On Ergodic Theory in Non-Archimedean Settings

Thesis submitted in accordance with the requirements of the

University of Liverpool

for the degree of

Doctor of Philosophy

by

Alena Jaššová

November 2014

Abstract

On Ergodic Theory in Non-Archimedean Settings

Alena Jaššová

November 2014

In this thesis we use tools from ergodic theory to study ergodic and metric properties of Schneider's continued fraction map in non-Archimedean settings. We show that the natural extension of this map is isomorphic to a Bernoulli shift with entropy $\frac{\#(k)}{\#(k)-1}\log(\#(k))$. Results about various averages and moving averages of partial quotients of this expansion are proved as well.

We also use ergodic theoretic methods to prove the uniform distribution of $\boldsymbol{\beta}$-adic Halton sequences in the form $(\phi_{\boldsymbol{\beta}}(k_j))_{j\geq 1}$ where a sequence of non-negative integers $(k_j)_{j=1}^{\infty}$ is Hartman uniformly distributed and L^2-good universal and the bases $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_s)$ are special Pisot-Vijayaraghavan numbers.

Acknowledgements

I would like to thank my supervisor, Dr. Radhakrishnan Nair, for all his guidance, help and support throughout my PhD studies and during writing this thesis. I also wish to thank my office colleagues, especially Poj Lertchoosakul, for their support and for creating a nice and friendly environment in our postgraduate student office. I would like to give special thanks to Dr. Joel Haddley for his continued support and patience during work on this thesis.

Contents

Introduction

In this thesis we use tools in ergodic theory to study ergodic and metric properties of a continued fraction map in non-Archimedean settings and to investigate the distribution behaviour of multidimensional $\boldsymbol{\beta}$-adic Halton subsequences.

Motivation to study continued fractions in non-Archimedean settings, especially on p-adic numbers, is given by many well-known results explored for continued fractions on real numbers. Naturally one of the first things one might try to do is to explore the extent to which, theorems true for continued fractions on the real numbers, extend to the p-adic numbers. For the most part, the regular continued fraction expansion and its properties can be extended to the field of formal Laurent series over a finite field in a relatively trouble free manner. In the context of the p-adic numbers, the direct analogue of the regular continued fraction is the Ruban continued fraction [59]. However, there are problems with this algorithm. Although it is possible to define a sequence of rationals analogous to the convergents of the regular continued fractions, their convergence to the number they are supposed to represent is not assured. This deficiency can be avoided using a system of weights which leads to Schneider's continued fraction expansion [63]. However, this brings some other issues: we are dealing with a non-Archimedean field so one should expect to have a theory which is not as satisfactory as that one in the case of the regular continued fraction expansions. For instance, in the case of Schneider's continued fraction expansion, the sequence of convergents does not necessarily provide a sequence of best approximations to the p-adic number they approximate, unlike in the real case of the regular continued fraction expansion.

Despite the fact that Schneider's continued fraction map does not possess as many desirable properties as the regular continued fraction map, it is worthy of study as it can be a useful and powerful tool in many situations. For instance, it is sometimes very useful in delicate constructions on the p-adic numbers, as observed in [9] where it is used to construct numbers that distinguish between the Mahler and Koksma schemes of approximation to a specified degree.

Another interesting application of Schneider's continued fraction is to determine the algebraic independence of a set of p-adic numbers. See [11] and [38] for details.

It is well known that the restriction of the Gauss map to the rational numbers gives the Euclidean algorithm. If we set $p = 2$ and restrict the Schneider's continued fraction map to the rational numbers, we obtain the Binary Euclidean algorithm. This is another method of calculating the greatest common divisor of two integers, particularly well adapted to efficient implementation on binary machines.

The theory of Schneider's continued fraction has been of interest to many authors. Some p-adic analogues of standard properties of the regular continued fraction for the real numbers were partially recovered, for example, in [2], [3], [4], [10], [69] where authors investigate when a p-adic continued fraction is either finite or periodic. Recently, some ergodic and metric properties of this map were developed in [21] and [24].

In this thesis we will study Schneider's continued fraction map in greater generality and prove some new ergodic and metric properties of this general algorithm.

The other half of this thesis is devoted to studying uniform distribution of $\boldsymbol{\beta}$-adic Halton sequences. The importance of uniform distribution theory is as follows. A standard problem in numerical analysis is estimating the integral of a function using its value at a finite number of points $(x_n)_{n=1}^{N}$. In the case of stochastic sequences $(x_n)_{n=1}^{N}$, this is known as Monte Carlo integration. In the case of deterministic sequences $(x_n)_{n=1}^{N}$ we talk about Quasi-Monte Carlo integration. This is encapsulated in the famous Koksma-Hlawka inequality

[25] which bounds the integration error in Quasi-Monte Carlo integration

$$\left| \frac{1}{N} \sum_{n=1}^{N} f(\{\mathbf{x}_n\}) - \int_{[0,1]^s} f(\mathbf{x}) d\mathbf{x} \right| \leq V(f) D_N^\star.$$

Here $V(f)$ denotes the variation of a function f on $[0,1)^s$ (in the sense of Hardy and Krause) and $D_N^\star$ is the star-discrepancy of any finite set of points $\{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$ in $[0,1)^s$ defined by

$$D_N^\star = D_N^\star(\mathbf{x}_1, \ldots, \mathbf{x}_N) = \sup_{\mathbf{u} \in [0,1)^s} \left| \frac{1}{N} \#\{1 \leq n \leq N : \mathbf{x}_n \in [\mathbf{0}, \mathbf{u})\} - \lambda_s([\mathbf{0}, \mathbf{u})) \right|.$$

Here λ_s denotes s-dimensional Lebesgue measure. To estimate $\int_{[0,1]^s} f(\mathbf{x}) d\mathbf{x}$ sufficiently precisely, we need a good bound for $D_N^\star$ and a serviceable bound for $V(f)$, which is usually straight forward. To be useful this sequence $\{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$ must be uniformly distributed modulo one. The discrepancy can be viewed as a quantitative measure of uniformity of distribution. In particular, the sequence $(\mathbf{x}_n)_{n \geq 1}$ is uniformly distributed modulo one if and only if $D_N^\star \to 0$ as $N \to \infty$. In a sense the faster $D_N^\star$ decays as a function of N, the better uniformly distributed the sequence $(\mathbf{x}_n)_{n \geq 1}$ is. There are limitations to the uniformity of distribution of an arbitrary sequence which is one of the fundamental issues in this subject. Another is the complementary problem of constructing sequences with discrepancy as small as possible in order to minimise the integration error in Quasi-Monte Carlo integration.

The properties of one-dimensional β-adic Halton sequences have been investigated by a number of authors, for example [1], [48], [66]. The properties of the multidimensional version have also been explored in [26]. In this thesis we will be interested in a more general form of the multidimensional β-adic Halton sequence.

The thesis is organised as follows. The first chapter serves as a brief introduction to ergodic theory, non-Archimedean fields and other subjects which are used later in the thesis. In the second chapter, the ergodic and metric theory of the generalised Schneider's continued fraction map is studied. The third chapter is devoted to the distributional behaviour of multidimensional

$\boldsymbol{\beta}$-adic Halton sequences. The last chapter summarises the main results in this thesis and gives ideas for future research problems.

Chapter 1

Background

The purpose of this chapter is to give the reader a brief introduction to ergodic theory and non-Archimedean fields useful for other chapters of this thesis.

1.1 Non-Archimedean Fields

In this section we describe the general setting in which we work for most of this thesis. Definitions and results of this section can be also found in [12], [17], [47], [57] and [70].

Let K denote a locally compact topological field. By this we mean that the field K is a locally compact group under addition with respect to a topology. This ensures that there is a translation invariant Haar measure μ on K, that is unique up to positive scalar multiplication. For an element $a \in K$, we are now able to define its absolute value, as

$$|a| = \frac{\mu(aF)}{\mu(F)},$$

for every μ measurable $F \subseteq K$ with finite positive measure μ.

In this thesis, we will assume the topology to be non-discrete. Sometimes a field which is a locally compact topological field with respect to a non-discrete topology is referred to as a *local field*.

Throughout the thesis, we use $\mathbb{N} = \{1, 2, 3, \ldots\}$ to denote the natural numbers and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. Further, $\mathbb{Z}, \mathbb{Q}$ and $\mathbb{R}$ denote the integers, the rational numbers and the real numbers respectively. The non-negative real numbers are denoted by $\mathbb{R}_{\geq 0}$ and the positive real numbers by $\mathbb{R}_+$.

Let us recall the standard definition of an absolute value.

Definition 1.1.1. A function $|.| : K \to \mathbb{R}_{\geq 0}$ is called an *absolute value* on a field K if it has the following properties:

1. $|a| = 0$ if and only if $a = 0$,
2. $|ab| = |a||b|$ for all $a, b \in K$,
3. $|a + b| \leq |a| + |b|$ for all $a, b \in K$.

