**Calkin-Wilf Dynamics**
Borsje, Tom

**T. Borsje**

# Calkin-Wilf dynamics

**Bachelor thesis**

**20 December 2023**

**Thesis supervisors:** **dr. M. Roelands**
**dr. C.C.C.J. Kalle**

**Leiden University**
**Mathematical Institute**

# Contents

# Introduction

Studying the size of infinite sets goes back to the 1880's, when the German mathematician Georg Cantor introduced the concepts of countable and uncountable sets. This was done in his series of articles called "Über unendliche, lineare Punktmannichfaltigkeiten" [3]. In these articles, he stated that the set of rational numbers is countable, meaning it belongs to a smaller class of infinity than the real numbers. He proved this by means of his famous diagonal argument, which shows that there does not exist a one-to-one correspondence between the integers and the real numbers. In order to prove that the rationals are countable, he showed how to construct an enumeration of the rational numbers. He did not, however, explicitly give one [3]. The first time an explicit enumberation was constructed was in 1999, by the American mathematicians Neil Calkin and Herbert Wilf. In their paper [2], they described an object called the Calkin-Wilf tree, which they then used to construct an explicit enumeration of the positive rational numbers. This enumeration, the Calkin-Wilf sequence, was based on Stern's sequence, which denotes the number of ways an integer $n$ can be written as a sum of powers of 2, each power being used at most twice [2]. These two sequences and their combinatorial properties have been studied extensively in the past, for example in [6, 10, 17]. In 2003, a paper by the American mathematician Donald Knuth sought to find a generating function of the Calkin-Wilf sequence, that is, a function $T$ such that $T(a_n) = a_{n+1}$ for all $n \geq 0$, where $(a_n)_n$ denotes the Calkin-Wilf sequence. The function in question depends on the so-called ruler function $\rho$, where $\rho(n) = k$ if $n$ is divisible by $2^k$ but not by $2^{k+1}$. Multiple proofs of this generating function are contained in this article [7], one by C.P. Rupert and one by Alex Smith and Richard Strong. The article ends by crediting the Israeli private mathematics teacher Moshe Newman with finding a way to eliminate the ruler function out of the equation of $T$, meaning he found a way to express $a_{n+1}$ in terms of $a_n$ [7].

Similar to this map $T$ is the Gauss map, which has been studied in the context of ergodic theory. This field has evolved from the field of dynamical systems, which has been around since the development of calculus by Isaac Newton and the formulation of his laws of motion [9]. By the end of the nineteenth century, the French mathematician Henri Poincaré changed the way we look at dynamics by considering the set of all solutions to a dynamical system, rather than specific solution curves. This qualitative approach could often yield more general information about the system [9]. Combining this line of thinking with measure theory and probability theory, American mathematicians George David Birkhoff and John von Neumann laid the foundations of the field of ergodic theory, by stating their ergodic theorems in 1931 [1, 13]. Using ergodic theory, it can be shown that the Gauss map is a particularly interesting generator of continued fractions. In this thesis, we will extend the generating function of the Calkin-Wilf sequence to a piecewise continuous map $T : \mathbb{R}_+ \to \mathbb{R}_+$, called the Calkin-Wilf map, and motivated by the similarity between this map and the Gauss function, we apply ergodic theory to establish a link between $T$ and minus continued fractions. As the main results in this thesis, we will show that the Calkin-Wilf map has no periodic points, and we will construct an absolutely continuous invariant measure which is ergodic for $T$.

The motivation of this thesis is to study the statistical properties of this specific dynamical system, and by doing this provide stepping stones to studying the same dynamical system as a generator of minus continued fractions. Much research has already been done in continued fractions, see for example [12, 14, 16]. In this thesis, we aim to provide sufficient preliminary work to continue this line of research.

We will start this thesis by collecting the necessary preliminaries and notations used within this thesis. In order to be able to apply ergodic number theory, we will need to establish several concepts from measure theory, ergodic theory and graph theory. This will be done in Chapter 1.

In Chapter 2, we will discuss some previous literature on the Calkin-Wilf tree. Specifically, we will review the definitions of the Calkin-Wilf tree, sequence and function as presented in [2]. In addition, we will see that the Calkin-Wilf sequence is an enumeration of the positive rationals and present the proof as given in [10]. We will then define the Calkin-Wilf function using the proof by C.P. Rupert given in [7].

After completing our review of the most relevant previous results, we will start presenting our own results in Chapter 3. In this chapter, we explore the iterative behaviour of $T : \mathbb{R}_+ \to \mathbb{R}_+$, providing some results for sets $T^k[n-1, n)$ for positive integers $k$ and $n$. Our first main result will appear in this chapter as well, as we show that $T$ admits no periodic points in $\mathbb{R}_+$.

The measure theoretic and ergodic aspects of this thesis will play a major role in Chapter 4. Here, we will study the behaviour of our system $(\mathbb{R}_+, T)$ and use Carathéodory's extension theorem to construct a measure $\mu$ on the Borel $\sigma$-algebra $\mathscr{B}$ such that the dynamical system $(\mathbb{R}_+, \mathscr{B}, \mu, T)$ is not only measure preserving, but also ergodic. This reflects the fact that any orbit under $T$ is dense in the state space $\mathbb{R}_+$.

Finally, in Chapter 5, we will provide the reader with a link between our ergodic dynamical system and minus continued fractions. Although this thesis will not contain any results regarding the convergence of such continued fractions, we will compute the arithmetic and geometric means of the digit sequences produced by $T$.

# 1 Preliminaries

In this section we collect the necessary preliminaries and notations used within this thesis.

Let $A \subseteq \mathbb{R}$ and $y \in \mathbb{R}$. The notation $A + y$ is used to denote the set $\{x + y : x \in A\}$. Also, we will use the shorthand notations $\mathbb{R}_+ = \{x \in \mathbb{R} : x \geq 0\}$ and $\mathbb{Q}_+ = \mathbb{R}_+ \cap \mathbb{Q}$. Throughout this thesis, we will denote a disjoint union by $\biguplus$.

## 1.1 Measure theory

The reader is assumed to be familiar with the basic notions of measure theory. Here, we repeat some important concepts as found in most standard text books on measure theory. We refer to [11] for most of these definitions and theorems. Paraphrasing Definition 5.1 from [11], we define a Dynkin system as follows.

**Definition 1.** *Let $X$ be a set with power set $\mathscr{P}(X)$. A **Dynkin system** $\mathscr{D} \subseteq \mathscr{P}(X)$ is a collection of sets such that*

*1.a.* $X \in \mathscr{D}$,

*1.b.* $D \in \mathscr{D}$ *implies* $D^C \in \mathscr{D}$, *and*

*1.c. If* $(D_n)_n \subseteq \mathscr{D}$ *are pairwise disjoint, then* $\biguplus_{n=0}^{\infty} D_n \in \mathscr{D}$.

For any collection $\mathscr{G} \subseteq \mathscr{P}(X)$ there exists a smallest Dynkin system $\delta(\mathscr{G})$ that contains $\mathscr{G}$ (see Proposition 5.3 in [11]). We denote the smallest $\sigma$-algebra containing $\mathscr{G}$ by $\sigma(\mathscr{G})$. The following theorem is stated as Theorem 5.5 in [11].

**Theorem 1.** *If $\mathscr{G} \subseteq \mathscr{P}(X)$ is stable under finite intersections (or $\cap$-**stable**), then $\delta(\mathscr{G}) = \sigma(\mathscr{G})$.*

On page 63 in [11], a semi-ring is defined. We summarise this in the following definition.

**Definition 2.** *Let $X$ be a set. A **semi-ring** is a family $\mathscr{S} \subseteq \mathscr{P}(X)$ such that*

*a.* $\emptyset \in \mathscr{S}$,

*b.* $\mathscr{S}$ *is* $\cap$*-stable,*

*c. for any* $S, T \in \mathscr{S}$ *there exist finitely many disjoint* $S_1, \ldots, S_n \in \mathscr{S}$ *such that* $S \setminus T = \biguplus_{k=1}^{n} S_n$.

Semi-rings are useful for constructing measures. This can be done by Carathéodory's Extension Theorem, which is stated in [11] as Theorem 6.1. In this thesis, we will apply this theorem only to define probability measures. Therefore, the theorem below is written with this in mind.

**Theorem 2.** *If $\mathscr{S} \subseteq \mathscr{P}(X)$ is a semi-ring and $\mu^* : \mathscr{S} \to [0, \infty]$ a **pre-measure**, that is, a set function such that* *a.* $\mu^*(\emptyset) = 0$ *and*

*b. if* $(S_n)_n \subseteq \mathscr{S}$ *are pairwise disjoint and* $\biguplus_{n=0}^{\infty} S_n \in \mathscr{S}$ *then* $\mu^*(\biguplus_{n=1}^{\infty} S_n) = \sum_{n=1}^{\infty} \mu^*(S_n)$*,*

*then* $\mu^*$ *has a unique extension to a measure* $\mu$ *on* $\sigma(\mathscr{S})$*.*

## 1.2 Ergodic theory

Let $X$ be a non-empty set and $T : X \to X$ a transformation. The map $T \circ T$ will be denoted by $T^2$, and similarly, $T^n$ is the $n^{\text{th}}$ iterate of $T$. For any $A \subseteq X$, we write $T^{-1}A = \{x \in X : T(x) \in A\}$ for the inverse image of $A$ under $T$. Finally, we define $T^{-n} = T^{-1}(T^{-n+1}A)$ recursively for any $n > 1$. We use the convention that $T^0$ is the identity map. The next definition is quoted verbatim from Definition 1.2.12 in [4].

**Definition 3.** *Let* $X$ *be a nonempty set,* $\mathscr{A}$ *a* $\sigma$*-algebra on* $X$ *and* $\mu$ *a probability measure, so that* $(X, \mathscr{A}, \mu)$ *is a probability space. A measurable transformation* $T : X \to X$ *is* ***measure preserving*** *with respect to* $\mu$ *(equivalently:* $\mu$ *is* $T$***-invariant****, or* $\mu$ *is an* ***invariant measure*** *for* $T$*), if* $\mu(T^{-1}A) = \mu(A)$ *for all* $A \in \mathscr{A}$*.*

Paraphrasing Definition 1.2.17 in [4], we define a dynamical system as follows.

**Definition 4.** *A* ***dynamical system*** *is a quadruple* $(X, \mathscr{A}, \mu, T)$*, where* $X$ *is a non-empty set,* $\mathscr{A}$ *is a* $\sigma$*-algebra on* $X$*,* $\mu$ *is a probability measure on* $(X, \mathscr{A})$ *and* $T : X \to X$ *is a measure preserving transformation with respect to* $\mu$*.*

Summarising page 16 in [4], we get the following definition.

**Definition 5.** *Let* $(X, \mathscr{A}, \mu, T)$ *be a dynamical system and* $x \in X$*. The set*

$$\{x, T(x), T^2(x), \dots\}$$

*is called the* $T$***-orbit*** *of* $x$*. In case* $T^{-n}(x)$ *exists for all* $n \geq 1$*, the* ***two sided*** $T$***-orbit*** *of* $x$ *is*

$$\{\dots, T^{-1}(x), x, T(x), \dots\}.$$

Definition 5.3 in [15] states the following.

