positive disperion exponent $E > 0$. The more they disperse, the larger the dispersion exponent will be. We can state that a dynamical system is more chaotic for larger values of E.

Definition 10. Let (M, f) be a dynamical system such that M is a complete metric space and f is continuous. An evolution x is called chaotic if the dispersion exponent E is positive.

For some maps, we can calculate the dispersion exponent by hand. In other cases, we can find it numerically with the matlab functions in the appendix A.2.2. First, the matlab function `compute_sup.m` saves a segment of an orbit of length N with initial value x_0. Then it calculates $E(s,\varepsilon)$ for a value of s and ε. The matlab function `plot_Ese.m` calls on `compute_sup.m` for different values of s ε and returns a graph of $\log E(s)$ as a function of s for different values of ε. Since we cannot let s really go to infinity with a computer simulation, we will have to make an approximation. The slope is given by $\frac{\Delta \ln E(s)}{\Delta s}$. Note that $E(0) = 1$, because for $s = 0$, $E(s,\varepsilon)$ the numerator and denominator will be equal. Calculating the slope between $s = 0$ and a sufficiently large s, will then approximate the limit as $s \to \infty$:
$$\begin{aligned} E &= \lim_{s\to\infty} \frac{1}{s} \ln E(s) \\ &= \frac{\ln E(s)}{s}\Big|_{s=\infty} \\ &\approx \frac{\ln E(s) - \ln E(0)}{s - 0} \\ &= \frac{\Delta \ln E(s)}{\Delta s} \end{aligned}$$
In the following sections, we will try to find the dispersion exponent for the doubling map, the logistic map and for T.

2.4.1 The Doubling map

For the doubling map, we can calculate the dispersion exponent explicitly. The distance between two points θ_n and θ_m on an orbit is given by:
$$d(\theta_n, \theta_m) = 2^m - 2^n \mod \pi.$$
Suppose we have $\varepsilon < \frac{1}{2^s}$ and two points θ_n and θ_m on an orbit such that $d(x_n.x_m) < \varepsilon$. Then:
$$\begin{aligned} E(s,\varepsilon) &= \sup_{0<n<m;\, d(\theta_n,\theta_m)<\varepsilon} \frac{d(\theta_n, \theta_m)}{d(\theta_{n+s}, \theta_{m+s})} \\ &= \frac{2^m - 2^n}{2^{m+s} - 2^{n+s}} = 2^s. \end{aligned}$$

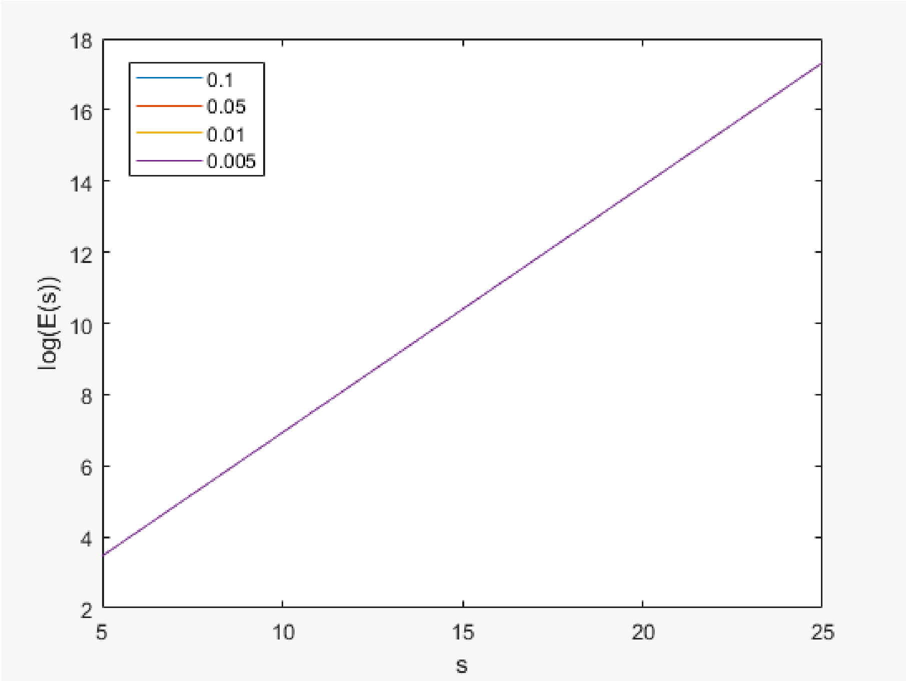


Figure 16: For the doubling map: $\log E(s)$ for different values of ε.

This is not dependent on ε, so $E(s) = 2^s \mod \pi$. The limit $s \to \infty$ gives:

$$E = \lim_{s\to\infty} \frac{1}{s} \log 2^s = \log 2$$

We can also calculate the dispersion exponent numerically. The resulting graph is found in figure 16. Note that we see only one line, even though we calculated it for 4 values of ε. This is because the dispersion exponent is not dependent on ε, as also seen in the calculation. The slope of this line is 0.6931..., which corresponds to the dispersion exponent $E = \log 2$.

2.4.2 The Logistic map

The dispersion exponent for the logistic map can also be calculated numerically. In figure 17, we see that for $1 \leq s \leq 10$, we have a nice straight line with slope 1.3858... $\approx \log 4$. For $s > 10$, the graph is not a nice straight line anymore. This is because at some point the distance $d(x_{n+s}, x_{m+s})$ has reached its maximum, it cannot be larger than 1. As s increases, this distance can decrease and increase a bit, but it cannot follow the same slope. This shows in figure 17, as the line goes up and down. Note that matlab has a limit for how small we can make ε. If we take ε smaller, it will take longer for $d(x_{n+s}, x_{m+s})$ to reach its maximum.

We can explain this value for the dispersion exponent. The derivative of the logistic map is $F_4'(x) = 4 - 8x$ and $|F_4'(x)| \leq 4 \, \forall x \in [0,1]$. By the mean value theorem, for any closed interval $[a,b] \subseteq [0,1]$, there exists $c \in [a,b]$ such that:

$$\begin{aligned}
f'(c) &= \frac{f(b)-f(a)}{b-a} \\
\left|\frac{f(b)-f(a)}{b-a}\right| &= |f'(c)| \leq 4 \\
|f(b)-f(a)| &\leq 4|b-a| \\
|f^s(b)-f^s(a)| &\leq 4|f^{s-1}(b)-f^{s-1}(a)| \leq \cdots \leq 4^s|b-a|.
\end{aligned}$$

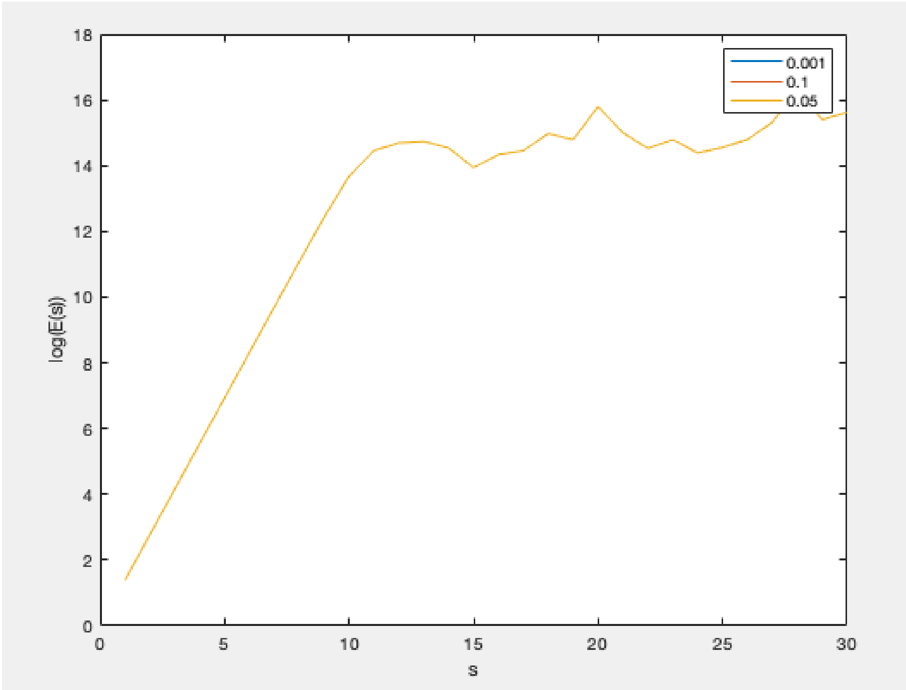


Figure 17: For the Logistic map: $\log E(s)$ for different values of ε.