Definition 1.1.2. An absolute value on K is called *non-Archimedean* when it has the additional property that

$3^\star$. $|a + b| \leq \max\{|a|, |b|\}$ for all $a, b \in K$.

This property is also called the *ultrametric inequality* and fields with an absolute value satisfying the property $3^\star$ are called *non-Archimedean fields*. Other fields are called *Archimedean*. An absolute value gives rise to a metric defined by $d(a, b) = |a - b|$ with $a, b \in K$, whose topology coincides with the original topology on the field K.

Definition 1.1.3. Let K be a field with an absolute value $|.|$. Let $a \in K$ and $r \geq 0$ be a real number. The *open ball* of radius r centered at a is the set

$$B(a, r) = \{x \in K : |x - a| < r\}.$$

The *closed ball* of radius r and center a is the set

$$\overline{B}(a, r) = \{x \in K : |x - a| \leq r\}.$$

These structures are standard in any metric space. However, in a non-Archimedean space, these balls have some surprising properties as one can see in the following proposition.

Proposition 1.1.4. *Let K be a non-Archimedean field.*

- *If $b \in B(a,r)$, then $B(a,r) = B(b,r)$ which means that every point that lies in an open ball is a center of that ball.*
- *If $b \in \overline{B}(a,r)$, then $\overline{B}(a,r) = \overline{B}(b,r)$ which means that every point that lies in a closed ball is a center of that ball.*
- *The set $B(a,r)$ is both open and closed.*
- *If $a, b \in K$ and $r > 0, s > 0$, then $B(a,r) \cap B(b,s) \neq \emptyset$ if and only if $B(a,r) \subset B(b,s)$ or $B(b,s) \subset B(a,r)$ which means that any two open balls are either disjoint or one ball is contained in the other ball.*

Proof: The proof can be found in [17] on page 34. $\square$

In Chapter 2 of this thesis we will work with non-Archimedean local fields.

Another approach to defining a non-Archimedan local field is via discrete valuations.

Definition 1.1.5. Let K be a field and $K^* = K\backslash\{0\}$. A map $v : K^* \to \mathbb{R}$ is called a *valuation* if it satisfies

- $v(K^*) \neq \{0\}$,
- $v(xy) = v(x) + v(y)$ for all $x, y \in K$,
- $v(x + y) \geq \min\{v(x), v(y)\}$ for all $x, y \in K$.

The valuation v then determines a non-Archimedean absolute value as follows. Fix some real number $\alpha > 1$ and put

$$|x| = \begin{cases} \alpha^{-v(x)} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

Conversely, if we are given a non-Archimedean absolute value $|.|$ then for a real number $\alpha > 1$ we put $v(x) = \log_\alpha |x|$. So, $v(x)$ is a valuation – an additive version of $|x|$. One can extend v to K formally by letting $v(0) = \infty$. We exclude the trivial valuation given by $v(x) = 0$ for all $x \in K^*$. We say that two valuations v_1 and v_2 are *equivalent* if $v_1 = cv_2$ for some real constant $c > 0$.

Definition 1.1.6. A valuation v is *discrete* if $v(K^*) = s\mathbb{Z}$ for some positive real number s. If $v(K^*) = \mathbb{Z}$, we call v a *normalised* discrete valuation.

Let $v : K^* \to \mathbb{R}$ be a valuation corresponding to the non-Archimedean absolute value $|.| : K \to \mathbb{R}_{\geq 0}$. Then

$$\mathcal{O} = \mathcal{O}_v := \{x \in K : v(x) \geq 0\} = \mathcal{O}_K := \{x \in K : |x| \leq 1\}$$

is a ring, called the *valuation ring* of v or the *ring of integers* and K is its field of fractions. The set of units in $\mathcal{O}$ is

$$\mathcal{O}^\times = \{x \in K : v(x) = 0\} = \{x \in K : |x| = 1\}$$

and

$$\mathcal{M} = \{x \in K : v(x) > 0\} = \{x \in K : |x| < 1\}$$

is the unique maximal ideal in $\mathcal{O}$. Note that $\mathcal{O} = \mathcal{O}^\times \cup \mathcal{M}$. Since $\mathcal{M}$ is a maximal ideal, we have that $k = \mathcal{O}/\mathcal{M}$ is a field called the *residue field* of $\mathcal{O}$.

There is an alternative definition of a non-Archimedean local field which is as follows.

Definition 1.1.7. If K is a field which is complete with respect to a discrete valuation and its residue field is finite, then K is called a (non-Archimedean) *local field*.

Definition 1.1.8. Suppose the valuation $v : K^* \to \mathbb{Z}$ is normalised and discrete. An element $\pi \in \mathcal{M}$ such that $v(\pi) = 1$ is called a *uniformiser*.

Every $x \in K^*$ can be written uniquely as $x = u\pi^n$ with $u \in \mathcal{O}^\times$ and $n \in \mathbb{Z}$. In particular, every $x \in \mathcal{M}$ can be written uniquely as $x = u\pi^n$ for a unit $u \in \mathcal{O}^\times$ and $n \geq 1$.

Proposition 1.1.9. *Let K be a field and $v : K^* \to \mathbb{Z}$ be a normalised and discrete valuation. Let $A \subseteq \mathcal{O}$ be a system of representatives for $\mathcal{O}/\mathcal{M}$ such that $0 \in A$ and $\pi \in \mathcal{O}$ is a uniformiser. Then every $x \in K^*$ can be expressed uniquely in the form*

$$x = \sum_{n=k}^{\infty} a_n \pi^n$$

with $a_n \in A$ for $n = k, k+1, \ldots$ and $a_k \neq 0, n \in \mathbb{Z}$. The field K is complete if and only if every sum $\sum_{n=k}^{\infty} a_n \pi^n$ converges.

Proof: The proof can be found in [47] on page 126. □

Examples of local fields:

- Archimedean local fields: real numbers $\mathbb{R}$ and complex numbers $\mathbb{C}$.
- Non-Archimedean local fields:
  - with characteristic 0: the finite extensions of the field of p-adic numbers $\mathbb{Q}_p$ for some prime number p,
  - with positive characteristic p (p is a prime number): the field of formal Laurent series over a finite field.

The above examples are the only types of local fields. This is formally stated in the following theorem.

Theorem 1.1.10. *Given any local field k, we have that*

- *if k has characteristic 0, then k is $\mathbb{R}$, $\mathbb{C}$ or a finite extension of $\mathbb{Q}_p$;*
- *if k has characteristic $p > 0$, then k is non-Archimedean and isomorphic to the field of formal Laurent series in one variable over a finite field.*

Proof: The proof can be found in [56] on page 140. □

The rest of this section will be devoted to a more detailed description of the field of p-adic numbers and the field of formal Laurent series.

1.1.1 p-adic Numbers

Let p be a prime. Any non-zero rational number a can be written in the form $a = p^{\alpha}(r/s)$ where $\alpha \in \mathbb{Z}, r, s \in \mathbb{Z}$ and $p \nmid r, p \nmid s$. Sometimes, α is denoted by $v_p(a)$ and is called the *p-adic valuation*.

Definition 1.1.11. The *p-adic absolute value* of $a \in \mathbb{Q}$ is defined by

$$|a|_p = p^{-\alpha}$$

and $|0|_p = 0$.

It can be shown that $|.|_p$ is a non-Archimedean absolute value on $\mathbb{Q}$.

It is a fact that the real numbers $\mathbb{R}$ is the completion of the rationals $\mathbb{Q}$ with the respect to the standard absolute value. Similarly, the field $\mathbb{Q}_p$ of p-adic numbers is constructed by completing the rationals with the respect to the p-adic absolute value.

Definition 1.1.12. The metric on $\mathbb{Q}$ defined by $d_p(x, y) = |x - y|_p$ is called the *p-adic metric*.

This metric satisfies the ultrametric inequality, i. e. for any $x, y, z \in \mathbb{Q}_p$ we have

$$|x - z|_p \leq \max\{|x - y|_p, |y - z|_p\}.$$

The topology of $\mathbb{Q}_p$ is generated by the metric induced by the p-adic absolute value. So, as in Definition 1.1.3 we define open and closed balls for p-adic numbers.

Definition 1.1.13. Let $a \in \mathbb{Q}_p$ and $r \geq 0$ be a real number. The *open ball* of radius r centered at a is the set

$$B(a, r) = \{x \in \mathbb{Q}_p : |x - a|_p < r\}.$$

The *closed ball* of radius r and center a is the set

$$\overline{B}(a, r) = \{x \in \mathbb{Q}_p : |x - a|_p \leq r\}.$$

As the field of p-adic numbers is non-Archimedean, these balls have all properties described in Proposition 1.1.4.

