Consideraciones Geométricas, Dinámicas y Topológicas de Funciones Tipo Herradura.

Alejandro Cardona Zapata

Universidad Nacional de Colombia

Facultad de Ciencias, Escuela de Matemáticas

Medellín, Colombia

2023

Consideraciones Geométricas, Dinámicas y Topológicas de Funciones Tipo Herradura.

Alejandro Cardona Zapata

Tesis presentada como requisito parcial para optar al título de:

Magíster en ciencias - Matemáticas

Director:

Prof. Óscar Iván Giraldo Galeano

Codirector:

Prof. José Gregorio Rodríguez Nieto

Universidad Nacional de Colombia

Facultad de Ciencias, Escuela de Matemáticas

Medellín, Colombia

2023

Geometric, Dynamical, and Topological Considerations of Horseshoe-Type Functions.

Alejandro Cardona Zapata

Universidad Nacional de Colombia

Facultad de Ciencias, Escuela de Matemáticas

Medellín, Colombia

2023

Agradecimientos

En primer lugar, quiero agradecer a mis asesores, Óscar Iván Giraldo y Jose Gregorio Rodríguez, por su dedicación y confianza. Al profesor Sergio Alejandro Carrillo, jurado de este trabajo, por su orientación y consejos. Y a todos los profesores del departamento de matemáticas que hicieron parte de mi formación profesional.

Resumen

En esta tesis se abordará el estudio de una versión de la **función herradura de Smale** y la **función de baker**, así como sus propiedades dinámicas, desde una perspectiva topológica. Con este fin, se presentarán algunas nociones necesarias sobre **geometría hiperbólica**, **sistemas dinámicos** y **superficies de Riemann**.

Posteriormente, definiremos las funciones herradura y Baker partiendo de la composición de dos transformaciones geométricas definidas sobre el cuadrado unitario, denotado por Q. Estas transformaciones inducirán una serie de identificaciones en la frontera de Q y usaremos las nociones introducidas sobre superficies de Riemann para estudiar las propiedades geométricas y topológicas del espacio cociente generado por dichas identificaciones.

Por último, introducimos algunos conceptos de la **teoría ergódica** para estudiar la dinámica del sistema de la función herradura.

Palabras clave: sistema dinámico, mapa herradura, superficies generadas dinámicamente, automorfismos de superficies, caos.

Abstract

In this thesis we will address the study of a version of the **Smale's horseshoe function** and the **baker's function**, along with their dynamical properties, from a topological perspective. To this end, fundamental notions regarding **hyperbolic geometry**, **dynamical systems**, and **Riemann surfaces** will be introduced.

Subsequently, we will define the horseshoe and Baker functions by composing two geometric transformations defined on the unit square, denoted as Q. These transformations will induce a series of identifications on the boundary of Q. To study the geometric and topological properties of the resulting quotient space, we will employ the notions introduced about Riemann surfaces.

Finally, we introduce some concepts from **ergodic theory** to study the dynamics of the horseshoe function system.

Keywords: dynamical system, horseshoe map, dynamically generated surfaces, surfaces automorphisms, chaos.

Contents

Introduction

One of the earliest appearances of chaos in mathematics is evident in the works of the French mathematician **Henri Poincaré** on celestial mechanics, specifically in the well-known three-body problem, which states the following: "*Let A, B, and C be three bodies moving in space solely under the influence of their mutual gravitational attraction. Given the initial positions and velocities of each body, determine their future motion at any given time t.*"

In November 1890, Poincaré believed he had found the solution to the problem. However, shortly afterward, he realized not only that his solution was incorrect but also that his "mistake" revealed the impossibility of solving the equations that modeled the motion of the three bodies. What Poincaré discovered was the existence of a point in phase space whose orbit connects an equilibrium point with itself, meaning a point whose orbit tends toward a certain equilibrium as time increases and also tends toward that same equilibrium as time moves backward towards the past, as shown in Figure 1. Poincaré named these states **homoclinic points**.

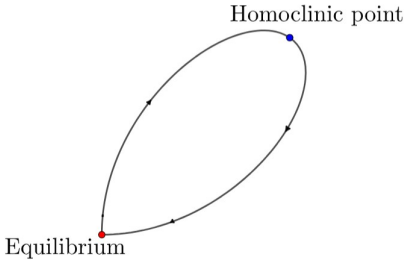


Figure 1

In an arbitrary dynamical system, the existence of homoclinic points implies chaotic behavior.

Thus, the **horseshoe function** was conceived by the American mathematician **Steve Smale**, as he himself mentions in [13], while strolling along the beaches of Rio de Janeiro, Brazil. It emerged as an attempt to provide a simple explanation for the intricate dynamics exhibited by dynamical systems that possess homoclinic points.

Smale presents the horseshoe function as follows:

Initially, we consider the unit square $ABCD$, which is compressed along the vertical axis by a factor of λ, where $0 < \lambda < 1/2$, while simultaneously expanding along the horizontal axis by a factor of $1/\lambda$, resulting in a rectangle $A^*B^*C^*D^*$. To introduce nonlinearity to

the system, the rectangle $A^*B^*C^*D^*$ is folded onto itself over the square $ABCD$, yielding a horseshoe-shaped geometric figure as depicted in Figure 2.

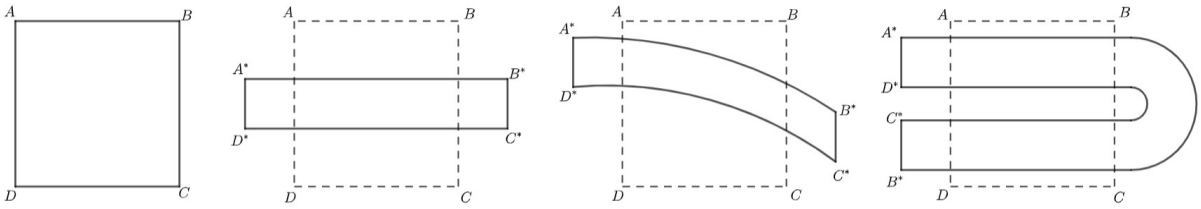


Figure 2

Thus, the dynamics of the function are described by taking an arbitrary point on the square $ABCD$ and moving it to its image on the horseshoe. For instance, corner A is moved to point A^* in one unit of time. Continuing in this manner, the movement of an arbitrary point x in the square is described as an infinite sequence of points $\{x_0, x_1, x_2, ...\}$ called an **orbit**, where $x_0 = x$, x_1 is the state after one unit of time, x_2 is the state after two units of time, and so on. Smale proves that any dynamical system possessing a homoclinic point also contains a horseshoe, as shown in Figure 3. Thus, the horseshoe function is established as the simplest and fundamental tool for defining chaos in mathematics.

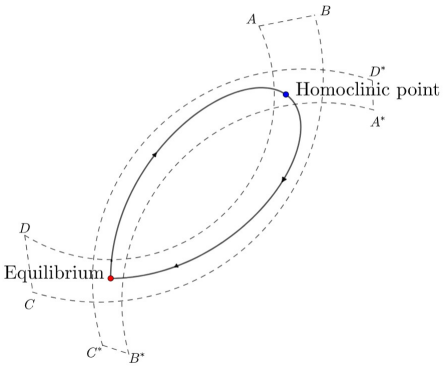


Figure 3

The main goal of this work is to apply the known theory on **discrete dynamical systems, hyperbolic geometry, Riemann surfaces, pseudo-Anosov automorphisms** and **ergodic theory** to the two maps presented in [4].

We will study a slightly different version of the horseshoe function presented by André de Carvalho, as stated in [4], which allows us to generalize the concept to the **theory of surface automorphisms**. That is, the successive iteration of the horseshoe function, denoted by h, on the square $ABCD$, induces a series of identifications on its boundary such that the resulting quotient space is homeomorphic to a Riemann surface denoted as X, and the function h defines an automorphism, H, on the surface X. The construction of this automorphism, along with the surface it defines and its topological properties, is presented in detail in Chapter 2. Additionally, the connections with another similar type of function known as the Baker function are also studied.

Chapter 1 of this work introduces the necessary basic concepts on discrete dynamical systems, including the definition provided in [6] for chaos. Subsequently, an introduction

to symbolic dynamics is given, which is employed in Chapter 3 to understand the behavior of the orbits of each point on the surface X under the iteration of the automorphism H. Here, the concepts of **sequence space** and **shift function**, along with their main dynamic properties, are introduced. These classic results have been largely taken from the books: "*An introduction to chaotic dynamical systems*" [6] and "*Dynamical systems: stability, symbolic dynamics and chaos*" [12].

Chapter 1 also contains a brief summary to **hyperbolic geometry**, where the hyperbolic metric defined on the upper half-plane, $\mathbb{H}$, is introduced, as well as the concept of a **Fuchsian group** as a discrete subgroup of the orientation-preserving isometries of $\mathbb{H}$, denoted as $PSL(2,\mathbb{R})$. The results of this section were taken from the book "*Fuchsian groups*" [9]. This directly connects with the theory of Riemann surfaces and the **uniformization theorem for Riemann surfaces**, which will allow us to study the structure of the surface X in Chapter 2. Thus, the classic results from the theory of Riemann surfaces and extremal length were taken from the book "*Quasiconformal Teichmüller theory*" [7].

Chapter 1

Preliminaries

In this chapter, we introduce the notion of a **discrete dynamical system** as the iteration of a function f that sends a set to itself, and we present some concepts and results from **symbolic dynamics**, an essential tool for understanding the evolution of a discrete dynamical system. Symbolic dynamics allows us to encode the "trajectory" of a point and study how it evolves under each iteration of f.

The central concept of symbolic dynamics is the **space of sequences**, denoted by Σ. Each element of Σ is an infinite sequence of 0's and 1's in both directions, and it is mainly used to simplify the study of the complex behavior that certain types of dynamical systems, called **chaotic**, can exhibit. In Chapter 3, we will use the tools presented in this section in the context of **differential topology** and the **theory of surface automorphisms** to exhibit the chaotic behavior of a certain automorphism that we will call the **horseshoe map**.

On the other hand, in this chapter, we also give a brief overview of the theory of **Riemann surfaces** and one of its central theorems, the **uniformization theorem for Riemann surfaces**, which allows us to classify all Riemann surfaces into three types: flat, elliptic, and hyperbolic. This concept will then be essential for the second chapter.

1.1 Discrete Dynamical Systems

In this section, we introduce the basic concepts of discrete dynamics. The definitions and results are mostly taken from the book [6] and included in this section for the sake of completeness.

Let (X, d) be a metric space and $f : X \longrightarrow X$ an automorphism of X. A **discrete dynamical system** is a sequence of elements $..., x_{i-2}, x_{i-1}, x_i, x_{i+1}, x_{i+2}, ...$ of X such that $x_{i+1} = f(x_i)$ for all $i \in \mathbb{Z}$. In this way, for any $x \in X$ the **orbit** of x is the set:

$$\mathcal{O}(x) := \{f^n(x) \ / \ n \in \mathbb{Z}\}.$$

Given any element $x \in X$ and a number $n \in \mathbb{N}$, we say that x is a **periodic point** of period n or n **- periodic point**, if it satisfies the statements:

1. $f^j(x) \neq x$ for all $j = 1, 2, ..., n-1$.
2. $f^{m_1}(x) = f^{m_2}(x)$ for all $m_1, m_2 \in \mathbb{Z}$ such that $m_1 \equiv m_2 \ \mathrm{mod}(n)$.

We denote the set of all periodic points of period n as $Per(f, n)$.

If $x \in X$ is a periodic point of period 1 we say that x is a **fixed point**. We denote the set of all fixed points of f as $Fix(f)$.

Chaos is one of the most interesting property that a dynamical system can have and because of the complexity that this property shows, there exists different ways to try to catch its essence in a definition. In order to describe the chaos, we will give some preliminary notions about the topology of metric spaces.

Consider a subset D of a metric space (X, d) and a function $f : D \longrightarrow D$. If for any pair $x, y \in D$ and any $\epsilon > 0$, there exists $z \in D$ and $n \in \mathbb{N}$ such that $d(z, x) < \epsilon$ and $d(f^n(z), y) < \epsilon$ then f is called **topologically transitive**.

The following proposition allows characterizing a topologically transitive function based on its periodic points.

Proposition 1.1.1 *Let D be a subset of a metric space X and let $f : D \longrightarrow D$ be a map. If the periodic points of f are dense in D and there exists a point $x \in D$ whose orbit is dense in D, then f is topologically transitive on D.*

Proof. Suppose that the periodic points of f are dense in D and let $x_0 \in D$ be a point such that for every open set $V \subseteq X$ intersecting D, there exists $k \in \mathbb{Z}$ such that $f^k(x_0) \in V$. Such $x_0 \in D$ exists, since the existence of points whose orbit is dense in D is guaranteed by hypothesis.

Consider the following observations:

1. Suppose that D is a finite set, as the orbit of x_0 is dense in D we have that $\mathcal{O}(x_0) = D$, and therefore, f is topologically transitive.
2. If D is an infinite set, then no finite subset of D can be dense in D, so no periodic point of D can have a dense orbit in D.

Let $x, y \in D$ be two arbitrary points and let $\epsilon > 0$ be given. We will show that there exists a point $z \in D$ and a number $n \in \mathbb{N}$ such that $d(z, x) < \epsilon$ and $d(f^n(z), y) < \epsilon$.

Suppose that D is infinite. Since the orbit $\mathcal{O}(x_0)$ is dense in D, there exists $k \in \mathbb{Z}$ such that $d(f^k(x_0), x) < \epsilon$. To prove the result, it suffices to show the existence of a number $m \in \mathbb{Z}$ greater than k such that $d(f^m(x_0), y) < \epsilon$. Thus, taking $z = f^k(x_0)$ and $n = m - k$, we have $d(z, x) < \epsilon$ and $f^n(z) = f^{m-k}(f^k(x_0)) = f^m(x_0)$, so that $d(f^n(z), y) < \epsilon$ and the result follows.

To prove the existence of such $m \in \mathbb{Z}$ it suffices to prove the existence of an infinite number of iterates of x_0, $f^i(x_0)$ with $i = 0, 1, 2, ...$ that satisfy $d(f^i(x_0), y) < \epsilon$, in this way, some i will be greater than k and the assertion will be proven. Reasoning by contradiction, suppose that there exists a finite number of iterates of x_0 that satisfy the inequality.

$$d(f^i(x_0), y) < \epsilon. \tag{1.1}$$

Let p be a periodic point of f such that $d(p, y) < \epsilon$, such p exists since periodic points of f are dense in D. By **observation 2.**, we have that p does not belong to the orbit of x_0, thus $d(f^i(x_0), p) > 0$ for any iterate $f^i(x_0)$ satisfying (1.1). As the number of iterates satisfying inequality 1.1 is finite, we can define $\delta = \min d(f^i(x_0), p) \ / \ d(f^i(x_0), y) < \epsilon$.

Then, the intersection of open balls $B(y,\epsilon) \cap B(p,\delta)$ is non-empty and is an open set in X contained in D, therefore $B(y,\epsilon) \cap B(p,\delta) \cap D \neq \emptyset$.

Now, since the orbit of x_0 is dense in D, there must be a number $j \in \mathbb{N}$ such that

$$f^j(x_0) \in B(p,\delta) \cap B(y,\epsilon).$$

Therefore, $d(f^j(x_0), y) < \epsilon$. However, it must be the case that $d(f^j(x_0), p) \geq \delta$, which implies that $f^j(x_0) \notin B(y,\epsilon) \cap B(p,\delta)$. This is a contradiction, and so the ball $B(y,\epsilon)$ contains infinitely many iterates of x_0.

□

Some physical phenomena that evolve over time, such as the motion of a double pendulum or the weather, are characterized by their sensitivity to initial conditions. This means that small changes in the initial parameters will result in large changes in their state over a long period, thus complicating our ability to make predictions about such phenomena. This sensitivity to initial conditions is the main characteristic of a chaotic dynamical system.

Now, let us consider a subset D of a metric space (X,d) and let $f : D \longrightarrow D$ be a function. If there is $\delta > 0$ such that for every $x \in D$ and any $\epsilon > 0$ there is $y \in D$ and $n \in \mathbb{N}$ such that $d(x,y) < \epsilon$ and $d(f^n(x), f^n(y)) > \delta$, then we say that f depends **sensitively on initial conditions**.

Furthermore, a function $f : D \longrightarrow D$ is **chaotic** if it satisfies the following three conditions:

1. The periodic points of f are dense in D.
2. f is topologically transitive.
3. f depends sensitively on initial conditions.

1.2 The Sequence Space Σ

A two-sided infinite sequence of 0's and 1's is a function $\mathbf{s} : \mathbb{Z} \longrightarrow \{0,1\}$. In a more visual way, $\mathbf{s}$ looks like an infinite tuple of 0's and 1's, for example:

$$\mathbf{s} = (\cdots 1,0,0,0,1|1,1,1,0,1,\cdots).$$

We use the symbol | to separate the -1-nth position from the 0-nth position of the sequence $\mathbf{s}$.

We define the **sequence space** Σ, as the set of all two-sided infinite sequence of 0's and 1's, i.e.,

$$\Sigma := \{(\cdots, s_{-2}, s_{-1}|s_0, s_1, s_2, \cdots) \ / \ s_i \in \{0,1\} \text{ for all } i \in \mathbb{Z}\}.$$

If $\mathbf{s} = (\cdots, s_{-2}, s_{-1}|s_0, s_1, s_2, \cdots)$ and $\mathbf{t} = (\cdots, t_{-2}, t_{-1}|t_0, t_1, t_2, \cdots)$ are elements of Σ, it is possible to define a distance between $\mathbf{s}$ and $\mathbf{t}$ given by:

$$d[\mathbf{s},\mathbf{t}] := \sum_{i\in\mathbb{Z}} \frac{|s_i - t_i|}{2^{|i|}}. \tag{1.2}$$

The function $d : \Sigma \times \Sigma \longrightarrow \mathbb{R}$ satisfies the metric definition hypothesis, i.e., for any $\mathbf{x}, \mathbf{t}, \mathbf{u} \in \Sigma$ we have:

1. $d[\mathbf{s}, \mathbf{t}] \geq 0$.
2. $d[\mathbf{s}, \mathbf{t}] = d[\mathbf{t}, \mathbf{s}]$.
3. $d[\mathbf{s}, \mathbf{t}] \leq d[\mathbf{s}, \mathbf{u}] + d[\mathbf{u}, \mathbf{t}]$.

Note that $0 \leq |s_i - t_i| \leq 1$ for every $i \in \mathbb{Z}$. Hence, the metric d satisfies:

$$0 \leq d[\mathbf{s}, \mathbf{t}] \leq \sum_{i \in \mathbb{Z}} \frac{1}{2^{|i|}} = 3 \text{ for any } \mathbf{s}, \mathbf{t} \text{ in } \Sigma.$$

The above means that the greatest separation between any two elements of Σ is 3 and it is achieved when $s_i \neq t_i$, for all i.

It is possible to visualize the sequence space considering a square of side length 1 divided into two parts as depicted in Figure 1.1

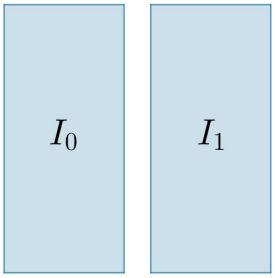


Figure 1.1: Sequence Space

The set I_0 contains the sequences $\mathbf{s} = (\cdots, s_{-2}, s_{-1} | s_0, s_1, \cdots)$ such that $s_0 = 0$ and I_1 contains those sequences such that $s_0 = 1$. Following in this way, the Figure 1.2a shows the sets:

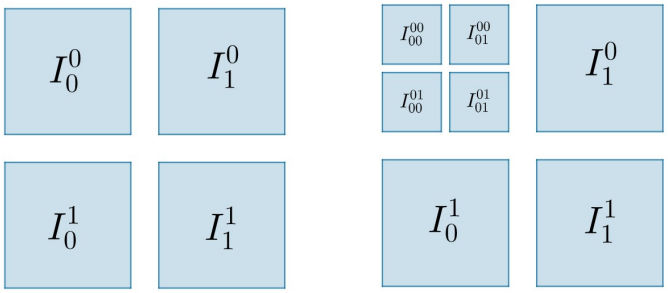
$$I_0^0 = \{\mathbf{s} \in \Sigma \ / \ s_{-1} = 0 \ ; \ s_0 = 0\} \qquad I_0^1 = \{\mathbf{s} \in \Sigma \ / \ s_{-1} = 1 \ ; \ s_0 = 0\}.$$

$$I_1^0 = \{\mathbf{s} \in \Sigma \ / \ s_{-1} = 0 \ ; \ s_0 = 1\} \qquad I_1^1 = \{\mathbf{s} \in \Sigma \ / \ s_{-1} = 1 \ ; \ s_0 = 1\}.$$

Which divides I_0 and I_1. Furthermore, Figure 1.2b shows the sets:

$$I_{00}^{00} = \left\{ \mathbf{s} \in \Sigma \ \middle/ \ \begin{matrix} s_{-1} = 0 & ; & s_{-2} = 0 \\ s_0 = 0 & ; & s_1 = 0 \end{matrix} \right\} \qquad I_{01}^{00} = \left\{ \mathbf{s} \in \Sigma \ \middle/ \ \begin{matrix} s_{-1} = 0 & ; & s_{-2} = 0 \\ s_0 = 0 & ; & s_1 = 1 \end{matrix} \right\}.$$

$$I_{00}^{01} = \left\{ \mathbf{s} \in \Sigma \ \middle/ \ \begin{matrix} s_{-1} = 0 & ; & s_{-2} = 1 \\ s_0 = 0 & ; & s_1 = 0 \end{matrix} \right\} \qquad I_{01}^{01} = \left\{ \mathbf{s} \in \Sigma \ \middle/ \ \begin{matrix} s_{-1} = 0 & ; & s_{-2} = 1 \\ s_0 = 0 & ; & s_1 = 1 \end{matrix} \right\}.$$



(a) (b)

Figure 1.2: Sequence Space

In a more general sense, we define the following set:

$$I_{s_0,s_1,s_2,\ldots,s_n}^{s_{-1},s_{-2},s_{-3},\ldots,s_{-n}} = \left\{ \mathbf{t} \in \Sigma \;\middle/\; \begin{array}{cccc} t_{-n} = s_{-n}, & \cdots, & t_{-2} = s_{-2}, & t_{-1} = s_{-1}, \quad t_0 = s_0, \\ t_1 = s_1, & t_2 = s_2, & \cdots, & t_n = s_n \end{array} \right\}.$$

Now, let $n \in \mathbb{N} \cup \{0\}$ and let $\mathbf{x}, \mathbf{y} \in I_{s_0,s_1,s_2,\ldots,s_n}^{s_{-1},s_{-2},s_{-3},\ldots,s_{-n}}$ be two sequences in Σ. Then, we have $x_i = y_i = s_i$ for all $i \in \mathbb{Z}$ such that $|i| \leq n$, and therefore, $|x_i - y_i| = 0$ for all $|i| \leq n$. Calculating the distance between $\mathbf{x}$ and $\mathbf{y}$ we obtain:

$$d[\mathbf{x}, \mathbf{y}] = \sum_{i \in \mathbb{Z}} \frac{|x_i - y_i|}{2^{|i|}} = \sum_{i=n+1}^{\infty} \frac{|x_{-i} - y_{-i}|}{2^i} + \sum_{i=n+1}^{\infty} \frac{|x_i - y_i|}{2^i} \leq 2 \sum_{i=n+1}^{\infty} \frac{1}{2^i} = \frac{1}{2^{n-1}}.$$

On the other hand, if there exists $j \in \mathbb{Z}$ such that $|j| \leq n$ for which $x_j \neq y_j$, then:

$$d[\mathbf{x}, \mathbf{y}] = \sum_{i \in \mathbb{Z}} \frac{|x_i - y_i|}{2^{|i|}} \geq \frac{1}{2^{|j|}} \geq \frac{1}{2^n}.$$

So, if $d[\mathbf{x}, \mathbf{y}] < \dfrac{1}{2^n}$ then $x_j = y_j$ for all $j \in \mathbb{Z}$ such that $|j| \leq n$. This proves the following theorem:

Theorem 1.2.1 *Let $n \in \mathbb{N} \cup \{0\}$ and $\boldsymbol{x}, \boldsymbol{y} \in \Sigma$. If $\boldsymbol{x}, \boldsymbol{y} \in I_{s_0,s_1,\ldots,s_n}^{s_{-1},s_{-2},\ldots,s_{-n}}$, then:*

$$d[\boldsymbol{x}, \boldsymbol{y}] \leq \frac{1}{2^{n-1}}.$$

On the other hand, if $d[\boldsymbol{x}, \boldsymbol{y}] < \dfrac{1}{2^n}$ then $\boldsymbol{x}, \boldsymbol{y} \in I_{s_0,s_1,\ldots,s_n}^{s_{-1},s_{-2},\ldots,s_{-n}}$.