**Definition 6.** *A point* $x \in X$ *is a* ***periodic point*** *of* $T$ *if* $T^n x = x$ *for some integer* $n > 0$*. The least positive* $n$ *with this property is called the* ***period*** *of* $x$*.*

If the period of $x$ is 1, we call $x$ a **fixed point**. The following definition can be found as Definition 3.1.4 in [4].

**Definition 7.** *Let $(X, \mathscr{A}, \mu, T)$ be a dynamical system. Then $T$ is called **ergodic** if for every $\mu$-measurable set $A$ satisfying $T^{-1}A = A$ (such a set is called $T$**-invariant**) one has that $\mu(A) = 0$ or $\mu(A) = 1$.*

Proposition 3.1.9 of the same book gives us the following useful characterisation of ergodicity.

**Proposition 1.** *Let $(X, \mathscr{A}, \mu, T)$ be a dynamical system. Then $T$ is ergodic if and only if every measurable function $f : X \to \mathbb{R}$ with $f = f \circ T$ is constant almost everywhere.*

Finally, we paraphrase the ergodic theorem which is stated in Theorem 3.1.7 in [4].

**Theorem 3.** *Let $(X, \mathscr{A}, \mu, T)$ be a dynamical system where $T : X \to X$ is ergodic. Then, for any $f$ in $L^1(X, \mathscr{A}, \mu)$, we have*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i(x)) = \int_X f \, d\mu$$

*for $\mu$-a.e. $x \in X$.*

## 1.3 Graph theory

In this section, we will outline some definitions from graph theory as found in [5]. We find the definition of a graph in Section 1.1.

**Definition 8.** *A **graph** is a pair of sets $(\mathscr{V}, \mathscr{E})$ such that $\mathscr{E} \subseteq \mathscr{V} \times \mathscr{V}$. We call $\mathscr{V}$ our set of **vertices** and $\mathscr{E}$ is the set of **edges**.*

We will need the notions of paths and cycles. The following definitions are paraphrased from Section 1.3 of [5].

**Definition 9.** *Let $G = (\mathscr{V}, \mathscr{E})$ be a graph. A **path** is a subgraph $P = (\mathscr{V}_P, \mathscr{E}_P) \subseteq G$ of the form*

$$\mathscr{V}_P = \{x_1, x_2, \ldots, x_n\}, \mathscr{E}_P = \{(x_1, x_2), (x_2, x_3), \ldots, (x_{n-1}, x_n)\}.$$

As the vertices $\mathscr{V}_P$ of $P$ can be deduced from the set of edges $\mathscr{E}_P$, we can instead use the shorthand notation $P = ((x_1, x_2), (x_2, x_3), \ldots, (x_{n-1}, x_n))$. The **empty path** will be denoted by $\epsilon$. We write the collection of all paths as

$$\mathscr{E}^* := \{P \subseteq G : P \text{ is a path}\}.$$

**Definition 10.** *Let $G = (\mathscr{V}, \mathscr{E})$ be a graph. A **cycle** of length $n$ is a path*

$$P = ((x_1, x_2), (x_2, x_3), \ldots, (x_{n-1}, x_n), (x_n, x_1).$$

*A **loop** is a cycle of length 1.*

The next two definitions are taken from Section 1.4 of [5].

**Definition 11.** *A graph $(\mathscr{V}, \mathscr{E})$ is called **connected** if for any $x, y \in \mathscr{V}$ there exists a path $P \in \mathscr{E}^*$ that starts at $x$ and ends at $y$.*

**Definition 12.** *A connected graph $(\mathscr{V}, \mathscr{E})$ is a **tree** if it contains no cycles.*

Finally, this definition is taken form Section 1.10 of [5].

**Definition 13.** *A graph $(\mathscr{V}, \mathscr{E})$ is called **undirected** if each edge $e \in \mathscr{E}$ is an unordered pair of vertices, and **directed** if each edge $e \in \mathscr{E}$ is an ordered pair of vertices.*

## 2 The Calkin-Wilf function

In the 1999 paper by Calkin and Wilf [2], in Section 1, the Calkin-Wilf sequence is constructed by building a binary tree, such that:

- the root of the tree is $\frac{1}{1}$ and
- every node $\frac{a}{b}$ of the tree has two children: a left child $\frac{a}{a+b}$ and a right child $\frac{a+b}{b} = \frac{a}{b} + 1$,

see Figure 1.

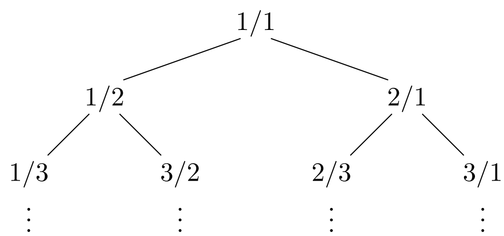



Figure 1: A visualisation of the Calkin-Wilf tree.

Now, if we were to read this graph from left to right, top to bottom, we get the following sequence:

$$\frac{1}{1}, \frac{1}{2}, \frac{2}{1}, \frac{1}{3}, \frac{3}{2}, \frac{2}{3}, \frac{3}{1}, \frac{1}{4}, \frac{4}{3}, \frac{3}{5}, \frac{5}{2}, \frac{2}{5}, \frac{5}{3}, \frac{3}{4}, \frac{4}{1}, \ldots .$$

**Definition 14.** *The* ***Calkin-Wilf sequence*** $(a_n)_n$ *is given by reading the Calkin-Wilf tree breadth-first, that is,* $a_1 = 1$ *is the root of the tree, and* $a_{n+1}$ *is the node to the right of* $a_n$*, or the leftmost node of the next depth if* $a_n$ *has no right neighbour. We put* $a_0 = 0$.

The depth of a node $a_n$, as used in the previous definition, refers to the number of edges in the path from $a_n$ to 1. The depth of $a_1$ is zero, and as the Calkin-Wilf tree is a binary tree, every depth level $d$ contains precisely $2^d$ nodes. Each number $n$ can be decomposed as $n = 2^m + b$ by taking $b = n \mod 2^m$, so we call $m$ the depth and $b$ the breadth of $a_n = a_{2^m+b}$.

The main result Calkin and Wilf found in Section 1 of [2] is the following.

**Theorem 4.** *The function that takes* $n$ *to* $a_n$ *is a bijection between the positive integers* $\mathbb{Z}_{\geq 0}$ *and the positive rationals* $\mathbb{Q}_+$.

There are several ways to prove this. An elegant method given in Sections 2.1 and 2.2 of [10] makes use of the following observation. Any map $F : S \to S$ of a countable set to itself can be seen as a directed graph, where the nodes are the elements of $S$ and the edges are arrows from $x \in S$ to $F(x)$. This graph has loops at all of the fixed points of $F$, and it might also have cycles. Making use of these concepts, the next result is proven in Theorem 2.1 of [10].

**Lemma 1.** *If $S$ is countable, $F : S \to S$ is a function with a set $S_0$ of fixed points and $\Phi : S \to \mathbb{Z}_{>0}$ is such that $\Phi(F(x)) < \Phi(x)$ whenever $x \notin S_0$, then by removing the loops at all of the fixed points of $F$ from the graph of $F$ as above we obtain a union of disjoint directed trees, and $S_0$ is the set of their roots.*

*Proof.* We let $\mathscr{E}$ be the set of edges $(x, F(x))$, where $x$ ranges over $S$. Consider the directed graph $G = (S, \mathscr{E})$. Fix $x \in S$. The sequence $(\Phi(F^n(x)))_n$ is decreasing by definition of $\Phi$. To see that, notice that $F^k(x) \neq F^{k-1}(x)$ whenever $F^{k-1}(x) \notin S_0$, and this gives us $\Phi(F^k(x)) < \Phi(F^{k-1}(x))$. On the other hand, if $F^k(x) = F^{k-1}(x)$, then $\Phi(F^k(x)) \leq \Phi(F^{k-1}(x))$.

However, as any subset of $\mathbb{Z}_{>0}$ has a smallest element, the sequence $(\Phi(F^n(x)))_n$ cannot be strictly decreasing. Thus, there must be some minimal $N > 0$ and a $z \in S_0$ such that for all $n \geq N$ we have $\Phi(F^n(x)) = \Phi(z)$. We see that there exists a path from $x$ to $z$ of length $N$. Also, as $F$ is well-defined, every $x \in S$ has no more than one outgoing edge. This means that the path from $x$ to $z$ is unique, and the connected component of $G$ containing $x$ must be a tree with root $z \in S_0$ after removing the loop in $z$. □

Using this theorem, the proof of Theorem 4 follows quickly.

*Proof of Theorem 4.* Let $S = \mathbb{Q}_+ \setminus \{0\}$, and let $F : S \to S$ and $\Phi : S \to \mathbb{Z}_{>0}$ be given by

$$F(x) = \begin{cases} \max\left\{\frac{x}{1-x}, x-1\right\} & x \neq 1 \\ 1 & x = 1 \end{cases}$$

and

$$\Phi\left(\frac{a}{b}\right) = a + b,$$

where we assume $\gcd(a, b) = 1$. Note that the set of fixed points is $S_0 = \{1\}$.

Further notice that for any element of $S$, we have

$$F^{-1}\left\{\frac{a}{b}\right\} \supseteq \left\{\frac{a}{a+b}, \frac{a+b}{b}\right\},$$

as

$$F\left(\frac{a}{a+b}\right) = \max\left\{\frac{\frac{a}{a+b}}{1-\frac{a}{a+b}}, \frac{-b}{a+b}\right\} = \frac{a}{b}$$

and

$$F\left(\frac{a+b}{b}\right) = \max\left\{\frac{\frac{a+b}{b}}{1-\frac{a+b}{b}}, \frac{a}{b}\right\} = \frac{a}{b}.$$

Since for $x \in (0,1) \cap \mathbb{Q}, F(x) = \frac{x}{1-x}$, it follows that $F((0,1) \cap \mathbb{Q}) \subseteq S$. Similarly, for we have $x \in [1, \infty) \cap \mathbb{Q}, F(x) = 1 - x$, so $F([1, \infty) \cap \mathbb{Q}) \subseteq S$. As $F$ is strictly increasing on $(0,1) \cap \mathbb{Q}$ and on $[1,\infty) \cap \mathbb{Q}$, this means that $F^{-1}\{\frac{a}{b}\}$ contains no more than two elements except if $\frac{a}{b} = 1$. We conclude for any $\frac{a}{b} \in S \setminus \{1\}$ that

$$F^{-1}\left\{\frac{a}{b}\right\} = \left\{\frac{a}{a+b}, \frac{a+b}{b}\right\}.$$

We let $\mathscr{E}$ be the set of edges $(x, F(x))$ for $x \in S$ and $G = (S, \mathscr{E})$. For $\frac{a}{b} \neq 1$, we have that $\Phi(F^{-1}\{\frac{a}{b}\}) = \Phi(\{\frac{a}{a+b}, \frac{a+b}{b}\}) = \{2a+b, a+2b\}$. Both $2a+b$ and $a+2b$ are larger than $\Phi(\frac{a}{b}) = a+b$, so $\Phi(x) < \min \Phi(F^{-1}\{x\})$. This means $\Phi(F(x)) < \Phi(x)$ and so by Lemma 1, $G$ is a union of disjoint trees with roots in $S_0 = \{1\}$ after removing the loop in 1. As $S_0$ contains only one element, $G$ must be a tree. Therefore, every $x \in S$ appears exactly once in $G$.