The field $\mathbb{Q}_p$ is fully determined by the facts that:

- $\mathbb{Q}$ is dense in $\mathbb{Q}_p$,

- the absolute value $|.|_p$ extends the p-adic norm from $\mathbb{Q}$ to $\mathbb{Q}_p$,
- $\mathbb{Q}_p$ is complete with respect to $|.|_p$, i.e. every Cauchy sequence in $\mathbb{Q}_p$ (with the respect to $|.|_p$) has a limit in $\mathbb{Q}_p$.

There exists a special subset of $\mathbb{Q}_p$ and that is the ring of p-adic integers.

Definition 1.1.14. The *ring of p-adic integers* is

$$\mathbb{Z}_p = \{x \in \mathbb{Q}_p : |x|_p \leq 1\}.$$

Let us notice that $\overline{B}(0,1) = \mathbb{Z}_p$. The ideas introduced for non-Archimedean fields specialise in the case of the p-adic numbers to:

- the field: $K = \mathbb{Q}_p$,
- the valuation ring: $\mathcal{O} = \mathbb{Z}_p$,
- the maximal ideal in $\mathcal{O}$: $\mathcal{M} = p\mathbb{Z}_p$,
- the uniformiser: $\pi = p$,
- the residue field of K: $k = \mathbb{Z}_p/p\mathbb{Z}_p$.

It can be shown that $\mathbb{Z}_p/p\mathbb{Z}_p \cong \mathbb{Z}/p\mathbb{Z}$ which means that k is a finite field. Since $\mathbb{Q}_p$ is complete with respect to the p-adic absolute value $|.|_p$ and k is finite, we can conclude that the field of p-adic numbers $\mathbb{Q}_p$ is a local field. Hence, $\mathbb{Q}_p$ is locally compact and so it comes endowed with a translation invariant Haar measure. The ring of p-adic integers $\mathbb{Z}_p$ is then compact [47], p. 135.

Using Proposition 1.1.9 we get the following representation of p-adic numbers. Every $x \in \mathbb{Z}_p$ can be expressed in the form

$$x = \sum_{n=0}^{\infty} b_n p^n \tag{1.1}$$

with b_n in $\{0, 1, \ldots, p-1\}$ and this represetation is unique.

A representation for $\mathbb{Q}_p$ can be then constructed as follows. Any element of $\mathbb{Q}_p$ can be written in the form z/p^m where $z \in \mathbb{Z}_p$. Notice that z in the form

(1.1) divided by p^m gives us a power series in p where some of the powers can be negative. Hence, every $x \in \mathbb{Q}_p$ can be expressed in the form

$$x = \sum_{n=k}^{\infty} b_n p^n$$

where k is an integer, b_n is in $\{0, 1, \ldots, p-1\}$ for $n = k, k+1, \ldots$ and $b_k \neq 0$. This representation is unique.

1.1.2 Formal Laurent Series

Let q be a power of the prime p and let $\mathbb{F}_q$ be the finite field with q elements. Denote by $\mathbb{F}_q[X]$ the ring of polynomials with coefficients in $\mathbb{F}_q$ and by $\mathbb{F}_q(X)$ the quotient field of $\mathbb{F}_q[X]$. Let $\mathbb{F}_q((X^{-1}))$ denote the field of formal Laurent series

$$\mathbb{F}_q((X^{-1})) = \left\{ g = a_n X^n + \cdots + a_0 + a_{-1} X^{-1} + \cdots : n \in \mathbb{Z}, a_i \in \mathbb{F}_q \right\}.$$

The field $\mathbb{F}_q((X^{-1}))$ is the completion of $\mathbb{F}_q(X)$ with respect to the non-Archimedean absolute value $|.|$ which is determined by $|g| = q^{-v(g)}$ for $g \in \mathbb{F}_q((X^{-1}))$ and $|0| = 0$. Here the valuation $v(g)$ is defined as

$$v(g) = -\deg(g) = \inf\{n \in \mathbb{Z}, a_n \neq 0\}$$

with $\deg(g)$ denoting the degree of $g \in \mathbb{F}_q((X^{-1}))$. One can see that $|g| = q^{\deg(g)}$ for all $g \in \mathbb{F}_q((X^{-1}))$ and $|0| = 0$. Also $d_q(x, y) = |x - y|$ for $x, y \in \mathbb{F}_q((X^{-1}))$ defines a metric on $\mathbb{F}_q((X^{-1}))$. We have a subset of $\mathbb{F}_q((X^{-1}))$ which we denote by $\mathbb{L}$ and it is

$$\mathbb{L} = \{x \in \mathbb{F}_q((X^{-1})) : |x| \leq 1\} = \{a_0 + a_{-1} X^{-1} + a_{-2} X^{-2} + \cdots, a_i \in \mathbb{F}_q\}.$$

In the case of the field of formal Laurent series over the finite field we have:

- the field: $K = \mathbb{F}_q((X^{-1}))$,
- the valuation ring: $\mathcal{O} = \mathbb{L}$,

- the maximal ideal in $\mathcal{O}$: $\mathcal{M} = X^{-1}\mathbb{L} = \{x \in \mathbb{F}_q((X^{-1})) : |x| < 1\} = \{a_0X^{-1} + a_{-1}X^{-2} + a_{-2}X^{-3} + \cdots, a_i \in \mathbb{F}_q\}$,
- the uniformiser: $\pi = X^{-1}$,
- the residue field of K: $k = \mathbb{L}/X^{-1}\mathbb{L} = \mathbb{F}_q$.

Since $\mathbb{F}_q((X^{-1}))$ is complete with respect to the non-Archimedean absolute value $|.|$ and $k = \mathbb{F}_q$ is finite, we again conclude that the field of formal Laurent series $\mathbb{F}_q((X^{-1}))$ is a local field. Thus, $\mathbb{F}_q((X^{-1}))$ is locally compact and so it comes endowed with a translation invariant Haar measure. The ring $\mathbb{L}$ is then compact [47], p. 135.

For any $a \in \mathbb{F}_q((X^{-1}))$ and any real number $r \geq 0$ we define a (open) ball of radius r centered at a by

$$B(a, r) = \{g \in \mathbb{F}_q((X^{-1})) : |g - a| < r\}.$$

Let a radius of a ball be $r = q^{-n}$ and a center $a \in \mathbb{F}_q((X^{-1}))$. Then the Haar measure on $\mathbb{F}_q((X^{-1}))$ is fully characterised by its value on the balls $B(a, q^{-n})$, i.e. $\mu(B(a, q^{-n})) = q^{-n}$, see [65], p. 65–70.

1.2 Ergodic Theory – Basic Definitions and Results

The aim of this section is to introduce basic definitions and results in ergodic theory which will be used in other chapters of this thesis. Definitions and results in this section can be also found in [13], [52], [51], [54], [55] and [68].

Ergodic theory studies the long-term average behaviour of dynamical systems. We will consider a dynamical system consisting of a space X and a map or a transformation $T : X \to X$ (for the case of discrete time). Let T^0 be the identity map of X. Then for $n \geq 1$ we have the n-th iterate of T given by $T^n = T \circ T^{n-1} = T(T^{n-1})$. If T is invertible, then the notation extends to all $n \in \mathbb{Z}$ such that $T^{-n} = (T^{-1})^n$.

Next, we will consider a probability space $(X, \mathcal{B}, \mu)$ where X is a set, $\mathcal{B}$ is a σ-algebra of subsets of X and μ is a probability measure defined on $\mathcal{B}$.

Definition 1.2.1. A transformation $T : X \to X$ is *measurable* if $T^{-1}(A) \in \mathcal{B}$ for all $A \in \mathcal{B}$.

Definition 1.2.2. T is a *measure-preserving transformation* with respect to μ (or the measure μ is T*-invariant*) if T is measurable and $\mu(T^{-1}(A)) = \mu(A)$ for all $A \in \mathcal{B}$.

Example 1.2.3. The doubling map $T : \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$ which is defined by $T(x) = 2x \pmod 1$ is measure-preserving with respect to Lebesgue measure.

Definition 1.2.4. Let $(X, \mathcal{B}, \mu)$ be a probability space and $T : X \to X$ be a measure-preserving transformation. T is said to be an *ergodic transformation* with respect to μ (or μ is an *ergodic measure*) if $\mu(A) = 0$ or 1 for any T-invariant set $A \in \mathcal{B}$, i. e. $T^{-1}(A) = A$.

We can think of ergodic transformations as being indecomposable into subsystems.

Example 1.2.5. The doubling map $T : \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$ which is defined by $T(x) = 2x \pmod 1$ is ergodic with respect to Lebesgue measure.

The following theorem is a useful characterisation of ergodicity.