1.2.1 The Shift Map

We will now define the **shift map** and study some of its most interesting properties. Let $\mathbf{x} = (\ldots, x_{-2}, x_{-1} | x_0, x_1, x_2, \ldots)$ be an element of the sequence space Σ. The shift map $\sigma : \Sigma \longrightarrow \Sigma$ takes each entry of the sequence $\mathbf{x}$ and shifts it to the left, that is:

$$\sigma(\ldots, x_{-2}, x_{-1} | x_0, x_1, x_2, \ldots) = (\ldots, x_{-1}, x_0 | x_1, x_2, x_3, \ldots).$$

Let $\epsilon > 0$ be given and let n be a natural number such that $\dfrac{1}{2^{n-1}} < \epsilon$. For any pair $\mathbf{x}, \mathbf{y} \in \Sigma$ take $\delta = \dfrac{1}{2^{n+1}}$. If $d[\mathbf{x}, \mathbf{y}] < \delta$ then by the theorem 1.2.1 we have that:

$$\mathbf{x}, \mathbf{y} \in I^{s_{-1},s_{-2},...,s_{-n-1}}_{s_0,s_1,...,s_{n+1}}.$$

Therefore $\sigma(\mathbf{x}), \sigma(\mathbf{y}) \in I^{s_0,s_{-1},...,s_{-n}}_{s_1,s_2,...,s_n}$. Again, the theorem 1.2.1 implies that:

$$d[\sigma(\mathbf{x}), \sigma(\mathbf{y})] \leq \frac{1}{2^{n-1}} < \epsilon.$$

The above means that the shift map $\sigma : \Sigma \longrightarrow \Sigma$ is continuous.

Theorem 1.2.2 *The shift map σ is chaotic on Σ.*

Proof. We must prove that σ satisfies the following three conditions:

1. The periodic points of σ are dense in Σ.
2. σ is topologically transitive.
3. σ depends sensitively on initial conditions.

(1): A k-periodic point of σ is a sequence of 0's and 1's in which the digits $x_0, x_1, ..., x_{k-1}$ repeat indefinitely in both directions:

$$\mathbf{x} = (..., x_0, ...x_{k-1}, x_0, ..., x_{k-1}|x_0, ..., x_{k-1}, x_0, ..., x_{k-1}, ...).$$

Let $\epsilon > 0$ be given and let $n \in \mathbb{N}$ such that $\frac{1}{2^{n-1}} < \epsilon$, take $\mathbf{y} \in I^{s_{-1},...,s_{-n}}_{s_0,...,s_n} \subset \Sigma$. Then the sequence

$$\mathbf{x} = (..., s_{-n}, ..., s_n, s_{-n}, ..., s_n, s_{-n}, ..., s_{-1}|s_0, ..., s_n, s_{-n}, ..., s_n, ..., s_{-n}, ..., s_n, ...).$$

in which the digits $s_{-n}, ..., s_{-1}, s_0, ..., s_n$ are repeated periodically, is a periodic point of Σ contained in $I^{s_{-1},...,s_{-n}}_{s_0,...,s_n}$. Hence, by the theorem 1.2.1 we have:

$$d[\mathbf{x}, \mathbf{y}] \leq \frac{1}{2^{n-1}} < \epsilon.$$

This proves that the periodic points of σ are dense in Σ.

To prove **(2)** it is enough to show that there exists $\mathbf{x} \in \Sigma$ whose orbit is dense in Σ. The above together with the statement **(1)** and the proposition 1.1.1 proves that σ is topologically transitive.

Let $\mathbf{y} \in I^{s_{-1},...,s_{-n}}_{s_0,...,s_n}$ be an arbitrary sequence in Σ and let $\mathbf{x} \in \Sigma$ be a sequence which consists of all possible blocks of 0's and 1's of length 1, followed by all such blocks of length 2, then length 3 and so forth in both directions, for example:

$$\mathbf{x} = (...010\ \ 100\ \ 000\ \ 11\ \ 01\ \ 10\ \ 00\ \ 1\ \ 0|0\ \ 1\ \ 00\ \ 01\ \ 10\ \ 11\ \ 000\ \ 001\ \ 010...).$$

Therefore, far to the right of the expression for $\mathbf{x}$, there is a block of length $2n+1$ that consists of the digits $s_{-n}, ..., s_{-1}, s_0, ..., s_n$ and therefore there exists $m \in \mathbb{N}$ such that $\sigma^m(\mathbf{x}) \in I^{s_{-1},...,s_{-n}}_{s_0,...,s_n}$. This proves that the orbit of $\mathbf{x}$ is dense in Σ and therefore σ is topologically transitive.

(3): Let $\mathbf{y} = (..., y_{-2}, y_{-1}|y_0, y_1, y_2, ...) \in \Sigma$ be a sequence and let $\epsilon > 0$ be given. Take $N \in \mathbb{N}$ such that $\frac{1}{2^N} < \epsilon$ and $\mathbf{x} = (..., x_{-2}, x_{-1}|x_0.x_1, x_2, ...) \in \Sigma$ that satisfies the following properties:

I $x_i = y_i$ for all $i \in \mathbb{Z}$ such that $|i| \leq N+1$.

II $x_i \neq y_i$ for all $i \in \mathbb{Z}$ such that $|i| > N+1$.

Then by the theorem 1.2.1 we have that $d[\mathbf{x},\mathbf{y}] \leq \frac{1}{2^N} < \epsilon$.

Furthermore, if $n > N+1$, by the property II, the j-nth entry of $\sigma^n(\mathbf{x})$ differs from j-nth entry of $\sigma^n(\mathbf{y})$ for $j \in \{0,1,2,3,...\}$, i.e., $x_{j+n} \neq y_{j+n}$. Hence, for all $j \in \{0,1,2,3,...\}$ we have that $|x_{j+n} - y_{j+n}| = 1$, therefore:

$$d[\sigma^n(\mathbf{x}),\sigma^n(\mathbf{y})] = \sum_{j\in\mathbb{Z}} \frac{|x_{j+n}-y_{j+n}|}{2^{|j|}} \geq \sum_{j=0}^{\infty} \frac{|x_{j+n}-y_{j+n}|}{2^j} = \sum_{j=0}^{\infty} \frac{1}{2^j} = 2.$$

Then, $\mathbf{x}$ satisfies that $d[\mathbf{x},\mathbf{y}] < \epsilon$ and $d[\sigma^n(\mathbf{x}),\sigma^n(\mathbf{y})] \geq 2$ for all $n > N+1$. This proves that σ depends sensitively on initial conditions, and therefore, σ is chaotic on Σ.

□

1.3 Fuchsian Groups

The main goal of this section is to study the geometry of the hyperbolic model or Lobachevsky plane, which is formed by the **upper half-plane** endowed with the metric:

$$ds = \frac{\sqrt{dx^2+dy^2}}{y}. \tag{1.3}$$

We will introduce particular subgroups of the isometries of the upper half-plane, known as **Fuchsian groups**.

First, let us see some basic notions of hyperbolic geometry.

1.3.1 A Brief Look at Hyperbolic Geometry

Let $\mathbb{C}$ be the complex plane. The Lobachevsky plane is formed by the upper half-plane:

$$\mathcal{H} = \{z \in \mathbb{C} \ / \ \mathrm{Im}(z) > 0\},$$

equipped with the metric 1.3, known as the **hyperbolic metric**. Let us consider two points $z, w \in \mathcal{H}$. To measure the distance between z and w given by the hyperbolic metric, we consider a differentiable curve defined by:

$$\begin{array}{rccc} \gamma: & [0,1] & \longrightarrow & \mathcal{H} \\ & t & \longmapsto & \gamma(t) = x(t) + iy(t)' \end{array}$$

so that $\gamma(0) = z$ and $\gamma(1) = w$. Then, the **hyperbolic length** of γ is given by:

$$h(\gamma) = \int_0^1 \frac{\sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}}{y(t)} dt = \int_0^1 \frac{|\frac{dz}{dt}|}{y} dt.$$

Thus, the **hyperbolic distance** $\rho(z,w)$ between z and w is defined as the minimum length of all possible paths γ that connect z and w, that is:

$$\rho(z,w) := \inf_{\gamma}\{h(\gamma)\}. \tag{1.4}$$

Recall that a Möbius transformation is a function from $\mathbb{C}$ to $\mathbb{C}$ given by:

$$\begin{array}{rccc} T: & \mathbb{C} & \longrightarrow & \mathbb{C} \\ & z & \longmapsto & \dfrac{az+b}{cz+d} \end{array} \text{ where } a,b,c,d \in \mathbb{C} \text{ and } ad-bc \neq 0.$$

If we consider a differentiable curve $\gamma : [0,1] \longrightarrow \mathcal{H}$ given by $z(t) = x(t) + iy(t)$ and a Möbius transformation T, then $T(\gamma)$ is given by:

$$w(t) = \frac{az(t)+b}{cz(t)+d} = u(t) + iv(t).$$

Then, by calculating the length of $T(\gamma)$, we have:

$$h(T(\gamma)) = \int_0^1 \frac{|\frac{dw}{dt}|}{v} dt = \int_0^1 \frac{|\frac{dw}{dz}||\frac{dz}{dt}|}{v} dt.$$

Now, $\left|\dfrac{dw}{dz}\right| = \left|\dfrac{ad-bc}{(cz+d)^2}\right|$ and $v(t) = \mathrm{Im}(w(t)) = \dfrac{w-\overline{w}}{2i}$, thus:

$$v(t) = \frac{\dfrac{az+b}{cz+d} - \dfrac{\overline{a}\overline{z}+\overline{b}}{\overline{c}\overline{z}+\overline{d}}}{2i} = \frac{(a\overline{c}-\overline{a}c)|z|^2 + (a\overline{d}-\overline{b}c)z - (\overline{a}d-b\overline{c})\overline{z} + b\overline{d} - \overline{b}d}{2i|cz+d|^2}.$$

So, if $a,b,c,d \in \mathbb{R}$ and $ad-bc=1$, we have that:

$$v(t) = \frac{z-\overline{z}}{2i|cz+d|^2} = \frac{\mathrm{Im}(z)}{|cz+d|^2} \text{ and therefore } \left|\frac{dw}{dz}\right| = \frac{\mathrm{Im}(w)}{\mathrm{Im}(z)} = \frac{v}{y}.$$

Then, the length of $T(\gamma)$ is given by:

$$h(T(\gamma)) = \int_0^1 \frac{\frac{v}{y}|\frac{dz}{dt}|}{v} dt = \int_0^1 \frac{|\frac{dz}{dt}|}{y} dt = h(\gamma).$$

Thus, if $T(z) = \dfrac{az+b}{cz+d}$ is a Möbius transformation, then T is an isometry of the upper half-plane $\mathcal{H}$ if and only if $a,b,c,d \in \mathbb{R}$ and $ad-bc=1$.

If $T(z) = \dfrac{az+b}{cz+d}$ is a Möbius transformation that preserves the hyperbolic length, then T can be identified with a matrix of the **special linear group** $SL(2,\mathbb{R})$.

$$SL(2,\mathbb{R}) = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \Big/ a,b,c,d \in \mathbb{R} \text{ and } ad-bc=1 \right\}.$$

Note that $M_1 = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $M_2 = \begin{bmatrix} -a & -b \\ -c & -d \end{bmatrix}$ are two matrices in $SL(2,\mathbb{R})$, both corresponding to the same Möbius transformation. Hence, by identifying M_1 with M_2, we obtain that the group of Möbius transformations that preserve the hyperbolic distance under the operation of function composition is isomorphic to the **projective special linear group**, denoted by:

$$PSL(2,\mathbb{R}) = SL(2,\mathbb{R}) \Big/ \{\pm \mathrm{id}\}.$$

Geodesics: The **geodesic curves** in $\mathcal{H}$ are the shortest trajectories between two points

in $\mathcal{H}$ with respect to the hyperbolic metric. Any pair of points in $\mathcal{H}$ can be joined by a unique geodesic, and the distance between those points is precisely the length of the geodesic that connects them.

Let $z_1 = ia$ and $z_2 = ib$ be two points in $\mathcal{H}$ with $b > a$. If $\gamma : [0,1] \longrightarrow \mathcal{H}$ is any differentiable path that connects z_1 and z_2, with $\gamma(t) = x(t) + iy(t)$, then its length is given by:

$$h(\gamma) = \int_0^1 \frac{|\frac{dz}{dt}|}{y(t)} dt \geq \int_0^1 \frac{|\frac{dy}{dt}|}{y(t)} = \int_a^b \frac{dy}{y} = \log\left(\frac{b}{a}\right).$$

As $\log\left(\frac{b}{a}\right)$ is the hyperbolic length of the line segment on the imaginary axis joining z_1 and z_2, then this segment is the geodesic that connects z_1 and z_2.

Now, for arbitrary $z_1, z_2 \in \mathcal{H}$, if z_1 and z_2 have different real parts and $\mathbf{S}$ is the unique Euclidean semicircle orthogonal to the real axis with center at $\alpha \in \mathbb{R}$ and radius r that passes through z_1 and z_2, then the transformation:

$$T(z) = \frac{z - \alpha - r}{z - \alpha + r} \in PSL(2, \mathbb{R})$$

maps $\mathbf{S}$ onto the imaginary axis. Since the elements of $PSL(2, \mathbb{R})$ are isometries, then T preserves the hyperbolic distance between z_1 and z_2. Therefore, assuming without loss of generality that $\mathrm{Im}(T(z_2)) > \mathrm{Im}(T(z_1))$, we have that:

$$\rho(z_1, z_2) = \rho(T(z_1), T(z_2)) = \log\left(\frac{\mathrm{Im}(T(z_2))}{\mathrm{Im}(T(z_1))}\right).$$

The section of the semicircle $\mathbf{S}$ that joins z_1 and z_2 is precisely the geodesic that joins both points.

Analogously, if z_1 and z_2 have the same real part, let $\mathbf{L}$ be the line orthogonal to the real axis that joins z_1 and z_2, and assume that $\mathbf{L}$ intersects the real axis at $\alpha \in \mathbb{R}$. Then the transformation:

$$S(z) = z - \alpha \in PSL(2, \mathbb{R})$$

maps $\mathbf{L}$ onto the imaginary axis. By the same reasoning as before, we conclude that the geodesic joining z_1 and z_2 is precisely the segment of $\mathbf{L}$ joining them. In other words:

*The geodesics in $\mathcal{H}$ are **semicircles** and **straight lines**, both orthogonal to the real axis.*

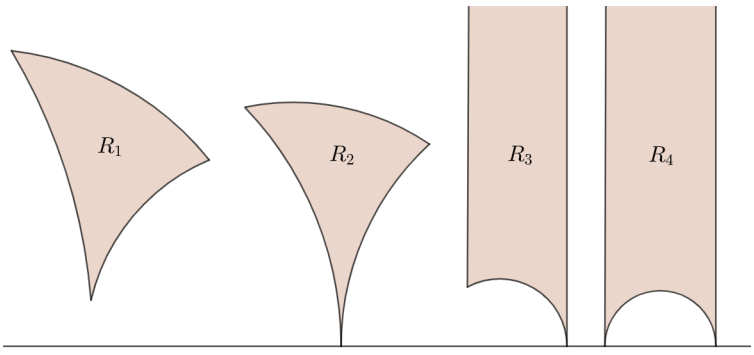


Figure 1.3: Hyperbolic Triangles

The figure 1.3 shows **hyperbolic triangles** whose sides are geodesics in $\mathcal{H}$.

If z and w are two distinct points in $\mathcal{H}$, then we denote by $[z, w]$ the geodesic segment that connects these points, and we have that:

$$\rho(z, w) = \rho(z, \beta) + \rho(\beta, w) \text{ if and only if } \beta \in [z, w].$$

Therefore, if $\beta \in [z, w]$ and T is a transformation in $PSL(2, \mathbb{R})$, since T preserves hyperbolic distance, it follows that $T(\beta) \in [T(z), T(w)]$, that is, T maps $[z, w]$ to $[T(z), T(w)]$, so every element of $PSL(2, \mathbb{R})$ sends geodesics to geodesics.

Hyperbolic Area: Let A be a subset of $\mathcal{H}$. We define the **hyperbolic area** of the set A, by

$$\mu(A) = \int_A \frac{dx\, dy}{y^2}.$$

in the case where this integral exists.

Just as with hyperbolic distance, it is possible to prove that the elements of $PSL(2, \mathbb{R})$ preserve hyperbolic area, in other words:

$$\textit{If } A \subseteq \mathcal{H},\ \mu(A) \textit{ exists, and } T \in PSL(2, \mathbb{R}),\textit{ then } \mu(T(A)) = \mu(A).$$

A **hyperbolic polygon with** n **sides** is a closed set in $\mathcal{H} \cup \mathbb{R} \cup \{\infty\}$ bounded by n hyperbolic geodesic segments. In the case of hyperbolic triangles, the **Gauss-Bonnet formula** shows that the hyperbolic area depends only on its angles.

Theorem 1.3.1 (Gauss-Bonnet) *Let Δ be a hyperbolic triangle with internal angles α, β, γ. Then:*

$$\mu(\Delta) = \pi - \alpha - \beta - \gamma.$$

See the proof in [9], theorem 1.4.2.

1.3.2 Discrete Subgroups of $PSL(2, \mathbb{R})$

It is possible to endow the group $PSL(2, \mathbb{R})$ with a topology as follows. Let $X \subseteq \mathbb{R}^4$ be a set defined as follows:

$$X = \{(a, b, c, d) \in \mathbb{R}^4 \ / \ ad - bc = 1\}.$$

The set X inherits the usual topology of $\mathbb{R}^4$, and each transformation $T \in SL(2, \mathbb{R})$ can be identified with an element of X through the map:

$$\begin{array}{rccc} \psi: & X & \longrightarrow & SL(2, \mathbb{R}) \\ & (a, b, c, d) & \longmapsto & \begin{bmatrix} a & b \\ c & d \end{bmatrix}. \end{array}$$

As a topological space, $SL(2, \mathbb{R})$ can be identified with the set X, and therefore, the projective special linear group $\mathrm{PSL}(2, \mathbb{R})$ can be identified with the quotient space $X \big/ \{-\mathrm{id}, \mathrm{id}\}$, where id is the identity map in X. Therefore, the topology of $\mathrm{PSL}(2, \mathbb{R})$ is the quotient topology.

We can define a **norm** on $PSL(2,\mathbb{R})$ as follows:

$$\|T\| = \sqrt{a^2+b^2+c^2+d^2}.$$

Note that $\|T\|$ is a well-defined function and therefore, $PSL(2,\mathbb{R})$ is a topological group with respect to the metric $\|T-S\|$.

A subgroup Γ of $PSL(2,\mathbb{R})$ is called a **Fuchsian group** if the induced topology on Γ is a discrete topology, i.e., a Fuchsian group is precisely a discrete subgroup of $PSL(2,\mathbb{R})$.

The following lemma will be useful to prove the main property of Fuchsian groups.

Lemma 1.3.1 *Let $z_0 \in \mathcal{H}$ and let K be a compact subset of $\mathcal{H}$. Then the set:*

$$E = \{T \in PSL(2,\mathbb{R}) \ / \ T(z_0) \in K\}.$$

is compact.

Proof. Since we can identify $SL(2,\mathbb{R})$ with the set $X = \{(a,b,c,d) \in \mathbb{R}^4 \ / \ ad-bc=1\}$ and $PSL(2,\mathbb{R})$ with the quotient space $X\big/\{-id, id\}$, we have a continuous map.

$$\begin{array}{rccc} \varphi: & X & \longrightarrow & PSL(2,\mathbb{R}) \\ & (a,b,c,d) & \longmapsto & T(z) = \dfrac{az+b}{cz+d}. \end{array}$$

If we show that $E_1 = \left\{(a,b,c,d) \in X \ \middle/ \ \dfrac{az_0+b}{cz_0+d} \in K\right\}$ is compact, then it follows that $E = \varphi(E_1)$ is compact. To see this, let us show that E_1 is a closed and bounded set as a subset of $\mathbb{R}^4$.

Let $\beta : X \longrightarrow \mathcal{H}$ be the function defined by

$$\begin{array}{rccc} \beta: & X & \longrightarrow & \mathcal{H} \\ & A=(a,b,c,d) & \longmapsto & \varphi(A)\big|_{z_0}. \end{array}$$

Since φ is continuous and every Möbius transformation is continuous, then β is also continuous. Since K is compact, it is closed and bounded, therefore $E_1 = \beta^{-1}(K)$ is closed and there exists $M > 0$ such that:

$$\left|\frac{az_0+b}{cz_0+d}\right| < M \text{ for all } (a,b,c,d) \in E_1. \tag{1.5}$$

Also, due to the compactness of K, there exists $L > 0$ such that $\operatorname{Im}\left(\dfrac{az_0+b}{cz_0+d}\right) > L$, otherwise K would not be bounded in $\mathcal{H}$. Now, since $\operatorname{Im}\left(\dfrac{az_0+b}{cz_0+d}\right) = \dfrac{\operatorname{Im}(z_0)}{|cz_0+d|^2}$, we have that:

$$\frac{\operatorname{Im}(z_0)}{|cz_0+d|^2} > L \Longrightarrow |cz_0+d| < \sqrt{\frac{\operatorname{Im}(z_0)}{L}}. \tag{1.6}$$

Combining 1.5 and 1.6, we obtain that $|az_0+b| < M\sqrt{\dfrac{\operatorname{Im}(z_0)}{L}}$. This shows that $a, b, c,$ and d are bounded, and therefore E_1 is bounded and $E = \varphi(E_1)$ is compact.

□

Let Γ be a Fuchsian subgroup of $PSL(2,\mathbb{R})$. Then, Γ is a discrete subset of $PSL(2,\mathbb{R})$, and by Lemma 1.3.1, we know that $\{T \in PSL(2,\mathbb{R}) \ / \ T(z) \in K\}$ is a compact set for any $z \in \mathcal{H}$ and any compact subset $K \subset \mathcal{H}$. Therefore, the set $\mathcal{O} = \{T \in \Gamma \ / \ T(z) \in K\}$ is the intersection of a compact set and a discrete set, and this implies that $\mathcal{O}$ is a finite set. That is, the intersection of the orbits of a Fuchsian group with a compact subset of $\mathcal{H}$ is a finite set.

If the action of a group $\Gamma \subseteq PSL(2,\mathbb{R})$ satisfies the property described above, then we say that the group acts **properly discontinuously** on $\mathcal{H}$.

Let us assume that Γ is a subgroup of $PSL(2,\mathbb{R})$ that acts properly discontinuously on $\mathcal{H}$. If $p \in \mathcal{H}$ is a fixed point of some transformation $T \in \Gamma$ other than the identity, then there must exist a neighborhood V of p such that no other point of V is fixed by an element of Γ other than the identity.

To see this, let us reason by contradiction. If there were fixed points of elements of Γ in every neighborhood of p, then there should exist a sequence $\{p_n\}$ of points in $\mathcal{H}$ and a sequence $\{T_n\}$ of transformations in Γ such that $p_n \longrightarrow p$ and $T_n(p_n) = p_n$ for all $n \in \mathbb{N}$. Now consider the hyperbolic disk:

$$\overline{B_\epsilon(p)} = \{z \in \mathcal{H} \ / \ \rho(z,p) \leq \epsilon\}.$$

Since the topology induced by the hyperbolic metric is the same as the usual topology, we have that $\overline{B_\epsilon(p)}$ is compact and Γ acts properly discontinuously on $\mathcal{H}$, then the set

$$\{T \in \Gamma \ / \ T(p) \in \overline{B_\epsilon(p)}\}$$

is a finite set. Therefore, there exists $N \in \mathbb{N}$ such that if $n > N$, then $T_n(p) \notin \overline{B_\epsilon(p)}$ while $p_n \in \overline{B_{\epsilon/2}(p)}$. By the triangle inequality, we have:

$$\rho(T_n(p), p) \leq \rho(T_n(p), T_n(p_n)) + \rho(T_n(p_n), p).$$

Furthermore, since $T_n(p_n) = p_n$ and the hyperbolic length is invariant under any transformation T in Γ, then:

$$\rho(T_n(p), T_n(p_n)) = \rho(p, p_n) \text{ and } \rho(T_n(p_n), p) = \rho(p_n, p).$$

Therefore,

$$\rho(T_n(p), p) \leq \rho(p, p_n) + \rho(p_n, p) < \epsilon/2 + \epsilon/2 = \epsilon.$$

However, this is a contradiction and therefore there must exist a neighborhood V of p such that $T(V) \cap V = \{p\}$ for every $T \in \Gamma \setminus \{id\}$.

The above implies the existence of points $s \in \mathcal{H}$ that are not fixed by any element other than the identity in Γ. Thus, if Γ were a non-discrete subgroup of $PSL(2,\mathbb{R})$ that also acts properly discontinuously on $\mathcal{H}$, then there should exist a sequence $\{T_n\}$ of elements in Γ such that $T_n(s) \longrightarrow id(s) = s$, where $s \in \mathcal{H}$ is not fixed by any element $T \in \Gamma \setminus \{id\}$.

That is, $\{T_n(s)\}$ is a sequence of points distinct from s that converges to s, therefore the intersection of every closed hyperbolic disk centered at s with the orbit of Γ on s is infinite, which contradicts that Γ acts properly discontinuously on $\mathcal{H}$, and thus Γ must be discrete.

In summary, $\Gamma \subset PSL(2,\mathbb{R})$ is a **Fuchsian group** if and only if Γ acts **properly and discontinuously** on $\mathcal{H}$.

1.3.3 Fundamental Regions

In a general context where X is a metric space and G is a group acting properly discontinuously on X, a closed region $F \subset X$ is called a **fundamental region** of G if it satisfies the following conditions:

1. $\bigcup_{T\in G} T(F) = X$.
2. $int(F) \cap T(int(F)) = \emptyset$ for all $T \in G \setminus \{id\}$.

The family $\{T(F) \ / \ T \in G\}$ is called a **tessellation** of the metric space X.

It is important to mention that for a group G, there may exist different regions that are fundamental regions of the group. However, the following theorem establishes that the area of a fundamental region, in the case where it is finite, is a numerical invariant of the group G.