Now consider one orbit of f and let $|f^n(x) - f^m(x)| < \varepsilon$ and let $b = f^n(x)$ and $a = f^m(x)$ such that $[a, b]$ (or $[b, a]$) is a closed interval. Then we can find an upper bound for E:

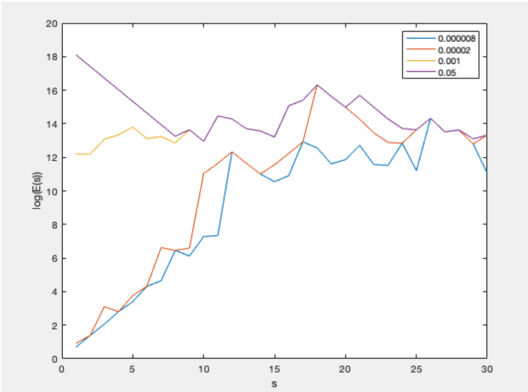
$$|f^{n+s}(b) - f^{m+s}(a)| \leq 4^s |f^n(b) - f^m(a)|$$
$$E(s, \varepsilon) \leq 4^s$$
$$E \leq \lim_{s\to\infty} \frac{1}{s} \log 4^s = \log 4.$$

Now all we have to show is that a dispersion exponent smaller than $\log 4$ cannot be the case. Consider x_n and x_m close to 0. The distance between them will expand by almost a factor 4. For example $x_n = 0.001$ and $x_m = 0.0011$ will give $d(x_n, x_m) = 0.0001$ and in the next iteration $d(x_{n+1}, x_{m+1}) = 0.000399$. As these points move closer to 0, this expansion factor will converge to 4. Recall that the logistic map is chaotic in Devaney's sense. The logistic map has a dense orbit [7]. This means that there must exist a point x_n on this dense orbit such that $d(x_n, 0) < \varepsilon$. If $\varepsilon \to 0$, this expansion factor will converge to 4 and thus, $E < \log 4$ is not possible, and $\log 4$ really is the dispersion exponent.

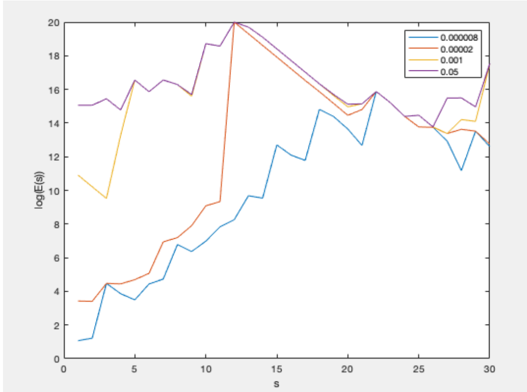
2.4.3 Map T.

Now it is time to turn our attention to our map $T(x) = \frac{1}{2}(x - \frac{1}{x})$. We were able to determine the dispersion exponent of the doubling map and logistic map. Unfortunately, the dispersion exponent is not generally conserved by topological conjugacies. For example, there exists a topological conjugacy between the doubling map and F_4 [7], while they have a different dispersion exponent. We cannot use the conjugacy with the doubling map to determine the dispersion exponent of T.

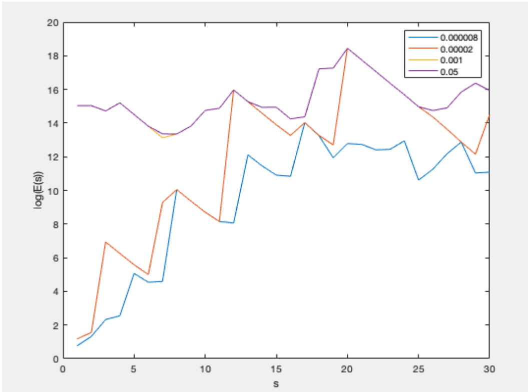
We can use equation 6 to express the distance between x_{n+s} and x_{m+s} and find an expression for $E(s, \varepsilon)$:



(a) Initial value 0.000004343.



(b) Initial value 0.23.



(c) Initial value 0.5442.

Figure 18: Graph of calculating dispersion exponent for T with $N = 1500$ for different initial conditions and values for ε.

$$
\begin{aligned}
d(x_{n+s}, x_{m+s}) &= |T^s(x_n) - T^s(x_m)| \\
&= |x_n - x_m| \prod_{i=0}^{s-1} \frac{1}{2} |1 + \frac{1}{T^i(x_n)T^i(x_m)}| \\
E(s, \varepsilon) = & \sup_{0<n<m;\, d(x_n, x_m)<\varepsilon} \prod_{i=0}^{s-1} \frac{1}{2} |1 + \frac{1}{T^i(x_n)T^i(x_m)}|.
\end{aligned}
$$

Note that for this map, $E(s, \varepsilon)$ does depend on ε. Changing ε has an effect on which x_n and x_m are chosen on the orbit which again has an effect on $E(s, \varepsilon)$. Because of this, it seems we cannot determine the limit $\lim_{\varepsilon \to 0} E(s, \varepsilon)$. Note that we have seen this product before for sensitive dependence on initial conditions in section 2.3. In that case, we expected this product to larger than 1. In that case, the logarithm will be larger than 0, and the dispersion exponent will be positive.

Let us see if this expectation is confirmed by the numerical results. Figure 18 shows some numerical results for different initial conditions and different values of ε. As expected, ε has a great effect on the results. Furthermore, there is not a neat straight line to calculate the slope from. This can be explained by the nature of the map. If two points are close together, the distance between them blows up by the $\frac{1}{x}$ term, but then it decreases by the $\frac{x}{2}$ term until it is small enough to be blown up again. Indeed, as s increases, $E(s)$ goes up and down. Even though we cannot determine the exact dispersion exponent, it is clear that there is a positive slope. In fact, the slope of graphs for $\varepsilon = 0.0002$ and $\varepsilon = 0.000008$ seem to all have approximately the same slope.

To conclude, we cannot determine the dispersion exponent for T exactly. Note that we ignored the requirement of a complete metric space, as it is not really needed in the calculations. However, a complete domain for T might have solved the problem. The real problem is that the domain is not compact. As a result, $T^n(x)$ can become arbitrarily large in the domain. Recall that for the doubling map, any iteration is at most twice as large as the one before. For the logistic map, each iteration is at most four times as large as the one before. The derivative of T has no upper bound, which shows that iterations can get arbitrarily large, which makes calculating the limit as $s \to \infty$ difficult. The good news is that we were able to write an expression for $E(s, \varepsilon)$. It was not helpful in calculating the dispersion exponent for T, but led to the expectation that the dispersion exponent must be positive. This expectation was confirmed by the numerical results. This positive dispersion exponent shows that T is chaotic.

3 Ergodic theory

So far, we have looked at dynamical systems from a topological viewpoint. Now we are going to investigate it using *ergodic theory*. Ergodic theory concerns the statistical properties of deterministic dynamical systems, in particular, the behavior of time averages of various functions along trajectories of dynamical systems. We will study (invariant) measures, the Birkhoff's ergodic theorem and Lyapunov exponents. We will do so for T and the logistic map for $\mu = 4$. From now on we will drop the subscript and just denote it F.

3.1 Invariant measure

A measure is a function that measures the size of a given (sub)set. For a set X, it is not always possible to measure the size of every subset in a consistent and meaningful way. Instead, we consider a collection of subsets. These subsets are called *measurable* sets and together form the $\sigma-$*algebra* $\mathcal{A}$. It satisfies the following three conditions:

1. $\emptyset, X \in \mathcal{A}$,
2. if $A \in \mathcal{A}$, then $X \backslash A \in \mathcal{A}$,
3. if $A_1, A_2, \ldots, A_n \in \mathcal{A}$, then $\cup_{n=1}^{\infty} A_n \in \mathcal{A}$.

A *measure* is a function $\mu : \mathcal{A} \to \mathbb{R}_{\geq 0}$ that "measures" the size of a measurable set by assigning a real, nonnegative number to it. A measure satisfies the following conditions:

1. $\mu(A) \in [0, \infty) \cup \{\infty\}$
2. $\mu(\emptyset) = 0$
3. if $A_1, A_2, \ldots$ are pairwise disjoint measurable subsets, then $\mu(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$

An example of a measure is the *Lebesgue measure*. It is the standard way of measuring subsets of $\mathbb{R}^n$. For $n = 1, 2, 3$, the Lebesgue measure is equal to the length, area and volume of $\mathbb{R}^n$.