Theorem 1.2.6. *Let $(X, \mathcal{B}, \mu)$ be a probability space and $T : X \to X$ be a measure-preserving transformation. The following statements are equivalent:*

- *T is ergodic;*
- *whenever $f \in L^1(X, \mathcal{B}, \mu)$ satisfies $(f \circ T)(x) = f(x)$ μ-a.e. then f is constant μ-a.e.*

Proof: The proof can be found in [68] on page 28. $\square$

There are properties that imply ergodicity and also many results follow from ergodicity. In many cases, it is easier to prove another property satisfied by a transformation T which then implies ergodicity than to prove directly that T is ergodic. We will now present some of these properties.

Definition 1.2.7. Let T be a measure-preserving transformation of $(X, \mathcal{B}, \mu)$.

- T is defined to be *strong-mixing* if for all $A, B \in \mathcal{B}$ we have

$$\lim_{n\to\infty} \mu(T^{-n}A \cap B) = \mu(A)\mu(B).$$

- T is said to be *weak-mixing* if for all $A, B \in \mathcal{B}$ we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} |\mu(T^{-j}A \cap B) - \mu(A)\mu(B)| = 0.$$

Proposition 1.2.8. *If T is a strong-mixing transformation on $(X, \mathcal{B}, \mu)$, then T is weak-mixing.*

Proof: The proof follows from definitions and using the fact that for any sequence $\{a_n\}$ of real numbers if $\lim_{n\to\infty} a_n = 0$ then $\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n} |a_i| = 0$. $\square$

Proposition 1.2.9. *If T is a weak-mixing transformation on $(X, \mathcal{B}, \mu)$, then T is ergodic.*

Proof: If T is weak-mixing then for all $A, B \in \mathcal{B}$ we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} |\mu(T^{-j}A \cap B) - \mu(A)\mu(B)| = 0.$$

This implies that

$$\lim_{n\to\infty} \left(\frac{1}{n} \sum_{j=0}^{n-1} \mu(T^{-j}A \cap B) \right) - \mu(A)\mu(B) = 0.$$

Let $E \in \mathcal{B}$ be an invariant set, i.e. $T^{-1}E = E, E \in \mathcal{B}$. Taking $A = B = E$, we get

$$\mu(E) = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \mu(E) = \mu^2(E).$$

Thus, $\mu(E) = \mu^2(E)$ and so $\mu(E) = 0$ or 1 which proves that T is ergodic. $\square$

The converse of this proposition is not true. It means that there exist transformations which are ergodic but not weak-mixing as we can see in the following example.

Example 1.2.10. Let $X = \mathbb{R}/\mathbb{Z}$ and define the rotation map $T : X \to X$ by $T(x) = x + a (\bmod 1)$ where $a \in \mathbb{R}$ is irrational. It is known that T is ergodic with respect to Lebesgue measure μ. One can check that T is not weak-mixing.

Definition 1.2.11. Let T be a measure-preserving transformation of $(X, \mathcal{B}, \mu)$. Let $\mathcal{N}$ be the trivial σ-algebra, i.e. $\mathcal{N} = \{A \in \mathcal{B} \,|\, A = \emptyset \text{ a.e. or } A = X \text{ a.e.}\}$. The transformation T is called *exact* if

$$\bigcap_{n=0}^{\infty} T^{-n}\mathcal{B} = \mathcal{N}.$$

The following lemma is a useful tool for proving exactness.

Lemma 1.2.12. *Let $(X, \mathcal{B}, \mu)$ be a probability space and assume $\mathcal{A} \subset \mathcal{B}$ is an algebra that generates $\mathcal{B}$. If there exists $K > 0$ such that*

$$K\mu(B)\mu(I) \leq \mu(B \cap I)$$

for all $I \in \mathcal{A}$ then $\mu(B) = 0$ or $\mu(B) = 1$.

Proof: Assume $\varepsilon > 0$. Since $\mathcal{B}$ is generated by $\mathcal{A}$, there is a set $I \in \mathcal{A}$ such that $\mu(B^C \triangle I) < \varepsilon$ where B^C denotes the complement of the set B. So, $|\mu(B^C) - \mu(I)| < \varepsilon$. Note that $\mu(B \cap I) < \varepsilon$. This is because $B \cap I \subset B^C \triangle I$. Thus we get

$$\begin{aligned} \mu(B)\mu(B^C) &\leq \mu(B)(\mu(I) + \varepsilon) \leq \mu(B)\mu(I) + \mu(B)\varepsilon \leq \frac{1}{K}\mu(B \cap I) + \varepsilon \\ &\leq \left(\frac{1}{K} + 1\right)\varepsilon. \end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, we have that $\mu(B)\mu(B^C) = 0$ and thus, $\mu(B) = 0$ or $\mu(B) = 1$. $\square$

There are several relationships between exactness, ergodicity and mixing as the following propositions show.

Proposition 1.2.13. *If T is an exact transformation of the probability space $(X, \mathcal{B}, \mu)$ then T is ergodic.*

Proof: Assume that $T^{-1}B = B, B \in \mathcal{B}$. When we iterate this, we have that $T^{-n}B = B$ for all $n \geq 0$. So, $B \in T^{-n}\mathcal{B}$ for all $n \geq 0$, which implies that $B \in \bigcap_{n=0}^{\infty} T^{-n}\mathcal{B}$. Since T is exact, it means that $B \in \mathcal{N}$ and so $\mu(B) = 0$ or 1 which concludes the proof. $\square$

Proposition 1.2.14. *If T is an exact transformation of the probability space $(X, \mathcal{B}, \mu)$ then T is strong-mixing.*

Proof: The proof can be found in [55] on page 125. $\square$

To sum up, from Propositions 1.2.8 – 1.2.14 we can conclude that for non-invertible transformations there is the following hierarchy:

$$\text{Exact} \Rightarrow \text{strong-mixing} \Rightarrow \text{weak-mixing} \Rightarrow \text{ergodic.}$$

The definition of exactness holds for non-invertible transformations. If T is invertible, we have $T^{-1}\mathcal{B} = \mathcal{B}$ and so $\bigcap_{n=0}^{\infty} T^{-n}\mathcal{B} = \mathcal{B}$. Hence an invertible transformation can never be exact (except for the trivial case when $\mathcal{B} = \mathcal{N}$). However, there is the following definition for invertible transformations.

Definition 1.2.15. An invertible measure-preserving transformation T of a probability space $(X, \mathcal{B}, \mu)$ is said to be a K-*automorphism* if there exists a sub-σ-algebra $\mathcal{A} \subset \mathcal{B}$ satisfying:

- $\mathcal{A} \subset T\mathcal{A}$,
- $\bigvee_{n=0}^{\infty} T^n\mathcal{A} = \mathcal{B}$,
- $\bigcap_{n=0}^{\infty} T^{-n}\mathcal{A} = \mathcal{N}$.

Similarly, as for non-invertible transformations, there is also the following hierarchy for invertible transformations (proofs are omitted):

$$K\text{-automorphism} \Rightarrow \text{strong-mixing} \Rightarrow \text{weak-mixing} \Rightarrow \text{ergodic.}$$

A basic question in ergodic theory asks about the convergence of averages $\frac{1}{n}\sum_{i=0}^{n-1} f(T^i x)$. This means we are interested in when $\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1} f(T^i x)$ exists in some sense. Ergodic theorems resolve this problem and fundamental examples are due to von Neumann and Birkhoff. There are many variants of the ergodic theorem which depend on the the type of function (whether f is integrable, L^2, or continuous for instance) and the type of convergence (pointwise, L^2, uniform and so on). We will state the first ergodic theorem relevant for this thesis which is Birkhoff's Theorem for $f \in L^1$ and pointwise convergence.

Theorem 1.2.16 (Birkhoff's Ergodic Theorem)**.** *Let $(X, \mathcal{B}, \mu)$ be a probability space. Assume $T : X \to X$ is an ergodic measure-preserving transformation and $f \in L^1(X, \mathcal{B}, \mu)$. Then*

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

for almost all $x \in X$.

Proof: The proof can be found in [68] on page 38. □

1.3 Entropy

The entropy theory of dynamical systems is an important part of ergodic theory. Entropy plays a big role in the 'isomorphism problem' because it is invariant under an isomorphism which will be defined later. Entropy also has wider applications to other fields such as information theory and communication theory. We first define information as a function on a probability space with respect to a finite or countable partition. We can then understand the entropy as the expected value of the information.

There are a number of 'stages' in the definition of the entropy of a transformation. These are entropy of a partition, conditional entropy, the entropy of a transformation relative to a partition and finally, the most important – the entropy of a transformation irrespective of partition. All of them will be defined in this section. For definitions and results in this section and for more information see also [13], [52], [51], [54], [55] and [68].

We will use logarithms to base 2 and we will also set $0 \cdot \log 0 = 0$.