Theorem 1.3.2 *Let F_1 and F_2 be two fundamental regions for a Fuchsian group Γ such that the hyperbolic area of F_1 is finite, that is, $\mu(F_1) < \infty$, and suppose that the boundaries of F_1 and F_2 have zero area. Then, $\mu(F_1) = \mu(F_2)$.*

Proof. Since the hyperbolic area of the boundaries of F_1 and F_2 is zero, we have that $\mu(int(F_1)) = \mu(F_1)$ and $\mu(int(F_2)) = \mu(F_2)$. Moreover:

$$\bigcup_{T\in\Gamma}(F_1 \cap T(int(F_2))) = F_1 \cap \left(\bigcup_{T\in\Gamma} T(int(F_2))\right) \subseteq F_1.$$

Since F_2 is a fundamental region, the sets $F_1 \cap T(int(F_2))$ are disjoint, and thus we have that:

$$\mu(F_1) \geq \sum_{T\in\Gamma} \mu(F_1 \cap T(int(F_2))) = \sum_{T\in\Gamma} \mu(T^{-1}(F_1) \cap int(F_2)) = \sum_{T\in\Gamma} \mu(T(F_1) \cap int(F_2)).$$

Now, since F_1 is also a fundamental region, we have $\bigcup_{T\in\Gamma} T(F_1) = \mathcal{H}$ and therefore we have:

$$\bigcup_{T\in\Gamma}(T(F_1) \cap int(F_2)) = int(F_2).$$

Then:

$$\sum_{T\in\Gamma} \mu(T(F_1) \cap int(F_2)) \geq \mu\left(\bigcup_{T\in\Gamma} T(F_1) \cap int(F_2)\right) = \mu(int(F_2)) = \mu(F_2).$$

And hence, $\mu(F_1) \geq \mu(F_2)$. Using the same argument, by interchanging F_1 and F_2, we obtain the inequality $\mu(F_2) \geq \mu(F_1)$, and therefore, $\mu(F_1) = \mu(F_2)$.

□

The goal now is to prove that every Fuchsian group has a connected and convex fundamental region called the **Dirichlet region**. This region is defined as follows: Let Γ be a

Fuchsian group and let $p \in \mathcal{H}$ be a point in the upper half-plane such that $T(p) \neq p$ for every $T \in \Gamma \setminus \{id\}$. We define the **Dirichlet region** for Γ centered at p as the set:

$$D_p(\Gamma) := \{z \in \mathcal{H} \ / \ \rho(z,p) \leq \rho(z,T(p)) \text{ for all } T \in \Gamma\}. \tag{1.7}$$

As the hyperbolic metric ρ is invariant under transformations of $PSL(2,\mathbb{R})$, we can also define the Dirichlet region as:

$$D_p(\Gamma) = \{z \in \mathcal{H} \ / \ \rho(z,p) \leq \rho(T(z),p) \text{ for all } T \in \Gamma\}.$$

For each fixed $T_1 \in \Gamma$, $H_p(T_1) = \{z \in \mathcal{H} \ / \ \rho(z,p) \leq \rho(z,T_1(p))\}$ is the set of points z in the upper half-plane that are closer to p than to $T_1(p)$ according to the hyperbolic metric.

Clearly, $p \in D_p(\Gamma)$ and since the Γ-orbit of p is a discrete set, $D_p(\Gamma)$ completely contains some neighborhood of p. To graphically represent the set $H_p(T_1)$, we join the points p and $T_1(p)$ with a geodesic segment and construct the line $L_p(T_1)$ given by the equation $\rho(z,p) = \rho(z,T_1(p))$, as shown in Figure 1.4.

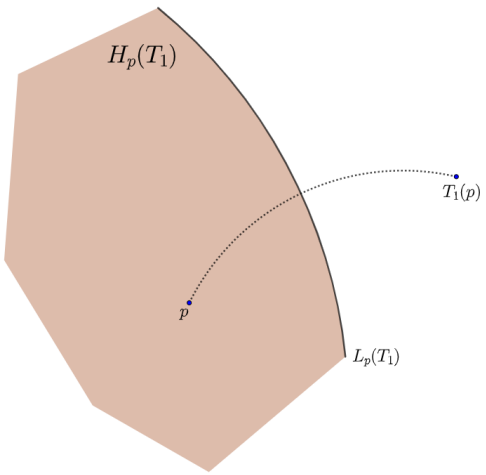


Figure 1.4: Set $H_p(T_1)$

Then the sets $H_p(T)$ are convex sets and moreover $D_p(\Gamma) = \bigcap_{T \in \Gamma \setminus \{id\}} H_p(T)$ and hence $D_p(\Gamma)$ is a convex region.

Theorem 1.3.3 *If p is not a fixed point of any element in $\Gamma \setminus \{id\}$, then $D_p(T)$ is a connected fundamental region for Γ.*

Proof. Let $z \in \mathbb{H}$ and Γz be the Γ-orbit of z. Since Γz is a discrete set, there exists a point $z_0 \in \Gamma z$ such that $\rho(z_0,p)$ is minimal, so $\rho(z_0,p) \leq \rho(T(z_0),p)$ for all $T \in \Gamma$, and therefore $z_0 \in D_p(\Gamma)$. Thus, $D_p(\Gamma)$ contains at least one point from each Γ-orbit.

Now, let $z_1, z_2 \in \text{int}(D_p(\Gamma))$. We claim that z_1 and z_2 cannot lie in the same Γ-orbit. Indeed, if $\rho(z,p) = \rho(T(z),p)$ for some $T \in \Gamma \setminus \{id\}$, then $\rho(z,p) = \rho(z,T^{-1}(p))$, so $z \in L_p(T^{-1})$, which implies that $z \notin D_p(\Gamma)$ or $z \in \partial D_p(\Gamma)$. Therefore, if $z \in \text{int}(D_p(\Gamma))$, then $\rho(z,p) < \rho(T(z),p)$ for all $T \in \Gamma \setminus \{id\}$.

Thus, if z_1 and z_2 belong to the same Γ-orbit, then $\rho(z_1,p) < \rho(z_2,p)$ and $\rho(z_2,p) < \rho(z_1,p)$, which is a contradiction. Therefore, $\text{int}(D_p(\Gamma))$ contains at most one point

from each Γ-orbit. Since $D_p(\Gamma)$ is the intersection of closed and convex sets, it is path-connected, which implies that it is connected.

□

1.4 Riemann Surfaces

In this section, we define the concept of Riemann surface, provide some classic examples of manifolds with Riemann surface structure, and state the uniformization theorem. The results and definitions are taken from sections 2.1 and 2.2 from the book [7].

A **Riemann surface** $\mathcal{S}$ is a connected Hausdorff topological space together with an open covering $\{U_\alpha\}_{\alpha\in I}$ such that for each U_α there exists a homeomorphism z_α from U_α to $\mathbb{C}$ such that the transition functions $f_{\beta\alpha} := z_\alpha \circ z_\beta^{-1} : z_\beta(U_\alpha \cap U_\beta) \longrightarrow z_\alpha(U_\alpha \cap U_\beta)$ are **holomorphic**. For a Riemann surface $\mathcal{S}$, each pair (U_α, z_α) is called a **chart**, and the system of charts $\{(U_\alpha, z_\alpha)\}_{\alpha\in I}$ is called an **atlas**.

Two **atlases** $\{(U_\alpha, z_\alpha)\}$ and $\{(V_\beta, w_\beta)\}$ are said to be **compatible** if the transition functions $f_{\beta\alpha} = w_\beta \circ z_\alpha^{-1} : z_\alpha(U_\alpha \cap V_\beta) \longrightarrow w_\beta(U_\alpha \cap V_\beta)$ are holomorphic. Note that the compatibility of atlases defines an equivalence relation, and the class formed by compatible atlases is called the **conformal structure** of the surface $\mathcal{S}$.

On the other hand, the map $z_\alpha : U_\alpha \longrightarrow \mathbb{C}$ defines an orientation on U_α by taking the preimage of the standard orientation of $\mathbb{C}$. Moreover, over $U_\alpha \cap U_\beta$ the **Jacobian** of the transition function is $|f'_{\beta\alpha}|^2 > 0$, which defines an **orientation** for the surface $\mathcal{S}$. In other words, a **Riemann surface** is an oriented two-dimensional manifold equipped with a conformal structure.

Let $\mathcal{S}_1$ and $\mathcal{S}_2$ be two Riemann surfaces with atlases $\{(U_\alpha, z_\alpha)\}$ and $\{(V_\beta, w_\beta)\}$ respectively. A homeomorphism $f : \mathcal{S}_1 \longrightarrow \mathcal{S}_2$ is conformal, if for every $p \in \mathcal{S}_1$ and every local charts $z_\alpha : p \in U_\alpha \longrightarrow z_\alpha(U_\alpha) \subset \mathbb{C}$ and $w_\beta : f(p) \in V_\beta \longrightarrow w_\beta(V_\beta) \subset \mathbb{C}$ we have that the function $w_\beta \circ f \circ z_\alpha^{-1}$ is conformal.

We say that two Riemann surfaces $\mathcal{S}_1$ and $\mathcal{S}_2$ are **conformally equivalent** if there exists a conformal homeomorphism from $\mathcal{S}_1$ to $\mathcal{S}_2$.

1.4.1 The Uniformization Theorem

Before stating the **uniformization theorem for Riemann surfaces**, let's take a look at some classic examples of Riemann surfaces.

- **Example 1**: Let $\mathcal{S} = \mathbb{C}$ with the single chart atlas $(\mathbb{C}, \mathrm{id})$, where id is the identity on $\mathbb{C}$. It is clear that $\mathbb{C}$ is a Riemann surface, since the compatibility of charts is trivial.
- **Example 2**: For each $n \in \mathbb{Z}$, let $g_n : \mathbb{C} \longrightarrow \mathbb{C}$ be the translation such that $z \longmapsto z+n$. Then $G = \{g_n \ / \ n \in \mathbb{Z}\}$ is a group that acts on $\mathbb{C}$ in such a way that the space of orbits $\mathbb{C}/G$ is topologically equivalent to the set $\{z \in \mathbb{C} \ / \ 0 \leq \mathrm{Re}(z) < 1\}$, where the point iy is identified with the point $1+iy$. Therefore, $\mathbb{C}/G$ is a **cylinder**.

  Let $V_\alpha \subset \mathbb{C}$ be an open subset of $\mathbb{C}$ contained strictly in the strip $\{z \in \mathbb{C} \ / \ 0 \leq \mathrm{Re}(z) < 1\}$. If $\pi : \mathbb{C} \longrightarrow \mathbb{C}/G$ is the canonical projection, then the functions

$\varphi_\alpha : {}^{\mathbb{C}}\!/_{G} \longrightarrow \mathbb{C}$ such that $\varphi_\alpha(\pi(V_\alpha)) = V_\alpha$ are the homeomorphisms that form an atlas for ${}^{\mathbb{C}}\!/_{G}$. Therefore, ${}^{\mathbb{C}}\!/_{G}$ is a Riemann surface. Moreover, the map $z \longmapsto e^{2\pi i z}$ is an isomorphism from ${}^{\mathbb{C}}\!/_{G}$ to $\mathbb{C} \setminus \{0\}$.

- **Example 3**: For each pair $n, m \in \mathbb{Z}$, consider the functions $h_{n,m} : \mathbb{C} \longrightarrow \mathbb{C}$ defined by $z \longmapsto z + in + m$, and the group $G = \{h_{n,m} \ / \ n, m \in \mathbb{Z}\}$. Similar to the previous example, the **torus** $\mathcal{S} = {}^{\mathbb{C}}\!/_{G}$ is a Riemann surface.

- **Example 4**: Let us consider the one-point compactification of $\mathbb{C}$, denoted by $\overline{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$, along with the charts (U_1, z_1) and (U_2, z_2), where $U_1 = \{z \in \mathbb{C} \ / \ |z| < 2\}$, $U_2 = \{z \in \mathbb{C} \ / \ |z| > 1\} \cup \{\infty\}$, and the homeomorphisms are $z_1 = z$ and $z_2 = 1/z$ if $z \neq \infty$, and $z_2 = 0$ at ∞. Then the transition function $f_{2,1}(z) = (z_2) \circ (z_1)^{-1}(z) = 1/z$ is holomorphic on $U_1 \cap U_2$. Therefore, $\overline{\mathbb{C}}$ is a Riemann surface called the **Riemann sphere**.

- **Example 5**: Consider a Fuchsian group Γ acting on the upper half-plane $\mathcal{H}$, and let $\mathcal{S} = {}^{\mathcal{H}}\!/_{\Gamma}$. Then for every $p \in \mathcal{H}$, there exists a neighborhood R such that $\gamma(R) \cap R = \emptyset$ for every $\gamma \in \Gamma \setminus \{id\}$. Therefore, $\mathcal{S}$ is a Hausdorff topological space and the projection $\pi : \mathcal{H} \longrightarrow \mathcal{S}$ is a local homeomorphism. Thus, $\mathcal{S}$ is a Riemann surface.

The uniformization theorem for Riemann surfaces arises historically as a generalization of the **Riemann mapping theorem**, which states that if Ω is a simply connected region completely contained in the complex plane $\mathbb{C}$ and $a \in \Omega$, then there exists a unique function $f(z)$ on Ω such that $f(a) = 0$, $f'(a) > 0$, and $f(z)$ maps Ω injectively onto the open unit disk $\mathbb{D} = \{z \in \mathbb{C} \ / \ |z| < 1\}$.

For example, if $\Omega = \mathcal{H}$ and $a = i$, then the function $f(z) = \dfrac{z - i}{z + i}$ maps the upper half-plane $\mathcal{H}$ conformally onto the unit disk $\mathbb{D}$.

On the other hand, the **Koebe's planarity theorem** states that if $\mathcal{S}$ is a Riemann surface such that every simple closed curve on $\mathcal{S}$ divides the surface into two components, then $\mathcal{S}$ is conformally equivalent either to the **Riemann sphere**, the **complex plane** $\mathbb{C}$, or a domain contained in $\mathbb{C}$.

So, from the **theory of covering spaces**, every surface $\mathcal{S}$ has a simply connected universal covering $\tilde{\mathcal{S}}$ with covering map $\pi : \tilde{\mathcal{S}} \longrightarrow \mathcal{S}$. π is a local homeomorphism and there exists a group G of **Deck transformations** acting freely on $\tilde{\mathcal{S}}$ such that $\tilde{\mathcal{S}}/G$ is homeomorphic to $\mathcal{S}$ and for every $p \in \tilde{\mathcal{S}}$ there exists a neighborhood R such that $\gamma(R) \cap R = \emptyset$ for every $\gamma \in G \setminus \{id\}$.

The conformal structure of $\mathcal{S}$ induces a conformal structure on $\tilde{\mathcal{S}}$, making $\tilde{\mathcal{S}}$ into a simply connected Riemann surface such that every simple closed curve divides $\tilde{\mathcal{S}}$ into two components. Then, the Koebe's planarity theorem and the Riemann mapping theorem imply that $\tilde{\mathcal{S}}$ is conformally equivalent to:

1. The Riemann sphere $\overline{\mathbb{C}}$.
2. The complex plane $\mathbb{C}$.
3. The open disk $\mathbb{D}$, or equivalently, to the upper half-plane $\mathcal{H}$.

In case 1, we have that $\mathcal{S}$ is of the type in **Example 4**. In case 2, we have that $\mathcal{S}$ is of the type in **Examples 1, 2, and 3**, and case 3 leads to $\mathcal{S}$ being of the type in **Example 5**. We can then state the uniformization theorem for Riemann surfaces as follows:

Theorem 1.4.1 (The Uniformization Theorem for Riemann Surfaces)

*Every Riemann surface $\mathcal{S}$ is **conformally equivalent** to exactly one of the above examples.*

1.4.2 Extremal Length

In this section, we define the extremal length of a family of curves, F, on a Riemann surface. This can be thought of as the average minimum length of the curves in F, and its primary application lies in the fact that it is invariant under conformal mappings.

Let $\mathcal{S}$ be a Riemann surface with atlas $\{(U_\alpha, z_\alpha)\}_{\alpha\in I}$, and let F be a family of curves on $\mathcal{S}$. A metric $\rho(z)|dz|$ is an **admissible metric** on $\mathcal{S}$ if it satisfies the following conditions:

- $\rho_1(z_1)|dz_1| = \rho_2(z_2)|dz_2|$ where ρ_1 and ρ_2 are representations of ρ in terms of the coordinates z_1 and z_2.
- ρ is locally L_2 and positive definite on all of $\mathcal{S}$.
- $A(\rho) = \int_{\mathcal{S}} \rho^2 dxdy \in \mathbb{R}^+$, i.e, $A(\rho)$ is different from 0 and ∞.

Let M be the set of all admissible metrics on $\mathcal{S}$. For $\rho \in M$, the length of a curve $\gamma \in F$ is defined as:

$$L_\gamma(\rho) = \int_\gamma \rho|dz|,$$

if ρ is measurable along γ, otherwise $L_\gamma(\rho) = +\infty$. We denote the infimum over all curves γ in F as $L(\rho) = L(\rho, F) = \inf_{\gamma\in F} L_\gamma(\rho)$. The **extremal length** of the family of curves F is defined as:

$$\Lambda(F) := \sup_{\rho\in M} \left\{ \frac{L(\rho)^2}{A(\rho)} \right\}. \tag{1.8}$$

Then $\Lambda(F)$ is the "average" minimum length of curves in F.

Example: Let $R = \{z \in \mathbb{C} \,/\, r_1 < |z| < r_2\}$ be an annulus, and let $\rho_0(z) = \dfrac{1}{2\pi|z|}$ be an admissible metric on R. Let F be a family of arcs in R that connect the inner and outer circles of the annulus. For $r \in (r_1, r_2)$, and for $\theta_0 \in [0, 2\pi)$ fixed, let $\gamma_0(r) = r(\cos(\theta_0) + i\,\mathrm{sen}(\theta_0))$ be a straight line segment connecting the outer and inner circles of R. For all $\gamma \in F$, we then have:

$$L_{\gamma_0}(\rho_0) = \int_{\gamma_0} \frac{|dz|}{2\pi|z|} \leq \int_\gamma \frac{|dz|}{2\pi|z|} = L_\gamma(\rho_0).$$

Therefore, as θ_0 is fixed and r varies, we have that:

$$L_{\gamma_0}(\rho_0) = \int_{r_1}^{r_2} \frac{dr}{2\pi r} \leq L_\gamma(\rho_0) \Longrightarrow \frac{1}{2\pi}\log(r_2/r_1) \leq L_\gamma(\rho_0).$$

And therefore:

$$L(\rho_0) \geq \frac{1}{2\pi}\log(r_2/r_1).$$

On the other hand, the area of R with respect to the metric ρ_0 is:

$$A(\rho_0) = \int_R \frac{rdrd\theta}{4\pi^2 r^2} = \frac{1}{2\pi}\log(r_2/r_1).$$

Then, $\dfrac{L(\rho_0)^2}{A(\rho_0)} \geq \dfrac{1}{2\pi}\log(r_2/r_1)$, which implies that the extremal length of F satisfies the inequality $\Lambda(F) \geq \dfrac{1}{2\pi}\log(r_2/r_1)$.

Now, for a given admissible metric ρ, we have that $L(\rho) \leq \displaystyle\int_{r_1}^{r_2} \rho(z)dr$ and therefore:

$$2\pi L(\rho) = \int_0^{2\pi} L(\rho)d\theta \leq \int_0^{2\pi}\int_{r_1}^{r_2} \rho(z)drd\theta = \int_R \frac{1}{\sqrt{r}}\rho(z)\sqrt{r}drd\theta.$$

Raising both sides of the inequality to the power of 2 and using the Cauchy-Schwarz inequality:

$$4\pi^2 L(\rho)^2 \leq \left(\int_R \frac{1}{\sqrt{r}}\rho(z)\sqrt{r}drd\theta\right)^2 \leq \left(\int_R \frac{1}{r}drd\theta\right)\left(\int_R \rho^2(z)\cdot r\ drd\theta\right).$$

Calculating each integral, we get:

$$\frac{L(\rho)^2}{A(\rho)} \leq \frac{1}{2\pi}\log(r_2/r_1) \Longrightarrow \Lambda(F) \leq \frac{1}{2\pi}\log(r_2/r_1).$$

And therefore, $\Lambda(F) = \dfrac{1}{2\pi}\log(r_2/r_1)$. We will call this quantity the **modulus**, $m(R)$, of the annulus R.

Note that if $r_1 \to 0$, then R becomes a disk without a point and the modulus $m(R)$ tends to infinity.

Proposition 1 in the book **Quasiconformal Teichmüller Theory** [7] implies that the extremal length, and thus the modulus of an annulus R, is an invariant quantity under conformal mappings and therefore under homeomorphisms of Riemann surfaces that preserve the complex structure.

If Ω is a domain in the complex plane that is the union of two subdomains Ω_1 and Ω_2, then a problem of extremal length for a family of curves F in Ω becomes a problem of extremal length for two families of curves F_1 in Ω_1 and F_2 in Ω_2. To illustrate this fact in the context of interest, suppose that R_1 and R_2 are two topological annuli such that the inner boundary of R_1 coincides with the outer boundary of R_2. Let $R = R_2 \cup R_2$ be the topological annulus whose outer boundary is the outer boundary of R_1 and whose inner boundary is the inner boundary of R_2, and let F_j be a family of curves in R_j that, as in the previous example, connect the inner and outer boundaries of R_j. Thus, if F is a family of curves in R that connect the outer and inner boundaries of R, then:

$$\Lambda(F) \geq \Lambda(F_1) + \Lambda(F_2) \text{ and therefore } m(R) \geq m(R_1) + m(R_2).$$

This fact is a direct consequence of the approximation property of the supremum applied to the definition 1.8 of extremal length.

Chapter 2

Topological Properties of the Horseshoe Maps

In this chapter, we define two geometric transformations on the unit square Q whose successive application induces a series of identifications defined on the boundary of Q. These transformations were presented by Chamanara, Gardiner and Lakic in [4], Where, in turn, they mention that this construction was shown to them by André de Carvalho as a topological version of the well-known **Smale horseshoe map**. The main goal is to apply the theory seen in the Chapter 1 to these particular examples.

The identifications induced by both applications define two quotient spaces, X and Y, shown in detail in sections 2.1 and 2.2. The main objective of section 2.3 is to study the topological properties of X and Y using the theory seen in Chapter 1 on Riemann surfaces, Fuchsian groups, and extremal length.

On the other hand, section 2.4 introduces some concepts from the **theory of surface automorphisms**, such as **foliation of a surface**, **transversal measure**, and **pseudo-Anosov automorphism**, applied in the context of the horseshoe map and the surface induced by the identifications defined on the boundary of Q.

2.1 The Horseshoe Map

In this section, we will provide a detailed definition of the horseshoe map presented in the article [4] and the surface on which it operates.

In order to define the **Horseshoe map**, we consider the unit square in the complex plane:

$$Q = \{z \in \mathbb{C} \ / \ 0 \leq \mathrm{Re}(z) \leq 1 \text{ and } 0 \leq \mathrm{Im}(z) \leq 1\}.$$

Let $\mathcal{S}$ be the following set of points on the boundary of Q that accumulate at the origin.

$$\mathcal{S} := \left\{ \cdots, \frac{1}{8}, \frac{1}{4}, \frac{1}{2}, 1, 1+i, i, \frac{i}{2}, \frac{i}{4}, \frac{i}{8}, \cdots \right\} \cup \{0\}. \tag{2.1}$$

We shall consider the set $\tilde{Q} := Q \setminus \mathcal{S}$ partitioned into two non-disjoint subsets:

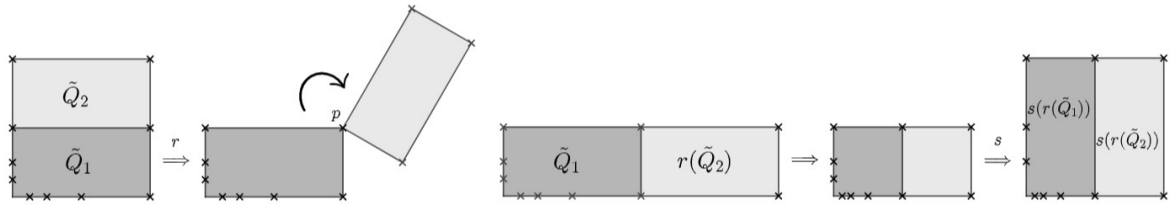
$$\tilde{Q}_1 = \left\{ z \in \tilde{Q} \ / \ 0 \leq \mathrm{Im}(z) \leq \frac{1}{2} \right\}.$$

$$\tilde{Q}_2 = \left\{ z \in \tilde{Q} \ / \ \frac{1}{2} \leq \mathrm{Im}(z) \leq 1 \right\}.$$

Let $\tilde{H}$ be a transformation defined on $\tilde{Q}$ given by the composition of the following geometric maps:

- Let $\tilde{Q}_2$ be rotated 180 degrees about the point $p = 1 + \frac{i}{2}$ in the counterclockwise direction by the transformation $r(z) = -z+2+i$, while the set $\tilde{Q}_1$ remains invariant. (See Figure 2.1a)
- Let $\tilde{Q}_1 \cup r(\tilde{Q}_2)$ be compressed by a factor of $\frac{1}{2}$ along the real axis and expanded by a factor of 2 along the imaginary axis by the transformation $s(x+iy) = \frac{x}{2} + i(2y)$. (See Figure 2.1b)

Similarly, the inverse transformations consist of a compression by a factor of $\frac{1}{2}$ along the imaginary axis, an expansion by a factor of 2 along the real axis, and a rotation of the set $\left\{ z \in \mathbb{C} \ / \ 1 \leq \mathrm{Re}(z) \leq 2 \text{ and } 0 \leq \mathrm{Im}(z) \leq \frac{1}{2} \right\}$, 180° clockwise around the point $p = 1 + \frac{i}{2}$.



(a) (b)

Figure 2.1

We will use these transformations to define an ordered pair (X, H) where X is a topological space, which we will construct by defining an equivalence relation $\sim_{\tilde{Q}}$ on $\tilde{Q}$ that identifies certain points on the boundary of $\tilde{Q}$. And H is an automorphism of the space X, which is called the **Horseshoe Map**.