We say that a map $S : X_1 \to X_2$ is *measurable* if $S^{-1}(E) \subseteq X_1$ is measurable for every measurable $E \subset X_2$. It is possible to have different measures for X_1 and X_2. Let μ_i be a measure on the space X_i. We say that S is *measure preserving* if $\mu_1(S^{-1}(E)) = \mu_2(E)$. If $X_1 = X_2$ and $\mu_1 = \mu_2$, we say that S is a *transformation*. Lastly, if a measurable transformation $S : X \to X$ preserves μ, then μ is S*-invariant*.

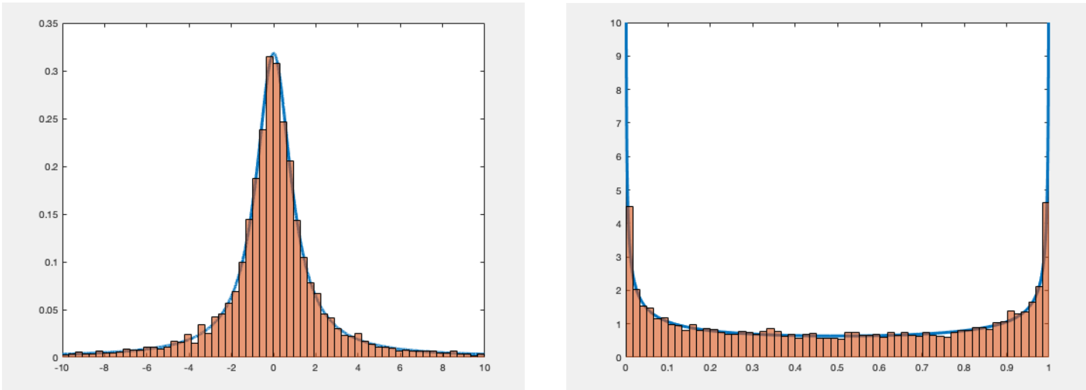
The map T and the logistic map are measurable transformations. There exist invariant measures μ_T and μ_F such that for all measurable sets E in $\mathbb{R} \backslash K_T$ and $[0, 1]$ respectively, these measures satisfy:

$$\mu_T(T^{-1}(E)) = \mu_T(E), \qquad \mu_F(F^{-1}(E)) = \mu_F(E).$$

To find these invariant measures we need the following definition of a *probability density function*:

Definition 11. If a measurable function $\rho(x) \geq 0$ satisfies $\mu(E) = \int_E \rho(x)dx$ for any measurable subset $E \subseteq X$, then ρ is called a *density function*. If μ is a probability measure (i.e. $\mu(X) = 1$), ρ is a *probability density function (pdf)*.

A probability density function $\rho(x)$ says something about the probability that an orbit of a map will be at value x. If $\rho(x)$ is large for some value of x, we would expect an orbit to visit x more often than some point y with a smaller $\rho(y)$. Let us consider an orbit of length N and let us divide the domain in a number of equal parts ('n_{bins}'). Then we will count how many times the



(a) Tx with $n_{\text{bins}} = 20001$ (b) F with $n_{\text{bins}} = 60$.

Figure 19: Histogram and density functions in the same figure for $N = 10000$ and $x_0 = 0.38$.

orbit visits each 'bin' and make a histogram. As $N \to \infty$, and $n_{\text{bins}} \to \infty$, the histogram will take on the shape of the density function. Such histograms are found in figure 19. For T, the orbit has most points around 0. For the logistic map, the orbit has most points close to 0 and 1. The figures also show the density functions, and indeed, we see that the histograms have the same shape.

The pdf and invariant measure of T are given by [6]:

$$\begin{aligned}
\rho_T(x) &= \frac{1}{\pi(x^2+1)} \\
\mu_T([a,b]) &= \int_a^b \frac{1}{\pi(x^2+1)} dx \\
&= \frac{1}{\pi}(\arctan b - \arctan a)
\end{aligned}$$

The pdf and invariant measure for F are given by:

$$\begin{aligned}
\rho_F(x) &= \frac{1}{\pi\sqrt{x(1-x)}} \\
\mu_F([a,b]) &= \int_a^b \frac{1}{\pi\sqrt{x(1-x)}} dx \\
&= -\frac{2}{\pi}(\arcsin\sqrt{1-b} - \arcsin\sqrt{1-a}).
\end{aligned}$$

The computations to check that these are indeed pdf's, and to check that these measures are invariant can be found in the appendix A.3.1. Now that we have these invariant measures, we need to consider ergodicity:

Definition 12. Let $(X, \mathcal{A}, \mu)$ be a probability space. Suppose $S : X \to X$ is $\mu-$invariant. Then S is said to be *ergodic* if for every measurbale $E \in \mathcal{A}$ satisfying $S^{-1}E = E$, we have $\mu(E) = 0$ or $\mu(E) = 1$.

The logistic map is ergodic [17]. Intuitively, a set E with $F^{-1}(E) = E$ can be the whole domain $[0, 1]$ or a set of periodic points. The whole domain has measure $\mu_F[0, 1] = 1$. Recall that F has a double inverse, every point has a pre-image of two points. So the pre-image of each periodic point x consists of this point x and another point not equal to x. Suppose we take the union $\cup_{i=1}^{\infty} F_{\pm}^{-i}(x)$. Then the pre-image of this union equals this union. The measure of this set is $\mu_F(\cup_{i=1}^{\infty} F_{\pm}^{-i}(x)) = \sum_{i=1}^{\infty} \mu(F_{\pm}^{-i}(x)) = \sum 0 = 0$. Indeed, F is ergodic. The same reasoning can be applied to the map T. We can also prove that these two maps are ergodic using the following proposition:

Proposition 9. *If S has a unique probability density function ρ and has $\rho(x) > 0$ almost everywhere, then S is ergodic.*

Proof. [11] □

The map T has a unique pdf. To derive it, we first need to use the *Perron-Frobenius operator* to find the pdf of the doubling map and then use a change of coordinates to find the pdf of T. This Perron-Frobenius operator P defines a sequence of density functions:

$$\rho_{n+1} = P(\rho_n) = \sum_{y \in f^{-1}(x)} \frac{\rho_n(y)}{|f'(y)|}.$$

The density function of f is then found with the limit $k \to \infty$:

$$\rho(x) = \lim_{k\to\infty} \frac{1}{k} \sum_{n=0}^{k-1} \rho_n(x). \tag{7}$$

To find the pdf of the doubling map, let $\rho_0 = 1$. The pre-image of θ is $\{\frac{1}{2}\theta, \frac{1}{2}(\theta+\pi)\}$. Then the density function is easily found:

$$\begin{aligned}
\rho_1 &= \sum_{y \in \{\frac{1}{2}\theta, \frac{1}{2}(\theta+\pi)\}} \frac{1}{2} = 1 \\
\rho_n &= 1 \quad \forall n \\
\rho &= \lim_{k\to\infty} \frac{1}{k} \sum_{n=0}^{k=1} 1 = 1.
\end{aligned}$$

Note that ρ can be any constant to be a density function for the doubling map, but there is one unique pdf. Normalizing ρ such that it is a pdf gives us $\rho_g = \frac{1}{\pi}$. A *change of variable* [5] then gives us the pdf of T.

Theorem 4. Let X and Y be random variables such that $Y = \varphi(X)$, where φ is a monotonic function. Define $\mathcal{X} = \{x : \rho_X(x) > 0\}$ and $\mathcal{Y} = \{y : y = \varphi(x) \text{ for some } x \in \mathcal{X}\}$. Suppose pdf ρ_X is continuous on $\mathcal{X}$ and that $\varphi^{-1}(y)$ has a continuous derivative on $\mathcal{Y}$. Then the pdf of Y is given by:

$$\rho_Y(y) = \begin{cases} \rho_X(\varphi^{-1}(y)) \left|\frac{d}{dy}(\varphi^{-1}(y))\right| & y \in \mathcal{Y} \\ 0 & \text{otherwise} \end{cases}$$

Proof. [5] □

Consider random variables $X = g$ and $Y = T$. Function $\varphi = h(\theta) = \cot\theta$ is monotonic and continuous. Since $\rho_g > 0$ everywhere, $\mathcal{X}$ is the whole domain of g and $\mathcal{Y}$ is the whole domain of T. All requirements of the theorem are met, so the pdf of T is given by:

$$\rho_T(x) = \rho_g(h^{-1}(x)) \left|\frac{d}{dx} h^{-1}(x)\right| \quad x \in \mathbb{R}\backslash K_T. \tag{8}$$

We know that $h^{-1}(x) = \operatorname{arccot} x$, and its derivative is $-\frac{1}{x^2+1}$. Combining everything gives us exactly the pdf:

$$\rho_T = \frac{1}{\pi} \cdot \frac{1}{x^2+1}.$$

The map T is ergodic, since its density function is unique and larger than 0 everywhere. We can apply Birkhoff's ergodic theorem to ergodic functions. This will be investigated in the next section.