As each node $a/b \in S$ except for 1 has one edge connecting it to $F(a/b)$ and two edges connecting it to $a/(a+b)$ and $(a+b)/b$, $G$ is the Calkin-Wilf tree by construction, and the enumeration $S \to \mathbb{Z}_{>0}$ obtained by reading $G$ breadth-first is precisely the Calkin-Wilf sequence. $\square$

The following lemma will prove to be useful later.

**Lemma 2.** *Let $n \geq 1$. The left child of $a_n$ is $a_{2n}$ and the right child of $a_n$ is $a_{2n+1}$.*

*Proof.* For $n \geq 1$, let $m \geq 0$ and $0 \leq b < 2^m$ be integers such that $n = 2^m + b$. Let $k$ be the index of the left child of $n$. Notice that $b$ is the number of nodes to the left of $n$. As each of the $a_{2^m}, \ldots, a_{2^m+b-1}$ has two children of its own, $a_k$ has $2b$ nodes to its left. Also, we have that the depth level of $a_k$ is $m+1$, so that $k = 2^{m+1} + 2b = 2n$. The right child of $a_n$ has index $k+1 = 2n+1$. $\square$

**Corollary 1.** *Let $n \geq 1$ be an integer and suppose $n = a_k$ and $\frac{1}{n} = a_l$ for some $k, l \geq 1$. We have $l = 2^n$ and $k = 2^n - 1$.*

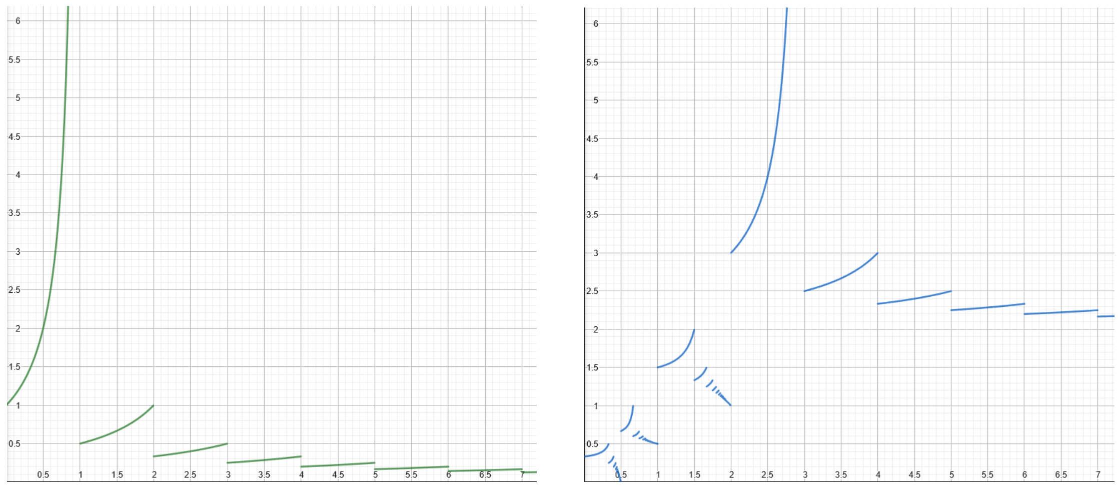
*Proof.* For $n = 1$, we have $n = a_1$, and so the base case follows. Now suppose $n \geq 1$ is such that $n = a_{2^n-1}$ and $\frac{1}{n} = a_{2^{n+1}}$. As 1 and $n$ are coprime, the left child of $\frac{1}{n}$ is $\frac{1}{n+1}$, so $\frac{1}{n+1} = a_{2^{n+1}}$ by Lemma 2. Similarly, the right child of $\frac{n}{1}$ is $\frac{n+1}{1} = n+1$, so $n+1 = a_{2^{n+1}-1}$ by the same lemma. The statement follows by induction. $\square$

The main subject of study in this thesis will be the generating map of the Calkin-Wilf sequence. This map, which we will call the **Calkin-Wilf map**, is given by $T : \mathbb{R}_+ \to \mathbb{R}_+$, with

$$T(x) = \frac{1}{2\lfloor x \rfloor - x + 1}.$$

See Figure 2a for its graph.

A proof of the following theorem is given by C.P. Rupert in [7]. We repeat the proof here, as it gives us some insight in the fundamental properties of $T$.



(a) Graph of $T$ (b) Graph of $T^4$

Figure 2: Two graphs showing the iterative behaviour of $T$.

**Theorem 5.** *The restriction $T|_{\mathbb{Q}_+}$ is such that $T(a_n) = a_{n+1}$ for all $n \geq 0$, or equivalently, $a_n = T^n(0)$, where $(a_n)_n$ is the Calkin-Wilf sequence.*

*Idea of the proof.* To prove the above theorem, we need to relate the paths in the Calkin-Wilf tree to the values on its nodes. As this proof will be quite involved, we first outline its structure in four steps.

*Step 1.* Define a labeling of the paths $g : \mathscr{E}^* \to \mathbb{Z}$ of the Calkin-Wilf tree, and write the labels as their binary expansions. We provide some recurrence relations for $g$, which we will use later in Step 3.

*Step 2.* Describe the path from 1 to any rational $a_n = \frac{x_n}{y_n}$ in the Calkin-Wilf tree, in terms of some sequence $(\rho(n))_n$.

*Step 3.* Write $x_n$ in terms of $x_{n-1}$, and write $y_n$ in terms of $y_{n-1}$ and $\rho(n)$. By assuming that $\gcd(x_n, y_n) = 1$ for all $\frac{x_n}{y_n}$ in the Calkin-Wilf tree, this will give us an expression of $\frac{x_n}{y_n}$ in terms of $\frac{x_{n-1}}{y_{n-1}}$ and $\rho(n)$.

*Step 4.* Finally, as $\frac{x_n}{y_n}$ is equal to $a_n$ by assumption, we now only need to express $\rho(n)$ as a function of $a_n$.

After the last step, we will end up with an expression of $a_n$ in terms of $a_{n-1}$ for all $n \geq 1$, and we will be able to conclude that $a_n = T(a_{n-1})$ for all $n \geq 1$.

*Proof.* First, we introduce the sequence $(\rho(n))_n$ which we will need in Step 2. Let $\rho : \mathbb{Z}_{>0} \to \mathbb{Z}_{\geq 0}$ be defined by $\rho(n) = \max\{k \in \mathbb{Z}_{\geq 0} : 2^k \text{ divides } n\}$. For any $n$, we have $\rho(2n+1) = 0$ and

$$\rho(2n) = \max\{k \in \mathbb{Z}_{\geq 0} : 2^k \text{ divides } 2n\} = \max\{k \in \mathbb{Z}_{\geq 0} : 2^{k-1} \text{ divides } n\} = \rho(n) + 1.$$

Now, let us proceed by following the steps outlined above.

*Step 1.* For any node in the Calkin-Wilf tree, we define an edge labeling as follows. Let $\mathscr{E}$ be the set of edges in the Calkin-Wilf tree. We define $\ell : \mathscr{E} \to \{0, 1\}$ by

$$\ell\left(\left(\frac{a}{b}, \frac{a}{a+b}\right)\right) = 0,$$
$$\ell\left(\left(\frac{a}{b}, \frac{a+b}{b}\right)\right) = 1.$$

By Theorem 4, for every $m \in \mathbb{Q}_+$ a corresponding path from 1 to $m$ exists in the Calkin-Wilf tree. We define a function $g : \mathscr{E}^* \to \mathbb{Z}_{>0}$ by

$$g(e_1, e_2, \ldots, e_n) = 2^n + \sum_{k=0}^{n-1} 2^k \ell(e_{n-k}).$$

As the empty path $\epsilon$ leads to 1, we have $g(\epsilon) = 1$, and $g$ is well defined. Also notice $g$ is bijective, as the $2^n$ nodes on depth $n$ all have different labels. Let $\pi = (e_1, e_2, \ldots, e_n)$ be the path from 1 to $\frac{a}{b}$, then we have two recurrence relations

$$g\left(\pi, \left(\frac{a}{b}, \frac{a}{a+b}\right)\right) = 2^{n+1} + \sum_{k=1}^{n} 2^k \ell(e_{n-k}) + \ell\left(\left(\frac{a}{b}, \frac{a}{a+b}\right)\right) = 2g(\pi)$$

and

$$g\left(\pi, \left(\frac{a}{b}, \frac{a+b}{b}\right)\right) = 2^{n+1} + \sum_{k=1}^{n} 2^k \ell(e_{n-k}) + \ell\left(\left(\frac{a}{b}, \frac{a+b}{b}\right)\right) = 2g(\pi) + 1,$$

which will be useful later. We will denote the binary representation of $g(\pi)$ by

$$g(\pi) = (\ell(e_1)\ell(e_2)\ldots\ell(e_n))_2.$$

*Step 2.* Now for any $n \geq 1$, let $\frac{x_n}{y_n}$ be the rational such that $g^{-1}(n)$ is the path from 1 to $\frac{x_n}{y_n}$. Notice that $\frac{x_n}{y_n} = a_n$ by Definition 14. If $n$ is a power of 2, we have $n = 2^{\rho(n)}$, so

$$n = 2^{\rho(n)} = (1\underbrace{00\dots00}_{\rho(n)\text{ zeroes}})_2,$$

so we can follow the path from 1 to $\frac{x_n}{y_n}$ by taking the left child $\rho(n)$ times, so $\frac{x_n}{y_n} = \frac{1}{1+\rho(n)}$.

In addition, we have

$$n-1 = 2^{\rho(n)} - 1 = (\underbrace{111\dots11}_{\rho(n)\text{ ones}})_2.$$

So the path from 1 to $\frac{x_{n-1}}{y_{n-1}}$ can be followed by taking the right child $\rho(n)-1$ times, and we find $\frac{x_{n-1}}{y_{n-1}} = \frac{\rho(n)}{1} = \rho(n)$.