Note that if $w = x + \frac{i}{2} \in \tilde{Q}$, then w has two images after applying the transformation $\tilde{H}$:

- If we take w as an element of $\tilde{Q}_1$, then $\tilde{H}(w) = \frac{x}{2} + i$.
- Whereas, if we consider w as an element of $\tilde{Q}_2$, then $\tilde{H}(w) = \left(1 - \frac{x}{2}\right) + i$.

Therefore, to ensure the well-definition and continuity of H, we identify, through a rotation around the point $\frac{1}{2} + i$, the points $x_1 = \left(\frac{1}{2} - t\right) + i$ with $x_2 = \left(\frac{1}{2} + t\right) + i$ for $0 < t < \frac{1}{2}$ (see Figure 2.2a).

Moreover, note that for $0 < t < \frac{1}{2}$, the points $w_1 = 1 + i\left(\frac{1}{2} + t\right)$ and $w_2 = 1 + i\left(\frac{1}{2} - t\right)$ have the same image under the transformation $\tilde{H}$:

$$\tilde{H}(w_1) = \tilde{H}(w_2) = \frac{1}{2} + i(1 - 2t) \quad ; \quad 0 < t < \frac{1}{2}.$$

Therefore, in order for H to be an injective transformation, we must identify both points through a rotation around $1 + \frac{i}{2}$, (see Figure 2.2b).

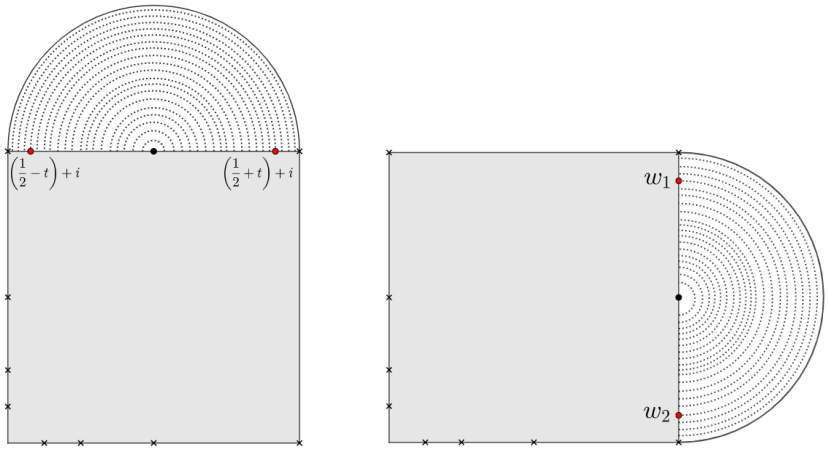


Figure 2.2: Identifications on the upper and right sides of $\tilde{Q}$.

On iterating the transformation $\tilde{H}$, we observe that the line segment connecting the points i and $1 + i$, denoted by $[i; 1 + i]$, is mapped to the line segment $\left[\frac{1}{2}; 1\right]$, which, in turn, is mapped to the line segment $\left[\frac{1}{4}; \frac{1}{2}\right]$. By iterating $\tilde{H}$ indefinitely, we obtain the following sequence of images:

$$I = [i; 1 + i] \longmapsto I_1 = \left[\frac{1}{2}; 1\right] \longmapsto I_2 = \left[\frac{1}{4}; \frac{1}{2}\right] \longmapsto \cdots \longmapsto I_k = \left[\frac{1}{2^k}; \frac{1}{2^{k-1}}\right] \longmapsto \cdots$$

This implies that in order to ensure the well-definition of H, we must make the following identifications through rotations:

- In I_1, we identify $x_1 = \frac{3}{4} - t$ with $x_2 = \frac{3}{4} + t$ for $0 < t < \frac{1}{4}$.
- In I_2, $x_1 = \frac{3}{8} - t$ is identified with $x_2 = \frac{3}{8} + t$ for $0 < t < \frac{1}{8}$.

In general.

- In I_k, we identify $x_1 = \frac{3}{2^{k+1}} - t$ with $x_2 = \frac{3}{2^{k+1}} + t$ for $0 < t < \frac{1}{2^{k+1}}$.

Similarly, for every $k \in \mathbb{N}$, by iterating H, we obtain that the vertical line segment $J_k = \left[\frac{i}{2^k}; \frac{i}{2^{k-1}}\right]$ maps successively onto $J_{k-1} = \left[\frac{i}{2^{k-1}}; \frac{i}{2^{k-2}}\right]$ until $J_1 = \left[\frac{i}{2}; i\right]$, which, in turn, maps onto $J = [1; 1 + i]$. Thus, we obtain the following sequence of images of H:

$$\cdots \mapsto J_k = \left[\frac{i}{2^k}, \frac{i}{2^{k-1}}\right] \longmapsto J_{k-1} = \left[\frac{i}{2^{k-1}}, \frac{i}{2^{k-2}}\right] \mapsto \cdots \mapsto J_1 = \left[\frac{i}{2}, i\right] \longmapsto J = [1+i, 1].$$

So in each segment J_k, the points $i\left(\frac{3}{2^{k+1}} - t\right)$ and $i\left(\frac{3}{2^{k+1}} + t\right)$ for $0 < t < \frac{1}{2^{k+1}}$ are identified by a rotation.

These identifications define a relation $\sim_{\tilde{Q}}$ on $\tilde{Q}$ as follows:

- If $z \in int(\tilde{Q})$ then $z \sim_{\tilde{Q}} z$; that is, if z is in the interior of $\tilde{Q}$, then z is only related to itself.
- $\left(\frac{1}{2} - t\right) + i \sim_{\tilde{Q}} \left(\frac{1}{2} + t\right) + i$ for $0 < t < \frac{1}{2}$ on the upper side of $\tilde{Q}$.
- $1 + i\left(\frac{1}{2} - t\right) \sim_{\tilde{Q}} 1 + \left(\frac{1}{2} + t\right)$ for $0 < t < \frac{1}{2}$ on the right side of $\tilde{Q}$.
- On the lower side of $\tilde{Q}$, for each $k \in \mathbb{N}$, $\frac{3}{2^{k+1}} - t \sim_{\tilde{Q}} \frac{3}{2^{k+1}} + t$ for $0 < t < \frac{1}{2^{k+1}}$.
- Finally, on the left side of $\tilde{Q}$, for each $k \in \mathbb{N}$:

$$i\left(\frac{3}{2^{k+1}} - t\right) \sim_{\tilde{Q}} i\left(\frac{3}{2^{k+1}} + t\right) \text{ para } 0 < t < \frac{1}{2^{k+1}}.$$

The relation $\sim_{\tilde{Q}}$ is an equivalence relation defined on $\tilde{Q}$ (see Figure 2.3). The points marked with $\times$ represent the endpoints of each line segment and are precisely the points that form the set $\mathcal{S}$. On the other hand, the red points represent the midpoints of each segment and are called "**hinge points**". The transformation shown in Figure 2.1 defines an automorphism H of the topological space $X := \tilde{Q}/\sim_{\tilde{Q}}$. This automorphism H is called the **Horseshoe Map** on X.

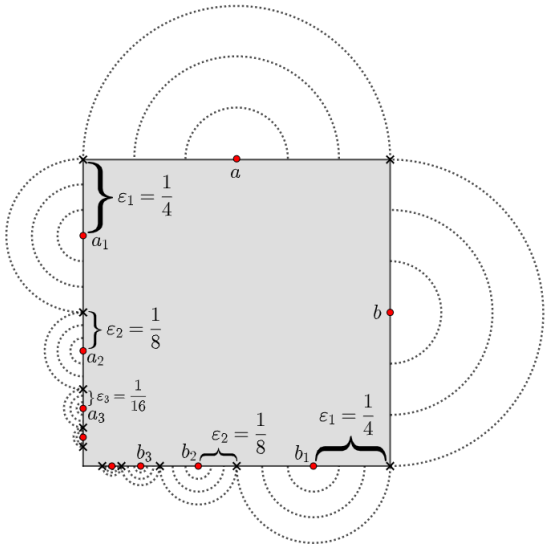


Figure 2.3: Identifications on the boundary of Q

Let $\varepsilon_n = \frac{1}{2^{n+1}}$ be half the length of the segments $I_n = \left[\frac{1}{2^n}; \frac{1}{2^{n-1}}\right]$ and $J_n = \left[\frac{i}{2^n}; \frac{i}{2^{n-1}}\right]$.

We have that the sum $\sum_{j=1}^{\infty} 2\varepsilon_j$ converges to 1, and for any $k \in \mathbb{N}$:

$$\sum_{n=k}^{\infty} \varepsilon_n = \sum_{n=k}^{\infty} \frac{1}{2^{n+1}} = \sum_{n=1}^{\infty} \frac{1}{2^{n+k}} = \frac{1}{2^k} \sum_{n=1}^{\infty} \frac{1}{2^n} = \frac{1}{2^k} = \frac{2}{2^{k+1}} = 2\varepsilon_k.$$

And therefore:

$$\sum_{n=k}^{\infty} \varepsilon_n = 2\varepsilon_k. \tag{2.2}$$

The property of the sequence $\{\varepsilon_j\}$ is called **geometric decay** and will be used later on.

2.2 The Baker's Map

In this section, just as in section 2.1, we will construct in detail the baker's map and the surface on which it operates.

As in the case of the horseshoe map, we start with the unit square in the complex plane:

$$Q = \{z \in \mathbb{C} \ / \ 0 \leq \mathrm{Re}(z) \leq 1 \text{ and } 0 \leq \mathrm{Im}(z) \leq 1\}$$

Now, we consider the set $\mathcal{S}' = \mathcal{A} \cup \mathcal{B}$ of points on the boundary of Q, where $\mathcal{A}$ is a set of points on the left and bottom sides of Q that accumulate at the origin, and $\mathcal{B}$ is a set of points on the right and top sides of Q that accumulate at the point $1 + i$ given by:

$$\mathcal{A} := \left\{ \cdots, \frac{1}{8}, \frac{1}{4}, \frac{1}{2}, 1, \frac{i}{2}, \frac{i}{4}, \frac{i}{8}, \cdots \right\}.$$

$$\mathcal{B} := \left\{ \cdots, \frac{7}{8} + i, \frac{3}{4} + i, \frac{1}{2} + i, i, 1 + \frac{i}{2}, 1 + \frac{3i}{4}, 1 + \frac{7i}{8}, \cdots \right\}.$$

Let us consider the set $P := Q \setminus \mathcal{S}'$ divided into two non-disjoint subsets:

$$P_1 = \left\{ z \in P \ / \ 0 \leq \mathrm{Im}(z) \leq \frac{1}{2} \right\}.$$

$$P_2 = \left\{ z \in P \ / \ \frac{1}{2} \leq \mathrm{Im}(z) \leq 1 \right\}.$$

Instead of a counterclockwise rotation of 180 degrees, as in the case of the horseshoe map, we will consider a translation of the set P_2 given by $L(z) = z + 1 - \frac{i}{2}$. $L(z)$ shifts the set P_2 one unit to the right and $\frac{1}{2}$ units down.

Then, we apply again the transformation $s(x + iy) = \frac{x}{2} + i(2y)$ to the set $P_1 \cup L(P_2)$ as shown in Figure 2.4.

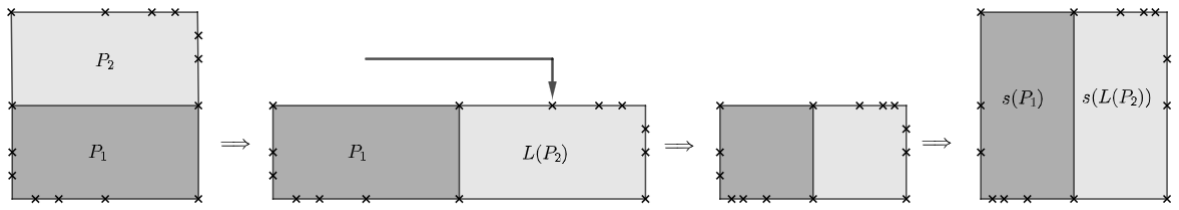


Figure 2.4

Let us consider the transformation $\tilde{B}$ defined on P as the composition of the two aforementioned geometric maps. Once again, our aim is to construct a topological space Y through identifications on the boundary of Q such that the transformation $\tilde{B}$ defines an automorphism B, which we will call the **baker's map**.

Similar to the example of the horseshoe map, if $w = x + \dfrac{i}{2} \in P$, then w is mapped to two different points:

- If we consider w as an element of P_1, then $\tilde{B}(w) = \dfrac{x}{2} + i$. That is, the line segment $x + \dfrac{i}{2}$ with $0 < x < 1$ is mapped to the line segment $I_1 = \left[i; \dfrac{1}{2} + i\right]$.
- Alternatively, if we take w as an element of P_2, then $\tilde{B}(w) = \dfrac{x+1}{2}$. That is, in this case, the line segment $x + \dfrac{i}{2}$ with $0 < x < 1$ is mapped onto the line segment $\tilde{I}_1 = \left[\dfrac{1}{2}; 1\right]$.

In order to ensure the well-definition and continuity of B, we must identify the line segments I_1 and $\tilde{I}_1$ by a translation.

Furthermore, let $J_1 = \left[\dfrac{i}{2}; i\right]$ and $\tilde{J}_1 = \left[1; 1 + \dfrac{i}{2}\right]$. It is clear from Figure 2.4 that $\tilde{B}$ maps both segments onto the line segment $\dfrac{1}{2} + iy$ with $0 < y < 1$. Therefore, to ensure the injectivity of B, the intervals J_1 and $\tilde{J}_1$ are also identified through a translation.

Thus, by considering successive compositions of $\tilde{B}$ with itself, as in the example of the horseshoe map, we obtain the following identifications on the boundary of Q:

For each $k \in \mathbb{N} \cup 0$, let $I_k = \left[\left(1 - \dfrac{1}{2^k}\right) + i; \left(1 - \dfrac{1}{2^{k+1}} + i\right)\right]$ and $\tilde{I}_k = \left[\dfrac{1}{2^{k+1}}; \dfrac{1}{2^k}\right]$ be line segments on the top and bottom sides of Q, respectively.

On the other hand, let $J_k = \left[\dfrac{i}{2^{k+1}}; \dfrac{i}{2^k}\right]$ and $\tilde{J}_k = \left[1 + i\left(1 - \dfrac{1}{2^k}\right); 1 + i\left(1 - \dfrac{1}{2^{k+1}}\right)\right]$ be line segments on the left and right sides of Q, respectively.

The identifications on the boundary of P given by the successive iteration of the transformation $\tilde{B}$, shown in Figure 2.4, are:

$$I_k \sim_B \tilde{I}_k \text{ for } k = 0, 1, 2, 3, ...$$

$$J_k \sim_B \tilde{J}_k \text{ for } k = 0, 1, 2, 3, ...$$

If, in addition, for $z \in int(P)$ it holds that $z \sim_B z$, that is, z is only identified with itself, then $\sim_B$ is an equivalence relation given by translations as shown in Figure 2.5.

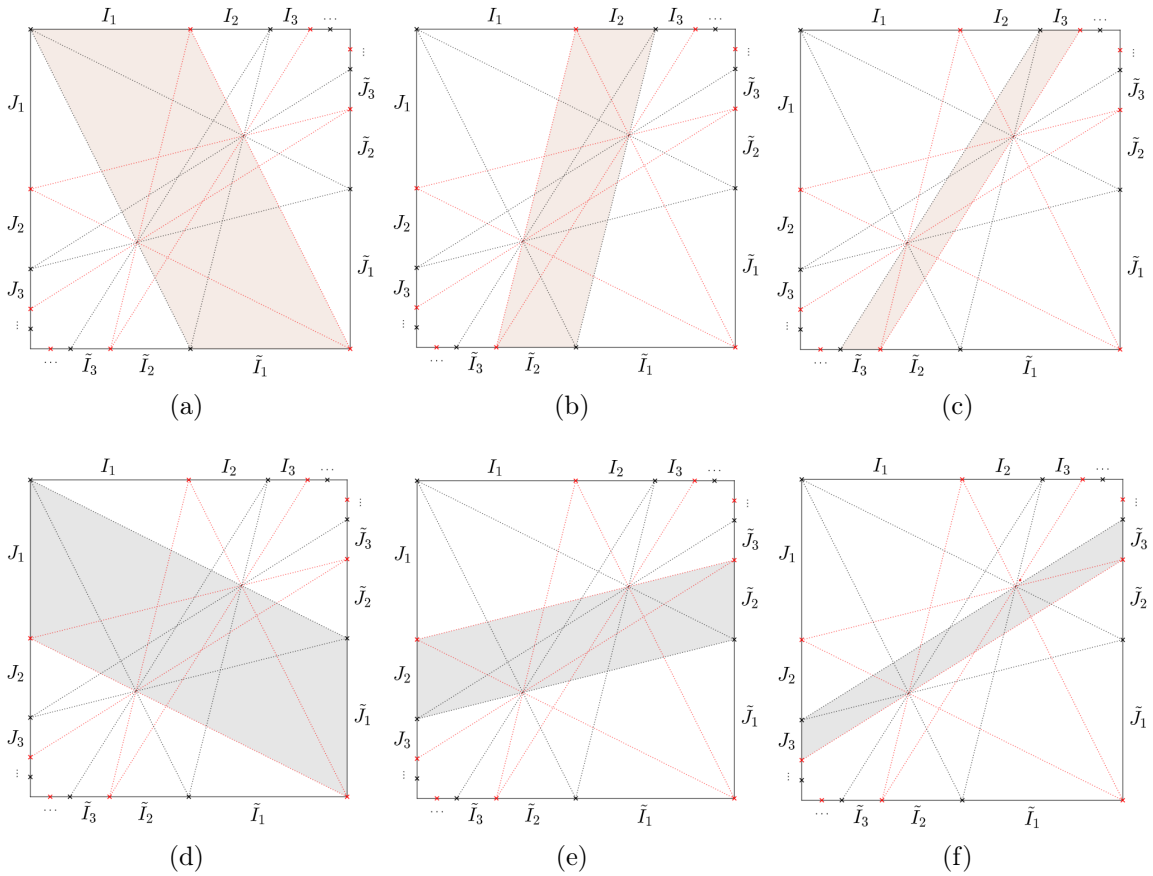


Figure 2.5

The points marked with $\times$ are precisely the points that belong to the set $\mathcal{S}'$ and are excluded from the boundary of the square Q.

The transformation $\tilde{B}$ then defines an automorphism B on the topological space constructed with the translations identifications shown in Figure 2.5, $Y = {}^{P}\!\big/\!\sim_B$. This automorphism is called the **baker's map**.

2.3 Dynamically Generated Surfaces

In this section, we will apply the concepts seen in **Chapter 1** on Riemann surfaces, uniformization theorem, Fuchsian groups, and extremal length to study the structure of the topological spaces $X = {}^{\tilde{Q}}\!\big/\!\sim_{\tilde{Q}}$ and $Y = {}^{P}\!\big/\!\sim_B$ defined in Sections 2.1 and 2.2.

2.3.1 Surface Generated by the Horseshoe Map

Let us consider the topological space $X = {}^{\tilde{Q}}\!\big/\!\sim_{\tilde{Q}}$ formed by the identifications on the boundary of Q shown in Figure 2.3. In addition to the set $\mathcal{S}$ of points removed from the boundary of Q, we will also consider the set of all midpoints of the intervals I, J, I_k, and J_k as shown in Figure 2.3, denoted as follows:

$$\mathcal{K} := \{\cdots, a_k, \cdots, a_1, a, b, b_1, \cdots, b_k, \cdots\}.$$

Let us note that the rotations defining the relation $\sim_{\tilde{Q}}$ are performed precisely with respect to the points in $\mathcal{K}$, which means that these points are identified only with themselves. In other words, the equivalence classes of points in $\mathcal{K}$ consist of a single element.

Let us denote by X^* the topological space $X \setminus \mathcal{K}$. It is easy to define and visualize holomorphic charts on X^* that are compatible with the chart $(int(Q);\ \mathrm{id}(z) = z)$ on the interior of the unit square. Let U_1 and U_2 be two open semidiscs in Q whose centers are located at $\frac{1}{4} + i$ and $\frac{3}{4} + i$, respectively, and let $U = U_1 \cup U_2$. Then the following chart:

$$\left(U; \varphi_U(z) = \left(z - \frac{1}{2} - i\right)^2\right)$$

is holomorphic and is compatible with $(int(Q);\ \mathrm{id}(z) = z)$ since the function φ_U is polynomial. Similarly, the union, V, of the open semidiscs in Q, V_1 and V_2, centered at the points $1 + \frac{i}{4}$ and $1 + \frac{3i}{4}$ respectively, together with the function $\varphi_V(z) = \left(z - 1 - \frac{i}{2}\right)^2$ form a holomorphic chart compatible with the chart $(int(Q);\ \mathrm{id}(z) = z)$.

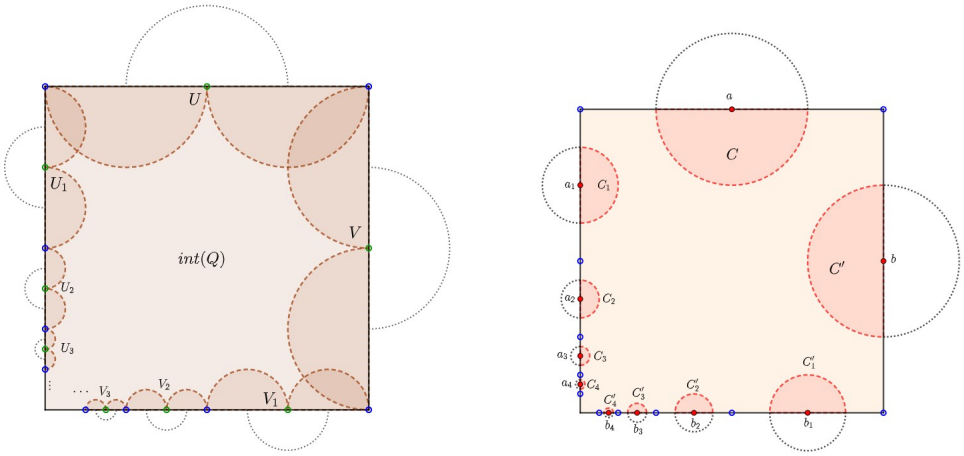
In general, for each $k \in \mathbb{N}$, let U_k be a sequence of subsets of Q that are unions of open semidiscs in Q centered at the points $u_1^k = i\left(\frac{3}{2^{k+1}} + \frac{1}{2^{k+2}}\right)$ and $u_2^k = i\left(\frac{3}{2^{k+1}} - \frac{1}{2^{k+2}}\right)$ respectively. And let $\{V_k\}$ be a sequence of subsets of Q that are unions of open semidiscs in Q centered at the points $v_1^k = \frac{3}{2^{k+1}} + \frac{1}{2^{k+2}}$ and $v_2^k = \frac{3}{2^{k+1}} - \frac{1}{2^{k+2}}$ (see Figure 2.6a). These sets U_k and V_k together with the functions:

$$\varphi_{U_k}(z) = \left(z - \frac{3i}{2^{k+1}}\right)^2 \text{ and } \varphi_{V_k}(z) = \left(z - \frac{3}{2^{k+1}}\right)^2.$$

Form a system of holomorphic charts compatible with $(int(Q);\ \mathrm{id}(z) = z)$.

Therefore, X^* has the structure of a Riemann surface.

On the other hand, the topological space X can be seen as $X^* \cup \mathcal{K}$. Therefore, for each $k \in \mathbb{N}$, the functions $(z-a)^2$, $(z-b)^2$, $(z-a_k)^2$, and $(z-b_k)^2$ provide holomorphic charts at the points a, b, a_k, b_k (see Figure 2.6b) and allow us to extend the Riemann surface structure of X^* to X.



(a) (b)

Figure 2.6

Let us recall that the equivalence relation $\sim_{\tilde{Q}}$ was defined on the set $\tilde{Q} = Q \setminus \mathcal{S}$, where $\mathcal{S}$ is a set of points on the boundary of Q that accumulate at the origin given in Section 2.1. We can then extend the relation $\sim_{\tilde{Q}}$ so that all points in $\mathcal{S}$ form a new equivalence class $[\mathcal{S}]$, and by adding $[\mathcal{S}]$ to the space X, we define $\hat{X} := X \cup \{[\mathcal{S}]\}$. The quotient space $\hat{X}$ is then a **topological sphere**, and its open subsets X and X^* are Riemann surfaces of the **planar** type, that is, any simple closed curve embedded in X and X^* divides both surfaces into two components. Therefore, the **Koebe's Planarity Theorem** guarantees that X and X^* can be conformally embedded in the Riemann sphere $\overline{\mathbb{C}}$.

Let us now see that every conformal embedding of X into the Riemann sphere $\overline{\mathbb{C}}$ extends to a homeomorphism of $\hat{X}$ onto $\overline{\mathbb{C}}$. The proof of the following result is an adaptation of Proposition 8.1 from [11] using the identifications on the boundary of Q generated by the Horseshoe map.

Theorem 2.3.1 *Every holomorphic embedding of the Riemann surface X into the Riemann sphere $\overline{\mathbb{C}}$ extends to a homeomorphism of $\hat{X}$ onto $\overline{\mathbb{C}}$. Moreover, under this homeomorphism, the images of the points in $\mathcal{K}$ form a sequence of points converging to the image of the equivalence class $[\mathcal{S}]$.*

Proof. To prove the theorem, we must construct a coordinate chart around point $[\mathcal{S}]$. To do this, we will build a sequence of topological annuli A_k, homeomorphic to an annulus $R = \{z \in \mathbb{C} \,/\, r_1 \leq |z| \leq r_2\} \subset \mathbb{C}$, contained in $\hat{X}$, that surround the point $[\mathcal{S}] \in \hat{X}$ and whose intersection with X is isomorphic to a punctured disc in $\mathbb{C}$.