3.2 Birkhoff's ergodic theorem

Recall that the density function has the same shape as the histogram produced by one orbit. One important consequence of this is that if we want to calculate the average of a map along an orbit, we may instead calculate the average over the whole space X. That is, for a map f, we have that for almost all x, the average of $f(T^i(x))$ as i runs from 1 to n converges to the integral of $f(x)\rho_T(x)$ over X, where $\rho_T(x)$ is the invariant density function. Note that $\int_X f(x)\rho_T(x)dx = \int_X f(x)d\mu_T$. This is written more formally as the *Birkhoff Ergodic theorem*, which, according to [6], is the most fundamental fact in ergodic theory:

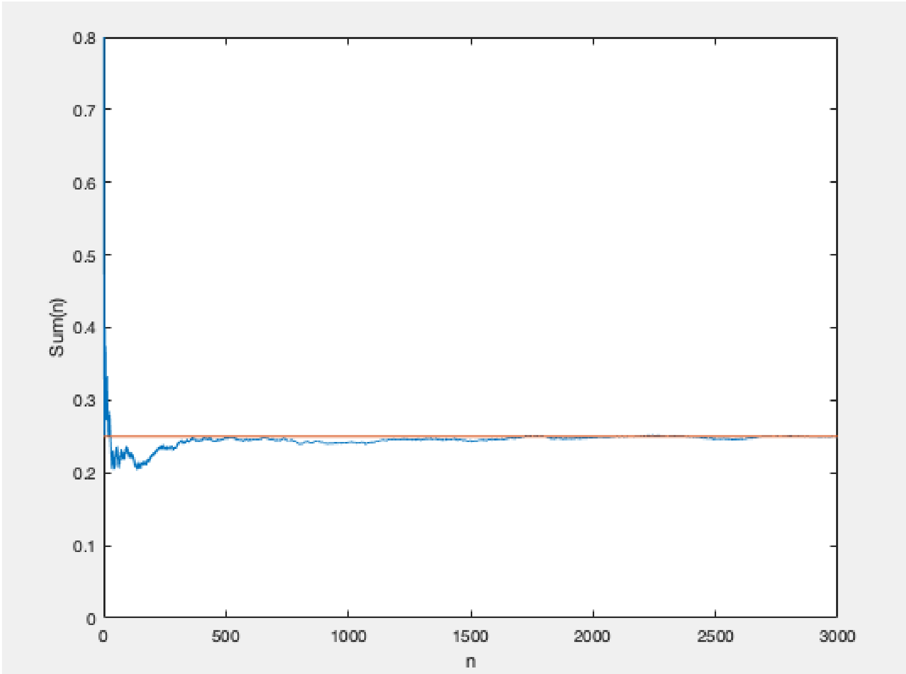


Figure 20: Initial value $x_0 = \pi - 3$. Convergence of $\text{Sum}(n) = \frac{1}{n}\sum_{k=0}^{n-1}\mathbf{1}_{(0,1)}(T^k x)$ to the orange line at $\mu_T\{(0,1)\} = \frac{1}{4}$.

Theorem 5. Let (X, μ) be a probability space. If S is $\mu-$invariant and f is integrable, then

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1} f(S^k x) = f^*(x)$$

for some $f^* \in L^1(X, \mu)$ with $f^*(Tx) = f^*(x)$ for almost every x. Furthermore, if S is ergodic, then f^* is constant and

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1} f(S^k x) = \int_X f d\mu \tag{9}$$

Recall that f is integrable if:

$$\int_X |f| d\mu < \infty.$$

Let $f(x) = \mathbf{1}_E$ be the indicator function of a measurable set E, such that $f(x) = 1$ if $x \in E$ and $f(x) = 0$ otherwise. Clearly, f is integrable. With this f, the left-hand side of equation 9 is a time-average, it says how often the orbit of x lies in E. The right-hand side is the average of f in the state space X. This means that for any $\mu-$invariant, ergodic function S we have:

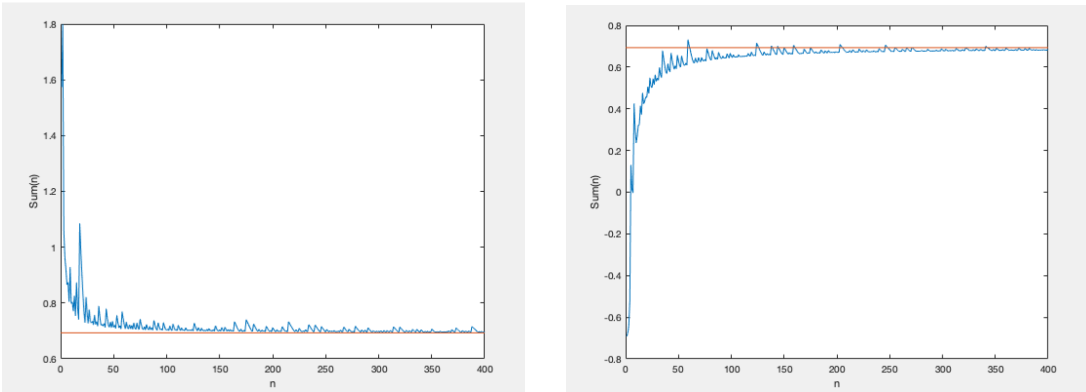
$$\lim_{n\to\infty}\sum_{k=0}^{n-1}\mathbf{1}_E(S^k x) = \int_X \mathbf{1}_E d\mu = \mu(E). \tag{10}$$

In other words, the sequence $\{\frac{1}{n}\sum_{k=1}^{n-1}\mathbf{1}_E(s^k x)\}$ converges to $\mu(E)$ as n goes to infinity. An example of this is shown in figure 20. For $T(x) = \frac{1}{2}(x - \frac{1}{x})$, this convergence is seen for $E = (0,1)$. The measure $\mu_T(0,1) = \frac{1}{4}$, and indeed, we see the sequence converging nicely to it.

3.3 Lyapunov exponent

An application of Birkhoff's theorem is the computation of the Lyapunov exponent. It is a way to measure sensitive dependence on initial conditions. Let x, y be two points close together. Then:

$$|S(x) - S(y)| \approx |S'(x)| \cdot |x - y|,$$



(a) $x_0 = 0.34$ (b) $x_0 = 23$

Figure 21: Convergence of $\text{Sum}(n) = \frac{1}{n}\sum_{i=1}^{n-1} \log |T'(T^i x)|$ to the orange line at $\log 2$

which gives by induction:

$$|S^n(x) - S^n(y)| \approx \prod_{i=0}^{n-1} |S'(S^i x)| \cdot |x - y|.$$

Then we can write this as:

$$\frac{1}{n} \log |S^n(x) - S^n(y)| \approx \frac{1}{n} \sum_{i=0}^{n-1} \log |S'(S^i x)|.$$

Note that we can drop the term $\frac{1}{n} \log |x - y|$ since it converges to 0 as $n \to \infty$. Using Birkhoff's ergodic theorem, the right-hand side converges to $\int_X \log |S'| d\mu$ for ergodic, $\mu-$invariant transformations S.

Definition 13. The number $\int_X \log |S'| d\mu$ is called the *Lyapunov exponent* of S.

This Lyapunov exponent measures the exponent of the speed of divergence of x and y [6]. A positive Lyapunov exponent usually indicates that a map has sensitive dependence on initial conditions [3]. The logistic map and the map T are ergodic $\mu-$invariant transformations so we can determine their Lyapunov exponents. In appendix A.4 are the computations for the Lyapunov exponent for F and T. Both have a lyapunov exponent of $\lambda = \log 2$. The convergence of $\frac{1}{n}\sum_{i=0}^{n-1} \log |T'(T^i(x))|$ also shows that the Lyapunov exponent for T is equal to $\log 2$. Figure 21 shows this convergence for initial conditions $x_0 = 0.34$ and $x_0 = 23$. Both converge nicely to the orange line at $\log 2$.