Now suppose $n$ is not a power of 2. Then there exists some maximal $i$ such that the $i^{\text{th}}$ digit of the binary representation of $n$ is 1. We can write

$$n = (1\alpha_1\alpha_2\dots\alpha_{i-1}1\underbrace{0\dots0}_{\rho(n)\text{ zeroes}})_2$$

and

$$n-1 = (1\alpha_1\alpha_2\dots\alpha_{i-1}0\underbrace{1\dots1}_{\rho(n)\text{ ones}})_2.$$

We see that the last common ancestor of $\frac{x_n}{y_n}$ and $\frac{x_{n-1}}{y_{n-1}}$ is given by $\frac{x_m}{y_m}$, where $m = (1\alpha_1\alpha_2\dots\alpha_{i-1})_2$. The path from $\frac{x_m}{y_m}$ to $\frac{x_{n-1}}{y_{n-1}}$ is given by taking the left child first, followed by taking the right child $\rho(n)$ times. The left child of $\frac{x_m}{y_m}$ is $\frac{x_m}{x_m+y_m}$, so we find

$$\frac{x_{n-1}}{y_{n-1}} = \frac{x_m + \rho(n)(x_m+y_m)}{x_m+y_m}.$$

On the other hand, the path from $\frac{x_m}{y_m}$ to $\frac{x_n}{y_n}$ is given by taking a right child, and $\rho(n)$ left children after. This yields

$$\frac{x_n}{y_n} = \frac{x_m+y_m}{y_m + \rho(n)(x_m+y_m)}.$$

*Step 3.* Assuming that $\gcd(x_m, y_m) = 1$, all of the above fractions are reduced. To write $x_n$ and $y_n$

in terms of $x_{n-1}$ and $y_{n-1}$, we need to solve the following system of equations.

$$\begin{cases} x_n = x_m + y_m \\ y_n = y_m + \rho(n)(x_m + y_m) \\ x_{n-1} = x_m + \rho(n)(x_m + y_m) \\ y_{n-1} = x_m + y_m. \end{cases}$$

Solving this gives us

$$\frac{x_n}{y_n} = \frac{y_{n-1}}{y_{n-1}(1 + 2\rho(n)) - x_{n-1}} = \frac{1}{2\rho(n) - \frac{x_{n-1}}{y_{n-1}} + 1}.$$

As $\frac{x_n}{y_n} = a_n$ for all $n \geq 1$, we can write

$$a_n = \frac{1}{2\rho(n) - a_{n-1} + 1},$$

which leads us to the final step in this proof.

*Step 4.* Now let us consider the behaviour of $\lfloor a_n \rfloor$ for $n \geq 1$. Let $a_n = \frac{a}{b}$ with $a, b \geq 1$ coprime. Recall that the left child of $a_n$ is $\frac{a}{a+b} < 1$, and the right child is $\frac{a+b}{b} = \frac{a}{b} + 1$. By Lemma 2, we have that $\frac{a}{a+b} = a_{2n}$ and $\frac{a+b}{b} = a_{2n+1}$. This means that $\lfloor a_{2n} \rfloor = 0$ and $\lfloor a_{2n+1} \rfloor = \lfloor a_n \rfloor + 1$, with $\lfloor a_1 \rfloor = 1$. We see that the sequence $(\lfloor a_n \rfloor)_n$ satisfies the same recurrence relation as $(\rho(n+1))_n$, so we find $\lfloor a_{n-1} \rfloor = \rho(n)$.

Finally, we conclude

$$a_n = \frac{1}{2\rho(n) - a_{n-1} + 1} = \frac{1}{2\lfloor a_{n-1} \rfloor - a_{n-1} + 1} = T(a_{n-1}).$$

□

Notice that $T(a_n) = a_{n+1}$ implies that $T^k(a_n) = a_{n+k}$. This leads us to the following corollary.

**Corollary 2.** *For any integer $n \geq 1$, we have $n + 1 = T^{2^n}(n)$ and $\frac{1}{n+1} = T^{2^n}(\frac{1}{n})$.*

*Proof.* Let $n \geq 1$ be an integer. By Corollary 1 we have $n = a_{2^n - 1}$ and $\frac{1}{n} = a_{2^n}$. This gives us

$$T^{2^n}(n) = T^{2^n}(a_{2^n-1}) = a_{2^{n+1}-1} = n + 1$$

and

$$T^{2^n}\left(\frac{1}{n}\right) = T^{2^n}(a_{2^n}) = a_{2^{n+1}} = \frac{1}{n+1}.$$

$\square$

Some observations are immediately clear from the graphs in Figure 2. We list them below.

(i) $T : \mathbb{R}_+ \to \mathbb{R}_+ \setminus \{0\}$ is invertible and its inverse is given by $T^{-1}(x) = 2\lceil 1/x \rceil - 1/x - 1$. We have $T^{-1} = r \circ T \circ r$ on $\mathbb{R}_+ \setminus \mathbb{Z}$, where $r : (0, \infty) \to \mathbb{R}_+$ is the map $r(x) = \frac{1}{x}$. In other words, we have the equality $\frac{1}{T^{-1}(x)} = T\left(\frac{1}{x}\right)$.

(ii) $T$ is smooth on $\mathbb{R}_+ \setminus \mathbb{Z}$, and the derivative can be calculated to be $T'(x) = \frac{1}{(2\lfloor x \rfloor - x + 1)^2} = (T(x))^2$ whenever it exists.

(iii) $T$ has no fixed points, as we have $T[0, 1) = [1, \infty)$ and $T[1, \infty) = (0, 1)$. We also have $T(0, 1) = (1, \infty)$.

# 3 Recurrence

As we can see in Figure 2, some interesting patterns occur when taking iterates $T^k$, where $k$ is a positive integer. In this section we will explore this behaviour and prove some results concerning the sets $T^k(n-1,n)$, where $k$ and $n$ are positive integers.

We already noted before that $T$ has no fixed points. In addition, a close inspection of the graph in Figure 2b reveals that $T^4$ does not appear to have any fixed points either. At the end of this section, we will show that $T$ has no periodic points.

**Lemma 3.** *We have $T^{2n}(x) = T^n(x-1)+1$ for all $x > 1$ and $n > 0$.*

*Proof.* We prove the statement by induction. For $n = 1$ and $x \in (1,\infty)$ it holds that

$$\begin{aligned} T(x-1)+1 &= \frac{1}{2\lfloor x-1\rfloor - (x-1)+1} + 1 \\ &= \frac{1}{2\lfloor x\rfloor - x} + 1 \\ &= \frac{2\lfloor x\rfloor - x + 1}{2\lfloor x\rfloor - x}. \end{aligned}$$

On the other hand, since $T(x) < 1$, we have

$$\begin{aligned} T^2(x) &= \frac{1}{2\lfloor T(x)\rfloor - \frac{1}{2\lfloor x\rfloor - x+1} + 1} \\ &= \frac{2\lfloor x\rfloor - x+1}{2\lfloor x\rfloor - x} = T(x-1)+1. \end{aligned}$$

Now suppose $T^{2n}(x) = T^n(x-1)+1$ for some $n \geq 1$ and all $x \in (1,\infty)$. Then for $n+1$ we have

$$\begin{aligned} T^{2(n+1)}x &= T^{2n}T^2x \\ &= T^{2n}(T(x-1)+1) \\ &= T^n(T(x-1))+1 \\ &= T^{n+1}(x-1)+1. \end{aligned}$$

The claim follows by induction. ☐

A notable consequence of this lemma is the following.

**Corollary 3.** *We have $T^{2^{n-1}}(n-1,n) = (n,\infty)$ and $T^{2^{n-1}}[n-1,n) = [n,\infty)$ for all $n \geq 1$.*

*Proof.* From Figure 2, we know that $T(0,1) = (1,\infty)$. Suppose we have $T^{2^{n-1}}(n-1,n) = (n,\infty)$

for some $n \geq 1$, then by Lemma 3 we have

$$T^{2^n}(n, n+1) = T^{2^{n-1}}(n-1, n) + 1 = (n, \infty) + 1 = (n+1, \infty),$$

so the first statement follows by induction. In particular, by Corollary 2 we have $T^{2^{n-1}}(n-1) = n$ for any integer $n \geq 1$, so

$$T^{2^{n-1}}[n-1, n) = T^{2^{n-1}}(n-1, n) \cup \{T^{2^{n-1}}(n-1)\} = (n, \infty) \cup \{n\} = [n, \infty).$$

□

If we wish to work more with the images of intervals, it will prove useful to show some more basic properties of $T^k(n-1, n)$ for all integers $n \geq 1$ and $k \geq 0$.

**Lemma 4.** *For all $n \geq 1$, we have the following properties for $T$.*

*(i) For all $0 \leq k, l < 2^n$ such that $k \neq l$, $T^k[n-1, n) \cap T^l[n-1, n) = \emptyset$.*

*(ii) For all $0 \leq k < 2^n$ with $k \neq 2^{n-1}$, there exists some $j \geq 1$ such that $T^k[n-1, n) \subseteq [j-1, j)$.*

*(iii) For all $0 \leq k < 2^n$, $T^k[n-1, n)$ is equal to an interval.*

*Proof.* (i) We first cover the base case where $n = 1$. As $0 \leq k, l < 2^1$, it follows by $T[0, 1) = [1, \infty)$ that $T[0, 1) \cap [0, 1) = \emptyset$.

Now suppose (i) holds for some $n \geq 1$. We show that this implies (i) for $n + 1$ with $l = 0$. As $[n, n+1) \subseteq [1, \infty)$, we have that $T^{2k+1}[n, n+1) \subseteq [0, 1)$ for any integer $0 < 2k+1 < 2^{n+1}$ by property (iii) on page 15. In particular, this means that $T^{2k+1}[n, n+1) \cap [n, n+1)$ is empty. Similarly, it follows from Lemma 3 that

$$\begin{aligned} T^{2k}[n, n+1) \cap [n, n+1) &= [T^k[n-1, n) + 1] \cap (n, n+1) \\ &= (T^k[n-1, n) \cap [n-1, n)) + 1 \\ &= \emptyset \end{aligned} \tag{1}$$

by the induction hypothesis, so for all $0 < 2k < 2^{n+1}$ we have $T^{2k}(n, n+1) \cap (n, n+1) = \emptyset$. Hence for any $n \geq 1$ and any $0 < k < 2^{n+1}$, we find that $T^k(n, n+1) \cap (n, n+1) = \emptyset$.

Now choose $l > 0$ and assume $k > l$ without loss of generality. Then some $m > 0$ exists such that $k = l + m$, and by injectivity of $T$,

$$T^{m+l}(n, n+1) \cap T^l(n, n+1) = T^l(T^m(n, n+1) \cap (n, n+1)) = T^l(\emptyset) = \emptyset,$$

where the penultimate equality follows from (1). We conclude that

$$T^k(n, n+1) \cap T^l(n, n+1) = \emptyset$$

for all $k \neq l$ and $k, l < 2^{n+1}$.

(ii) For our base case $n = 1$, we only have to check $k = 0$. This follows as $T^0[0, 1) \subseteq [0, 1)$. Now suppose (ii) holds for some $n \geq 1$. We show that this implies (ii) for $n + 1$.

First, we show the case for even iterations of $T[n, n+1)$. For any $0 \leq 2k < 2^{n+1}$, we have

$$\begin{aligned} T^{2k}[n, n+1) &= \{T^{2k}(x) : x \in [n, n+1)\} \\ &= \{T^k(x-1) + 1 : x \in [n, n+1)\} \\ &= \{T^k(x) + 1 : x \in [n-1, n)\} \\ &= T^k[n-1, n) + 1, \end{aligned}$$

where the second equality follows from Lemma 3. As $T^k[n-1, n) \subseteq [j-1, j)$ for some $j \geq 1$ and $k \neq 2^{n-1}$ by assumption, we see that $T^{2k}[n, n+1) \subseteq [j, j+1)$ if $2k \neq 2^n$.

For the odd iterations, we have for all $0 \leq 2k+1 < 2^{n+1}$ that $T^{2k+1}[n-1, n) \subseteq [0, 1)$ by property (iii) on page 15. We have now proven that (ii) holds for $n + 1$ for all $0 \leq k < 2^{n+1}$ with $k \neq 2^n$, and the statement follows by induction.