The nested rings A_k on the surface $\hat{X}$ are shown in Figure 2.7. The blue points correspond to the points that form the equivalence class $[\mathcal{S}]$, and it can be observed how each ring A_k surrounds the point $[\mathcal{S}]$ by following the identifications defined on the boundary of Q.

For each $k \in \mathbb{N}$, let $\delta_k = \dfrac{\varepsilon_k}{2}$ be the width of the annuli A_k, where ε_k is the Euclidean length of the line segments $J_k = \left[\dfrac{i}{2^{k-1}}, \dfrac{i}{2^k}\right]$ and $I_k = \left[\dfrac{1}{2^{k-1}}, \dfrac{1}{2^k}\right]$ on the boundary of Q shown in Figure 2.3. Thus, by equation 2.2, we have:

$$\sum_{j=k}^{\infty} \delta_j = \sum_{j=k}^{\infty} \frac{\varepsilon_j}{2} = \varepsilon_k. \tag{2.3}$$

This ensures that the rings A_k can be embedded into each segment I_k and J_k without overlapping, as shown in figure 2.7.

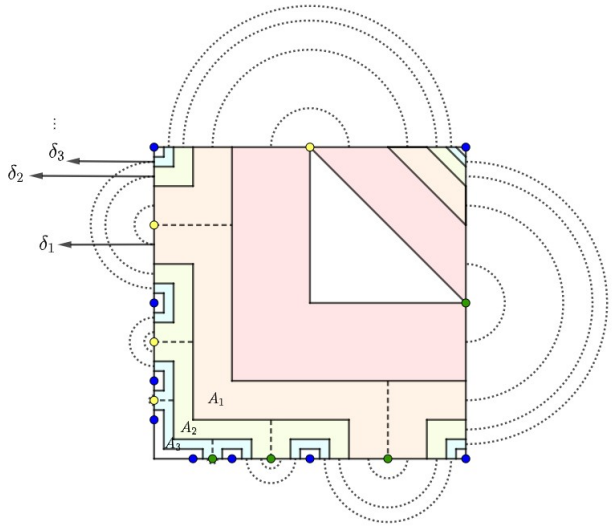


Figure 2.7

For each A_k, divide $\hat{X}$ into two components, the component of the complement A_k^c that contains the point $[\mathcal{S}]$ is called the **inner component**. As seen in subsection 1.4.2 of chapter 1, the extremal length of a family of curves, and therefore the modulus of a ring, is an invariant quantity under conformal mappings. This implies that if the sum of the moduli of the rings A_k diverges, then the interiors of the inner components of the rings A_k provide the required neighborhood of the point $[\mathcal{S}]$.

To calculate the **extremal length** of each ring A_k, we divide it into three parts as follows:

- A **central** component, denoted by $(A_k)^c$, that surrounds the lower-left corner of Q.
- Three **corner components**, denoted by $(A_k)^e$, that surround the upper-left, upper-right, and lower-right corners of the unit square Q.
- $2(k-1)$ **side components**, denoted by $(A_k)^s$, located on the left and bottom sides of Q and which surround the first $2(k-1)$ points belonging to the set $\mathcal{S}$.

We will use the **Euclidean metric** to obtain a lower bound for the extremal length of a family of arcs, F, in A_k that connect its boundary components, we will denote the area of A_k by $\mu(A_k)$.

For each $k \in \mathbb{N}$, the minimum length that an arc in F can have is precisely the width of the ring A_k, thus we obtain an upper bound for the **Euclidean area** of the ring A_k by approximating the area of each component of the ring and summing these quantities.

- $\mu((A_k)^c) \leq \delta_k \left(8 \sum_{j=k}^{\infty} \delta_j \right)$.

- $\mu((A_k)^e) \leq \delta_k \left(2 \sum_{j=k}^{\infty} \delta_j \right)$.
- $\mu((A_k)^s) \leq \delta_k \left(4 \sum_{j=k}^{\infty} \delta_j \right)$.

The Euclidean area of the ring A_k is therefore less than or equal to:

$$\delta_k \left(8 \sum_{j=k}^{\infty} \delta_j + 6 \sum_{j=k}^{\infty} \delta_j + 8(k-1) \sum_{j=k}^{\infty} \delta_j \right).$$

Using equation 2.2, we have that $\sum_{j=k}^{\infty} \delta_j = \sum_{j=k}^{\infty} \frac{\varepsilon_j}{2} = \varepsilon_k = 2\delta_k$. Therefore, the Euclidean area of A_k satisfies the inequality $\mu(A_k) \leq \delta_k^2(12 + 16k)$.

On the other hand, the minimum length of an arc in F is δ_k. This implies that the extremal length of the family of arcs F, and in turn the modulus $m(A_k)$ of each ring A_k, satisfies the following inequality:

$$m(A_k) = \Lambda(F) \geq \frac{\delta_k^2}{\delta_k^2(12 + 16k)} = \frac{1}{12 + 16k}. \tag{2.4}$$

Now, as the series $\sum_{k=1}^{\infty} \frac{1}{12 + 16k}$ diverges, we have that $\sum_{k=1}^{\infty} m(A_k)$ also diverges. Therefore, it follows that the interior of the component of A_k containing $[\mathcal{S}]$, which we call the **inner component**, provides the required neighborhood of $[\mathcal{S}]$ homeomorphic to an open disk in $\mathbb{C}$.

The above demonstrates that every holomorphic embedding of the Riemann surface X in the Riemann sphere $\overline{\mathbb{C}}$ extends to a homeomorphism of $\hat{X}$ onto $\overline{\mathbb{C}}$. Under this homeomorphism, the images of the points a_k and b_k form sequences of points converging to the image of $[\mathcal{S}]$.

□

By Theorem 1.4.1 of **uniformization of Riemann surfaces** stated in section 1.4, we have that the surfaces X, $\tilde{X}$, and X^* are conformally equivalent to one of the following three examples of Riemann surfaces:

- The Riemann Sphere $\overline{\mathbb{C}}$.
- A quotient of the form $\mathbb{C}/G$, where G is a discrete group of conformal isometries of the complex plane $\mathbb{C}$.
- A quotient of the form $\mathcal{H}/\Gamma$, where Γ is a discrete group of conformal isometries of the hyperbolic plane $\mathcal{H}$, which we define as a **Fuchsian group**.

Theorem 2.3.1 implies that $\hat{X}$ is homeomorphic to the Riemann sphere $\overline{\mathbb{C}}$. On the other hand, since X and X^* are of planar type, and X can be seen as the sphere $\hat{X}$ with one point removed, which corresponds to the image of $[\mathcal{S}]$, then X is conformally equivalent to the complex plane $\mathbb{C}$.

Finally, since $X^* = X \setminus \mathcal{K}$, where $\mathcal{K}$ is a set with more than one element, then X^* cannot be conformally equivalent to either the Riemann sphere or the complex plane. Therefore, the uniformization theorem implies that X^* is conformally equivalent to the hyperbolic plane $\mathcal{H}$ factored by a Fuchsian group $\Gamma \subset PSL(2,\mathbb{R})$. The group Γ is a discrete group of isometries of the hyperbolic plane that acts on $\mathcal{H}$ without fixed points, and whose elements identify the sides of a certain fundamental region of the upper half-plane homeomorphic to the set $Q^* = Q \setminus (\mathcal{S} \cup \mathcal{K})$.

Figure 2.8 displays a region $\mathcal{R} \subset \mathcal{H}$ homeomorphic to Q^*, where the points belonging to $\mathcal{S}$ are represented in Q^* with blue $\times$ and the points belonging to $\mathcal{K}$ are represented with red $\times$. Moreover, note that the points of $\mathcal{S}$ and $\mathcal{K}$ are mapped onto the subset of the real axis $\left\{..., -\frac{1}{8}, -\frac{1}{4}, -\frac{1}{2}, -1, 1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, ...\right\}$, hence fall outside of $\mathcal{H}$.

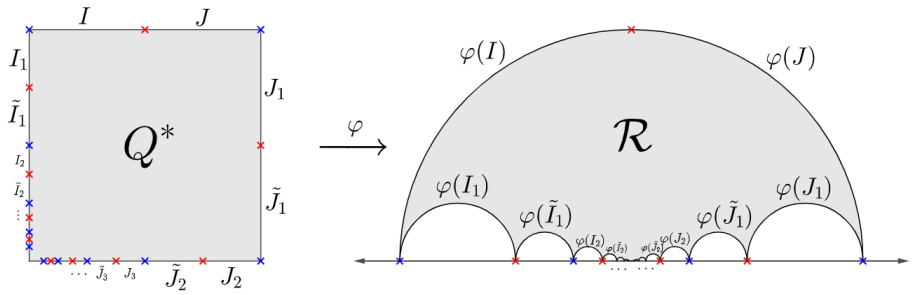


Figure 2.8: Fundamental region for Γ.

The Fuchsian group Γ is then constructed by considering the identifications of line segments on the boundary of Q^* presented in section 2.1. Thus, as in Q^*, the segments I and J are identified, there must be a transformation $T \in PSL(2,\mathbb{R})$ that maps the segment $\varphi(I) \subset \partial\mathcal{R}$ onto $\varphi(J) \subset \partial\mathcal{R}$ and vice versa. On the other hand, for each $k \in \mathbb{N}$, the segment $I_k \subset \partial Q$ is identified with the segment $\tilde{I}_k \subset \partial Q$ by a rotation and therefore, for each $k \in \mathbb{N}$, there must exist a transformation $\alpha_k \in PSL(2,\mathbb{R})$ that maps the segment $\varphi(I_k)$ onto $\varphi(\tilde{I}_k)$. Finally, the segment $J_k \subset \partial Q$ is identified with the segment $\tilde{J}_k \subset \partial Q$ and therefore there must be a transformation $\beta_k \in PSL(2,\mathbb{R})$ that identifies the segment $\varphi(J_k) \subset \partial\mathcal{R}$ with the segment $\varphi(\tilde{J}_k) \subset \partial\mathcal{R}$.

The transformations T, α_k y β_k are given by:

$$T(z) = -\frac{1}{z} \quad ; \quad \alpha_k(z) = \frac{-\frac{3}{\sqrt{2}}z + \frac{10}{2^{2k+1/2}}}{-2^{2k+1/2}z + 3\sqrt{2}} \quad ; \quad \beta_k(z) = \frac{-\frac{3}{\sqrt{2}}z - \frac{10}{2^{2k+1/2}}}{2^{2k+1/2}z + 3\sqrt{2}}.$$

The group Γ is an infinitely generated group, and its generators are precisely the transformations T, α_k, and β_k.

$$\Gamma := \langle T(z),\ \alpha_k(z),\ \beta_k(z)\rangle \subset PSL(2,\mathbb{R}). \tag{2.5}$$

The region $\mathcal{R}$ is a **fundamental region** of Γ, and therefore, the following statements hold:

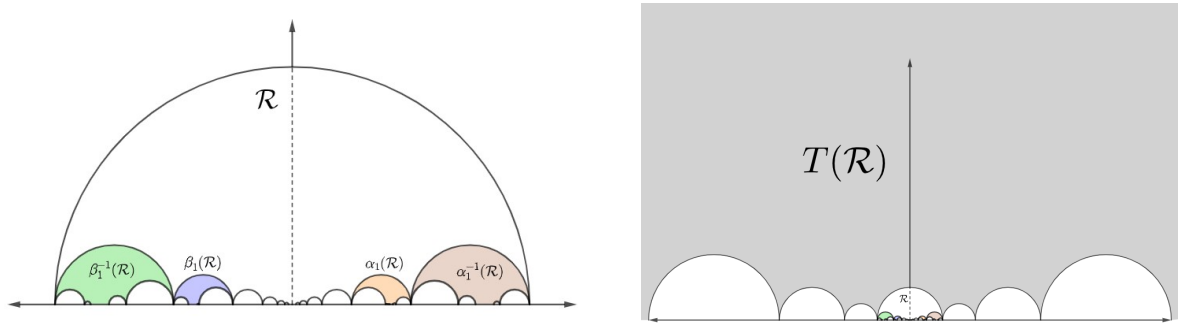
1. For all $\gamma \in \Gamma \setminus \{id\}$, we have that $\gamma(\mathcal{R}) \cap \mathcal{R} = \emptyset$.

2. $\bigcup_{\gamma\in\Gamma} \gamma(\mathcal{R}) = \mathcal{H}$.

Then, the set $\{\gamma(\mathcal{R}) \ / \ \gamma \in \Gamma\}$ is a tessellation of the upper half-plane $\mathcal{H}$.

Since the upper half-plane $\mathcal{H}$ is equipped with the hyperbolic metric, whose curvature is constant and negative, and the elements of Γ are isometries under the hyperbolic metric, we have that $\mathcal{H}/\Gamma$ inherits the canonical hyperbolic structure of $\mathcal{H}$, therefore X^* also has a hyperbolic structure.

Figure 2.9 depicts the tessellation of $\mathcal{H}$ generated by the region $\mathcal{R}$ and all its images under the isometries contained in Γ. Figure 2.9a shows the images of $\mathcal{R}$ under the elements $\alpha_n, \beta_m \in \Gamma$ and their inverses, while Figure 2.9b displays the image of $\mathcal{R}$ under the transformation T.



(a) (b)

Figure 2.9: Tessellation of the upper half-plane.

2.3.2 Surface Generated by the Baker's Map

Let us now consider the Baker map and the topological space $Y = P/\sim_B$ defined in section 2.2 by identifications on the boundary of the set $P = Q \setminus (\mathcal{A} \cup \mathcal{B})$, where $\mathcal{A}$ is a set of points on ∂Q that accumulate at the origin, and $\mathcal{B}$ is a set of points on ∂Q that accumulate at the point $1 + i$.

Figure 2.10 shows a hyperbolic polygon $\mathcal{T}$ in the plane $\mathcal{H}$ that is homeomorphic to the set P, such that the points contained in $\mathcal{A}$ and $\mathcal{B}$ are mapped onto the real axis and thus do not belong to $\mathcal{H}$.

Let us construct a group $\tilde{\Gamma}$ of isometries of $\mathcal{H}$ considering the identifications made on the boundary of P, so that if I_k is identified with $\tilde{I}_k$ by a translation, then $\Psi(I_k)$ is identified with $\Psi(\tilde{I}_k)$ by an isometry of $\tilde{\Gamma}$, and analogously for each segment J_k that is identified with the segment $\tilde{J}_k$ by a translation.

For each $k \in \mathbb{N}$, the transformation of $\tilde{\Gamma}$ that identifies the segments $\Psi(I_k)$ and $\Psi(\tilde{I}_k)$ is given by:

$$\tilde{\beta}_k(z) = \begin{cases} \dfrac{z}{3} & \text{If } k = 1 \\ \dfrac{(2^{k+2}-3)z - \dfrac{2^{k+3}}{3} + \dfrac{5}{3\cdot 2^k}}{-3\cdot 2^{k+1}z + (2^{k+2}+3)} & \text{If } k > 1 \end{cases} \tag{2.6}$$

On the other hand, the transformations in $\tilde{\Gamma}$ that identify the segments $\Psi(J_k)$ and $\Psi(\tilde{J}_k)$ are given by:

$$\tilde{\alpha}_k(z) = \frac{(2^{k+2}-3)z + 2^{k+3} - \dfrac{5}{2^k}}{2^{k+1}z + 2^{k+2} + 3}. \tag{2.7}$$

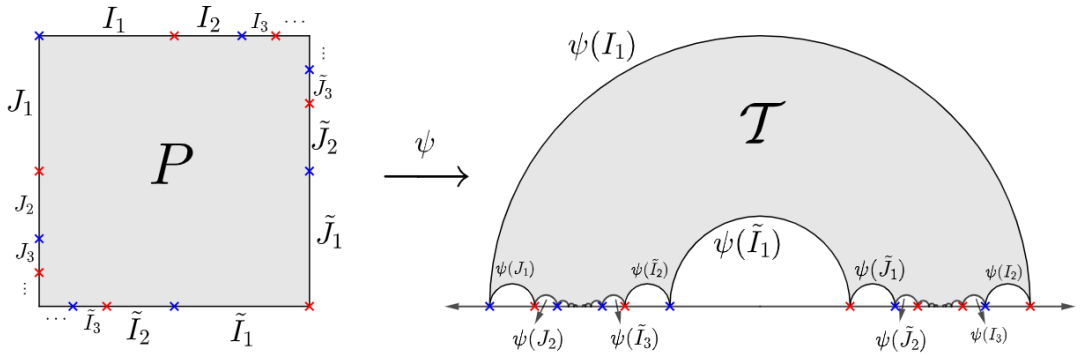


Figure 2.10

The group $\tilde{\Gamma} := \left\langle \tilde{\beta}_k;\ \tilde{\alpha}_k \right\rangle \subset PSL(2,\mathbb{R})$ is a Fuchsian group with a fundamental region $\mathcal{T}$ whose images under the transformations of $\tilde{\Gamma}$ tile the entire hyperbolic plane. The topological space $\mathcal{H}/\tilde{\Gamma}$ has the structure of a Riemann surface.

The following two results provide a more detailed proof of theorem 2 from [4].

Theorem 2.3.2 *Y is a Riemann surface of infinite genus.*

Proof. In order to establish this, we will construct two families of closed and simple curves, $\{\omega_i\}_{i\in\mathbb{N}}$ and $\{\tau_i\}_{i\in\mathbb{N}}$, on the surface Y, such that for any pair of curves, they are not homotopic. Furthermore, for $j,k \in \mathbb{N}$, ω_j intersects the curve τ_k at a single point only if $j = k$. Lastly, if $j \neq k$, then the curves ω_j and ω_k do not intersect the curves τ_j and τ_k, respectively.

Let a_i and $\tilde{a}_i$ denote the midpoints of the segments I_i and $\tilde{I}_i$, respectively, and let b_i and $\tilde{b}_i$ denote the midpoints of J_i and $\tilde{J}_i$. Since the segments I_i and J_i are identified with the segments $\tilde{I}_i$ and $\tilde{J}_i$, respectively, via translations, we denote by A_i the equivalence class $\{a_i, \tilde{a}_i\}$ in Y, and by B_i the class $\{b_i, \tilde{b}_i\}$. Thus, ω_1 is the curve connecting the points a_1 and $\tilde{a}_1$, while τ_1 connects the points b_1, a_1, $\tilde{a}_1$, and $\tilde{b}_1$. Consequently, both curves are closed in Y and intersect only at the point $A_1 \in Y$.

The curve ω_2 must connect the points a_2 and $\tilde{a}_2$ without intersecting ω_1 and τ_1. Thus, it starts from a_2 towards the point $1 + \dfrac{3i}{8} \in \tilde{J}_1$, then translates to the point $\dfrac{7i}{8} \in J_1$,

connects with $\frac{1}{8} + i \in I_1$, further translates to $\frac{5}{8} \in \tilde{I}_1$, and finally closes at the point $\tilde{a}_2 \in \tilde{I}_2$. On the other hand, the curve τ_2 starts from b_2 towards $\tilde{a}_2$, translates to a_2, and closes at $\tilde{b}_2$. In this way, ω_2 and τ_2 only intersect at $A_2 \in Y$ without intersecting ω_1 or τ_1.

Similarly, ω_3 must connect the points a_3 and $\tilde{a}_3$ without intersecting ω_1, ω_2, τ_1, and τ_2. Thus, it starts from a_3 towards $1 + \frac{11i}{16} \in \tilde{J}_2$, then translates to $\frac{7i}{16} \in J_2$, connects with $\frac{9i}{16} \in J_1$, further translates to $1 + \frac{i}{16} \in \tilde{J}_1$, connects with $\frac{15}{16} \in \tilde{I}_1$, then translates to $\frac{7}{16} + i \in I_1$, connects with $\frac{9}{16} + i \in I_2$, further translates to $\frac{5}{16} \in \tilde{I}_2$, and finally closes at $\tilde{a}_3$. The curve τ_3 simply starts from b_3 towards $\tilde{a}_3$, translates to a_3, and closes at $\tilde{b}_3$, intersecting ω_3 at the point $A_3 \in Y$.

The curves $\omega_1, \omega_2, \omega_3, \tau_1, \tau_2$, and τ_3 are depicted in the following Figure 2.11. The dashed lines represent the curves ω_i and τ_i, while the solid lines represent the translations:

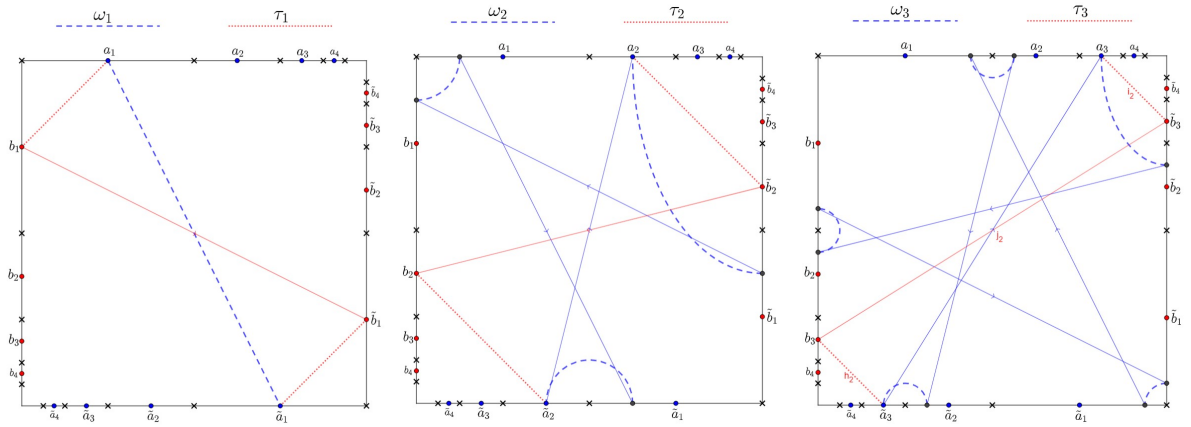


Figure 2.11

Following this pattern indefinitely, we construct a family of curves with the required properties. Therefore, Y is a Riemann surface of infinite genus.

□

Theorem 2.3.3 *Let $c : \mathbb{C} \longrightarrow \mathbb{C}$ be the rotation of 180 degrees around the point $\frac{1}{2}(1+i)$ given by $c(z) = -z + 1 + i$. Then c induces an automorphism c^* on the surface Y such that the quotient space Y/c^* is homeomorphic to X, and the Baker function defines an automorphism B^* on Y/c^* that commutes with the following diagram.*

$$\begin{array}{ccc} Y/c^* & \xrightarrow{B^*} & Y/c^* \\ \downarrow & & \downarrow \\ X & \xrightarrow[H]{} & X \end{array}$$

Proof. In order to see that c induces an automorphism on the surface Y, let's observe that it preserves the identifications defined by the relation $\sim_B$.

Let $p \in I_1$ and $\tilde{p} \in \tilde{I}_1$ such that $p \sim_B \tilde{p}$. This implies that $Im(p) = 1$ and $Im(\tilde{p}) = 0$. Furthermore, since the translation that identifies I_1 with $\tilde{I}_1$ consists of a downward shift of one unit parallel to the imaginary axis and a rightward shift of $\frac{1}{2}$ unit parallel to the real axis, we have $\mathrm{Re}(\tilde{p}) = \frac{1}{2} + \mathrm{Re}(p)$.

Now, as $c(x + iy) = (1 - x) + i(1 - y)$, we have:

- $\mathrm{Im}(c(p)) = 0$ and $\mathrm{Re}(c(p)) = 1 - \mathrm{Re}(p)$.
- $\mathrm{Im}(c(\tilde{p})) = 1$ and $\mathrm{Re}(c(\tilde{p})) = 1 - \mathrm{Re}(\tilde{p})$.

In this way:

$$\mathrm{Re}(c(p)) = 1 - \left(\mathrm{Re}(\tilde{p}) - \frac{1}{2}\right) = 1 - \left(1 - \mathrm{Re}(c(\tilde{p})) - \frac{1}{2}\right) = \mathrm{Re}(c(\tilde{p})) + \frac{1}{2}.$$

Therefore $c(p) \sim_B c(\tilde{p})$.

If $k \geq 2$, then the segments I_k and $\tilde{I}_k$ are identified through a translation of the form $t_k(z) = z + i + \dfrac{1}{2^k}$, and by a similar reasoning as before, it can be proven that if $p \in I_k$ and $\tilde{p} \in \tilde{I}_k$ are identified by $\sim_B$, then $c(p)$ and $c(\tilde{p})$ are also identified. Similarly, this holds for the segments J_k and $\tilde{J}_k$ located on the vertical sides of the square.

On the other hand, if t_k is a translation that identifies the segment I_k with $\tilde{I}_k$, then $c \circ t_k$ is one of the rotations that generate the relation $\sim_{\tilde{Q}}$ described in Section 2.1. For example, if $k = 1$, then $t_1(z) = z - \dfrac{1}{2} + i$, and therefore:

$$(c \circ t_1)(z) = -z + \frac{1}{2} - i + 1 + i = -z + \frac{3}{2},$$

which is a rotation of 180° around the point $\dfrac{3}{4}$.

For $k \geq 2$, we have $t_k(z) = z + \dfrac{1}{2^k} + i$, thus:

$$(c \circ t_k)(z) = -z - \frac{1}{2^k} - i + 1 + i = -z + 1 - \frac{1}{2^k},$$

which is a rotation of 180° around the point $\dfrac{3}{2^{k+1}}$.