4 Family N_c

Up till now, we have investigated the dynamics of T. Recall that $T(x) = N_1(x)$ is one member of this family:

$$N_c(x) = \frac{1}{2}(x - \frac{c}{x}), \quad c > 0$$

In this section, we will return to this family. Figure 22 shows that N_c has the same shape for all c. None of these maps have fixed points, and we expect to see chaotic iterations for all values of c. For T, we found a topological conjugacy, with the doubling map. This topological conjugacy can be extended to all N_c with a $\sqrt{c}$ factor:

$$h_c(\theta) = \sqrt{c} \cot \theta$$

Indeed:

$$\begin{aligned} N_c \circ h_c &= \frac{1}{2}\left(\sqrt{c}\cot\theta - \frac{c}{\sqrt{c}\cot\theta}\right) \\ &= \frac{\sqrt{c}}{2}\left(\cot\theta - \frac{1}{\cot\theta}\right) \\ &= \sqrt{c}\cot 2\theta = h_c \circ g. \end{aligned}$$

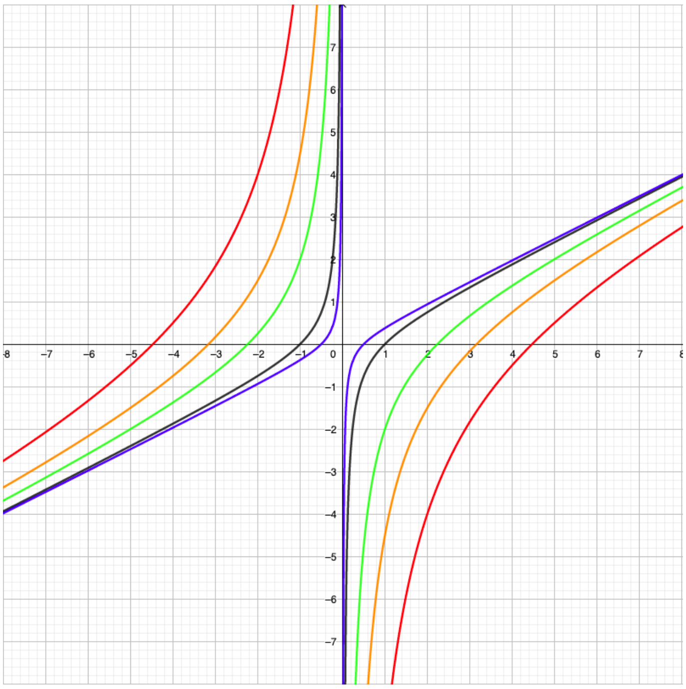


Figure 22: Map N_c for $c = \frac{1}{4}, 1, 5, 10, 20$.

Because of this conjugacy, we will indeed find similar dynamics for N_c. For example, if θ is a periodic point of the doubling map, then $x = \sqrt{c}\cot\theta$ is a periodic point of N_c. For ergodic theory, we also find similar results. The pdf and invariant measure for N_c are:

$$\begin{aligned}
\rho_{N_c}(x) &= \frac{\sqrt{c}}{\pi(x^2+c)}, \\
\mu_{N_c}([a,b]) &= \int_a^b \frac{\sqrt{c}}{\pi(x^2+c)}dx, \\
&= \frac{\sqrt{c}}{\pi}(\arctan\frac{b}{\sqrt{c}} - \arctan\frac{a}{\sqrt{c}}).
\end{aligned}$$

Surprisingly, the Lyapunov exponent of N_c is equal to $\log 2$ for all c, see appendix A.4. This indicates that all maps N_c are equally sensitive to initial conditions.

Conclusion

The Newton function of the family $z^2 + c$ is given by $N_c(x) = \frac{1}{2}(x - \frac{c}{x})$. In this thesis, the dynamics of the map $N_1(x) = T(x)$ have been investigated. There is a topological conjugacy with the well-known doubling map. Using this conjugacy, we were able to find all the periodic points and prove that they are repelling. Furthermore, the exact domain of T has been identified.

The iterations of the map seemed chaotic. They do not converge to any attracting periodic point. Different definitions of chaos have been considered, starting with Li-Yorke chaos. The map T is not defined on an interval, which was a requirement for this theorem. For the logistic map, this theorem can be applied and indeed, there is Li-Yorke chaos for parameters larger than $\mu_c \approx 3.56995$. Next, we considered an expansion on Li-Yorke chaos, namely Marotto's snap-back repeller. The requirements for the domain were a bit unclear as Marotto does not mention the need for an interval. As expected, the logistic map has a snap-back repeller. For T, we also found a snap-back repeller.

Another well-known definition for chaos is the one by Devaney. Using the conjugacy with the doubling map, we were able to prove that T is chaotic according to Devaney's definition. The last definition of chaos that we considered was the dispersion exponent. The dispersion exponent works with the l'historire se répète principle. This way of predicting future prediction does not work for a chaotic dynamical system. For the doubling map and the logistic map, we were able to calculate the dispersion exponent. For the map T, it was a bit harder. Iterations of T can move from very small values to very large values, which gives very irregular numerical results. However, there was a clear upward trend that suggests a positive dispersion exponent.

From these four notions of chaos, we can conclude that the map T is chaotic. It was not always easy to prove this. For Devaney, we gave an exact proof, and for the other notions of chaos, we have made it likely that T is chaotic. A compact domain would have been more convenient. In that case, Li and Yorke's theorem and Marotto can be applied with no problem. For the dispersion exponent, a compact domain would mean that the increase between iterations is bounded.

After the topological approach to the map T, we turned to ergodic theory. We found a probability density function for the logistic map and for the map T. This led to an invariant measure with which we proved that the maps are ergodic. Next, we took a look at Birkhoff's ergodic theorem. This important theorem of ergodic theory was used to calculate the Lyapunov exponent. The Lyapunov exponent of both T and the logistic map is $\log 2$. In fact, we were able to calculate the Lyapunov exponent for all N_c. It is equal to $\log 2$ for all c. This positive Lyapunov exponent implies sensitive dependence on initial conditions.

This thesis covered many aspects of the dynamics of T, but there is always much left to be investigated. For instance, other notions of chaos, or the link between the different notions of chaos. There is for example an overlap between Li and Yorke's chaos and Devaney's chaos. It would be interesting to find out what this overlap is. There are also more ergodic properties that can be investigated, like entropy and mixing.

References

[1] John Banks, Jeffrey Brooks, Grant Cairns, Gary Davis, and Peter Stacey. On Devaney's definition of chaos. *The American mathematical monthly*, 99(4):332–334, 1992.

[2] Paul Blanchard et al. The dynamics of Newton's method. In *Proc. Symp. Appl. Math*, volume 49, pages 139–154, 1994.

[3] Michael Brin and Garrett Stuck. *Introduction to dynamical systems.* Cambridge university press, 2002.

[4] Henk Broer and Floris Takens. *Dynamical systems and chaos*, volume 172. Springer Science & Business Media, 2010.

[5] George Casella and Roger L Berger. *Statistical inference.* Cengage Learning, 2021.

[6] Geon Ho Choe. *Computational ergodic theory*, volume 13. Springer Science & Business Media, 2006.

[7] Robert L. Devaney. *An Introduction to Chaotic Dynamical Systems Second Edition.* Second edition, 1989.

[8] William B Gordon. Period three trajectories of the logistic map. *Mathematics magazine*, 69(2):118–120, 1996.

[9] Peter Grassberger and Itamar Procaccia. Measuring the strangeness of strange attractors. In *The Theory of Chaotic Attractors*, pages 170–189. Springer, 2004.

[10] Wen Huang and Xiangdong Ye. Devaney's chaos or 2-scattering implies Li–Yorke's chaos. *Topology and its Applications*, 117(3):259–272, 2002.

[11] Andrzej Lasota and Michael C Mackey. *Chaos, fractals, and noise: stochastic aspects of dynamics*, volume 97. Springer Science & Business Media, 2013.

[12] Tien-Yien Li and James A Yorke. Period three implies chaos. *The American Mathematical Monthly*, 82(10):985–992, 1975.

[13] Edward N Lorenz and K Haman. The essence of chaos. *Pure and Applied Geophysics*, 147(3):598–599, 1996.

[14] Frederick R. Marotto. Snap-back repellers imply chaos in $\mathbb{R}^n$. *Journal of mathematical analysis and application*, 63:199–223, 1978.

[15] Frederick R Marotto. On redefining a snap-back repeller. *Chaos, Solitons & Fractals*, 25(1):25–28, 2005.

[16] Heinz-Otto Peitgen, Dietmar Saupe, and F v Haeseler. *Cayley's problem and Julia sets.* Universität Bremen. Fachbereiche Mathematik/Informatik, Elektrotechnik ..., 1983.

[17] Valérie Poulin. *A Generalization of the logistic equation.* PhD thesis, Carleton University, 2000.

[18] Yuming Shi and Guanrong Chen. Discrete chaos in Banach spaces. *Science in China Series A: Mathematics*, 48(2):222–238, 2005.