(iii) The restriction of $T$ to any interval $[j-1, j)$ with $j \geq 1$ a positive integer is given by a continuous and monotone function, namely $T(x) = \frac{1}{2(j-1)-x+1}$ for $x \in [j-1, j)$. Therefore, if $A \subseteq [j-1, j)$ is an interval, then so is $T(A)$.

For $n = 1$, we have $T[0, 1) = (1, \infty)$, so $T^k[0, 1)$ is an interval for all $0 \leq k < 2^1$. Now choose $n > 1$. By (ii), $T^k[n-1, n)$ is contained in an interval $[j-1, j)$ for all $0 \leq k < 2^n$ with $k \neq 2^{n-1}$, so as $[n-1, n)$ is an interval we see that $T[n-1, n), \ldots, T^{2^{n-1}-1}[n-1, n)$ are intervals as well. By Corollary 3, $T^{2^{n-1}}[n-1, n) = [n, \infty)$ is an interval as well. We have

$$\begin{aligned} T^{2^{n-1}+1} &= T[n, \infty) \\ &= \bigcup_{k=n}^{\infty} T[k, k+1) \\ &= \bigcup_{k=n}^{\infty} \left[\frac{1}{k+1}, \frac{1}{k}\right) \\ &= (0, 1/n). \end{aligned}$$

Thus, both $T^{2^{n-1}}[n-1, n)$ and $T^{2^{n-1}+1}[n-1, n)$ are intervals, and $T^{2^{n-1}+1}[n-1, n) \subseteq [0, 1)$.

We see that $T^{2^{n-1}+2}[n-1,n), \ldots, T^{2^n-1}[n-1,n)$ are also intervals, so (iii) follows for all $n \geq 1$.

$\square$

The restrictions $0 \leq k, l < 2^n$ in (i) and (ii) of Lemma 4 are necessary, as $T^2(0,1) \cap (0,1)$ is not empty. In fact, as we will show here, each of the intervals $(n-1,n)$ is recurrent, with a period of $2^n$.

**Lemma 5.** *For any $n \geq 1$ we have $T^{2^n}[n-1,n) = (n-1,n)$.*

*Proof.* For $n=1$, we have

$$T^2[0,1) = T[1,\infty) = (0,1).$$

Now suppose $T^{2^n}[n-1,n) = (n-1,n)$ for some $n \geq 1$. Then for $n+1$ we have by Lemma 3 that

$$\begin{aligned} T^{2^{n+1}}[n,n+1) &= \{T^{2^{n+1}}(x) : x \in [n,n+1)\} \\ &= \{T^{2^n}(x-1)+1 : x \in [n,n+1)\} \\ &= \{T^{2^n}(x)+1 : x \in [n-1,n)\} \\ &= T^{2^n}[n-1,n)+1 = (n,n+1). \end{aligned}$$

The statement follows by induction. $\square$

We see that while $[n-1,n)$ is almost recurrent, the difference $[n-1,n) \setminus T^{2^n}[n-1,n)$ is nonempty. As we will see in the next lemma, we must generally be careful with our equalities.

**Lemma 6.** *We have*

$$\biguplus_{m=0}^{2^n-1} T^m[n-1,n) = \mathbb{R}_+ \setminus \{a_0, \ldots, a_{2^{n-1}-2}\}$$

*for all $n \geq 1$.*

*Proof.* For $n=1$, $\{a_0, \ldots, a_{2^{n-1}-2}\}$ is empty, and indeed $[0,1) \cup T[0,1) = [0,1) \cup [1,\infty) = \mathbb{R}_+$ as follows from the properties on page 15.

For $n=2$, we have

$$\bigcup_{m=0}^{3} T^m[1,2) = [1,2) \cup [1/2,1) \cup [2,\infty) \cup (0,1/2) = (0,\infty) = \mathbb{R}_+ \setminus \{a_0\}$$

and $2^{2-1}-2=0$. Let $n \geq 2$ and suppose the we have $\biguplus_{m=0}^{2^n-1} T^m[n-1,n) = \mathbb{R}_+ \setminus \{a_0, \ldots, a_{2^{n-1}-2}\}$.

Then for $n+1$ we have

$$\begin{aligned}
\bigcup_{m=0}^{2^{n+1}-1} T^m[n,n+1) &= \left(\bigcup_{m=0}^{2^n-1} T^{2m}[n,n+1)\right) \cup \left(\bigcup_{m=0}^{2^n-1} T^{2m+1}[n,n+1)\right) \\
&= \left(\bigcup_{m=0}^{2^n-1} (T^m[n-1,n)+1)\right) \cup T\left(\bigcup_{m=0}^{2^n-1} T^m([n-1,n)+1)\right) \\
&= (1+\mathbb{R}_+ \setminus \{a_0,\dots,a_{2^{n-1}-2}\}) \cup T(1+\mathbb{R}_+ \setminus \{a_0,\dots,a_{2^{n-1}-2}\}) \\
&= ([1,\infty) \setminus \{a_0+1,\dots,a_{2^{n-1}-2}+1\}) \cup T([1,\infty) \setminus \{a_0+1,\dots,a_{2^{n-1}-2}+1\}),
\end{aligned}$$

where the second equality follows from Lemma 3 and the third equality follows from the induction hypothesis. Notice that

$$T([1,\infty) \setminus \{a_0+1,\dots,a_{2^{n-1}-2}+1\}) = ((0,1) \setminus \{T(a_0+1),\dots,T(a_{2^{n-1}-2}+1)\}.$$

Since the right child of $a_k$ is $a_k+1 = a_{2k+1}$ by Lemma 2, it follows that

$$\{a_0+1,\dots,a_{2^{n-1}-2}+1\} = \{a_{2k+1}\}_{k=0}^{2^{n-1}-2} = \{a_1,a_3,\dots,a_{2^n-3}\},$$

so

$$\{T(a_0+1),\dots,T(a_{2^{n-1}-2}+1)\} = \{a_{2k+2}\}_{k=0}^{2^{n-1}-2} = \{a_2,a_4,\dots,a_{2^n-2}\}.$$

As $\{a_0+1,\dots,a_{2^{n-1}-2}+1\} \subseteq [1,\infty)$ and $\{T(a_0+1),\dots,T(a_{2^{n-1}-2}+1)\} \subseteq (0,1)$, it follows that

$$\bigcup_{m=0}^{2^{n+1}-1} T^m[n,n+1) = \mathbb{R}_+ \setminus \{a_0,\dots,a_{2^n-2}\}$$

and this union is in fact disjoint by (i) from Lemma 4. By induction, we conclude that

$$\biguplus_{m=0}^{2^n-1} T^m[n-1,n) = \mathbb{R}_+ \setminus \{a_0,\dots,a_{2^{n-1}-2}\}$$

for all $n \geq 1$. $\square$

Finally, we have enough tools to show the first of the main results of this thesis.

**Theorem 6.** *$T$ has no periodic points.*

*Proof.* Let $x \in \mathbb{R}_+$. If $x$ is rational, suppose some $p > 0$ exists such that $T^p(x) = x$. Then the function that takes $n$ to $a_n$ is not injective, and this is in contradiction with Theorem 4.

Assume $x$ is irrational. In Lemma 6, we established for each $n \geq 1$ that

$$\mathbb{R}_+ \setminus \biguplus_{m=0}^{2^n-1} T^m[n-1, n) \subset \mathbb{Q}_+$$

meaning that for any $n \geq 1$ there exists a $0 \leq k_n < 2^n$ such that $x \in T^{k_n}[n-1, n)$. As $T$ is injective, we have

$$T^m \left(T^{k_n}[n-1, n)\right) \cap T^{k_n}[n-1, n) = T^{k_n} \left(T^m[n-1, n) \cap [n-1, n)\right) = \emptyset$$

for any $0 < m < 2^n$ by (i) from Lemma 4. Now choose any $n \geq 1$. If $x$ is a periodic point, its period $p$ must satisfy $p \geq 2^n$. As $n$ is arbitrary, we find that no such $p$ can exist, so $x$ is not a periodic point. Therefore, we conclude that $T$ cannot have any periodic points. $\square$

# 4 Measure theoretic properties

So far, we have studied the behaviour of the Calkin-Wilf function and its iterates. Despite every interval of the form $T^k(n-1,n)$ for positive integers $k$ and $n$ returning to itself after $2^n$ iterations, $T$ has no periodic points whatsoever.

In this section, we will study $\mathbb{R}_+$ equipped with the transformation $T$ through a measure theoretic lens. By constructing measures related to $T$, we might not be able to make claims about the individual orbits under $T$, but we will have the tools to discuss more general behaviour of the system.

Our first step is to construct an invariant measure for $T$.

**Proposition 2.** *Let $\tau$ be the counting measure on $\mathbb{R}_+ \setminus \mathbb{Q}_+$. On the measure space $(\mathbb{R}_+, \mathscr{P}(\mathbb{R}_+), \tau)$, $T$ is a measure preserving transformation.*

*Proof.* Suppose $A$ is a subset of $\mathbb{R}_+ \setminus \mathbb{Q}_+$. As $T : \mathbb{R}_+ \to \mathbb{R}_+ \setminus \{0\}$ is invertible, and as we have $T(\mathbb{R}_+ \setminus \mathbb{Q}_+) = T(\mathbb{R}_+) \setminus T(\mathbb{Q}_+) = \mathbb{R}_+ \setminus \mathbb{Q}_+$ by injectivity and Theorem 4, there exists a bijection between $A$ and its preimage $T^{-1}A$ (namely $T$). By bijectivity we have $\tau(T^{-1}A) = \tau(A)$. $\square$

Given any irrational $x$, we can construct an invariant measure by restricting $\tau$ to the two-sided $T$-orbit $\Gamma_x$ of $x$. Notice that for a rational number $y \in \mathbb{Q}_+$ we have $0 \in \Gamma_y$, so $\tau|_{\Gamma_y}(\{0\}) = 1$, while on the other hand $\tau|_{\Gamma_y}(T^{-1}\{0\}) = \tau|_{\Gamma_y}(\emptyset) = 0$. Thus, the restriction of $\tau$ to $\Gamma_y$ is not an invariant measure if $y$ is rational.

## 4.1 An ergodic measure

Although we have just built an invariant measure, it is infinite and therefore not as workable as one would like. To remedy this, we will construct a different measure using Theorem 2.

First, let $\mathscr{S}_0 = \{\emptyset\} \cup \{\{x\} : x \in \mathbb{Q}_+\}$ and define

$$\mathscr{S}_n = \{[n-1,n), T[n-1,n), \ldots, T^{2^n-1}[n-1,n)\}$$

for $n \geq 1$. We let $\mathscr{S} = \bigcup_{n=0}^{\infty} \mathscr{S}_n$ and show this is a semi-ring. Before we can do this, we show some properties of the elements of $\mathscr{S}$.