Similarly, if t_k^* is a translation that identifies the segment J_k with $\tilde{J}_k$ of P, then $c \circ t_k^*$ is a rotation of 180° around $1 + \frac{i}{4}$ (in the case $k = 1$), and around $\dfrac{3i}{2^{k+1}}$ (in the case $k \geq 2$).

Finally, note that for $0 < r < \dfrac{1}{2}$, we have $c\left(\left(\dfrac{1}{2} \pm r\right) + \dfrac{i}{2}\right) = \left(\dfrac{1}{2} \mp r\right) + \dfrac{i}{2}$, so these points are identified through a rotation of 180° around the point $m = \dfrac{1}{2}(1 + i)$. The following image 2.12 shows the identifications induced by c.

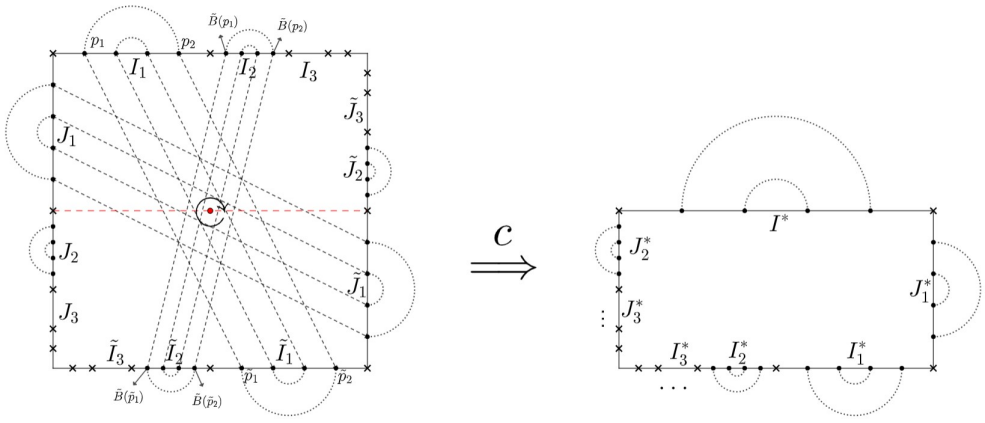


Figure 2.12

In this way, the quotient space Y/c^* is naturally homeomorphic to the surface X.

Now let us see that $B : Y \longrightarrow Y$ preserves the identifications defined by c, i.e., if $p_1, p_2 \in I_i$ and $\tilde{p}_1, \tilde{p}_2 \in \tilde{I}_i$ are in the same equivalence class of Y/c^*, then their images after applying the transformations on P that define B as introduced in section 2.2 also belong to the same equivalence class of Y/c^*.

To show this, let us note that if $p_1, p_2 \in I_i$ and $\tilde{p}_1, \tilde{p}_2 \in \tilde{I}_i$ are in the same equivalence class of Y/c^*, then we have $p_1 \sim_B p_2$ and $\tilde{p}_1 \sim_B \tilde{p}_2$. Furthermore, the points p_1 and p_2 are symmetric with respect to the midpoint of I_i, while the points $\tilde{p}_1$ and $\tilde{p}_2$ are symmetric with respect to the midpoint of the segment $\tilde{I}_i$.

Since $\tilde{B}$ is the composition of a translation, a compression along the real axis, and an expansion along the imaginary axis, as shown in Section 2.2, such that $\tilde{B}(I_i) = I_{i+1}$ and $\tilde{B}(\tilde{I}_i) = \tilde{I}_{i+1}$, it follows that $\tilde{B}(p_1)$ and $\tilde{B}(p_2)$ are symmetric with respect to the midpoint of I_{i+1}, while $\tilde{B}(\tilde{p}_1)$ and $\tilde{B}(\tilde{p}_2)$ are symmetric with respect to the midpoint of $\tilde{I}_{i+1}$. Furthermore, we have $\tilde{B}(p_1) \sim_B \tilde{B}(\tilde{p}_1)$ and $\tilde{B}(p_2) \sim_B \tilde{B}(\tilde{p}_2)$. Therefore, $\tilde{B}(p_1)$, $\tilde{B}(p_2)$, $\tilde{B}(\tilde{p}_1)$, and $\tilde{B}(\tilde{p}_2)$ belong to the same equivalence class of Y/c^*. A similar process can be done for the segments J_i and $\tilde{J}_i$, with the difference being that in this case we have $\tilde{B}(J_{i+1}) = \tilde{J}_i$ and $\tilde{B}(\tilde{J}i + 1) = \tilde{\tilde{J}}_i$. Thus, the Baker function B preserves the identifications defined by c^* and lifts to an automorphism B^* of Y/c^*. Figure 2.12 illustrates the process for two points in I_1.

If we denote by I_i^* and J_i^* the equivalence classes in Y/c^* that result from the identifications made by $\sim_B$ and c^* on the segments $I_i, \tilde{I}_i, J_i$, and $\tilde{J}_i$ of P, then it is clear that B^* acts on Y/c^* in the same way that H acts on X.

□

2.4 Pseudo – Anosov Automorphisms

Thurston's Theorem classify the automorphisms of a closed surface in three classes: **periodic, reducible and pseudo-Anosov**, in this section, we apply the definitions presented in section 6 from [3] to prove that the Horseshoe map is a pseudo-Anosov automorphism of the surface X. The calculations in this section are original work.

A **singular foliation** $\mathcal{F}$ on a surface S is a decomposition of S into a disjoint union of **leaves**. For any point $x \in S$ outside of a finite set A, there exists a chart $\varphi : U \longrightarrow \mathbb{C}$ that maps the components of $U \cap leaf$ to horizontal intervals.

For $x \in A$, there exists a chart $\varphi : U \longrightarrow \mathbb{C}$ such that $\mathcal{F} \cap U$ is mapped to W_k, where W_k is the standard "k-prong singularity". The set A is the **singular set** of $\mathcal{F}$.

Two foliations $\mathcal{F}^+$ and $\mathcal{F}^-$ are **transverse** if they have the same singular set, and at all other points the leaves are transverse.

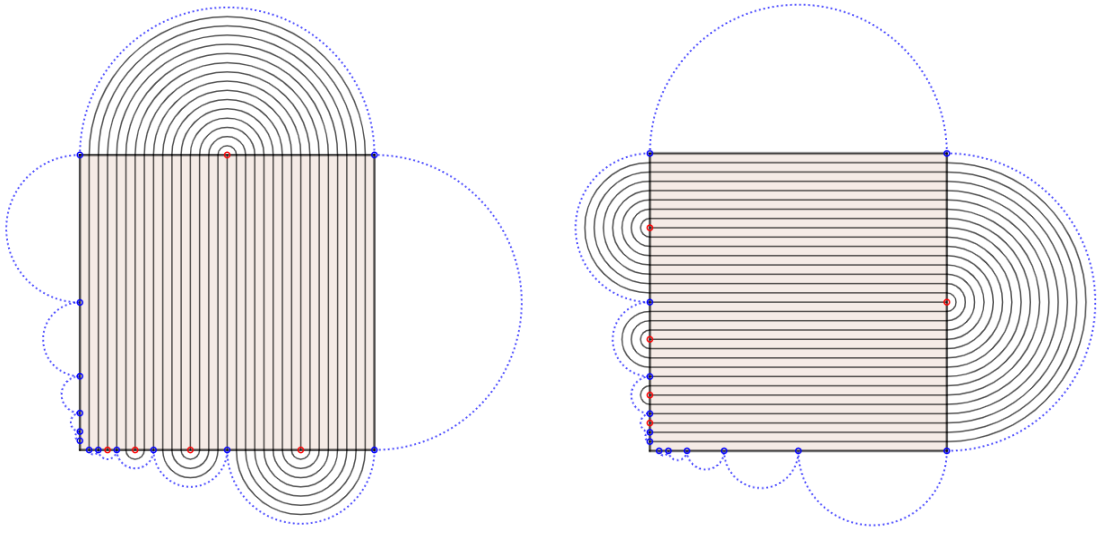
For the Horseshoe surface X. Consider the following line segments:

$$\mathcal{L}^+ := \{z \in X \ / \ \mathrm{Re}(z) = 1/2\} \text{ and } \mathcal{L}^- := \{z \in X \ / \ \mathrm{Im}(z) = 1/2\}.$$

The iterates of H and H^{-1} defines two transverse foliations on X, $\mathcal{F}^+$ and $\mathcal{F}^-$, given by:

$$\mathcal{F}^+ := \bigcup_{n \in \mathbb{N} \cup \{0\}} H^n(\mathcal{L}^+) \text{ and } \mathcal{F}^- := \bigcup_{n \in \mathbb{N} \cup \{0\}} H^{-n}(\mathcal{L}^-).$$

For each $m \in \mathbb{Z}$, $H^m(\mathcal{L}^+)$ is a **leaf** of the foliation $\mathcal{F}^+$ and $H^m(\mathcal{L}^-)$ is a **leaf** of the foliation $\mathcal{F}^-$. Let $\mathcal{B}(\mathcal{F}^+)$ and $\mathcal{B}(\mathcal{F}^-)$ be the sets of leaves of the foliations $\mathcal{F}^+$ and $\mathcal{F}^-$ respectively. (See Figure 2.11)



(a) Foliation $\mathcal{F}^+$ (b) Folaition $\mathcal{F}^-$

Figure 2.13

Definition 2.4.1 *A **transverse measure** μ for a singular foliation $\mathcal{F}$ defines on each arc α transverse to $\mathcal{F}$ a non-negative **Borel measure** $\mu(\alpha)$ with the following properties:*

1. *If β is a subarc of α, then $\mu(\beta)$ is the restriction of $\mu(\alpha)$.*

2. *If α_0 and α_1 are arcs transverse to $\mathcal{F}$ related by a homotopy $F : [0,1] \times [0,1] \longrightarrow X$ such that $F([0,1] \times \{0\}) = \alpha_0$, $F([0,1] \times \{1\}) = \alpha_1$ and $F(\{a\} \times [0,1])$ is contained in a leaf of $\mathcal{F}$ for all $a \in [0,1]$, then $\mu(\alpha_0) = \mu(\alpha_1)$.*

An automorphism h of a closed orientable surface is **pseudo-anosov** if there exist transverse singular foliations $\mathcal{F}^+$ and $\mathcal{F}^-$ such that h expands the foliation $\mathcal{F}^+$ by a factor of λ and contracts $\mathcal{F}^-$ by a factor of λ^{-1}. This means that there exist transverse measures μ^u and μ^s with the following properties:

$$h(\mathcal{F}^+, \mu^u) = (\mathcal{F}^+, \lambda\mu^u).$$
$$h(\mathcal{F}^-, \mu^s) = (\mathcal{F}^-, \lambda^{-1}\mu^s).$$

For some $\lambda > 1$.

A **rectangle** R is a map $\rho : [0,1] \times [0,1] \longrightarrow X$ such that ρ is an embedding on the interior $int(\tilde{X})$. For all $a, b \in [0,1]$, there exist two leaves $\ell \in \mathcal{B}(\mathcal{F}^+)$ and $\ell' \in \mathcal{B}(\mathcal{F}^-)$ such that:

$$\rho(\{a\} \times [0,1]) \subset \ell \text{ and } \rho([0,1] \times \{b\}) \subset \ell'.$$

For a rectangle R, denote $\rho(\{0,1\} \times [0,1])$ by $\partial^+ R$ and $\rho([0,1] \times \{0,1\})$ by $\partial^- R$.

Consider the **extremal leaves** of $\mathcal{F}^+$ and $\mathcal{F}^-$ defined as:

$$\lambda^+ := \lim_{n\to\infty} H^n(\mathcal{L}^+) \in \mathcal{B}(\mathcal{F}^+) \text{ and } \lambda^- := \lim_{n\to\infty} H^{-n}(\mathcal{L}^-) \in \mathcal{B}(\mathcal{F}^-).$$

There exists a decomposition of the surface $\tilde{X}$ into a finite union of rectangles R_1, R_2, R_3 and R_4 with the following properties:

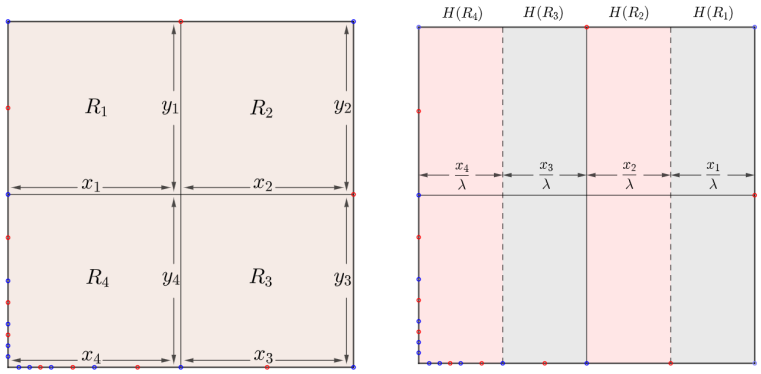
1. $\partial^+ R_i \subset \mathcal{L}^+ \cup \lambda^+$.
2. $\partial^- R_i \subset \mathcal{L}^- \cup \lambda^-$.
3. $H\left(\bigcup_{i=1}^{4} \partial^+ R_i\right) \subset \bigcup_{i=1}^{4} \partial^+ R_i$.
4. $H^{-1}\left(\bigcup_{i=1}^{4} \partial^- R_i\right) \subset \bigcup_{i=1}^{4} \partial^- R_i$.

Such a decomposition is called **Markov partition** for H and it is shown in Figure 2.12a.

Let us construct the transverse measures μ^u and μ^s for the foliations $\mathcal{F}^-$ and $\mathcal{F}^+$ such that $H(\mu^u) = \lambda\mu^u$ and $H(\mu^s) = \lambda^{-1}\mu^s$ for some $\lambda > 1$. The measure μ^u will assign to each rectangle R_i a "height", y_i, which will be the measure of any "vertical" cross-section of R_i and similarly, μ^s will assign to R_i a "width", x_i, which will be the measure of any "horizontal" cross-section of R_i.

In order to establish the necessary conditions for x_i and y_i, for $i, j \in \{1,2,3,4\}$, consider the set $N_{i,j} = H(R_i) \cap R_j$. $N_{i,j}$ is empty, or it consists of a finite union of subrectangles $s_1, ..., s_{l(i,j)}$ of R_j with $\partial^- s_k \subset \partial^- R_j$.

If $N_{i,j} \neq \emptyset$ then $l(i,j)$ denotes how many times $H(R_i)$ intersects R_j as shown in Figure 2.14b.



(a) (b)

Figure 2.14

Set $a_{i,j} = 0$ if $N_{i,j} = \emptyset$, otherwise set $a_{i,j} = l(i,j)$, therefore we have that:

$$a_{1,1} = 0;\ a_{1,2} = 1;\ a_{1,3} = 1;\ a_{1,4} = 0.$$

$$a_{2,1} = 0;\ a_{2,2} = 1;\ a_{2,3} = 1;\ a_{2,4} = 0.$$

$$a_{3,1} = 1;\ a_{3,2} = 0;\ a_{3,3} = 0;\ a_{3,4} = 1.$$

$$a_{4,1} = 1;\ a_{4,2} = 0;\ a_{4,3} = 0;\ a_{4,4} = 1.$$

Since R_j is the union of the sets $H(R_i) \cap R_j$ with $1 \leq i \leq 4$, we have that:

$$x_1 = a_{1,1}\lambda^{-1}x_1 + a_{2,1}\lambda^{-1}x_2 + a_{3,1}\lambda^{-1}x_3 + a_{4,1}\lambda^{-1}x_4 = \lambda^{-1}x_3 + \lambda^{-1}x_4.$$

$$x_2 = a_{1,2}\lambda^{-1}x_1 + a_{2,2}\lambda^{-1}x_2 + a_{3,2}\lambda^{-1}x_3 + a_{4,2}\lambda^{-1}x_4 = \lambda^{-1}x_1 + \lambda^{-1}x_2.$$

$$x_3 = a_{1,3}\lambda^{-1}x_1 + a_{2,3}\lambda^{-1}x_2 + a_{3,3}\lambda^{-1}x_3 + a_{4,3}\lambda^{-1}x_4 = \lambda^{-1}x_3 + \lambda^{-1}x_4.$$

$$x_4 = a_{1,4}\lambda^{-1}x_1 + a_{2,4}\lambda^{-1}x_2 + a_{3,4}\lambda^{-1}x_3 + a_{4,4}\lambda^{-1}x_4 = \lambda^{-1}x_1 + \lambda^{-1}x_2.$$

Let M be the transition matrix

$$M = [a_{i,j}] = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}.$$

M has eigenvalues $\lambda = 0$ with multiplicity 3 and $\lambda = 2$ whose eigenvector is:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

Similarly, the same procedure is used to determine the values of y_i by considering the set $H^{-1}(R_i) \cap R_j$. The transition matrix remains the same, and therefore $\mathbf{y} = \mathbf{x}$.

Let α be a closed arc transverse to the foliation $\mathcal{F}^+$. For each $m \in \mathbb{N}$, α is the union of the sets $\alpha \cap H^m(R_i)$; Thus, if $u_{i,m}$ is the number of components of $\alpha \cap H^m(R_i)$, then the transversal measure μ^s is defined as:

$$\mu^s(\alpha) := \lim_{m\to\infty} \sum_{i=1}^{4} \lambda^{-m} x_i u_{i,m} = \lim_{m\to\infty} \frac{1}{2^m} \sum_{i=1}^{4} u_{i,m}. \tag{2.8}$$

Analogously, if α is a closed arc transverse to $\mathcal{F}^-$, for each $m \in \mathbb{N}$, α would be the union of sets $\alpha \cap H^{-m}(R_i)$; if $v_{i,m}$ is the number of components of $\alpha \cap H^{-m}(R_i)$, we define the measure μ^u as:

$$\mu^u(\alpha) := \lim_{m\to\infty} \sum_{i=1}^{4} \lambda^{-m} y_i v_{i,m} = \lim_{m\to\infty} \frac{1}{2^m} \sum_{i=1}^{4} v_{i,m}. \tag{2.9}$$

Note that if α is an arc transverse to $\mathcal{F}^+$ and 2^{-m} is the "width" of the rectangles $H^m(R_i)$ containing α, then upon applying the map H, the width of these rectangles is divided by 2. Therefore, we have:

$$\mu^s(H(\alpha)) = 2^{-1}\mu^s(\alpha).$$

Similarly, if α is a closed arc transversal to $\mathcal{F}^-$ and 2^{-m} is the "height" of the rectangles $H^{-m}(R_i)$ containing α, then applying H multiplies the "height" by 2, and thus:

$$\mu^u(H(\alpha)) = 2\mu^u(\alpha).$$

That is, H is a pseudo-Anosov automorphism of the surface X.

Chapter 3

Dynamical Properties of the Horseshoe Map

Chaos is one of the most interesting properties that a dynamical system can have. It first appeared in the works of the French mathematician **Henri Poincaré** on the stability of the solar system, where he realized that small variations in the initial conditions of the system produce large changes in the positions and velocities of the planets in the long term. Within the mathematical theory of chaos, the concept of the **entropy** of a function arises, introduced by the Russian mathematician **Andrei Kolmogorov** as a way to measure how the orbits of a dynamical system diverge. If the entropy of a system is greater than zero, then that system exhibits chaotic dynamics.

In Section 3.1, we will begin by applying some known concepts from **ergodic theory** [14]. Using the definition of transverse measure of a foliation to assign a measure to each rectangle defined on the surface X, in order to calculate the **entropy** generated by the dynamical system (X, H).

Subsequently, in Section 3.2, we will use the concepts seen in Chapter 1 regarding **symbolic dynamics, sequence space, and shift function** to explicitly understand the dynamics of the function H on the surface X.

3.1 Entropy of Horseshoe map

In this section, we will use the transversal measures defined in Section 2.4 to establish a measure on the surface X and apply some classic concepts from ergodic theory to compute the entropy of the horseshoe map. The definitions and the Kolmorov-Sinai theorem, stated later, were taken from the book [14], the entropy calculations are original work.

Recall that a **rectangle** R of X is an embedding $\rho : [0,1] \times [0,1] \longrightarrow X$ such that the sets $\partial^+ R = \rho(\{0,1\} \times [0,1])$ and $\partial^- R = \rho([0,1] \times \{0,1\})$ are contained in leaves of the transversal foliations $\mathcal{F}^+$ and $\mathcal{F}^-$ respectively.

Both connected components of $\partial^+ R$ satisfy property 2 of the definition 2.4.1 of transverse measure and therefore have the same measure under μ^u, which we denote by $\mu^u(\partial^+ R)$. Similarly, both components of $\partial^- R$ have the same measure under μ^s, denoted by $\mu^s(\partial^- R)$.

Let $\beta := \{R \subset X \ / \ R \text{ is a } \textbf{rectangle} \text{ of } X\}$. Then we can assign a measure, $\mu(R)$, to each $R \in \beta$, as follows:

$$\mu(R) := \mu^u(\partial^+ R) \cdot \mu^s(\partial^- R).$$

It is easy to see that the horseshoe map preserves the measure μ, since if $R \in \beta$ is a rectangle, then $H(R) \in \beta$ and its measure under μ satisfies:

$$\mu(H(R)) = \mu^u(\partial^+ H(R)) \cdot \mu^s(\partial^- H(R)) = 2\mu^u(\partial^+ R)\left(\frac{1}{2}\mu^s(\partial^- R)\right) = \mu(R).$$

For example, let $R_0 \in \beta$ be the rectangle of X that satisfies:

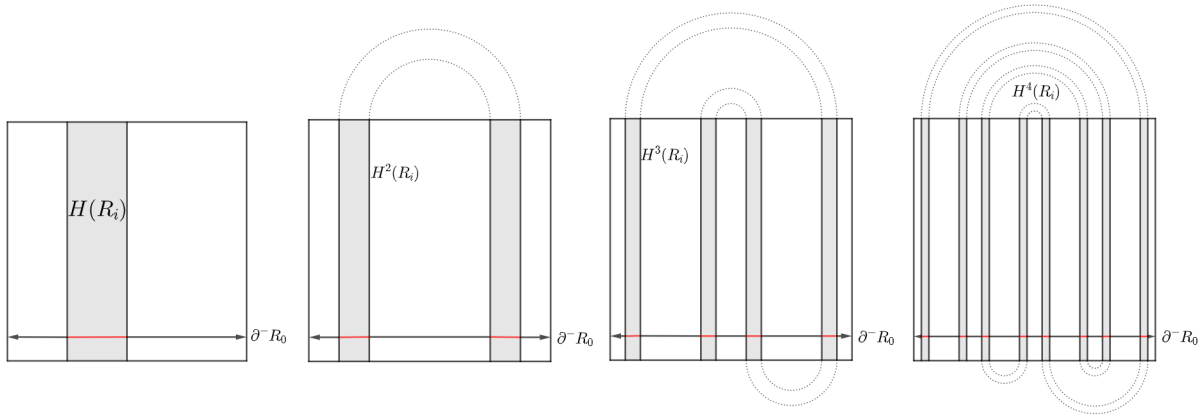
1. $\partial^+ R_0 \subset \lim_{n\to\infty} H^n(\mathcal{L}^+)$.
2. $\partial^- R_0 \subset \lim_{n\to\infty} H^{-n}(\mathcal{L}^-)$.
3. If $R \in \beta$ then $R \subseteq R_0$.

This is the "largest" rectangle that can be embedded in X. Let us again consider the Markov partition R_1, R_2, R_3, and R_4 as defined in Section 2.4 (see Figure 2.12a.), then $R_0 = R_1 \cup \cdots \cup R_4$. Let us calculate the measures $\mu^s(\partial^- R_0)$ and $\mu^u(\partial^+ R_0)$ using equations (2.8) and (2.9) given in Section 2.4:

$$\mu^s(\partial^- R_0) = \lim_{m\to\infty} \frac{1}{2^m} \sum_{i=1}^{4} u_{i,m}.$$

Here, $u_{i,m}$ is the number of components of $\partial^- R_0 \cap H^m(R_i)$ for $i = 1, ..., 4$. Therefore, we have:

$$u_{i,1} = 1 \ ; \ u_{i,2} = 2 \ ; \ u_{i,3} = 4 \ ; \ u_{i,4} = 8 \ ; \ \cdots \ ; \ u_{i,m} = 2^{m-1} \text{ (See Figure 3.1)}$$



(a) $u_{i,1} = 1$ (b) $u_{i,2} = 2$ (c) $u_{i,3} = 4$ (d) $u_{i,4} = 8$

Figure 3.1

So, we have $\mu^s(\partial^- R_0) = \lim_{m\to\infty} \dfrac{4 \cdot 2^{m-1}}{2^m} = 2$, and similarly we have $\mu^u(\partial^+ R_0) = 2$. Therefore, $\mu(R_0) = 4$.

We can then redefine the measure of a rectangle R as $m(R)$ such that for all $R \in \beta$, $0 \leq m(R) \leq 1$, by setting:

$$m(R) = \frac{\mu(R)}{\mu(R_0)}.$$

Before defining the entropy of the horseshoe map, let us introduce some necessary concepts in the context we are interested in.

A **partition** of R_0 is a **finite** collection of rectangles in X such that their union is R_0. Moreover, if $\xi = \{A_1, ..., A_n\}$ and $\eta = \{C_1, ..., C_k\}$ are two partitions of R_0, then we define their **union** as:

$$\xi \vee \eta := \{A_i \cap C_j \ / \ 1 \leq i \leq n \text{ and } 1 \leq j \leq k\}.$$

If $\xi = \{A_1, ..., A_n\}$ is a partition of R_0, then for the horseshoe map H and each $n \in \mathbb{N}$, we define the partition $H^{-n}(\xi) = \{H^{-n}(A_1), ..., H^{-n}(A_n)\}$.