Definition 1.3.1. A *partition*, denoted by $\alpha = \{A_1, A_2, \ldots\}$, of a probability space $(X, \mathcal{B}, \mu)$ is a collection of elements of $\mathcal{B}$ such that $A_i \cap A_j = \emptyset$ for $i \neq j$ and $\bigcup_i A_i = X$.

We will consider finite or countable partitions.

Definition 1.3.2. Given two finite partitions $\alpha = \{A_1, A_2, \ldots, A_n\}$ and $\beta = \{B_1, B_2, \ldots, B_k\}$ of $(X, \mathcal{B}, \mu)$ we define their *join* to be the partition

$$\alpha \vee \beta = \{A_i \cap B_j : 1 \leq i \leq n, 1 \leq j \leq k\}.$$

Definition 1.3.3. Let α be a finite or countable partition of $(X, \mathcal{B}, \mu)$. The *information function* $I(\alpha) : X \to \mathbb{R}$ is defined by

$$I(\alpha)(x) = -\sum_{A \in \alpha} \chi_A(x) \log \mu(A).$$

Definition 1.3.4. Let α be a finite or countable partition of $(X, \mathcal{B}, \mu)$. The *entropy* of α is the number

$$H(\alpha) = \int I(\alpha) d\mu = -\sum_{A \in \alpha} \mu(A) \log \mu(A).$$

One can see that the entropy of a partition is the expected value of the information.

Besides the entropy of a partition there is also conditional entropy which is not necessary for defining the entropy of a transformation but is useful in deriving properties of entropy. Moreover, we will use it to derive an alternative definition of the entropy of a transformation relative to a partition. As before, we define conditional entropy using the conditional information function.

Definition 1.3.5. Given a finite or countable partition $\alpha = \{A_1, \ldots\}$ of $(X, \mathcal{B}, \mu)$ and a sub-σ-algebra $\mathcal{A} \subset \mathcal{B}$ we define the *conditional information* *function of α given $\mathcal{A}$* by

$$I(\alpha|\mathcal{A})(x) = -\sum_{A\in\alpha} \chi_A(x) \log \mu(A|\mathcal{A}).$$

Here $\mu(A|\mathcal{A})$ is the *conditional probability* of A given $\mathcal{A}$ and is defined as a function $\mu(A|\mathcal{A}) = E(\chi_A|\mathcal{A})$ where the function $E(\chi_A|\mathcal{A})$ is called the *conditional expectation.*

If $\mathcal{A} \subset \mathcal{B}$ is a sub-σ-algebra and $f \in L^1(X, \mathcal{B}, \mu)$ then the conditional expectation of f given $\mathcal{A}$, denoted by $E(f|\mathcal{A})$, is the unique $\mathcal{A}$-measurable function such that $\int_A f d\mu = \int_A E(f|\mathcal{A}) d\mu$ for all $A \in \mathcal{A}$.

We can think of the conditional information as follows. If we know which element of $\mathcal{A}$ a given point $x \in X$ lies in, then I$(\alpha|\mathcal{A})$ indicates the amount of additional information we get from knowing which element of the partition α the point x is in. The conditional entropy can be then defined as the expected value of the conditional information.

Definition 1.3.6. Let α be a finite or countable partition of $(X, \mathcal{B}, \mu)$ and let $\mathcal{A}$ be a sub-σ-algebra. The *conditional entropy of α given $\mathcal{A}$* is

$$H(\alpha|\mathcal{A}) = \int I(\alpha|\mathcal{A}) d\mu = -\sum_{A\in\alpha} \mu(A|\mathcal{A}) \log \mu(A|\mathcal{A}).$$

If β is a finite or countable partition of $(X, \mathcal{B}, \mu)$, then β generates a sub-σ-algebra, say $\hat{\beta}$, of $\mathcal{B}$ formed by the collection of all elements of $\mathcal{B}$ which are unions of elements of β. We will use the notational convention $E(f|\beta)$ for $E(f|\hat{\beta})$, $\mu(A|\beta)$ for $\mu(A|\hat{\beta})$ or $H(\alpha|\beta)$ for $H(\alpha|\hat{\beta})$ etc. For computations we will use the fact that the conditional probability of a set $A \in \mathcal{B}$ given β is

$$\mu(A|\beta) = \sum_{B\in\beta} \chi_B \frac{\mu(A \cap B)}{\mu(B)}.$$

There are a number of basic properties for the conditional entropy, e.g. if α, β, γ are countable partitions, then we have [68], p. 81:

- $H(\alpha \vee \beta|\gamma) = H(\alpha|\gamma) + H(\beta|\alpha \vee \gamma)$,

- $H(\alpha \vee \beta) = H(\alpha) + H(\beta|\alpha)$,
- $H(\alpha) \geq H(\alpha|\beta)$,
- $H(\alpha \vee \beta|\gamma) \leq H(\alpha|\gamma) + H(\beta|\gamma)$.

Definition 1.3.7. Let $T : X \to X$ be a measure-preserving transformation of the probability space $(X, \mathcal{B}, \mu)$ and α be a countable partition of X. The *entropy of T relative to α* is defined by

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

It can be shown that the above limit always exists. There is an alternative formula for $h_\mu(T, \alpha)$ which is frequently used in calculating entropy and that is

$$h_\mu(T, \alpha) = \lim_{n\to\infty} H\left(\alpha|\bigvee_{i=1}^{n} T^{-i}\alpha\right) = H\left(\alpha|\bigvee_{i=1}^{\infty} T^{-i}\alpha\right). \tag{1.2}$$

Finally, we can give the definition of the measure-theoretic entropy of T with respect to the measure μ (irrespective of α).

Definition 1.3.8. Let $T : X \to X$ be a measure-preserving transformation of the probability space $(X, \mathcal{B}, \mu)$. Then the *(measure-theoretic) entropy of T with respect to μ* is

$$h_\mu(T) = \sup h_\mu(T, \alpha)$$

where the supremum is taken over all finite or countable partitions α with $H(\alpha) < +\infty$.

To compute the entropy from the definition can be very difficult. However, there is a method for practical computation which is given by the following theorems and definitions. First, we state a theorem where the supremum in Definition 1.3.8 is replaced by a limit.

Theorem 1.3.9 (Abramov)**.** *Let $\alpha_1 \leq \alpha_2 \leq \ldots \uparrow \mathcal{B}$ be countable partitions such that $H(\alpha_n) < \infty$ for all $n \geq 1$. Then*

$$h_\mu(T) = \lim_{n\to\infty} h_\mu(T, \alpha_n).$$

Remark 1.3.10. Given two countable partitions we say that α_2 is *a refinement of* α_1 and we write $\alpha_1 \leq \alpha_2$ if every element of α_1 is a union of elements of α_2.

Proof: The proof can be found in [55] on page 87. $\square$

A way to generate the increasing partitions is stated by the following definition.

Definition 1.3.11. A countable partition α is a *strong generator* for the probability space $(X, \mathcal{B}, \mu)$ if

$$\bigvee_{i=0}^{n-1} T^{-i}\alpha \to \mathcal{B}$$

as $n \to \infty$. A countable partition α is a *generator* for the probability space $(X, \mathcal{B}, \mu)$ if T is invertible and

$$\bigvee_{i=-(n-1)}^{n-1} T^{-i}\alpha \to \mathcal{B}$$

as $n \to \infty$.

The following theorem is the main tool in computing the entropy of a transformation.

Theorem 1.3.12 (Sinai). *If α is a (strong) generator and $H(\alpha) < \infty$ then*

$$h_\mu(T) = h_\mu(T, \alpha).$$

Proof: The proof can be found in [68] on page 95. $\square$

Example 1.3.13. The entropy of the doubling map $T : \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$ defined by $T(x) = 2x \pmod 1$ is $\log 2$.

1.4 Isomorphism Problem and Bernoulli property

In mathematics, it is natural to ask when two mathematical objects of the same class are in some sense 'the same', which can be referred to as the isomorphism problem. In ergodic theory, this problem is to decide when two measure-preserving transformations are isomorphic. Since this is very hard, the usual way we deal with this is to look for invariants, i.e. quantities that do not change under isomorphism. One of these invariants is entropy which was defined earlier.

Definitions and results in this section can be also found in [13] and [68].