[19] P Straffin Jr. Newton's method and fractal patterns. *Applications of Calculus*, 3:68–84, 1991.

[20] Gilbert Strang. A chaotic search for i. *The College Mathematics Journal*, 22(1):3–12, 1991.

[21] Wilson A Sutherland. *Introduction to metric and topological spaces.* Oxford University Press, 2009.

A Appendix

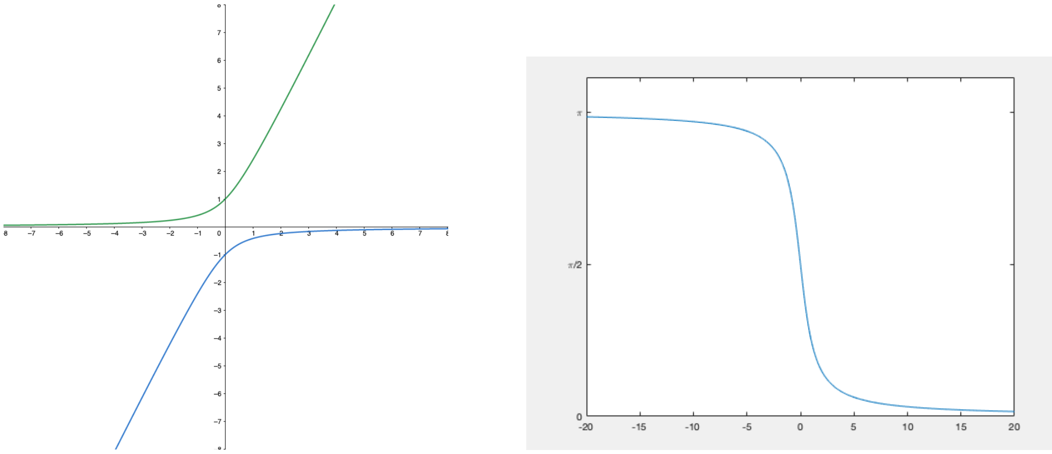
A.1 Inverse functions

A.1.1 Inverse T^{-1}

T has a double inverse, see figure 23a:

$$T_{\pm}^{-1}x = x \pm \sqrt{x^2+1}$$

For an inverse $T^{-1} \circ T(x) = T \circ T^{-1}(x) = x$ has to be true. Identities $T \circ T_{+}^{-1}(x) = T_{+}^{-1} \circ T(x) = x$ and $T \circ T_{-}^{-1}(x) = x$ are easily checked. For $T_{-}^{-1} \circ Tx$, we need to be careful:



(a) Green is $T_+^{-1}(x)$ and blue is $T_-^{-1}(x)$ (b) $\operatorname{arccot}\theta$

Figure 23: Inverse functions.

$$\begin{aligned}
T_-^{-1} \circ T(x) &= \frac{1}{2}(x - \frac{1}{x}) - \sqrt{(\frac{1}{2}(x - \frac{1}{x}))^2 + 1} \\
&= \frac{1}{2}(x - \frac{1}{x}) - \sqrt{\frac{1}{4}(x^2 - 2 + \frac{1}{x^2}) + 1} \\
&= \frac{1}{2}(x - \frac{1}{x}) - \sqrt{\frac{1}{4}(x + \frac{1}{x})^2} \\
&= \frac{1}{2}(x - \frac{1}{x}) - |\frac{1}{2}(x + \frac{1}{x})| \\
&= \begin{cases} x & \text{if } x > 0 \\ \frac{1}{x} & \text{if } x < 0 \end{cases}
\end{aligned}$$

$T_-^{-1} \circ T(x) = x$ only holds for $x > 0$ since:

$$\sqrt{\frac{1}{4}(x + \frac{1}{x})^2} = |\frac{1}{2}(x + \frac{1}{x})| = \begin{cases} \frac{1}{2}(x + \frac{1}{x}) & \text{if } x > 0 \\ -\frac{1}{2}(x + \frac{1}{x}) & \text{if } x < 0 \end{cases}$$

So T_-^{-1} and T_+^{-1} act as a right-inverse to T. For $x > 0$, we can use T_+^{-1} as a left-inverse and for $x < 0$ we can use T_-^{-1} as a left-inverse.

A.1.2 Inverse cotangent

The inverse cotangent is a function $h^{-1} : \mathbb{R} \to (0, \pi)$ given by:

$$h^{-1}(x) = \operatorname{arccot} x.$$

By definition $\cot \circ \operatorname{arccot}\theta = \operatorname{arccot} \circ \cot\theta = \theta$. We will consider $\operatorname{arccot} : \mathbb{R} \to [0, \pi]$. Figure 23b shows this function. Note that $\lim_{x\to\infty} \operatorname{arccot} x = 0$ and $\lim_{x\to-\infty} \operatorname{arccot} x = \pi$. This function is everywhere continuous.

A.1.3 T and the logistic map

In this section, we find the more elaborate computations of the relation between T and F_4. The fist substitution is $x_{n+1} = T(x_n)$ into $y_n = x_n^2 + 1$:

$$
\begin{aligned}
y_{n+1} &= x_{n+1}^2 + 1 \\
&= \frac{1}{4}(x_n - \frac{1}{x_n})^2 + 1 \\
&= \frac{1}{4}(x_n^2 - 2 + \frac{1}{x_n^2}) + 1 \\
&= \frac{1}{4}(x_n^2 + 2 + \frac{1}{x_n^2}) \\
&= \frac{1}{4}\frac{(x_n^2+1)^2}{x_n^2} \\
&= \frac{1}{4}\frac{y_n^2}{y_n - 1}
\end{aligned}
$$

The second substitution $z = \frac{1}{y}$ gives us exactly F_4:

$$
\begin{aligned}
y_{n+1} &= \frac{1}{4}\frac{y_n^2}{y_n - 1} \\
\frac{1}{y_{n+1}} &= \frac{4(y_n - 1)}{y_n^2} \\
&= \frac{4}{y_n}(1 - \frac{1}{y_n}) \\
z_{n+1} &= 4z_n(1 - z_n)
\end{aligned}
$$

This relation between T and F_4, given by $z_n = \frac{1}{x_n^2+1}$ is not a topological conjugacy, since z_n is not injective. The relation is shown below in figure 24 and in this equation:

$$
\begin{aligned}
z_n \circ T &= F_4 \circ z_n \\
\frac{1}{\frac{1}{4}(x_n - \frac{1}{x_n})^2 + 1} &= 4(\frac{1}{x_n^2+1})(1 - \frac{1}{x_n^2+1}) \\
\frac{4}{(x_n + \frac{1}{x_n})^2} &= \frac{4x_n^2}{(x_n^2+1)^2} = \frac{4}{(x_n + \frac{1}{x_n^2})}.
\end{aligned}
$$

$$
\begin{tikzcd}
\mathbb{R} \arrow[r, "T"'] \arrow[d, "z_n"'] & \mathbb{R} \arrow[d, "z_n"] \\
{[0,1]} \arrow[r, "F_4"'] & {[0,1]}
\end{tikzcd}
$$

Figure 24: Link with the Logistic map

A.2 Chaos

A.2.1 Li and Yorke

Lemma 3. A point $x \in J$ is asymptotically periodic if and only if there exists a periodic point $p \in J$ such that $|F^n(x) - F^n(p)| \to 0$ as $n \to \infty$.