The **entropy** of the partition ξ is then defined as follows:

$$E(\xi) := -\sum_{i=1}^{n} m(A_i) \cdot \log(m(A_i)). \tag{3.1}$$

Following this way, the **entropy** of the horseshoe map H with respect to a partition ξ of R_0 is defined as:

$$e(H, \xi) := \lim_{n\to\infty} \frac{1}{n} E\left(\bigvee_{i=0}^{n-1} H^{-i}(\xi)\right), \tag{3.2}$$

where $\bigvee_{i=0}^{n-1} H^{-i}(\xi)$ is the partition $\xi \vee H^{-1}(\xi) \vee \cdots \vee H^{-(n-1)}(\xi)$.

Finally, the **entropy** of the horseshoe map H is defined as $e(H) := \sup_{\xi} e(H, \xi)$, where the supremum is taken over all possible partitions of R_0.

A method for explicitly computing the entropy of the horseshoe map is provided by the **Kolmogorov-Sinai theorem** (see theorem 4.17 from [14]), which states that if T is an invertible transformation that preserves measure on a probability space X with σ-algebra τ and measure m, and if $\mathcal{A}$ is a finite subalgebra of τ such that $\bigvee_{n=-\infty}^{\infty} T^n(\mathcal{A}) = \tau$, then $e(T) = e(T, \mathcal{A})$.

In the context of the horseshoe map, if τ is the σ-algebra of R_0 generated by the set β of all rectangles in X, and $\mathcal{A}$ is the finite subalgebra generated by the Markov partition $\xi = \{R_1, R_2, R_3, R_4\}$ defined in Section 2.4, then by the Kolmogorov-Sinai theorem we have:

$$e(H) = e(H, \xi) = \lim_{n\to\infty} \frac{1}{n} E\left(\bigvee_{i=0}^{n-1} T^{-i}(\xi)\right).$$

Let us calculate the value of $\frac{1}{n} E\left(\bigvee_{i=0}^{n-1} T^{-i}(\xi)\right)$ for some values of n:

- For the case $n = 1$, we have that $\bigvee_{i=0}^{0} H^{-i}(\xi) = \xi$ is the Markov partition shown in Figure 3.2. Furthermore, note that for each $i = 1, 2, 3, 4$, the transversal measure $\mu^s(\partial^- R_i)$ and $\mu^u(\partial^+ R_i)$ satisfy:

$$\mu^s(\partial^- R_i) = \mu^u(\partial^+ R_i) = \frac{\mu^s(\partial^- R_0)}{2} = \frac{\mu^u(\partial^+ R_0)}{2} = 1.$$

Therefore, $\mu(R_i) = 1$ for $i = 1, 2, 3, 4$, and consequently, $m(R_i) = \frac{1}{4}$.

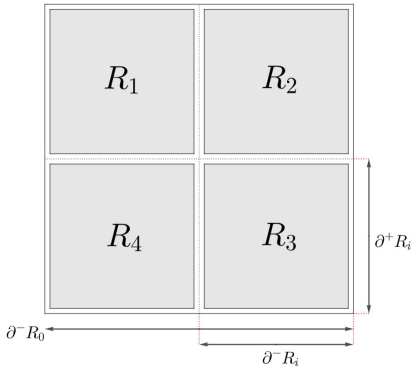


Figure 3.2: Partition ξ.

We thus have:

$$E(\xi) = -\sum_{i=1}^{4} m(R_i) \cdot \log(m(R_i)) = -\frac{1}{4} \cdot \log(1/4) \cdot 4 = 2 \cdot \log(2).$$

- If $n = 2$, then $H^{-1}(\xi) = \{H^{-1}(R_1), H^{-1}(R_2), H^{-1}(R_3), H^{-1}(R_4)\}$, and thus

$$\bigvee_{i=0}^{1} H^{-i}(\xi) = \xi \vee H^{-1}(\xi).$$

is the partition shown in Figure 3.3:

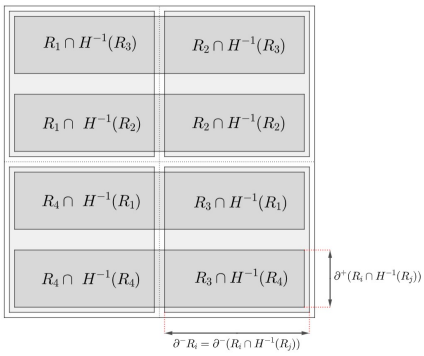


Figure 3.3: Partition $\xi \vee H^{-1}(\xi)$.

Now, since H is a **pseudo-Anosov** automorphism that contracts the foliation $(\mathcal{F}^-, \mu^s)$ and expands the foliation $(\mathcal{F}^+, \mu^u)$ by a factor of 2, then H^{-1} contracts $\mathcal{F}^+$ and expands $\mathcal{F}^-$. Thus, we have:

$$\mu^u(\partial^+ H^{-1}(R_j)) = \mu^u(\partial^+(R_i \cap H^{-1}(R_j))) = \frac{\mu^u(\partial^+ R_j)}{2} = \frac{1}{2}.$$

On the other hand, the components of $\partial^-(R_i \cap H^{-1}(R_j))$ are homotopic to the components of $\partial^- R_i$ along the foliation $\mathcal{F}^+$ and satisfy property 2 of definition 2.4.1 of **transversal measure**. Therefore, we have:

$$\mu^s(\partial^-(R_i \cap H^{-1}(R_j))) = \mu^s(\partial^- R_i) = 1.$$

Therefore, $\mu(R_i \cap H^{-1}(R_j)) = \frac{1}{2}$ and $m(R_i \cap H^{-1}(R_j)) = \frac{1}{8}$. Thus, we have:

$$\frac{1}{2}E(\xi \vee H^{-1}(\xi)) = -\sum_{i=1}^{4}\left(\sum_{j=1}^{4} m(R_i \cap H^{-1}(R_j)) \cdot \log(m(R_i \cap H^{-1}(R_j)))\right)$$
$$= \frac{1}{2}\left(-\frac{1}{8} \cdot \log(1/8) \cdot 8\right) = \frac{3}{2} \cdot \log(2).$$

- If $n = 3$, then $\bigvee_{i=0}^{2} H^{-i}(\xi) = \xi \vee H^{-1}(\xi) \vee H^{-2}(\xi)$ is the partition shown in Figure 3.4.

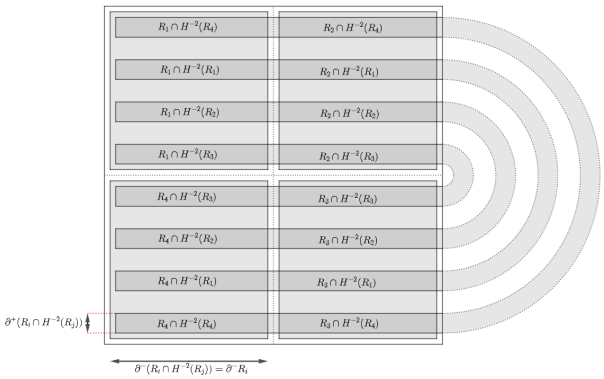


Figure 3.4: Partition $\xi \vee H^{-1}(\xi)$.

Again, since H^{-1} contracts the foliation $\mathcal{F}^+$ and expands the foliation $\mathcal{F}^-$, we have:

$$\mu^u(\partial^+(R_i \cap H^{-2}(R_j))) = \mu^u(\partial^+ H^{-2}(R_j)) = \frac{\mu^u(\partial^+ R_j)}{2} = \frac{1}{4}.$$

Now, similarly to the case $n = 2$, the components of $\partial^-(R_i \cap H^{-2}(R_j))$ are homotopic to the components of $\partial^- R_i$ along the foliation $\mathcal{F}^+$, thus they have the same transversal measure:

$$\mu^s(\partial^-(R_i \cap H^{-2}(R_j))) = \mu^s(\partial^- R_i) = 1.$$

Thus, $\mu(R_i \cap H^{-2}(R_j)) = \frac{1}{4}$, and therefore, $m(R_i \cap H^{-2}(R_j)) = \frac{1}{16}$. Hence, we have:

$$\frac{1}{3}E(\xi \vee H^{-1}(\xi) \vee H^{-2}) = -\sum_{j=1}^{4}\left(\sum_{i=1}^{4} m(R_j \cap H^{-2}(R_i)) \cdot \log(m(R_j \cap H^{-2}(R_i)))\right)$$
$$= \frac{1}{3}\left(-\frac{1}{16} \cdot \log(1/16) \cdot 16\right) = \frac{4}{3} \cdot \log(2).$$

In general, for $n \in \mathbb{N}$ we have that $E\left(\bigvee_{i=0}^{n-1} H^{-i}(\xi)\right) = (n+1)\cdot\log(2)$ and therefore, the entropy of the horseshoe map is:

$$e(H) = e(H,\xi) = \lim_{n\to\infty} \frac{n+1}{n}\cdot\log(2) = \log(2).$$

As the entropy of a system measures how orbits diverge, this result implies that the dynamical system given by successive iterations of the horseshoe map on the surface X exhibits some sensitivity to initial conditions. In the next section, we will study this behavior in detail.

3.2 Chaos

In the previous section, it was shown that the entropy of the horseshoe map is $\log(2)$. This result is sufficient to guarantee that the horseshoe map, as an automorphism of the surface X, exhibits chaotic dynamics.

In this section, we will demonstrate this result from a different perspective.

3.2.1 Cantor Sets on the Square

Let $Q = \{z \in \mathbb{C} \,/\, 0 \leq \mathrm{Re}(z) \leq 1 \text{ and } 0 \leq \mathrm{Im}(z) \leq 1\}$ be the unit square. On each side of Q, we consider a **Cantor ternary set** denoted by $\mathcal{C}^d$ for the bottom side, $\mathcal{C}^u$ for the top side, $\mathcal{C}^l$ for the left side, and $\mathcal{C}^r$ for the right side.

Since $\mathcal{C}^u$ is the translation of $\mathcal{C}^d$ by one unit upwards, each point $x \in \mathcal{C}^d$ is connected to the point $x+i \in \mathcal{C}^u$ by a line segment. Now, since the set $\mathcal{C}^u$ is symmetric with respect to the point $\frac{1}{2}+i$, it is possible to connect each point $x+i \in \mathcal{C}^u$ to the point $(1-x)+i \in \mathcal{C}^u$ by a semicircle contained in the exterior of Q.

Now, we divide the set $\mathcal{C}^d$ into two parts: $\mathcal{C}_0^d$ contained in the segment $\left[0;\frac{1}{2}\right]$ and $\mathcal{C}_1^d$ contained in the segment $\left[\frac{1}{2};1\right]$. The set $\mathcal{C}_1^d$ is symmetric with respect to $\frac{5}{6}$, so each pair of symmetric points in $\mathcal{C}_1^d$ is joined by a semicircle contained in the exterior of Q. Similarly, $\mathcal{C}_0^d$ is divided into two parts, a left part contained in the segment $\left[0;\frac{1}{6}\right]$ and a right part contained in the segment $\left[\frac{1}{6};\frac{1}{3}\right]$. Again, the right part is symmetric with respect to the point $\frac{5}{18}$, so as in the previous step, each pair of symmetric points is connected by a semicircle contained in the exterior of Q. Continuing this process on $\mathcal{C}^d$ and joining the respective segments and semicircles, we obtain a curve that we denote by $\mathcal{K}^+$.

Similarly, each point $iy \in \mathcal{C}^l$ is connected to the point $1+iy \in \mathcal{C}^r$ by a straight line segment, and since $\mathcal{C}^r$ is symmetric with respect to $1+\frac{i}{2}$, each pair of symmetric points is connected by a semicircle contained in the exterior of Q.

On the other hand, $\mathcal{C}^l$ is divided into two parts, an upper part contained in the segment

$\left[\frac{i}{2}; i\right]$ and a lower part contained in $\left[0; \frac{i}{2}\right]$. The upper part is symmetric with respect to the point $\frac{5i}{6}$, so each pair of symmetric points can be connected by a semicircle exterior to Q. Following the same procedure as with $\mathcal{C}^d$ and connecting the respective semicircles with the segments, we obtain the set $\mathcal{K}^-$. Figure 3.5 shows both sets $\mathcal{K}^+$ and $\mathcal{K}^-$.

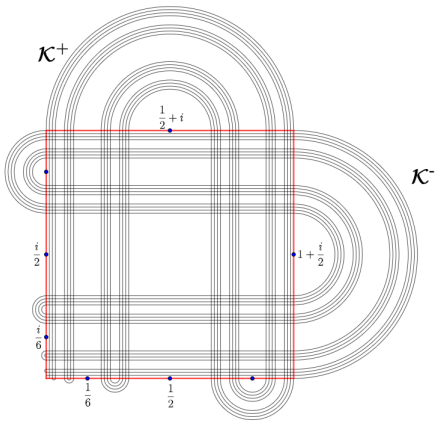


Figure 3.5: Sets $\mathcal{K}^+$ y $\mathcal{K}^-$.

Recall that a **continuum** is a compact and connected metric space. Moreover, we say that a continuum C is indecomposable if for any pair of not necessarily disjoint subcontinua A and B of C such that $C = A \cup B$, it holds that $A = C$ or $B = C$.

In [2], it is proven that the sets $\mathcal{K}^+$ and $\mathcal{K}^-$ are indecomposable continua known as **Knaster continua**.

The horseshoe map H acts on $\mathcal{K} = \mathcal{K}^+ \cup \mathcal{K}^-$ as a rotation around the point $1 + \frac{i}{2}$ followed by a horizontal contraction and a vertical stretching as shown in the figure 3.6.

Let $I_0 := \left\{ z \in Q \ / \ 0 \leq Re(z) \leq \frac{1}{2} \right\}$ and $I_1 := \left\{ z \in Q \ / \ \frac{1}{2} \leq Re(z) \leq 1 \right\}$ be two non-disjoint sub-rectangles of Q such that $Q = I_0 \cup I_1$. Based on this, we consider the following sets:

$$I_{0,0} = \left[0, \frac{1}{2}\right] \times \left[0, \frac{1}{2}\right]; \ I_{0,1} = \left[0, \frac{1}{2}\right] \times \left[\frac{1}{2}, 1\right]; \ I_{1,0} = \left[\frac{1}{2}, 1\right] \times \left[0, \frac{1}{2}\right]; \ I_{1,1} = \left[\frac{1}{2}, 1\right] \times \left[\frac{1}{2}, 1\right].$$

Then, the points of $\Lambda = \mathcal{K}^- \cap \mathcal{K}^+$ contained in $I_{0,0}$ are mapped to the set $\left[0, \frac{1}{6}\right] \times [0, 1] \subset I_0$, those in $I_{1,0}$ are mapped to $\left[\frac{1}{6}, \frac{1}{2}\right] \times [0, 1] \subset I_0$, those in $I_{1,1}$ are mapped to $\left[\frac{1}{2}, \frac{5}{6}\right] \times [0, 1] \subset I_1$, and those in $I_{0,1}$ are mapped to $\left[\frac{5}{6}, 1\right] \times [0, 1] \subset I_1$.

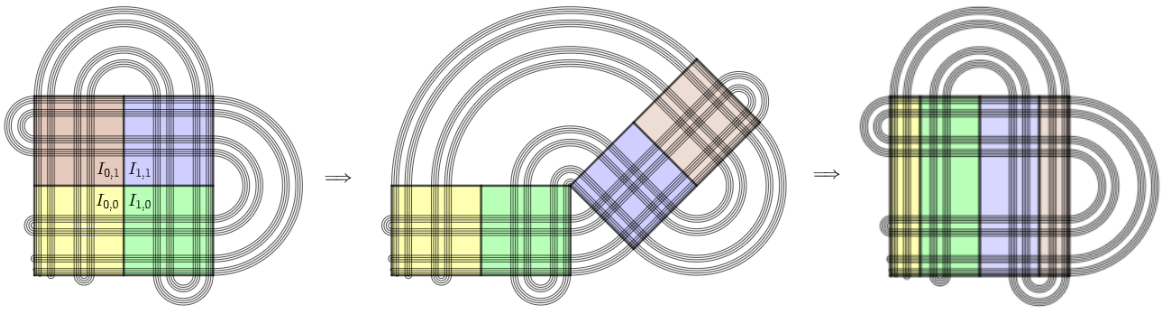


Figure 3.6: Horseshoe map on $\mathcal{K}$.

3.2.2 The Horseshoe map H is chaotic on the surface X

In this section, we will emulate the method presented in section 4.2 from [5] on the symbolic dynamics of the horseshoe. We will encode the trajectory followed by each point of the set $\mathcal{K}^- \cap \mathcal{K}^+$ when we apply the horseshoe map, H, with the goal of demonstrating that H inherits the dynamical properties of the shift map, which, by Theorem 1.2.2, is chaotic.

In order to study chaos as one of the main dynamic properties of the automorphism H defined in section 2.1, we first consider its action on the set Λ, defined as follows:

$$\Lambda := \mathcal{K}^- \cap \mathcal{K}^+.$$

To encode the orbit followed by a point $z \in \Lambda$ under iteration by H, we again consider the sets I_0 and I_1, and define the function $\Psi : \Lambda \longrightarrow \Sigma$ such that if $z \in \Lambda$, then:

$$\Psi(z) = (..., s_{-2}, s_{-1}|s_0, s_1, s_2, ...) \text{ where } s_i = 0 \text{ if } H^i(z) \in I_0 \text{ and } s_i = 1 \text{ if } H^i(z) \in I_1.$$

Let's prove that Ψ is a homeomorphism between Λ and Σ:

Before starting the proof, let's take the set of **midpoints** of the intervals removed in the construction of each Cantor set ($\mathcal{C}^d$ and $\mathcal{C}^l$) on the left and bottom sides of Q, that is:

$$M_{\mathcal{C}^d} := \left\{..., \frac{1}{54}, \frac{1}{18}, \frac{1}{6}, \frac{1}{2}, \frac{5}{6}, \frac{17}{18}, \frac{53}{54}, ...\right\} \cup \{0, 1\}.$$

$$M_{\mathcal{C}^l} := \left\{..., \frac{i}{54}, \frac{i}{18}, \frac{i}{6}, \frac{i}{2}, \frac{5i}{6}, \frac{17i}{18}, \frac{53i}{54}, ...\right\} \cup \{0, i\}.$$

Definition 3.2.1 *Let $r_1, r_2 \in M_{\mathcal{C}^d}$ and $t_1, t_2 \in M_{\mathcal{C}^l}$. We denote by $\mathbf{R}$ the set of all rectangles of the form:*

$$[r_1, r_2] \times [t_1, t_2] = \{z \in Q \ / \ r_1 \leq Re(z) \leq r_2 \ y \ t_1 \leq Im(z) \leq t_2\}.$$

So, starting from $I_{s_0} = I_0$ or $I_{s_0} = I_1$, we define I_{s_0,s_1} as the smallest rectangle in $\mathbf{R}$ that contains the set $\{z \in I_{s_0} \cap \Lambda \ / \ H(z) \in I_{s_1}\}$, i.e., for $s_0 \in \{0, 1\}$ and $s_1 \in \{0, 1\}$ we have $I_{s_0,s_1} \subset I_{s_0}$.

Inductively, given the rectangle $I_{s_0,s_1,...,s_n}$, we define $I_{s_0,s_1,...,s_n,s_{n+1}}$ as the smallest rectangle in $\mathbf{R}$ that contains the set $\{z \in I_{s_0,s_1,...,s_n} \cap \Lambda \ / \ H^{n+1}(z) \in I_{s_{n+1}}\}$.

Similarly, we define the set $I_{s_0}^{s_{-1}}$ as the smallest rectangle in $\mathbf{R}$ that contains the set:

$$\{z \in I_{s_0} \cap \Lambda \ / \ H^{-1}(z) \in I_{s_{-1}}\}.$$

And inductively, given the rectangle $I_{s_0}^{s_{-1},s_{-2},\ldots,s_{-(n-1)}}$, we define $I_{s_0}^{s_{-1},s_{-2},\ldots,s_{-n}}$ as the smallest rectangle in $\mathbf{R}$ that contains the set $\{z \in I_{s_0}^{s_{-1},s_{-2},\ldots,s_{-(n-1)}} \cap \Lambda \ / \ H^{-n}(z) \in I_{s_{-n}}\}$. (See Figure 3.7)

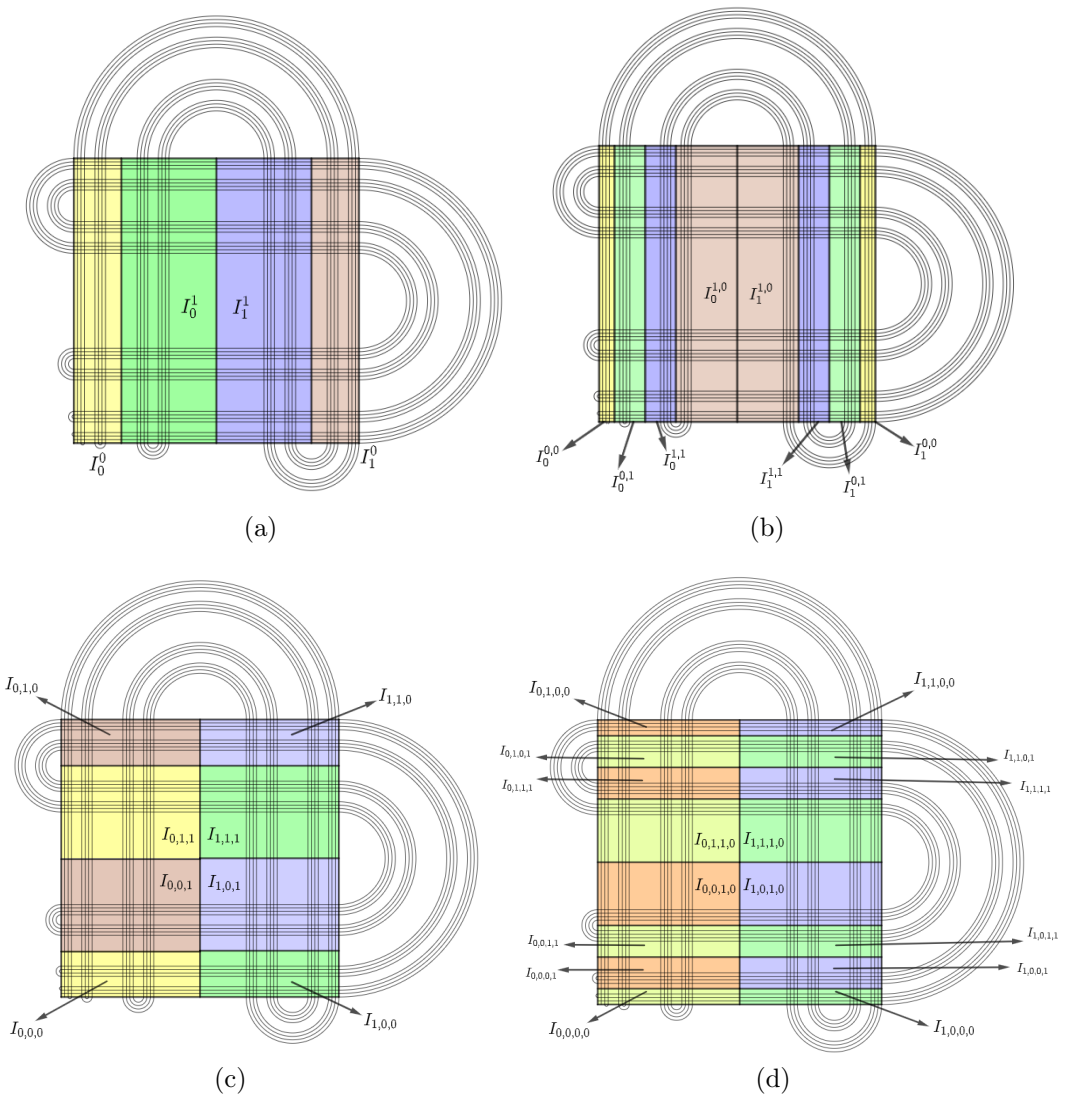


Figure 3.7

In this way, given the sequence $\{s_0, s_1, s_2, ...\}$ of 0's and 1's, we obtain a sequence of nested closed rectangles:

$$I_{s_0} \supset I_{s_0,s_1} \supset I_{s_0,s_1,s_2} \supset \cdots$$

In the same way, if for the sequence $\{\cdots, s_{-2}, s_{-1}, s_0\}$ we obtain the sequence of nested closed rectangles:

$$I_{s_0} \supset I_{s_0}^{s_{-1}} \supset I_{s_0}^{s_{-1},s_{-2}} \supset \cdots$$

The intersection between $I_{s_0,s_1,\ldots,s_n}$ and $I_{s_0}^{s_{-1},s_{-2},\ldots,s_{-n}}$ is denoted as $I_{s_0,s_1,\ldots,s_n}^{s_{-1},s_{-2},\ldots,s_{-n}}$.

These rectangles will be useful in proving the following theorem, with which we can assign a sequence of 0's and 1's from the sequence space Σ to each point in the set Λ. This

sequence will then serve to encode the orbit followed by each point of Λ when iterating the horseshoe function.

Theorem 3.2.1 *The function $\Psi : \Lambda \longrightarrow \Sigma$ is a homeomorphism.*

Proof.