Definition 1.4.1. Let $(X_1, \mathcal{B}_1, \mu_1)$ and $(X_2, \mathcal{B}_2, \mu_2)$ be probability spaces and let $T_1 : X_1 \to X_1, T_2 : X_2 \to X_2$ be measure-preserving transformations. We say that T_1 and T_2 are *isomorphic* if there exist $M_1 \in \mathcal{B}_1$ and $M_2 \in \mathcal{B}_2$ such that

- $\mu_1(M_1) = 1, \mu_2(M_2) = 1,$
- $T_1(M_1) \subseteq M_1, T_2(M_2) \subseteq M_2,$

and there exists an invertible measure-preserving transformation $\phi : M_1 \to M_2$ such that

$$\phi T_1(x) = T_2\phi(x) \text{ for all } x \in M_1.$$

For a better visualisation of the definition see the commutative diagram below.

$$\begin{array}{ccc} X_1 & \xrightarrow{T_1} & X_1 \\ \phi\big\downarrow & & \phi\big\downarrow \\ X_2 & \xrightarrow{T_2} & X_2 \end{array}$$

As we mentioned before, measure-theoretic entropy is invariant under isomorphism, meaning that if two measure-preserving transformations are isomorphic then they have the same entropy. However, in general, entropy is not a complete invariant by which we mean that two measure-preserving transformations with equal entropy are not necessarily isomorphic. However, there exists a class of measure-preserving transformations for which entropy is a complete isomorphism invariant. We will now define such transformations.

Definition 1.4.2. Suppose $(Y, \mathcal{C}, m)$ is a probability space and let $Y_n = (Y, \mathcal{C}, m)$ for each $n \in \mathbb{Z}$. Suppose that $(X, \mathcal{B}, \mu) = \prod_{n \in \mathbb{Z}} Y_n$ and $\sigma : X \to X$ is the shift $\sigma(\{x_j\}) = \{x_{j+1}\}, j \in \mathbb{Z}$. The shift σ is then an invertible measure-preserving transformation and is said to be the *Bernoulli shift* with state space $(Y, \mathcal{C}, m)$.

In this definition, $\{x_j\}$ denotes a bi-infinite sequence of elements of the set Y and the measure μ is the product measure generated by the measure m, that is $\mu = \otimes_{-\infty}^{\infty} m$, and is called *Bernoulli measure.* In 1969 D. S. Ornstein proved that entropy is a complete isomorphism invariant on the collection of all Bernoulli shifts. This is formally stated in the following fundamental theorem.

Theorem 1.4.3 (Ornstein)**.** *Any two Bernoulli shifts with the same entropy are isomorphic.*

Proof: The proof can be found in [49]. □

Remark 1.4.4. So far we have discussed two-sided Bernoulli shifts, i.e. invertible shifts. In the case, when $(X, \mathcal{B}, \mu) = \prod_{n \in \mathbb{Z}_0^+} Y_n$, the shift $\sigma : X \to X, \sigma(\{x_j\}) = \{x_{j+1}\}, j \in \mathbb{Z}_0^+$ is non-invertible and we talk about one-sided Bernoulli shifts. Ornstein's Theorem 1.4.3 is not true for one-sided shifts so entropy is not a complete invariant for one-sided shifts.

Definition 1.4.5. A measure-preserving transformation is said to be *Bernoulli* if it is isomorphic to a Bernoulli shift.

For a non-invertible transformation, the shift is understood to be one-sided, and for an invertible transformation the shift is two-sided.

Example 1.4.6. The doubling map $T : \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$ which is defined by $T(x) = 2x \pmod 1$ is isomorphic to the Bernoulli $\left(\frac{1}{2}, \frac{1}{2}\right)$-shift.

The Bernoulli property is even stronger than exactness for non-invertible transformations and stronger than K-automorphism for invertible transformations. This is stated in the following theorem.

Theorem 1.4.7. *If an invertible measure-preserving transformation is Bernoulli then it is also a K-automorphism. Similarly, if a non-invertible measure-preserving transformation is Bernoulli then it is exact.*

Proof: The proof can be found in [68] on page 107. $\square$

From Theorem 1.4.7 and Propositions 1.2.8 – 1.2.14 one can see that the Bernoulli property implies a number of strictly weaker properties which can be summarised in the following hierarchy:

For non-invertible transformations:

$$\text{Bernoulli} \Rightarrow \text{exact} \Rightarrow \text{strong-mixing} \Rightarrow \text{weak-mixing} \Rightarrow \text{ergodic.}$$

For invertible transformations:

$$\text{Bernoulli} \Rightarrow K\text{-automorphism} \Rightarrow \text{strong-mixing} \Rightarrow \text{weak-mixing} \Rightarrow \text{ergodic.}$$

To any non-invertible measure-preserving transformation $(X, \mathcal{B}_0, \mu_0, T_0)$, we can associate an invertible transformation $(X_{T_0}, \mathcal{B}, \mu, T)$ which may be constructed as follows. Set

$$X_{T_0} = \{(x_0, x_1, x_2, \ldots) : x_n = T_0(x_{n+1}), x_n \in X, n = 0, 1, 2, \ldots\},$$

and let $T : X_{T_0} \to X_{T_0}$ be given by the formula

$$T((x_0, x_1, \ldots,)) = (T_0(x_0), x_0, x_1, \ldots).$$

The map T is one-to-one on X_{T_0}. To transform X_{T_0} into a measure space, we define a measure μ on X_{T_0}, by defining μ on the cylinder sets of the form

$$C(A_0, A_1, \ldots, A_k) = \{(x_0, x_1, x_2, \ldots) : x_0 \in A_0, x_1 \in A_1, \ldots, x_k \in A_k\}$$

by

$$\mu(C(A_0, A_1, \ldots, A_k)) = \mu_0(T_0^{-k}(A_0) \cap T_0^{-k+1}(A_1) \cap \ldots \cap A_k),$$

for $k \geq 1$. By the Kolmogorov Extension Theorem, the measure μ on the cylinder sets can be extended to the measure μ defined on the σ-algebra $\mathcal{B}$.

One can check that the transformation $(X_{T_0}, \mathcal{B}, \mu, T)$ is measure-preserving as a consequence of the measure-preservation of the transformation $(X, \mathcal{B}_0, \mu_0, T_0)$. Given this construction, we will introduce the following definition.

Definition 1.4.8. The invertible measure-preserving transformation T of the space $(X_{T_0}, \mathcal{B}, \mu)$ is said to be the *natural extension* of the measure-preserving transformation T_0 of the space $(X, \mathcal{B}_0, \mu_0)$.

The existence and uniqueness of the natural extension is assured by the following theorem which was proved by V. A. Rokhlin.

Theorem 1.4.9. *Every non-invertible measure-preserving transformation has a natural extension and this extension is unique up to isomorphism.*

Proof: The proof can be found in [58]. □

There are relationships between the ergodic properties of T_0 and T which are summarised in the following theorem.

Theorem 1.4.10. *Let T_0 be a non-invertible measure-preserving transformation of $(X, \mathcal{B}_0, \mu_0)$ and let T be its natural extension as defined above. Then*

- *T is ergodic if and only if T_0 is ergodic.*
- *T is strong-mixing, resp. weak-mixing if and only if T_0 is strong-mixing, resp. weak-mixing.*
- *The entropy of T_0 equals to the entropy of its natural extension T.*
- *If T_0 is exact, then its natural extension T is a K-automorphism. If T is a K-automorphism, then T is a natural extension of some exact transformation.*

Proof: The proof can be found in [58]. □

1.5 Uniform Distribution Modulo One

In this section we give a brief relevant introduction to uniform distribution theory. For definitions and results in this section see also [35].

Definition 1.5.1. A sequence of real numbers $(x_n)_{n=1}^{\infty}$ is called *uniformly distributed modulo one* if for each interval $I \subseteq [0, 1)$, we have

$$\lim_{N\to\infty} \frac{1}{N} \#\{n \leq N : \{x_n\} \in I\} = |I|$$

where $|I|$ denotes the length of the interval I and for a real number x we denote its fractional part by $\{x\}$.

The definition says that a sequence of reals $(x_n)_{n=1}^{\infty}$ is uniformly distributed modulo 1 if for each interval I, the frequency with which the fractional parts of x_n lie in the interval I is the same as the length of the interval I.

A necessary and sufficient condition for the uniform distribution modulo 1 for $(x_n)_{n=1}^{\infty}$ is given by the following result.

Theorem 1.5.2 (Weyl's Criterion)**.** *The following statements are equivalent:*

- *the sequence of real numbers $(x_n)_{n=1}^{\infty}$ is uniformly distributed modulo 1;*
- *for any real-valued continuous function on $[0, 1]$, we have*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(\{x_n\}) = \int_0^1 f(x)\, dx;$$

- *for all integers $h \neq 0$, we have*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} e^{2\pi i h x_n} = 0.$$

Proof: The proof can be found in [35] on pages 2 and 7. $\square$

Example 1.5.3. The sequence $(\alpha n)_{n=1}^{\infty}$ where $\alpha \in \mathbb{R}\backslash\mathbb{Q}$ is uniformly distributed modulo 1.