**Lemma 7.** *Let $R, S \in \mathscr{S}$. If $S \cap R \neq \emptyset$, we either have $S \subseteq R$ or $R \subseteq S$.*

*Proof.* Let $n, m \geq 0$ be integers such that $S \in \mathscr{S}_n$ and $R \in \mathscr{S}_m$. If $m = 0$, $R$ contains at most a single element, so we either have $S \cap R = \emptyset$ or $R \subseteq S$. In case $n = m \geq 1$, we have that $S \cap R = \emptyset$ or $S = R$ by (i) from Lemma 4. Therefore, we assume $n > m > 0$ for the rest of this proof. There exist some $0 \leq k < 2^n$ so that $S = T^k[n-1,n)$ and some $0 \leq l < 2^m$ so that $R = T^l[m-1,m)$. We separately consider the cases $k = l$, $k < l$ and $k > l$.

Suppose $k = l$. Then

$$S \cap R = T^k[n-1, n) \cap T^k[m-1, m) = T^k([n-1, n) \cap [m-1, m))$$

by injectivity, and this is empty. Now, let us assume that $k < l$. Then

$$T^k[n-1, n) \cap T^l[n-1, n) = T^k([n-1, n) \cap T^{l-k}[m-1, m)).$$

Note that $0 < l-k \leq l < 2^m$. If $l-k = 2^{m-1}$, then we have $T^{l-k}[m-1, m) = [m, \infty)$ by Corollary 3. As $n > m$, we have $[n-1, n) \subseteq T^{l-k}[m-1, m)$, so $S = T^k[n-1, n) \subseteq T^l[m-1, m) = R$. If $l-k \neq 2^{m-1}$, we have by (ii) from Lemma 4 that $T^{l-k}[m-1, m) \subseteq [j-1, j)$ for some integer $j \geq 1$, meaning that either $[n-1, n) \cap T^{l-k}[m-1, m)$ is empty or $T^{l-k}[m-1, m) \subseteq [n-1, n)$, and so $R = T^l[m-1, m) \subseteq T^k[n-1, n) = S$.

Finally, suppose $l < k$, so that $T^k[n-1, n) \cap T^l[m-1, m) = T^l(T^{k-l}[n-1, n) \cap [m-1, m))$. Once again we notice $0 < k-l \leq k < 2^n$. If $k-l = 2^{n-1}$, then $T^{k-l}[n-1, n) = [n, \infty)$ by Corollary 3, so by $m < n$ we find $R \cap S = \emptyset$. If $k - l \neq 2^{m-1}$, we have by (ii) from Lemma 4 that $T^{k-l}[n-1, n) \subseteq [j-1, j)$ for some integer $j \geq 1$, so that either $T^{k-l}[n-1, n) \cap [m-1, m)$ is empty or $T^{k-l}[n-1, n) \subseteq [m-1, m)$, and so $S = T^k[n-1, n) \subseteq T^l[m-1, m) = R$. $\square$

**Lemma 8.** *Let $S \in \mathscr{S}_n$ for some $n > 0$ and suppose $m > n$ is an integer. Then $S$ is a disjoint union of $2^{m-n}$ sets in $\mathscr{S}_m$ and finitely many sets in $\mathscr{S}_0$.*

*Proof.* Let $S \in \mathscr{S}_n$ for some $n > 0$ and let $0 \leq k < 2^n$ be such that $S = T^k[n-1, n)$. For $m = n+1$, we can write $[n-1, n)$ as

$$[n-1, n) = (n-1, n-1/2) \cup [n-1/2, n) \cup \{n-1\}.$$

Suppose first that $0 \leq k < 2^{n-1}$. We have $\{n-1\} \in \mathscr{S}_0$, so we need to show that the intervals are contained in $\mathscr{S}_m$. For the first of these sets, we can use $n-1$ applications of Lemma 3 to obtain

$$\begin{aligned}
(n-1, n-1/2) &= (0, 1/2) + n - 1 \\
&= T^3[1, 2) + n - 1 \\
&= T^6[2, 3) + n - 2 \\
&\vdots \\
&= T^{3 \cdot 2^{n-1}}[n, n+1) + n - n \in \mathscr{S}_{n+1}.
\end{aligned}$$

We can apply the same process on the second set to find

$$\begin{aligned}[n-1/2,n) &= [1/2,1)+n-1 \\ &= T[1,2)+n-1 \\ &= T^2[2,3)+n-2 \\ &\ \ \vdots \\ &= T^{2^{n-1}}[n,n+1)+n-n \in \mathscr{S}_{n+1}.\end{aligned}$$

As $S = T^k[n-1,n)$, we can use the injectivity of $T$ to find

$$\begin{aligned}S &= T^k((n-1,n-1/2) \cup [n-1/2,n) \uplus \{n-1\}) \\ &= T^{k+2^{n-1}}[n,n+1) \cup T^{k+3\cdot 2^{n-1}}[n,n+1) \cup \{T^k(n-1)\}.\end{aligned}$$

These sets are disjoint by (i) from Lemma 4 as $0 \le k+2^{n-1} < k+3\cdot 2^{n-1} < 2^n$, so $S$ is a disjoint union of 2 sets in $\mathscr{S}_{n+1}$ and finitely many sets in $\mathscr{S}_0$. Now suppose $2^{n-1} \le k < 2^n$. We have $T^{2^{n-1}}[n-1,n) = [n,\infty)$ by Corollary 3, and

$$[n,\infty) = [n,n+1) \cup [n+1,\infty).$$

We have $[n+1,\infty) = T^{2^n}[n,n+1)$, so now

$$\begin{aligned}S &= T^k[n-1,n) \\ &= T^{k-2^{n-1}}[n,\infty) \\ &= T^{k-2^{n-1}}([n,n+1) \cup [n+1,\infty)) \\ &= T^{k-2^{n-1}}[n,n+1) \cup T^{k+2^{n-1}}[n,n+1).\end{aligned}$$

As $0 \le k-2^{n-1} \le k+2^{n-1} < 2^n$, these sets are disjoint by (i) from Lemma 4, so in this case $S$ is a disjoint union of 2 sets in $\mathscr{S}_{n+1}$ and finitely many sets in $\mathscr{S}_0$ as well. This shows that the statement holds for the base case $m = n+1$.

Now suppose $m > n$ is an integer such that any $S \in \mathscr{S}_n$ can be written as a disjoint union of $2^{m-n}$ sets $S_1, S_2, \ldots, S_{2^{m-n}} \in \mathscr{S}_m$ and finitely many sets in $\mathscr{S}_0$. Then by the above, we can write each of the $S_i$ as a finite disjoint union of two sets $S_{i,1}, S_{i,2} \in \mathscr{S}_{m+1}$ and finitely many sets in $\mathscr{S}_0$, so $S$ is a disjoint union of $2^{m+1-n}$ sets in $\mathscr{S}_{m+1}$ and finitely many sets in $\mathscr{S}_0$. The claim follows by induction. $\square$

Using Lemma 8, we can make a refinement to Lemma 7.

**Corollary 4.** *Suppose $0 < n < m$ are integers and let $S \in \mathscr{S}_n$ and $R \in \mathscr{S}_m$. Then either $R \subseteq S$ or $S \cap R = \emptyset$. Also, we have $R \neq S$.*

*Proof.* Let $S \in \mathscr{S}_n$ and $R \in \mathscr{S}_m$ for integers $0 < n < m$. By Lemma 8, $S$ can be written as a finite union of sets in $\mathscr{S}_m \cup \mathscr{S}_0$. Now as the sets in $\mathscr{S}_m$ are disjoint by (i) from Lemma 4, we either have $R \subseteq S$ or $S \cap R = \emptyset$. This proves the first claim.

Each of the sets in $\mathscr{S}_m$ is nonempty. As $S$ can be written as this finite disjoint union of at least $2^{m-n} \geq 2$ sets in $\mathscr{S}_m \cup \mathscr{S}_0$, the difference $S \setminus R$ is nonempty, and so $R \neq S$. □

We are now ready to prove the following theorem.

**Theorem 7.** *The collection $\mathscr{S}$ is a semi-ring.*

*Proof.* Definition 2 lists three conditions which we must show to hold true for $\mathscr{S}$.

a. By definition, we have $\emptyset \in \mathscr{S}$.

b. We show that $\mathscr{S}$ is $\cap$-stable. Let $R, S \in \mathscr{S}$ and suppose $S \cap R \neq \emptyset$. Then by Lemma 7, we either have $R \subseteq S$ or $S \subseteq R$. In the first case, we have $S \cap R = S \in \mathscr{S}$, and in the other case we have $S \cap R = R \in \mathscr{S}$. Thus, $\mathscr{S}$ is $\cap$-stable.

c. First, let $S \in \mathscr{S}_0$ and $R \in \mathscr{S}$. As $S$ contains at most one element, we either have $S \setminus R = S$ or $S \setminus R = \emptyset$, so $S \setminus R$ is a finite union of elements in $\mathscr{S}$.

   Now suppose $n, m > 0$ are integers such that $S \in \mathscr{S}_n$ and $R \in \mathscr{S}_m$. Without loss of generality, assume $n \leq m$. If $S \cap R = \emptyset$, then $S \setminus R = S$ and $R \setminus S = R$. Also, if $n = m$, then either $S = R$ or $S \cap R = \emptyset$ by (i) from Lemma 4, so we can assume $n < m$. Suppose $S \cap R \neq \emptyset$, meaning that $R \subseteq S$ by Corollary 4.

   In Lemma 8, we have seen that $S$ can be written as a finite disjoint union of $2^{m-n}$ sets in $\mathscr{S}_m$ and finitely many sets in $\mathscr{S}_0$. As all of the sets in $\mathscr{S}_m$ are disjoint by (i) from Lemma 4, $R$ must be one of those sets, so $S \setminus R$ is a finite union of $2^{m-n} - 1$ sets in $\mathscr{S}_m$ and finitely many sets in $\mathscr{S}_0$. This gives us that $S \setminus R$ is a finite disjoint union of sets in $\mathscr{S}$ for all $S, R \in \mathscr{S}$.

We have proven all properties, and conclude $\mathscr{S}$ is a semi-ring. □

Now let $\mu^* : \mathscr{S} \to [0, 1]$ be the set function defined by $\mu^*(A) = 2^{-n}$ if $A \in \mathscr{S}_n$ for some $n \geq 1$, and zero otherwise. In order to extend this function to a measure $\mu : \sigma(\mathscr{S}) \to [0, 1]$, we need to show $\mu^*$ is a premeasure. Notice that $\mu^*$ is well defined by Lemma 8, as there are no sets $A$ such that $A \in \mathscr{S}_n$ and $A \in \mathscr{S}_m$ for $n \neq m$.

By Lemma 8, we can write $A = R_l \cup \biguplus_{i=1}^{2^{l-k}} A_i$ for sets $A_1, A_2, \ldots, A_{2^{l-k}} \in \mathscr{S}_l$ for any integer $l \geq k$ and $R_l$ a finite union of sets in $\mathscr{S}_0$. This gives us

$$\begin{aligned} \mu^*(A) &= 2^{-k} \\ &= 2^{l-k} \cdot 2^{-l} \\ &= \sum_{n=1}^{2^{l-k}} \mu^*(A_n) + \sum_{x \in R_l} \mu^*(\{x\}). \end{aligned} \tag{2}$$

and so $\mu^*(A) = \sum_{n=1}^{m} \mu^*(A_n)$ if all $A_n$ are in $\mathscr{S}_l \cup \mathscr{S}_0$ for some $l \geq k$.