Proof. Assume that $x \in J$ is periodic, which means that there exists a period-k point $p \in J$ such that equation (5) holds. We claim that there exists an integer $0 \leq i \leq k-1$ such that

$$|F^n(x) - F^n(F^i(p))| \to 0.$$

If such an integer does not exist, then there exists $\varepsilon_i > 0$ such that $|F^n(x) - F^n(F^i(p))| \geq \varepsilon_i$ for infinitely many $n \in \mathbb{N}$. For $\varepsilon = \min\{\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_{k-1}\}$ we then have

$$|F^n(x) - F^n(F^i(p))| \geq \varepsilon$$

for infinitely many $n \in \mathbb{N}$ and all integers $0 \leq i \leq k-1$. But this contradicts equation (5). Therefore, we have that $|F^n(x) - F^n(q)| \to 0$ for some periodic point $q = F^i(p)$.
Conversely, let $p \in J$ be a period-k point such that $|F^n(x) - F^n(p)| \to 0$ as $n \to \infty$. Using division with remainder we can write any $n \in \mathbb{N}$ as $n = dk + r$ with $d \in \mathbb{N}$ and $0 \leq r < k-1$. This gives

$$\begin{aligned}|F^n(x) - F^n(p)| &= |F^n(x) - F^{dk+r}(p)| \\ &= |F^n(x) - F^r(p)| \\ &\geq d(F^n(x), \{F^i(p)\,;\, i = 0, \ldots, k-1\}).\end{aligned}$$

We conclude that equation (5) holds, which means that x is asymptotically periodic. □

A.2.2 Dispersion Exponent Matlab

doubling.m:

```
function g=doubling(x)
% doubling map on interval [0, pi)

g = 2*x;

while (g > pi)
    g = g - pi;
end;
while (g < 0)
    g = g + pi;
end;
```

funcT.m:

```
function y=funcT(x)
y=1/2.*(x-1/x);
```

logistic.m:

```
function y=logistic(mu,x)
y=4*x.*(1-x);
```

line_distance.m:

```
function d=line_distance(x, y)
% compute distance between two points on the real line

d  = abs(x-y);
```

circle_distance.m:

```
function d=circle_distance(x, y)
% compute distance between two points on a circle of length pi

d1 = abs(x-y);
d2 = pi-d1;
d  = min([d1, d2]);
```

compute_sup.m:

```
function Ese = compute_sup(x0, N, epsilon, s)
%function to calculate E(s,epsilon)

```

```
% compute orbit of length N+s (note: not length N!)
xarray = zeros(1,N+s);
xarray(1) = x0;
for i = 2:(N+s)
    xarray(i) = doubling(xarray(i-1));
end

% approximate E(s, epsilon) over a finite orbit segment
Ese=0;
for i = 1:N
    for j = (i+1):N
        d0 = circle_distance(xarray(i), xarray(j));
        if (d0 < epsilon)
            ds = circle_distance(xarray(i+s), xarray(j+s));
            q = ds/d0;
            % find the supremum
            if (q > Ese)
                Ese = q;
            end
        end
    end
end

% for the logistic map or T change the following
% 1) change "circle_distance" into "line_distance"
% 2) change "doubling" into the appropriate function name
```

plot_Ese.m:

```
function y=plot_Ese(x0, N, eps_array, s_array)
%function to make a graph of s vs log(E(s))

for i = 1:length(eps_array)
    epsilon = eps_array(i);
    y = zeros(1, length(s_array));
    for j = 1:length(s_array)       % y is sup of ratio so y=E(s,
       epsilon)
        s    = s_array(j);
        y(j) = compute_sup(x0, N, epsilon, s);
    end

    % plot log(E(s,epsilon)) as a function of s
    % for different fixed choices of epsilon
    plot(s_array, log(y))
    hold on
end

hold off
```

A.3 Ergodic Theory

A.3.1 Density functions and invariant measure T

The pdf of the map T is:

$$\rho_T(x) = \frac{1}{\pi(1+x^2)}$$

To show that this is in fact a pdf, its integral over the entire domain must equal 1. Note that K_T is a set of single points, and therefore $\int_{K_T} \rho(x)dx = \sum_{k\in K_T} \int_k \rho(x)dx = \sum 0 = 0$:

$$\int_{\mathbb{R}\setminus K_T} \rho(x)dx = \int_{\mathbb{R}} \rho(x)dx - \int_{K_T} \rho(x)dx$$
$$\int_{-\infty}^{\infty} \frac{1}{\pi(x^2+1)}dx = \frac{1}{\pi} \arctan x\Big|_{-\infty}^{\infty}$$
$$= \frac{1}{\pi}(\frac{\pi}{2} + \frac{\pi}{2}) = 1$$

Furthermore, $\mu_T(T^{-1}([a,b]))$ must equal $\mu_T([a,b])$ for an invariant measure:

$$\begin{aligned}\mu_T(T^{-1}([a,b])) =& \mu([a-\sqrt{a^2+1}, b-\sqrt{b^2+1}]) + \mu([a+\sqrt{a^2+1}, b+\sqrt{b^2+1}])\\ =& \frac{1}{\pi}(\arctan(b-\sqrt{b^2+1}) + \arctan(b+\sqrt{b^2+1}))\\ &- \frac{1}{\pi}(\arctan(a-\sqrt{a^2+1}) + \arctan(a+\sqrt{a^2+1}))\\ =& \frac{1}{\pi}(\arctan b - \arctan a).\end{aligned}$$

This last equality holds due tot the arctan sum rule. Let $\alpha = b - \sqrt{b^2+1}$ and $\beta = b + \sqrt{b^2+1}$ (The same works for a):

$$\arctan\alpha + \arctan\beta = \arctan\left(\frac{\alpha+\beta}{1-\alpha\beta}\right)$$
$$\arctan(b-\sqrt{b^2+1}) + \arctan(b+\sqrt{b^2+1}) = \arctan\left(\frac{2b}{1-(b^2-(b^2+1))}\right) = \arctan b.$$

A.3.2 Density function and invariant measure F

For the logistic map $F(x) = 4x(1-x)$, we have pdf:

$$\rho_F(x) = \frac{1}{\pi\sqrt{x(1-x)}}.$$

The integral over the whole domain is 1:

$$\int_0^1 \frac{1}{\pi\sqrt{x(1-x)}}dx = -\frac{2}{\pi}(\arcsin 0 - \arcsin 1) = -\frac{2}{\pi}(0 - \frac{\pi}{2}) = 1$$

The inverse of F is given by $F_{\pm}^{-1}(x) = \frac{1}{2} \pm \frac{1}{2}\sqrt{1-x}$.
To check that this measure is invariant, we need to check that:

$$\mu_F([a,b]) = \mu_F(F_{pm}^{-1}([a,b])).$$

Note that $[a,b]$ is mapped by the inverse to two intervals, namely $F^{-1}([a,b]) = [\frac{1}{2} - \frac{1}{2}\sqrt{1-a}, \frac{1}{2} - \frac{1}{2}\sqrt{1-b}] \cup [\frac{1}{2} + \frac{1}{2}\sqrt{1-b}, \frac{1}{2} + \frac{1}{2}\sqrt{1-a}]$. This gives:

$$\begin{aligned}\mu(F_{\pm}^{-1}([a,b])) &= \mu([\frac{1}{2} - \frac{1}{2}\sqrt{1-a}, \frac{1}{2} - \frac{1}{2}\sqrt{1-b}]) + \mu([\frac{1}{2} + \frac{1}{2}\sqrt{1-b}, \frac{1}{2} + \frac{1}{2}\sqrt{1-a}])\\ &= -\frac{2}{\pi}\left(\arcsin\sqrt{1-(\frac{1}{2}-\frac{1}{2}\sqrt{1-b})} - \arcsin\sqrt{1-(\frac{1}{2}-\frac{1}{2}\sqrt{1-a})}\right)\\ &- \frac{2}{\pi}\left(\arcsin\sqrt{1-(\frac{1}{2}+\frac{1}{2}\sqrt{1-a})} - \arcsin\sqrt{1-(\frac{1}{2}+\frac{1}{2}\sqrt{1-b})}\right)\\ &= -\frac{2}{\pi}\left(\arcsin\sqrt{\frac{1}{2}+\frac{1}{2}\sqrt{1-b}} - \arcsin\sqrt{\frac{1}{2}-\frac{1}{2}\sqrt{1-b}}\right)\\ &- \frac{2}{\pi}\left(\arcsin\sqrt{\frac{1}{2}-\frac{1}{2}\sqrt{1-a}} - \arcsin\sqrt{\frac{1}{2}+\frac{1}{2}\sqrt{1-a}}\right)\end{aligned}$$

We will apply the arcsin difference rule with $\alpha = \sqrt{\frac{1}{2}+\frac{1}{2}\sqrt{1-b}}$ and $\beta = \sqrt{\frac{1}{2}-\frac{1}{2}\sqrt{1-b}}$, note that $\sqrt{1-\alpha^2} = \beta$ and $\sqrt{1-\beta^2} = \alpha$ so:

$$\begin{aligned}\arcsin\alpha - \arcsin\beta &= \arcsin(\alpha\sqrt{1-\beta^2} - \beta\sqrt{1-\alpha^2}) \\ &= \arcsin(\alpha^2 - \beta^2)\end{aligned}$$

$$\arcsin\sqrt{\frac{1}{2}+\frac{1}{2}\sqrt{1-b}} - \arcsin\sqrt{\frac{1}{2}-\frac{1}{2}\sqrt{1-b}} = \arcsin\left(\frac{1}{2}+\frac{1}{2}\sqrt{1-b}-\frac{1}{2}+\frac{1}{2}\sqrt{1-b}\right) = \arcsin\sqrt{1-b}$$