We now introduce the concept of uniform distribution mod 1 in higher dimensions. Let $s \in \mathbb{Z}$ such that $s \geq 2$ and let $\mathbf{a} = (a_1, \ldots, a_s)$ and $\mathbf{b} = (b_1, \ldots, b_s)$ denote vectors with real components, i.e. $\mathbf{a}, \mathbf{b} \in \mathbb{R}^s$. The inequalities $\mathbf{a} < \mathbf{b}$ or $\mathbf{a} \leq \mathbf{b}$ mean that $a_j < b_j$ or $a_j \leq b_j$ for all $j = 1, \ldots, s$. We denote the set of points $\mathbf{x} \in \mathbb{R}^s$ such that $\mathbf{a} \leq \mathbf{x} < \mathbf{b}$ by $[\mathbf{a}, \mathbf{b})$ and call it a s-dimensional interval. Similarly, we define other s-dimensional intervals, e.g. $[\mathbf{a}, \mathbf{b}]$. The interval $[0, 1)^s$ is called s-dimensional unit cube. For the integral part of $\mathbf{x} = (x_1, \ldots, x_s)$ we write $[\mathbf{x}] = ([x_1], \ldots, [x_s])$ and for the fractional part of $\mathbf{x}$ we write $\{\mathbf{x}\} = (\{x_1\}, \ldots, \{x_s\})$.

Definition 1.5.4. A sequence $(\mathbf{x}_n)_{n=1}^{\infty}$ of vectors in $\mathbb{R}^s$ is called *uniformly distributed on* $[0, 1)^s$ if for all s-dimensional intervals $[\mathbf{a}, \mathbf{b}) \subseteq [0, 1)^s$, we have

$$\lim_{n \to \infty} \frac{1}{N} \#\{n \leq N : \{\mathbf{x}_n\} \in [\mathbf{a}, \mathbf{b})\} = \prod_{j=1}^{\infty} (b_j - a_j).$$

The definition says that a sequence $(\mathbf{x}_n)_{n=1}^{\infty} \in \mathbb{R}^s$ is uniformly distributed on $[0, 1)^s$ if for any s-dimensional cube, the frequency with which the fractional parts of $\mathbf{x}_n$ lie in the cube is the same as the s-dimensional volume of the cube.

Theorem 1.5.5 (Multidimensional Weyl's Criterion)**.** *The following statements are equivalent:*

- *the sequence* $(\boldsymbol{x}_n)_{n=1}^{\infty} = (x_n^1, \ldots, x_n^s)_{n=1}^{\infty}$ *of vectors in* $\mathbb{R}^s$ *is uniformly distributed on* $[0, 1)^s$*;*
- *for every continuous function on* $[0, 1)^s$*, we have*

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} f(\{\boldsymbol{x}_n\}) = \int_{[0,1)^s} f(\boldsymbol{x})\, d\boldsymbol{x};$$

- *for all* $\boldsymbol{h} = (h_1, \ldots, h_s) \in \mathbb{Z}^s, \boldsymbol{h} \neq \boldsymbol{0}$*, we have*

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} e^{2\pi i (h_1 x_n^1 + \cdots + h_s x_n^s)} = 0.$$

Proof: The proof can be found in [35] on page 48. □

Example 1.5.6. The sequence $(\alpha_1 n, \ldots, \alpha_s n)_{n=1}^{\infty} \in \mathbb{R}^s$ is uniformly distributed on $[0,1)^s$ if the real numbers $\alpha_1, \ldots, \alpha_s$ and 1 are linearly independent over the rational numbers.

1.6 Subsequence Ergodic Theory

This section will provide the arithmetic and number theoretic context in which the results of Section 2.6 are proved. We are interested in two issues that are determining which sequences of integers satisfy a pointwise ergodic theorem and calculating the limit of the ergodic averages in the instances where these limits exist. First we introduce some definitions to describe the framework in which this is done and then we state some known results. Definitions and results in this section can be also found in [21] and [42].

Definition 1.6.1. A sequence of integers $(a_n)_{n=1}^{\infty}$ is called *L^p-good universal* if for each dynamical system $(X, \mathcal{B}, \mu, T)$ and $f \in L^p(X, \mathcal{B}, \mu)$ we have that

$$\overline{f}(x) = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} f(T^{a_n - 1} x)$$

exists almost everywhere with respect to μ.

Example 1.6.2. We give some examples of L^p-good universal sequences for some $p \geq 1$:

1. *The positive integers:* The sequence $(n)_{n=1}^{\infty}$ is L^1-good universal. This follows from Birkhoff's Pointwise Ergodic Theorem 1.2.16.

2. *Polynomial like sequences:* Let $\phi(x)$ be a polynomial such that $\phi(\mathbb{N}) \subseteq \mathbb{N}$ and let p_n be n^{th} prime. Then $(\phi(n))_{n=1}^{\infty}$ and $(\phi(p_n))_{n=1}^{\infty}$ are L^p-good universal sequences for $p > 1$. See [7] and [41].

The following theorem enables us to calculate the limit of the ergodic averages for an L^p-good universal sequence. This theorem is used to make the calculations in Section 2.6 given an L^p-good universal sequence.

Theorem 1.6.3. *Suppose that $(a_n)_{n\geq 1}$ is L^2-good universal and $(\{a_n\gamma\})_{n=1}^{\infty}$ is uniformly distributed modulo one for each irrational number γ. If the dynamical system $(X, \mathcal{B}, \mu, T)$ is weak-mixing, then $\overline{f}(x)$ exists and $\overline{f}(x) = \int_X f d\mu$ almost everywhere with respect to μ.*

Proof: The proof can be found in [42]. $\square$

Remark 1.6.4. Note that Theorem 1.6.3 extends readily to $p > 1$ by approximation by L^2-functions.

1.7 Moving Averages

Let $(X, \mathcal{B}, \mu)$ be a probability space and T be an ergodic measure-preserving transformation from X onto itself. Given $(n_l, k_l)_{l=1}^{\infty}$ to be a sequence of pairs of positive integers, we can define the sequence of averaging operators

$$A_{l,f}(x) = \frac{1}{k_l}\sum_{i=0}^{k_l-1} f(T^{n_l+i}x). \tag{1.3}$$

In [5] Bellow, Jones and Rosenblatt gave necessary and sufficient conditions for the almost everywhere convergence of the sequence of averages (1.3). This result will be used for applications in Section 2.7. Definitions and theorems in this section can be also found in [5] and [21]. To proceed, we introduce some notation. Let Z be a collection of points in $\mathbb{Z} \times \mathbb{N}$ and let

$$Z^h = \{(n,k) : (n,k) \in Z \text{ and } k \geq h\},$$

$$Z_\alpha^h = \{(z,s) \in \mathbb{Z}^2 : |z-y| < \alpha(s-r) \text{ for some } (y,r) \in Z^h\}$$

and

$$Z_\alpha^h(\lambda) = \{n : (n,\lambda) \in Z_\alpha^h\}$$

where $\alpha > 0$ and λ is a positive integer. Geometrically we can think of Z_α^1 as the lattice points contained in the union of all solid cones with aperture α and vertex contained in $Z^1 = Z$.

Definition 1.7.1. A sequence of pairs of positive integers $(n_l, k_l)_{l=1}^{\infty}$ is called *Stoltz* if there exists a collection of points Z in $\mathbb{Z} \times \mathbb{N}$, and a function $h = h(t)$ tending to infinity with t such that $(n_l, k_l)_{l=t}^{\infty} \in Z^{h(t)}$ and there exist h_0, α_0 and $A > 0$ such that for all integers $\lambda > 0$ we have $|Z_{\alpha_0}^{h_0}(\lambda)| \leq A\lambda$.

This technical condition is important for the following theorem which was given in [5].

Theorem 1.7.2. *Let $(X, \mathcal{B}, \mu)$ be a probability space and $T : X \to X$ be a measure-preserving map. Suppose that the sequence of pairs of positive integers $(n_l, k_l)_{l=1}^{\infty}$ is Stoltz and $f \in L^1(X, \mathcal{B}, \mu)$. Then the limit*

$$A_f(x) = \lim_{l \to \infty} \frac{1}{k_l} \sum_{i=0}^{k_l - 1} f(T^{n_l + i} x)$$

exists almost everywhere with respect to the measure μ.

Proof: The proof can be found in [5]. □

The following theorem will be used for applications in Section 2.7.

Theorem 1.7.3. *Let $(X, \mathcal{B}, \mu)$ be a probability space and $T : X \to X$ be an ergodic map. Suppose that the sequence of pairs of positive integers $(n_l, k_l)_{l=1}^{\infty}$ is Stoltz and $f \in L^1(X, \mathcal{B}, \mu)$. Then*

$$\lim_{l \to \infty} \frac{1}{k_l} \sum_{i=0}^{k_l - 1} f(T^{n_l + i} x) = \int_X f d\mu$$

almost everywhere with respect to the measure μ.