**Theorem 8.** *$\mu^*$ is a premeasure.*

*Proof.* Recall the two conditions of Theorem 2. We will show both hold for $\mu^*$.

a. By definition, we have $\mu^*(\emptyset) = 0$.

b. Let $A \in \mathscr{S}_k$ for some integer $k > 0$. Suppose $A_1, A_2, \ldots, A_m \in \mathscr{S}$ are sets such that $A = \biguplus_{n=1}^{m} A_n$. We first show that $\mu^*(A) = \sum_{n=1}^{m} \mu^*(A_n)$.

   Let $k_n \geq 0$ be the integers such that $A_n \in \mathscr{S}_{k_n}$. We assume that $k \neq 0$. If there exists an $n$ such that $0 \neq k_n < k$ we find $A \subseteq A_n$, so $A = A_n$ which is not possible by Corollary 4. Thus, we have $k_n \geq k$ or $k_n = 0$ for all $n \geq 1$. As there are only finitely many $A_i$, we can define the maximum of the $k_n$ as $k_\infty = \max\{k_n : 1 \leq n \leq m\} \geq k$. We can write each of the $A_n$ as a finite disjoint union of $2^{k_\infty - k_n}$ sets in $\mathscr{S}_{k_\infty}$ and $M_n - 2^{k_\infty - k_n}$ sets in $\mathscr{S}_0$ for some $M_n < \infty$. Denote these sets by $B_{1,n}, B_{2,n}, \ldots, B_{M_n,n} \in \mathscr{S}_{k_\infty} \cup \mathscr{S}_0$. Then notice that

$$A = \bigcup_{n=1}^{m} \biguplus_{i=1}^{M_n} B_{i,n},$$

   and so by (2)

$$\mu^*(A) = \sum_{n=1}^{m} \sum_{i=1}^{M_n} \mu^*(B_{i,n}) = \sum_{n=1}^{m} \mu^*(A_n). \tag{3}$$

   Now let $(A_n)_n \subseteq \mathscr{S}$ be a sequence such that $\biguplus_{n=1}^{\infty} A_n = A \in \mathscr{S}$. For all $n \geq 1$, we define $k_n \geq 0$ such that $A_n \in \mathscr{S}_{k_n}$. Notice that for any $n$, we have $A_n \subseteq A$. Thus, if $k = 0$, we must have $k_n = 0$ for all $n$ as $A$ contains at most a single element. This means that

$$\mu^*(A) = 0 = \sum_{n=1}^{\infty} \mu^*(A_n).$$

We assume that $k \neq 0$. Once again notice that we have $k_n \geq k$ or $k_n = 0$ for all $n \geq 1$. For all $l \geq 0$ we define $N_l = \{n : k_n = k + l\}$. Now for any $l \geq 0$, we can write $A$ as a disjoint union of $2^l$ sets in $\mathscr{S}_{k+l}$ and $M_l - 2^l < \infty$ sets in $\mathscr{S}_0$ by Lemma 8. We denote these sets by $B_{1,l}, B_{2,l}, \ldots, B_{M_l,l}$ and write $A = \biguplus_{i=1}^{M_l} B_{i,l}$. Let $I_l = \{i : \exists\, n \in N_l \text{ such that } A_n = B_{i,l}\}$. We have

$$\mu^*(A) = \mu^*\left(\biguplus_{i=1}^{M_l} B_{i,l}\right) = \sum_{i=1}^{M_l} \mu^*(B_{i,l}) = \sum_{n \in N_l} \mu^*(A_n) + \sum_{i \notin I_l} \mu^*(B_{i,l}),$$

where the second equality follows from (3). Notice that $\biguplus_{i \notin I_l} B_{i,l} = \biguplus_{i \notin N_l} A_i$. Therefore, for any $n \geq 0$ we can define

$$J_n = \bigcap_{l=0}^{n} \biguplus_{i \notin I_l} B_{i,l} = \bigcap_{l=0}^{n} \biguplus_{i \notin N_l} A_i$$

and notice $J_{n+1} \subseteq J_n$ for all $n \geq 0$. Suppose $x \in \lim_{n\to\infty} J_n = \bigcap_{l=1}^{\infty} \biguplus_{i \notin N_l} A_i$. As $\bigcup_{l=1}^{\infty} N_l = \mathbb{Z}_{>0}$, we have that $x \notin A_i$ for any $i \geq 1$. Since $A = \biguplus_{n=1}^{\infty} A_n$ and $\lim_{n\to\infty} J_n \subseteq A$, this $x$ cannot exist, so $\lim_{n\to\infty} J_n = \emptyset$.

As $J_{n+1} \subseteq J_n$ for all $n \geq 0$, $J_n$ is a finite union of sets in $\mathscr{S}_{n+k} \cup \mathscr{S}_0$. We also have $J_n \subseteq A$ for all $n \geq 0$, so there exist finite index sets $K_n \subseteq \{1, 2, \ldots, M_n\}$ such that $J_n = \biguplus_{k \in K_n} B_{k,n}$ by Lemma 8 and

$$A = J_n \cup \bigcup_{l=0}^{n} \biguplus_{i \in N_l} A_i = \biguplus_{k \in K_n} B_{k,n} \cup \bigcup_{l=0}^{n} \biguplus_{i \in N_l} A_i.$$

Since this is a finite union, we can use (3) to obtain

$$\mu^*(A) = \sum_{l=1}^{n} \sum_{i \in N_l} \mu^*(A_i) + \sum_{k \in K_n} \mu^*(B_{k,n})$$

for any $n \geq 0$.

We have seen that $J_n \to \emptyset$ as $n \to \infty$. This means that $\#K_n \to 0$, so we finally conclude

$$\begin{aligned}\mu^*(A) &= \lim_{n\to\infty}\left(\sum_{l=1}^{n}\sum_{i\in N_l}\mu^*(A_i) + \sum_{k\in K_n}\mu^*(B_{k,n})\right)\\ &= \lim_{n\to\infty}\left(\sum_{l=1}^{n}\sum_{i\in N_l}\mu^*(A_i) + 2^{-n}\cdot\#K_n\right)\\ &= \sum_{l=1}^{\infty}\sum_{i\in N_l}\mu^*(A_i)\\ &= \sum_{i=1}^{\infty}\mu^*(A_n)\end{aligned}$$

by $\bigcup_{n=0}^{\infty} N_n = \mathbb{Z}_{>0}$.

We conclude that $\mu^*$ is a premeasure. $\square$

Finally, we can now use Theorem 2 to extend $\mu^*$ to a measure $\mu : \sigma(\mathscr{S}) \to [0,\infty]$. Notice that $\mu$ is a probability measure, as

$$\mu(\mathbb{R}_+) = \mu(\{0\}\cup(0,1)\cup\{1\}\cup(1,\infty)) = 2^{-1} + 2^{-1} = 1.$$

In order to prove ergodicity of $\mu$, we require one more lemma.

**Lemma 9.** *Let $x \in \mathbb{R}_+ \setminus \mathbb{Q}_+$, and suppose $(A_n)_n \subseteq \mathscr{S}$ is a sequence of sets such that $x \in A_n$ and $A_n \in \mathscr{S}_n$ for all $n \geq 1$. Then $\lim_{n\to\infty} A_n = \{x\}$.*

*Proof.* First, notice that there exists an $A_1 \in \mathscr{S}_1$ with $x \in A_1$, as $\mathscr{S}_1 = \{[0,1),[1,\infty)\}$. Now suppose there exists $A_n \in \mathscr{S}_n$ such that $x \in \mathscr{S}_n$. For any $n \geq 1$, we can rewrite $A_n$ as a disjoint union of sets in $\mathscr{S}_{n+1} \cup \mathscr{S}_0$ for any $n \geq 1$. As $x$ is irrational, there is no set $X \in \mathscr{S}_0$ such that $x \in X$, and thus there must exist a set $A_{n+1} \subseteq A_n$ with $A_{n+1} \in \mathscr{S}_{n+1}$ and $x \in A_{n+1}$. Therefore, $(A_n)_n$ exists as desired, and we have $A_1 \supseteq A_2 \cdots \supseteq A_n \supseteq \dots$.

Now let $y \in \mathbb{R}_+ \setminus \mathbb{Q}_+$. Without loss of generality, we let $y > x$. Let $q \in \mathbb{Q}_+$ be such that $x < q < y$, and $m \geq 1$ such that $q = a_m$. If we choose $N$ such that $m < 2^{N-1} - 2$, then $q \in \mathbb{R}_+ \setminus \bigcup_{k=0}^{2^N-1} T^k[N-1,N)$ by Lemma 6. As the $A_n$ are intervals by (iii) from Lemma 4, we must have that $y \notin A_n$ for all $n \geq N$. Hence, $\lim_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} A_n = \{x\}$. $\square$

**Theorem 9.** *The dynamical system $(\mathbb{R}_+, \sigma(\mathscr{S}), \mu, T)$ is measure preserving and ergodic.*

*Proof.* Let $\mathscr{D}_{inv} \subseteq \sigma(\mathscr{S})$ be defined by

$$\mathscr{D}_{inv} = \{A \in \sigma(\mathscr{S}) : \mu(A) = \mu(T^{-1}A)\}.$$

We show that this is a Dynkin system.

1.a As $T^{-1}\mathbb{R}_+ = \mathbb{R}_+$, we have $\mu(\mathbb{R}_+) = \mu(T^{-1}\mathbb{R}_+)$, so $\mathbb{R}_+ \in \mathscr{D}_{inv}$.

1.b If $D \in \mathscr{D}_{inv}$, we can use the fact that $\mu$ is a probability measure to find

$$\mu(D^C) = \mu(\mathbb{R}_+) - \mu(D) = \mu(\mathbb{R}_+) - \mu(T^{-1}D) = \mu((T^{-1}D)^C) = \mu(T^{-1}(\mathbb{R}_+ \setminus D)),$$

so $D^C \in \mathscr{D}_{inv}$.

1.c Suppose $(A_n)_n$ is a sequence of pairwise disjoint sets in $\mathscr{D}_{inv}$, then $(T^{-1}A_n)_n$ is a disjoint sequence as well and

$$\mu\left(\biguplus_{n=0}^{\infty} A_n\right) = \sum_{n=0}^{\infty} \mu(A_n) = \sum_{n=0}^{\infty} \mu(T^{-1}A_n) = \mu\left(\biguplus_{n=0}^{\infty} T^{-1}A_n\right).$$

Hence, $\mathscr{D}_{inv}$ is a Dynkin system.

Now let $A \in \mathscr{S}$. Then there exists some $n$ such that $A \in \mathscr{S}_n$, and there exists a $B \in \mathscr{S}_n$ so that $T^{-1}A \subseteq B$ and $\mu(B \setminus T^{-1}A) = 0$. This means that $\mu(T^{-1}A) = 2^{-n} = \mu(A)$, so $\mathscr{S} \subseteq \mathscr{D}_{inv}$.

Notice that $\mathscr{S}$ is $\cap$-stable by Theorem 7. We now have $\delta(\mathscr{S}) \subseteq \mathscr{D}_{inv} \subseteq \sigma(\mathscr{S})$ and by Theorem 1 and the $\cap$-stability of $\mathscr{S}$, we also see that $\sigma(\mathscr{S}) = \delta(\mathscr{S})$. Hence we have $\mu(A) = \mu(T^{-1}A)$ for all $A \in \sigma(\mathscr{S})$, and so $T$ is measure preserving.