1. Ψ **is bijective**: Given a sequence $\mathbf{s} = (\cdots, s_{-2}, s_{-1}|s_0, s_1, s_2, ...) \in \Sigma$, let us consider the sequences of nested rectangles

$$I_{s_0} \supset I_{s_0,s_1} \supset I_{s_0,s_1,s_2} \supset \cdots$$

$$I_{s_0} \supset I_{s_0}^{s_{-1}} \supset I_{s_0}^{s_{-1},s_{-2}} \supset \cdots$$

As each rectangle in the sequences is a non-empty closed set, by the **Cantor's nested intervals theorem**, we obtain that the intersections $I_{s_0,s_1,s_2,...}$ and $I_{s_0}^{s_{-1},s_{-2},s_{-3},\cdots}$ are non-empty. Moreover, since for a fixed s_0, $I_{s_0,s_1,s_2,...} \cap I_{s_0}^{s_{-1},s_{-2},\cdots} \neq \emptyset$, we have that there exists $x \in I_{s_0,s_1,s_2,...}^{s_{-1},s_{-2},s_{-3},\cdots} \cap \Lambda$ such that $\Psi(x) = \mathbf{s}$.

On the other hand, the rectangles $I_{s_0,s_1,s_2,...,s_n}^{s_{-1},s_{-2},...,s_{-n}}$ belong to $\mathbf{R}$ for all $n \in \mathbb{N}$, that is, there exist sequences of numbers $r_n, r_n' \in M_{\mathcal{C}^d}$ and $t_n, t_n' \in M_{\mathcal{C}^l}$ such that $|r_n - r_n'|$ converges to 0 and $|t_n - t_n'| \to 0$ as $n \to \infty$, and such that:

$$[r_n, r_n'] \times [t_n, t_n'] = I_{s_0,s_1,s_2,...,s_n}^{s_{-1},s_{-2},...,s_{-n}}.$$

Therefore:

$$I_{s_0,s_1,s_2,...,s_n,...}^{s_{-1},s_{-2},...,s_{-n},\cdots} = \{x\}.$$

Therefore, Ψ is a **bijective** function.

2. Ψ **is continuous**: Let $x \in \Lambda$ such that $\Psi(x) = (..., s_{-2}, s_{-1}|s_0, s_1, s_2, ...)$ and let $\epsilon > 0$ be given. Let $N \in \mathbb{N}$ such that $\dfrac{1}{2^N} < \epsilon$, then by theorem 1.2.1, we have:

$$x \in I_{s_0,s_1,s_2,...,s_N}^{s_{-1},s_{-2},...,s_{-N}} = [r_1, r_2] \times [t_1, t_2] \in \mathbf{R}.$$

So taking $\delta = \dfrac{\min\{|r_1 - r_2|, |t_1 - t_2|\}}{2}$ we have that the open ball of radius δ centered at x, $B(x, \delta)$, is contained in $I_{s_0,s_1,s_2,...,s_N}^{s_{-1},s_{-2},...,s_{-N}}$. If we take $y \in B(x, \delta)$ and $\Psi(y) = (..., u_{-2}, u_{-1}|u_0, u_1, u_2, ...)$ then $s_i = u_i$ for all $i \in \mathbb{Z}$ such that $|i| \leq N$. Again by theorem 1.2.1 we have that $d[\Psi(x), \Psi(y)] \leq \dfrac{1}{2^N}$ and therefore Ψ is a continuous function.

□

On the other hand, Theorem 1.2.2 guarantees that the shift function $\sigma : \Sigma \longrightarrow \Sigma$ satisfies the 3 items of the definition of **chaos**. To prove that the horseshoe map is also chaotic, let's see that the following diagram commutes:

$$\begin{array}{ccc} \Lambda & \xrightarrow{H} & \Lambda \\ {\scriptstyle\Psi}\downarrow & & \downarrow{\scriptstyle\Psi} \\ \Sigma & \xrightarrow[\sigma]{} & \Sigma \end{array}$$

To see this, let $z \in \Lambda$ such that:

$$\Psi(z) = (..., s_{-2}, s_{-1}|s_0, s_1, s_2, ...) \text{ and } \Psi(H(z)) = (..., t_{-2}, t_{-1}|t_0, t_1, t_2, ...)$$

Therefore, $H^{j+1}(z) \in I_{s_{j+1}}$ and since $H^{j+1}(z) = H^j(H(z))$, then $H^{j+1}(z) \in I_{t_j}$, so:

$$s_{j+1} = t_j \text{ for all } j \in \mathbb{Z}.$$

And therefore $\sigma(\Psi(z)) = \Psi(H(z))$, i.e., $\sigma \circ \Psi = \Psi \circ H$. Then the diagram commutes and therefore H exhibits chaotic dynamics over Λ.

3.2.3 Collapsing the Cantor Sets

Let us consider again the unit square Q, the Cantor sets $\mathcal{C}^d$, $\mathcal{C}^r$, $\mathcal{C}^u$, and $\mathcal{C}^l$ defined on the boundary of Q, the sets $\mathcal{K}^+$ and $\mathcal{K}^-$ defined in subsection 3.2.2. using these Cantor sets, and the set $\Lambda = \mathcal{K}^+ \cap \mathcal{K}^-$. It is clear that if $p + iq \in \Lambda$, then $p \in \mathcal{C}^d$ and $q \in \mathcal{C}^l$, so that Λ can be viewed as the Cartesian product of Cantor sets, $\mathcal{C}^d \times \mathcal{C}^l$.

In the previous section, it was shown that Λ is homeomorphic to the sequence space Σ. Now, as it is well known that the Cantor set is equipotent to the interval $[0, 1]$, let us see now that it is possible to "collapse" the spaces between points of the Cantor set $\mathcal{C}^d$, resulting in a space homeomorphic to the interval $[0, 1]$.

To do this, we will first define a continuous function that maps the Cantor set onto the interval $[0, 1]$:

Definition 3.2.2 *The **Cantor function** is defined as the limit of a sequence of continuous functions $\{f_n : [0,1] \longrightarrow [0,1]\}_{n\in\mathbb{N}}$ as follows:*

- *Let $f_1(x) = x$.*
- *For each $n \geq 1$, the function $f_{n+1} : [0,1] \longrightarrow [0,1]$ is defined in terms of f_n as follows:*

$$f_{n+1}(x) := \begin{cases} \frac{1}{2} \cdot f_n(3x) & \text{If} \quad 0 \leq x \leq \frac{1}{3} \\ \frac{1}{2} & \text{If} \quad \frac{1}{3} < x \leq \frac{2}{3} \\ \frac{1}{2} + \frac{1}{2} \cdot f_n(3x-2) & \text{If} \quad \frac{2}{3} < x \leq 1 \end{cases} \tag{3.3}$$

*The graphs of the functions f_1, f_2, f_3, and f_4 are shown in Figure 3.8. Thus, for $x \in [0,1]$, we define the **Cantor function** as $f(x) = \lim_{n\to\infty} f_n(x)$.*

Note that $\max_{x\in[0,1]} |f_{n+1}(x) - f_n(x)| \leq \frac{1}{2} \max_{x\in[0,1]} |f_n(x) - f_{n-1}(x)|$, $n \geq 2$. Therefore, for $n \geq 1$ we have that:

$$\max_{x\in[0,1]} |f(x) - f_n(x)| \leq \frac{1}{2^{n-1}} \max_{x\in[0,1]} |f_2(x) - f_1(x)|$$

This implies that $\{f_n\}_{n\in\mathbb{N}}$ converges uniformly to $f(x)$.

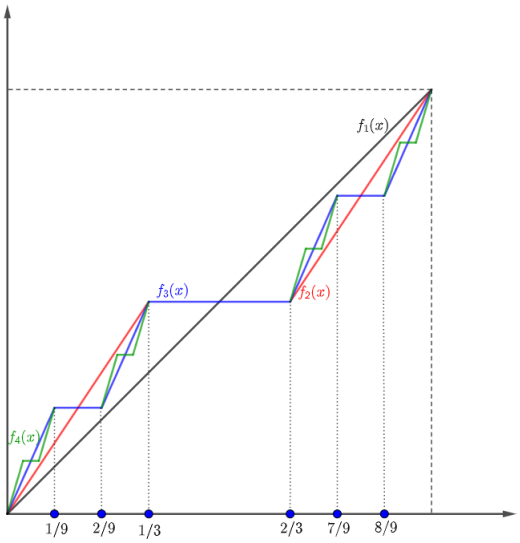


Figure 3.8: Graphs of f_1, f_2, f_3 and f_4.

As $f_n(x)$ is continuous for all $n \in \mathbb{N}$ and , then the Cantor function is continuous on $[0,1]$. Moreover, if $\mathcal{C}$ denotes the Cantor ternary set, then the Cantor function f restricted to $\mathcal{C}$ is continuous on $\mathcal{C}$ and surjective, i.e., for every $y \in [0,1]$, there exists $x \in \mathcal{C}$ such that $f(x) = y$.

We will now construct the **Cantor ternary set,** denoted by $\mathcal{C}$ step by step, as follows: Consider the unit interval $U = [0,1]$.

- In step $n = 1$, we extract from U the central interval $\mathcal{G}_1^1 = \left(\frac{1}{3}, \frac{2}{3}\right)$ of length $1/3$, resulting in the set:

$$C_1 = U \setminus \mathcal{G}_1^1.$$

- In step $n = 2$, we remove from C_1 the central intervals $\mathcal{G}_1^2 = \left(\frac{1}{9}, \frac{2}{9}\right)$ and $\mathcal{G}_2^2 = \left(\frac{7}{9}, \frac{8}{9}\right)$ of length $\frac{1}{3^2}$ each, resulting in the set:

$$C_2 = C_1 \setminus (\mathcal{G}_1^2 \cup \mathcal{G}_2^2).$$

- In step $n = 3$, we extract from the set C_2 the central intervals $\mathcal{G}_1^3 = \left(\frac{1}{27}, \frac{2}{27}\right)$, $\mathcal{G}_2^3 = \left(\frac{7}{27}, \frac{8}{27}\right)$, $\mathcal{G}_3^3 = \left(\frac{19}{27}, \frac{20}{27}\right)$, and $\mathcal{G}_4^3 = \left(\frac{25}{27}, \frac{26}{27}\right)$ of length $\frac{1}{3^3}$, resulting in the set:

$$C_3 = C_2 \setminus \left(\bigcup_{i=1}^{4} \mathcal{G}_i^3\right).$$

- In the $n = k$ step, we extract from the set C_{k-1} the complementary intervals $\mathcal{G}_1^k, ..., \mathcal{G}_{2^{k-1}}^k$ of length $\frac{1}{3^k}$, resulting in the set:

$$C_k = C_{k-1} \setminus \left(\bigcup_{i=1}^{2^{k-1}} \mathcal{G}_i^k \right).$$

The **Cantor ternary set** is defined as $\mathcal{C} := \bigcap_{i=1}^{\infty} C_i$. Let $\overline{\mathcal{G}}_i^k$ be the closure of the interval $\mathcal{G}i^k$. Then, the endpoints of all the closed intervals $\overline{\mathcal{G}}_i^k$ belong to $\mathcal{C}$, and we call these points the **Cantor numbers of generation** k. For example, $\frac{1}{3}$ and $\frac{2}{3}$ belong to the first generation, $\frac{1}{9}$, $\frac{2}{9}$, $\frac{7}{9}$ and $\frac{8}{9}$ belong to the second generation, and so on.

Note that all Cantor numbers belonging to some generation are rational numbers. However, there are numbers in $\mathcal{C}$ that do not belong to any generation, in particular 0, 1, and all irrational numbers that are in $\mathcal{C}$. Let us now define a relation $\sim_{\mathcal{C}}$ on the Cantor set $\mathcal{C}$.

Definition 3.2.3 *Let $x, y \in \mathcal{C}$, then $x \sim_{\mathcal{C}} y$ if and only if x and y belong to the same generation and are the endpoints of some interval $\overline{\mathcal{G}}_i^k$. On the other hand, if x does not belong to any generation, then x is related only to itself, that is, $x \sim_{\mathcal{C}} x$. For example,* 0 *and* 1 *are related only to themselves, as well as every irrational number in $\mathcal{C}$.*

Under this relation, we have that $\frac{1}{3} \sim_{\mathcal{C}} \frac{2}{3}$, $\frac{1}{9} \sim_{\mathcal{C}} \frac{1}{9} \sim_{\mathcal{C}} \frac{2}{9}$, etc.

Proposition 3.2.1 *The relation $\sim_{\mathcal{C}}$ is an equivalence relation defined on the Cantor set $\mathcal{C}$.*

Proof.

1. **Reflexivity**: Let $x \in \mathcal{C}$. If x does not belong to any generation, then by definition $x \sim_{\mathcal{C}} x$.

   On the other hand, if $x \in \mathcal{C}$ is a rational number and belongs to some generation, then it is an endpoint of some closed interval $\overline{\mathcal{G}}_i^k$. Hence $x \sim_{\mathcal{C}} x$.

2. **Symmetry**: Let $x, y \in \mathcal{C}$ such that $x \sim_{\mathcal{C}} y$. If x does not belong to any generation, then $y = x$ and the symmetry is evident. On the other hand, if x and y are rational numbers such that they belong to the same generation and are the endpoints of some closed interval $\overline{\mathcal{G}}_i^k$, hence $y \sim_{\mathcal{C}} x$.

3. **Transitivity**: Let $x, y, z \in \mathcal{C}$ such that $x \sim_{\mathcal{C}} y$ and $y \sim_{\mathcal{C}} z$, and suppose they are rational such that they belong to the same generation (if not, the statement is trivial). Then, x and y are the endpoints of some interval $\overline{\mathcal{G}}_i^k$, and y and z are the endpoints of the same interval $\overline{\mathcal{G}}_i^k$. Therefore, $x \sim_{\mathcal{C}} z$.

We then have that the relation $\sim_{\mathcal{C}}$ is an equivalence relation.

□

Let us now consider the quotient space $\mathcal{C}/\sim_{\mathcal{C}}$. We know that the Cantor set has a topology inherited from the interval $[0, 1]$, and therefore $\mathcal{C}/\sim_{\mathcal{C}}$ has a topology inherited from the Cantor set $\mathcal{C}$. We define a function $g : \mathcal{C}/\sim_{\mathcal{C}} \longrightarrow [0, 1]$ using the **Cantor function** from definition 3.2.2.

Let $\pi : \mathcal{C} \longrightarrow \mathcal{C}/\sim_{\mathcal{C}}$ be the projection that sends $x \in \mathcal{C}$ to its respective equivalence class $[x]$. The equivalence classes of $\mathcal{C}/\sim_{\mathcal{C}}$ are divided into two types:

- **Type 1**: Equivalence classes with a single element that corresponds to 0, 1 and the irrational numbers in $\mathcal{C}$.

- **Type 2**: Equivalence classes with two elements corresponding to the endpoints of some interval $\overline{\mathcal{G}}_i^k$.

Note that the Cantor function given in definition 3.2.2 is constant on each complementary interval $\overline{\mathcal{G}}_i^k$, so if $\overline{\mathcal{G}}_i^k = [x_1, x_2]$, then $f(x_1) = f(x_2)$.

We now define the function $g : \mathcal{C}/\sim_{\mathcal{C}} \longrightarrow [0,1]$ as follows:

- If $[x] \in \mathcal{C}/\sim_{\mathcal{C}}$ is an equivalence class of **Type 1**, then $[x] = \{x\}$ and thus $g([x]) = f(x)$.

- If $[x] \in \mathcal{C}/\sim_{\mathcal{C}}$ is an equivalence class of **Type 2**, then $[x] = \{x_1, x_2\}$ where x_1 and x_2 are the endpoints of some complementary interval $\overline{\mathcal{G}}_i^k$ and therefore $f(x_1) = f(x_2)$. Thus, it makes sense to define $g([x]) = f(x_1) = f(x_2)$.

The function $g : \mathcal{C}/\sim_{\mathcal{C}} \longrightarrow [0,1]$ is well-defined and inherits the properties of **continuity** and **surjectivity** of the Cantor function. Moreover, g is **injective**, since the points where f is not injective are identified by the relation $\sim_{\mathcal{C}}$.

On the other hand, Theorem 26.6 in the book [10] states that if $\varphi : X \longrightarrow Y$ is a continuous and bijective function between topological spaces where X is compact and Y is Hausdorff, then φ^{-1} is continuous. Thus, since $\mathcal{C}$ is compact and $\sim_{\mathcal{C}}$ is an equivalence relation, then $\mathcal{C}/\sim_{\mathcal{C}}$ is compact, and furthermore, $[0,1]$ is Hausdorff, so the function $g : \mathcal{C}/\sim_{\mathcal{C}} \longrightarrow [0,1]$ is a homeomorphism.

Intuitively, the equivalence relation $\sim_{\mathcal{C}}$ "collapses" the space between points in the Cantor set $\mathcal{C}$, yielding again the interval $[0,1]$.

Now, let us consider the set $\Lambda = \mathcal{K}^+ \cap \mathcal{K}^-$. We know that if $p + iq \in \Lambda$, then $p \in \mathcal{C}^d$ and $q \in \mathcal{C}^l$, both Cantor sets defined on the lower and left sides of the square Q. Therefore, we can consider Λ as the Cartesian product $\mathcal{C}^l \times \mathcal{C}^d$.

By "collapsing" each Cantor set defined on each side of the square Q using the respective relations $\sim_{\mathcal{C}^l}$, $\sim_{\mathcal{C}^d}$, $\sim_{\mathcal{C}^r}$, and $\sim_{\mathcal{C}^u}$, we obtain that these relations define an equivalence relation $\sim_\Lambda$ on Λ that "collapses" the spaces between points of Λ. Since $\mathcal{C}^d/\sim_{\mathcal{C}^d} \cong \mathcal{C}^l/\sim_{\mathcal{C}^l} \cong [0,1]$, then we have that $\Lambda/\sim_\Lambda = \mathcal{C}^d/\sim_{\mathcal{C}^d} \times \mathcal{C}^l/\sim_{\mathcal{C}^l} \cong [0,1] \times [0,1]$.

Figure 3.9 shows the set Λ represented by points on the square and the continua $\mathcal{K}^-$, $\mathcal{K}^+$ under the action of the relation $\sim_\Lambda$.

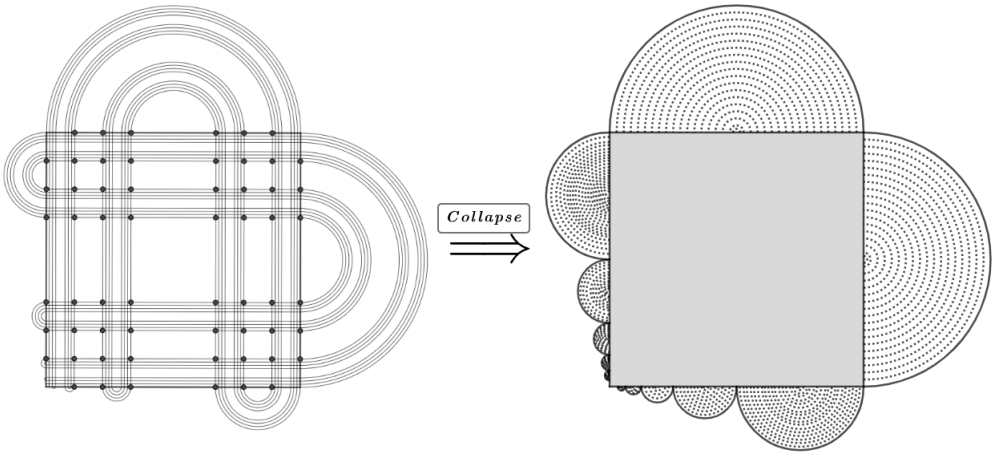


Figure 3.9: Collapsing Λ set.

It is clear then that we can obtain the surface X generated by the horseshoe map in section 2.3 starting from the set Λ, collapsing it and identifying points on the boundary that are connected by a path of semicircles. We will denote the space obtained after these identifications as ${}^{\Lambda^*}\!/\!\sim_\Lambda$. That is, ${}^{\Lambda^*}\!/\!\sim_\Lambda$ is homeomorphic to the surface X generated by the horseshoe function.

The function $H_\Lambda : \Lambda \longrightarrow \Lambda$ then defines a function $H^*_\Lambda : {}^{\Lambda^*}\!/\!\sim_\Lambda \longrightarrow {}^{\Lambda^*}\!/\!\sim_\Lambda$ such that if $H(u) = v$ then $H^*_\Lambda([u]) = [v]$. This function H^*_Λ is then equivalent to the horseshoe automorphism of the surface X defined in section 2.1. making the following diagram commute:

$$\begin{array}{ccc} {}^{\Lambda^*}\!/\!\sim_\Lambda & \xrightarrow{H^*_\Lambda} & {}^{\Lambda^*}\!/\!\sim_\Lambda \\ \downarrow & & \downarrow \\ X & \xrightarrow[H]{} & X \end{array}$$

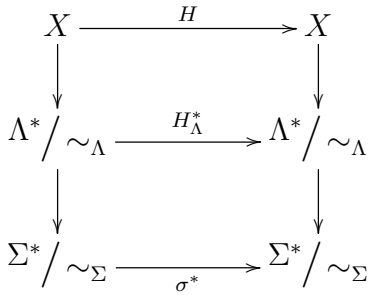
Now, since the function $\Psi : \Lambda \longrightarrow \Sigma$ defined in subsection 3.2.1 is a homeomorphism, we have that the relation $\sim_\Lambda$ defined on Λ defines in turn a relation $\sim_\Sigma$ defined on Σ such that:

$$\mathbf{s} \sim_\Sigma \mathbf{t} \text{ in } \Sigma \text{ if and only if } \Psi^{-1}(\mathbf{t}) \sim_\Lambda \Psi^{-1}(\mathbf{t}) \text{ in } \Lambda.$$

So, we conclude that ${}^{\Lambda}\!/\!\sim_\Lambda$ is homeomorphic to the space ${}^{\Sigma}\!/\!\sim_\Sigma$. Similarly, the points in ${}^{\Lambda}\!/\!\sim_\Lambda$ that are identified by paths of semicircles define identifications in the space ${}^{\Sigma}\!/\!\sim_\Sigma$ such that if $[x]$ and $[y]$ are connected in ${}^{\Lambda}\!/\!\sim_\Lambda$ by a path of semicircles, then $\Psi_{\sim_\Lambda}([x])$ is identified with $\Psi_{\sim_\Lambda}([y])$, where $\Psi_{\sim_\Lambda}$ is the homeomorphism between ${}^{\Lambda}\!/\!\sim_\Lambda$ and ${}^{\Sigma}\!/\!\sim_\Sigma$. Let ${}^{\Sigma^*}\!/\!\sim_\Sigma$ be the space obtained from ${}^{\Sigma}\!/\!\sim_\Sigma$ after these identifications. Then we have that ${}^{\Lambda^*}\!/\!\sim_\Lambda$ is homeomorphic to ${}^{\Sigma^*}\!/\!\sim_\Sigma$, and therefore X is homeomorphic to ${}^{\Sigma^*}\!/\!\sim_\Sigma$.

As with H_Λ, the shift map $\sigma : \Sigma \longrightarrow \Sigma$ also defines a function $\sigma^* : {}^{\Sigma^*}\!/\!\sim_\Sigma \longrightarrow {}^{\Sigma^*}\!/\!\sim_\Sigma$ on ${}^{\Sigma^*}\!/\!\sim_\Sigma$ such that if $\sigma(\mathbf{s}) = \mathbf{t}$ then $\sigma^*([\mathbf{s}]) = [\mathbf{t}]$. The function σ^* inherits the dynamical properties of the shift map σ and is also a chaotic function on ${}^{\Sigma^*}\!/\!\sim_\Sigma$.

As a consequence of all the above, we have that the following diagram commutes:



And therefore, the horseshoe map automorphism H of the surface X defined in section 2.1 has chaotic dynamics on X.

Bibliography

[1] Lars Valerian Ahlfors. *Conformal invariants: topics in geometric function theory*, volume 371. American Mathematical Soc., 2010.

[2] Javier Camargo and Rafael Isaacs. Continuos tipo Knaster y sus modelos geométricos. *Revista Colombiana de Matemáticas*, 47(1):67–81, 2013.

[3] Andrew J Casson and Steven A Bleiler. *Automorphisms of surfaces after Nielsen and Thurston*. Number 9 in London Mathematical Society Student Texts. Cambridge University Press, 1988.

[4] Reza Chamanara, Frederick P Gardiner, and Nikola Lakic. A hyperelliptic realization of the horseshoe and baker maps. *Ergodic Theory and Dynamical Systems*, 26(6):1749–1768, 2006.

[5] André de Carvalho and Toby Hall. Unimodal generalized pseudo-Anosov maps. *Geometry & Topology*, 8(3):1127–1188, 2004.

[6] Robert L Devaney. *An introduction to chaotic dynamical systems*. CRC Press, 2021.

[7] Frederick P Gardiner and Nikola Lakic. *Quasiconformal Teichmüller theory*. Number 76 in Mathematical Surveys and Monographs. American Mathematical Soc., 2000.

[8] Gollakota Hemasundar. Koebe's general uniformisation theorem for planar riemann surfaces. In *Annales Polonici Mathematici*, volume 1, pages 77–85, 2011.

[9] Svetlana Katok. *Fuchsian groups*. University of Chicago Press, 1992.

[10] James R Munkres. *Topology; a first course [by] James R. Munkres*. Prentice-Hall, 1974.

[11] John R Quine and Peter Sarnak. *Extremal Riemann Surfaces: From the Proceedings of the AMS Special Session with Related Papers January 4-5, 1995, San Francisco, California*, volume 201. American Mathematical Soc., 1997.

[12] Clark Robinson. *Dynamical systems: stability, symbolic dynamics, and chaos*. CRC Press, 1998.

[13] Steve Smale. Finding a horseshoe on the beaches of Rio. *The Mathematical Intelligencer*, 20(1):39–44, 1998.

[14] Peter Walters. *An introduction to ergodic theory*, volume 79. Springer Science & Business Media, 2000.