Apply the arcsin difference rule again with $\alpha = \sqrt{\frac{1}{2}-\frac{1}{2}\sqrt{1-a}}$. Again, $\sqrt{1-\alpha^2} = \beta$ and $\sqrt{1-\beta^2} = \alpha$, so:

$$\begin{aligned}\arcsin\sqrt{\frac{1}{2}-\frac{1}{2}\sqrt{1-a}} - \arcsin\sqrt{\frac{1}{2}+\frac{1}{2}\sqrt{1-a}} &= \arcsin\left(\frac{1}{2}-\frac{1}{2}\sqrt{1-a}-\frac{1}{2}-\frac{1}{2}\sqrt{1-a}\right) = \arcsin(-\sqrt{1-a}) \\ &= -\arcsin\sqrt{1-a}.\end{aligned}$$

And then finally:

$$\mu(F^{-1}[([a,b])) = -\frac{2}{\pi}(\arcsin\sqrt{1-b} - \arcsin\sqrt{1-a}) = \mu([a,b]).$$

A.4 Lyapunov Exponent

A.4.1 Logistic map F

The Lyapunov exponent of the logistic map F is given by the following integral. In this section it is solved.

$$\begin{aligned}\lambda_F &= \int_X \log|F'|\,d\mu_F \\ &= \int_0^1 \log|4-8x|\,d\mu_F \\ &= \int_0^1 \log 4 + \log|1-2x|\,d\mu_F \\ &= 2\log 2 + \int_0^1 \log|1-2x|\,d\mu_F \\ &= 2\log 2 + \int_0^1 \frac{\log|1-2x|}{\pi\sqrt{x(1-x)}}\,dx\end{aligned}$$

The part in the integral is symmetric around $x = \frac{1}{2}$, so:

$$\lambda_F = 2\int_0^{\frac{1}{2}} \frac{\log|1-2x|}{\pi\sqrt{x(1-x)}}\,dx + 2\log 2$$

Now put $1-2x = \sin\phi$, $0 \le \phi \le \frac{\pi}{2}$. Then $x = -\frac{1}{2}\sin\phi + \frac{1}{2}$ and $dx = -\frac{1}{2}\cos\phi\,d\phi$. Note furthermore that:

$$\sqrt{x(1-x)} = \sqrt{(-\frac{1}{2}\sin\phi+\frac{1}{2})(\frac{1}{2}+\frac{1}{2}\sin\phi)} = \sqrt{\frac{1}{4}\cos^2\phi} = \frac{1}{2}\cos\phi.$$

Lastly, $x = 0 \Rightarrow \phi = \frac{\pi}{2}$ and $x = \frac{1}{2} \Rightarrow \phi = 0$ and $\sin\phi > 0$ for these values of ϕ, so we find:

$$\begin{aligned}\lambda_F &= \frac{2}{\pi}\int_{\frac{\pi}{2}}^0 \frac{\log(\sin\phi)}{\frac{1}{2}\cos\phi} \cdot -\frac{1}{2}\cos\phi\,d\phi + 2\log 2 \\ &= \frac{2}{\pi}\int_0^{\frac{\pi}{2}} \log(\sin\phi)\,d\phi + 2\log 2\end{aligned}$$

Using the mean value theorem for harmonic functions we find [6]:

$$\begin{aligned}\int_0^{\frac{\pi}{2}} \log(\sin\phi)\,d\phi &= -\frac{2}{\pi}\log 2 \\ \implies \lambda_F &= -\log 2 + 2\log 2 = \log 2.\end{aligned}$$

A.4.2 Family N_c

The Lyapunov exponent of the family N_c is given by the following integral. Note that $c = 1$ gives the Lyapunov for T.

$$\begin{aligned}
\lambda_N &= \int_X \log|N_c'|\, d\mu_N \\
&= \int_{-\infty}^{\infty} \log|\frac{1}{2}(1+\frac{c}{x^2})| d\mu_N \\
&= \int_{-\infty}^{\infty} \log\frac{1}{2}\, d\mu_N + \int_{-\infty}^{\infty} \log|1+\frac{c}{x^2}|\, d\mu_N \\
&= -\log 2 + \frac{\sqrt{c}}{\pi}\int_{-\infty}^{\infty} \frac{\log|1+\frac{c}{x^2}|}{(x^2+c)}\, dx
\end{aligned}$$

Use the transformation

$$x = \frac{\sqrt{c}}{\tan(-\sqrt{c}t)}$$

Note that we get new bounds at $x = \infty$ at $t = 0$ and $x = 0$ at $t = -\frac{\pi}{2\sqrt{c}}$. Also:

$$dx = \frac{c}{\sin^2(-\sqrt{c}t)} dt.$$

$$\begin{aligned}
\lambda_N &= -\log 2 + \frac{2\sqrt{c}}{\pi}\int_{-\frac{\pi}{2\sqrt{c}}}^{0} \frac{\log|1+\tan^2(-\sqrt{c}t)|}{\frac{c}{\tan^2(-\sqrt{c}t)}+c}\cdot\frac{c}{\sin^2(-\sqrt{c}t)}\, dt \\
&= -\log 2 + \frac{2\sqrt{c}}{\pi}\int_{-\frac{\pi}{2\sqrt{c}}}^{0} \frac{\log|1+\tan^2(-\sqrt{c}t)|}{\cos^2(\sqrt{c}t)+\sin^2(\sqrt{c}t)}\, dt \\
&= -\log 2 + \frac{2\sqrt{c}}{\pi}\int_{-\frac{\pi}{2\sqrt{c}}}^{0} \log|1+\tan^2(-\sqrt{c}t)|\, dt \\
&= -\log 2 + \frac{2\sqrt{c}}{\pi}\int_{-\frac{\pi}{2\sqrt{c}}}^{0} \log|\frac{\cos^2(-\sqrt{c}t)+\sin^2(-\sqrt{c}t)}{\cos^2(-\sqrt{c}t)}|\, dt \\
&= -\log 2 - \frac{4\sqrt{c}}{\pi}\int_{-\frac{\pi}{2\sqrt{c}}}^{0} \log|\cos(-\sqrt{c}t)|\, dt
\end{aligned}$$

Another quick transformation: $t = -\frac{s}{\sqrt{c}}$ gives new bounds at $s = 0$ and $s = \frac{\pi}{2}$ and $dt = -\frac{1}{\sqrt{c}}ds$:

$$\begin{aligned}
\lambda_N &= -\log 2 + \frac{4\sqrt{c}}{\pi}\int_0^{\frac{\pi}{2}} \log|\cos(s)|\cdot -\frac{1}{\sqrt{c}}ds \\
&= -\log 2 - \frac{4}{\pi}\int_0^{\frac{\pi}{2}} \log|\cos(s)|ds
\end{aligned}$$

Now consider the analytic function $f(z) = \log(1+z)$. Its real part $u(z) = \log|1+z|$ is harmonic. Then by the Mean value theorem for harmonic functions:

$$u(0) = \frac{1}{2\pi}\int_0^{2\pi} u(0+re^{i\phi})\, d\phi = \frac{1}{2\pi}\int_0^{2\pi} \log|1+re^{i\phi}|\, d\phi.$$

Let $r \to 1$, then we find that $|1+e^{i\phi}| = 2|\cos(\frac{\phi}{2})|$:

$$
\begin{aligned}
0 &= \frac{1}{2\pi}\int_0^{2\pi} \log|1+e^{i\phi}|\,d\phi \\
&= \frac{1}{2\pi}\int_0^{2\pi} \log 2|\cos\frac{\phi}{2}|\,d\phi \\
&= \frac{1}{2\pi}\int_0^{2\pi} \log 2\,d\phi + \frac{1}{2\pi}\int_0^{2\pi} \log|\cos\frac{\phi}{2}|\,d\phi \\
-\log 2 &= \frac{1}{2\pi}\int_0^{2\pi} \log|\cos\frac{\phi}{2}|\,d\phi
\end{aligned}
$$

Let $\phi = 2t$, then $d\phi = 2dt$. If $\phi = 0$, then $t = 0$, and if $\phi = 2\pi$, then $t = \pi$:

$$
\begin{aligned}
-\log 2 &= \frac{1}{2\pi}\int_0^{2\pi} \log|\cos\frac{\phi}{2}|\,d\phi \\
-\pi\log 2 &= \int_0^{\pi} \log|\cos(t)|\,dt \\
-\frac{\pi}{2}\log 2 &= \int_0^{\frac{\pi}{2}} \log|\cos t|\,dt
\end{aligned}
$$

Combining this with our earlier result gives us:

$$
\lambda_N = -\log 2 - \frac{4}{\pi}\cdot\frac{-\pi}{2}\log 2 = \log 2
$$