We now prove that $T$ is ergodic. Suppose $x, y \in \mathbb{R}_+ \setminus \mathbb{Q}_+$ are such that $x < y$. Define sequences $(A_n)_n, (B_n)_n \subseteq \mathscr{S}$ as in Lemma 9 such that $x \in A_n$, $y \in B_n$ and $A_n, B_n \in \mathscr{S}_n$ for all $n \geq 1$. Suppose $f : \mathbb{R}_+ \to \mathbb{R}$ is a measurable function such that $f = f \circ T$. As $A_n, B_n \in \mathscr{S}_n$, there exist integers $0 \leq k_n < 2^n$ such that either $A_n = T^{k_n}B_n$ or $B_n = T^{k_n}A_n$. Therefore, we have $f(A_n) = f(B_n)$ for all $n \geq 1$. Taking the limit, we find

$$\{f(x)\} = \lim_{n\to\infty} f(A_n) = \lim_{n\to\infty} f(B_n) = \{f(y)\},$$

so $T$ is ergodic by Proposition 1. □

We can prove that $\sigma(\mathscr{S}) = \mathscr{B}$, where $\mathscr{B}$ denotes the Borel $\sigma$-algebra. To do this, we first paraphrase the following from page 35 from Chapter 42 of [8].

**Theorem 10.** *Let $(X, \mathscr{B}, \mu)$ be a measure space where $\mathscr{B}$ denotes the Borel $\sigma$-algebra. Suppose $\Sigma$ is a $\sigma$-algebra such that the following properties hold.*

*10.a We have $\Sigma \subseteq \mathscr{B}$.*

*10.b There exists a countable collection of subsets $\mathscr{A} \subseteq \mathscr{P}(X)$ such that $\Sigma = \sigma(\mathscr{A})$.*

*10.c For all $x, y \in X$, there exist $A, B \in \Sigma$ such that $x \in A$, $y \in B$ and $A \cap B = \emptyset$.*

*Then $\Sigma = \mathscr{B}$.*

**Theorem 11.** *We have $\sigma(\mathscr{S}) = \mathscr{B}$.*

*Proof.* We show that all properties of Theorem 10 hold for $\sigma(\mathscr{S})$.

10.a As every set $A \in \mathscr{S}$ is a Borel set, we have $\sigma(\mathscr{S}) \subseteq \mathscr{B}$.

10.b We have that $\mathscr{S}$ is a countable collection of subsets of $\mathbb{R}_+$ by definition.

10.c Let $x, y \in \mathbb{R}_+$. By Lemma 9, there exists a sequence $(A_n)_n \in \mathscr{S}$ such that $A_n \in \mathscr{S}_n$, and $x \in A_n$ for all $n \geq 1$, and we have $\lim_{n\to\infty} A_n = \{x\}$. Therefore, there exists an $n \geq 1$ such that $y \in \mathbb{R}_+ \setminus A_n$.

Thus, we conclude that $\sigma(\mathscr{S}) = \mathscr{B}$. □

# 5 Continued fractions

As we now have an ergodic measure, we can make a link between $T$ and minus continued fractions. Let $x \in \mathbb{R}_+$. We can rearrange our equation as follows.

$$\begin{aligned} T(x) &= \frac{1}{2\lfloor x \rfloor - x + 1} \\ x &= 2\lfloor x \rfloor + 1 - \frac{1}{T(x)}. \end{aligned}$$

Similarly, we can use our expression of $T^n(x)$ to find the following equality for all $n \geq 1$.

$$\begin{aligned} T^n(x) &= \frac{1}{2\lfloor T^{n-1}(x) \rfloor - T^{n-1}(x) + 1} \\ T^{n-1}(x) &= 2\lfloor T^{n-1}(x) \rfloor + 1 - \frac{1}{T^n(x)}. \end{aligned}$$

For any $n \geq 0$, we define a function $b_n : \mathbb{R}_+ \to \mathbb{Z}$ by $b_n(x) = 2\lfloor T^n(x) \rfloor + 1$ and write

$$x = b_0(x) + \cfrac{-1}{b_1(x) + \cfrac{-1}{b_2(x) + \cfrac{-1}{\cdots + \cfrac{-1}{b_n(x) + \cfrac{-1}{T^{n+1}(x)}}}}}.$$

By removing the tail part, we can define a sequence $(p_n(x)/q_n(x))_n$ by

$$\frac{p_n(x)}{q_n(x)} = b_0(x) + \cfrac{-1}{b_1(x) + \cfrac{-1}{b_2(x) + \cfrac{-1}{\cdots + \cfrac{-1}{b_n(x)}}}},$$

where the integers $p_n(x)$ and $q_n(x)$ are coprime. In this thesis, we will not show whether $(p_n(x)/q_n(x))_n$ converges to $x$ for any $x \in \mathbb{R}_+$, and we write this chapter assuming $\lim_{n\to\infty} p_n(x)/q_n(x) = x$ for $\mu$-almost every $x$.

However, we can analyse some of the properties of the $b_n(x)$, and we will spend the remainder of this chapter computing the arithmetic and geometric means of the $b_n(x)$ for almost every $x \in \mathbb{R}_+$.

The arithmetic mean of the digits is equal to

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} b_i(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \left(2\lfloor T^i(x) \rfloor + 1\right).$$

Notice that the function $f : \mathbb{R}_+ \to \mathbb{R}_+$ given by $f(x) = 2\lfloor x \rfloor + 1$ is in $L^1(\mathbb{R}_+, \mathscr{B}, \mu)$, as

$$\int_{\mathbb{R}_+} |f| \, d\mu \leq \int_{\mathbb{R}_+} 2\lfloor x \rfloor + 1 \, \mu(dx) = 1 + \sum_{k=0}^{\infty} 2k\mu([k, k+1)) = 1 + 2\sum_{k=0}^{\infty} k2^{-k-1} < \infty.$$

As the measure $\mu$ is ergodic, we can use Theorem 3 to see that for $\mu$-almost every $x \in \mathbb{R}_+$,

$$\begin{aligned}\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} 2\lfloor T^i(x) \rfloor + 1 &= 2 \int \lfloor x \rfloor \, \mu(dx) + 1 \\ &= 1 + 2 \sum_{k=0}^{\infty} k\mu([k, k+1)) \\ &= 1 + 2 \sum_{k=1}^{\infty} k 2^{-k-1}.\end{aligned}$$

Notice that $\sum_{k=1}^{\infty} k2^{-k-1}$ is the derivative of $\sum_{k=1}^{\infty} x^k$ for $x = 1/2$. This gives us

$$\begin{aligned}\sum_{k=1}^{\infty} k2^{-k-1} &= \frac{d}{dx} \left[ \sum_{k=1}^{\infty} x^k \right]_{x=1/2} \\ &= \frac{d}{dx} \left[ \frac{1}{1-x} - 1 \right]_{x=1/2} \\ &= \frac{1}{(1-1/2)^2} = 4,\end{aligned}$$

and we find $\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} b_i(x) = 9$ $\mu$-almost everywhere.

Let us compute the geometric mean of the digits in the continued fraction of $x$. This is equal to $\lim_{n\to\infty} I_n(x)$, where $I_n(x) := \sqrt[n]{b_1(x)b_2(x)\cdots b_{n-1}(x)}$. We have

$$\lim_{n\to\infty} \log(I_n(x)) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \log(b_i(x)) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \log(2\lfloor x \rfloor + 1).$$

Define $g : \mathbb{R}_+ \to \mathbb{R}_+$ by $g(x) = \log(2\lfloor x \rfloor + 1)$. As $f \in L^1(\mathbb{R}_+, \mathscr{B}, \mu)$ and $|g| \leq |f|$, we have $\int |g| \, d\mu \leq \int |f| \, d\mu < \infty$, so $g \in L^1(\mathbb{R}_+, \mathscr{B}, \mu)$. We can now once again use Theorem 3 to find

$$
\begin{aligned}
\lim_{n\to\infty} \log I_n(x) &= \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \log b_i(x) \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \log(2\lfloor T^i(x) \rfloor + 1) \\
&= \int \log(2\lfloor x \rfloor + 1)\, \mu(dx) \\
&= \sum_{k=0}^{\infty} \log(2k+1)\mu([k, k+1)) \\
&= \sum_{k=0}^{\infty} \log(2k+1) 2^{-k}.
\end{aligned}
$$

With the use of an online calculator like Wolfram|Alpha, we can see that $\lim_{n\to\infty} \log(I_n(x))$ is approximately equal to 1.49133. We then find that

$$
\lim_{n\to\infty} I_n(x) = \lim_{n\to\infty} e^{\log(I_n(x))} \approx 4.44299
$$

for $\mu$-almost all $x \in \mathbb{R}_+$.

## References

[1] George Birkhoff. Proof of the ergodic theorem. *Proceedings of the National Academy of Sciences of the United States of America*, 17:656–660, 1931.

[2] Neil Calkin and Herbert S. Wilf. Recounting the rationals. 1999.

[3] Georg Cantor. Über unendliche, lineare punktmannichfaltigkeiten. *Mathematische Annalen*, 21:545–591, 1883.

[4] Karma Dajani and Cor Kraaikamp. *Ergodic Theory of Numbers*. The Mathematical Association of America, 2002.

[5] Reinhard Diestel. *Graph Theory*. Springer Berlin, Heidelberg, 2016.

[6] Edgar Dijkstra. *Selected Writings on Computing*. Springer Berlin, Heidelberg, 1982.

[7] Alex Smith Richard Strong Donald E. Knuth, C.P. Rupert. Recounting the rationals, continued. *The American Mathematical Monthly*, 110:642–643, 2003.

[8] David Fremlin. *Measure Theory, Volume 4*. Torres Fremlin, Colchester, 2013.

[9] Harry Furstenberg. *Recurrence in Ergodic Theory and Combinatorial Number Theory*. Princeton University Press, New Jersey, 1981.

[10] Sam Northshield. Re$^3$counting the rationals. 2019. preprint, `https://arxiv.org/pdf/1905.10369.pdf`.

[11] René L. Schilling. *Measures, Integrals and Martingales*. Cambridge University Press, second edition, 2017.

[12] Valentin Ovsienko Sophie Morier-Genoud. Farey boat: Continued fractions and triangulations, modular group and polygon dissections. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 121:91–136, 2019.

[13] John von Neumann. Proof of the quasi-ergodic hyppthesis. *Proceedings of the National Academy of Sciences of the United States of America*, 48:70–82, 1932.

[14] Hubert Stanley Wall. *Analytic Theory of Continued Fractions*. Courier Dover Publications, 2018.

[15] Peter Walters. *Ergodic Theory - Introductory Lectures*. Springer Berlin, Heidelberg, 1975.

[16] Moshe Israeli Yair Shapira, Avram Sidi. Optimal error bounds for convergents of a family of continued fractions. *Journal of Mathematical Analysis and Applications*, 197:767–773, 1996.

[17] Yasuhisa Yamada. A function from stern's diatomic sequence, and its properties. 2020. preprint, https://arxiv.org/pdf/2004.00278.pdf.