Proof: If we set

$$A_{l,f}(x) = \frac{1}{k_l} \sum_{i=0}^{k_l - 1} f(T^{n_l + i} x)$$

then

$$A_{l,f}(Tx) - A_{l,f}(x) = k_l^{-1}(f(T^{n_l + k_l}) - f(T^{n_l} x)).$$

So, we get that $A_f(Tx) = A_f(x)$ μ almost everywhere. From Theorem 1.2.6 we have that if T is ergodic and $A_f(Tx) = A_f(x)$ almost everywhere, then

$A_f(x)$ is constant almost everywhere. Because $\int_X f(x)d\mu = \int_X A_f(x)d\mu$, we have $A_f(x) = \int_X f d\mu$ almost everywhere. $\square$

Averages where $k_l = 1$ for all l are called *non-moving*.

Example 1.7.4. If we take $n_l = 2^{2^l}$ and $k_l = 2^{2^{l-1}}$, then $(n_l, k_l)_{l=1}^{\infty}$ is Stoltz.

1.8 Continuous Transformations on Compact Metric Spaces

In this section we will assume that X is a compact metric space. For definitions and results in this section see also [13], [54] and [68].

We will denote the σ-algebra of Borel subsets of X by $\mathcal{B}$ and the set of all Borel probability measures on $(X, \mathcal{B})$ by $M(X)$. We can see that $M(X)$ is convex, i.e. for $\mu_1, \mu_2 \in M(X)$ and $0 \le \alpha \le 1$, we have $\alpha\mu_1 + (1-\alpha)\mu_2 \in M(X)$.

There is a member δ_x of $M(x)$ defined for each $x \in X$ by

$$\delta_x(A) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases}$$

We call δ_x the *Dirac measure* at x. The map $x \to M(x) : x \to \delta_x$ is then a continuous embedding of X in $M(X)$. We also have that $\int f \, d\delta_x = f(x)$.

Theorem 1.8.1 (Riesz Representation Theorem)**.** *Let X be a compact metric space and let $J : C(X) \to \mathbb{C}$ be a functional such that:*

1. *J is bounded, i.e. for all $f \in C(X)$ we have $|J(f)| \le \|f\|_\infty$;*
2. *J is linear, i.e. $J(\lambda_1 f_1 + \lambda_2 f_2) = \lambda_1 J(f_1) + \lambda_2 J(f_2)$ where λ_1, λ_2 are complex numbers;*
3. *J is positive, i.e. if $f \ge 0$ then $J(f) \ge 0$;*
4. *J is normalised, i.e. $J(1) = 1$.*

Then there exists a Borel probability measure $\mu \in M(X)$ such that for all $f \in C(X)$ we have

$$J(f) = \int_X f \, d\mu.$$

Proof: The proof can be found in [53] on page 145. $\square$

Note that a linear functional is continuous, i.e. $f_n \in C(X)$ and $f_n \to f$ then $J(f_n) \to J(f)$, if and only if it is bounded. So, if a functional is linear, the condition 1 is equivalent to saying that J is continuous.

The Riesz Representation Theorem describes the relationship between elements of $M(X)$ and continuous linear positive normalised functionals. This is a useful method for constructing measures.

Now suppose that $T : X \to X$ is a continuous transformation of the compact metric space X. One can see that $T^{-1}\mathcal{B} \subset \mathcal{B}$ so T is measurable. There is a map on the set $M(X)$ induced by T which is $T_* : M(X) \to M(X)$ defined by

$$(T_*\mu)(B) = \mu(T^{-1}B).$$

The following lemma shows how to perform integration with respect to $T_*\mu$.

Lemma 1.8.2. *For every function $f \in C(X)$ we have*

$$\int f \, d(T_*\mu) = \int f \circ T \, d\mu.$$

Proof: The proof can be found in [68] on page 150. $\square$

There is a topology on $M(X)$ which is given by the following definition.

Definition 1.8.3. The smallest topology on $M(X)$ such that for each of the maps $\mu \to \int_X f \, d\mu$ ($f \in C(X)$) is continuous, is called the *weak* topology* on $M(X)$.

In the weak* topology, a sequence of probability measures μ_n converge to μ in $M(X)$ if for every $f \in C(X)$ we have

$$\int f \, d\mu_n \to \int f \, d\mu$$

as $n \to \infty$. This is called weak* convergence. The space $M(X)$ is metrisable in the weak* topology and it is also compact in this topology [68], p. 148–150.

Our interest will now be focused on probability measures in $M(X)$ which are invariant with respect to T. We know that the measure μ is T-invariant if and only if $T_*\mu = \mu$. We will denote the space of all Borel probability invariant measures by

$$M(X,T) = \{\mu \in M(X) | T_*\mu = \mu\}.$$

There is a useful lemma which tells us how one can check whether a measure is T-invariant.

Lemma 1.8.4. *Suppose that $T : X \to X$ is a continuous transformation of a compact metric space and $\mu \in M(X)$. Then $\mu \in M(X,T)$ if and only if for all $f \in C(X)$ we have*

$$\int f \circ T \, d\mu = \int f \, d\mu.$$

Proof: The proof can be found in [68] on page 151. □

The properties of the set $M(X,T)$ where $T : X \to X$ is a continuous transformation of a compact metric space X, are [68], p. 152:

- $M(X,T)$ is non-empty, i.e. there exists at least one T-invariant probability measure.
- $M(X,T)$ is a compact subset of $M(X)$.
- $M(X,T)$ is convex.
- The probability measure μ in $M(X,T)$ is ergodic if and only if μ is an extreme point of $M(X,T)$, i.e. whenever $\mu = \alpha\mu_1 + (1-\alpha)\mu_2$, where $\mu_1, \mu_2 \in M(X,T), \alpha \in (0,1)$ then $\mu_1 = \mu_2 = \mu$.
- There exists at least one ergodic measure in $M(X,T)$.

Now, we will look at transformations where $M(X,T)$ has only one member.

Definition 1.8.5. Let T be a continuous transformation on a compact metrisable space X. We say that T is *uniquely ergodic* if there is only one T-invariant Borel probability measure on X.

Example 1.8.6. Let $T : \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$ be the irrational rotation map defined by $T(x) = x + \alpha \pmod 1$ with α irrational. Then T is uniquely ergodic where the unique invariant probability measure is Lebesgue measure.

For uniquely ergodic dynamical systems there is a stronger variant of Birkhoff's Ergodic Theorem.

Theorem 1.8.7. *Suppose that T is a continuous transformation on a compact metrisable space X and μ is a T-invariant Borel probability measure on X. The following statements are equivalent:*

1. *T is uniquely ergodic.*
2. *For any function $f \in C(X)$ (the space of continuous functions on X) the time means $\frac{1}{n}\sum_{i=0}^{n-1} f(T^i x)$ converge uniformly to a constant.*
3. *For any $f \in C(X)$ $\frac{1}{n}\sum_{i=0}^{n-1} f(T^i x)$ converges pointwise to a constant.*
4. *For any $f \in C(X)$ and all $x \in X$, we have*

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

Proof: The proof can be found in [68] on page 160. □

Chapter 2

Schneider's Continued Fraction Map on a Non-Archimedean Local Field

In this chapter we study ergodic properties of a generalisation of Schneider's p-adic continued fraction map. Results in this chapter can be also found in [21] and [29].

2.1 Motivation

To give a motivation for studying Schneider's continued fraction map, we begin with a brief summary of terms and results for regular continued fractions relevant for further discussion. Any real number $x \in [0, 1]$ can be written in the form

$$x = \cfrac{1}{c_1 + \cfrac{1}{c_2 + \cfrac{1}{c_3 + \ddots}}} \tag{2.1}$$

where $c_1, c_2, \ldots$ are natural numbers. The expression (2.1) is called the *regular continued fraction expansion* which is also written more compactly as

$[c_1, c_2, \ldots]$. The terms $c_1, c_2, \ldots$ are called the *partial quotients* of the continued fraction expansion and the sequence of rational truncates

$$[c_1, c_2 \ldots, c_n] = \frac{p_n}{q_n}, \qquad (n = 1, 2, \ldots)$$

are called the *convergents* of the continued fraction expansion. It is well known that if x is rational, the continued fraction expansion (2.1) is finite and if x is irrational, the expansion (2.1) is infinite.

There is a transformation related to continued fractions called the *Gauss map*, defined on $[0, 1]$ by

$$T(x) = \begin{cases} \{\frac{1}{x}\} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

Notice that $c_n(x) = c_{n-1}(Tx)$ $(n = 1, 2, \ldots)$. We now consider the particular ergodic properties of the dynamical system $(X, \mathcal{L}, \gamma, T)$ where X denotes $[0, 1]$, $\mathcal{L}$ is the σ-algebra on X, γ is the Gauss measure on $(X, \mathcal{L})$ defined for any A in $\mathcal{L}$ by

$$\gamma(A) = \frac{1}{\log 2} \int_A \frac{dx}{x+1}$$

and T is the Gauss map. It is known that the Gauss map T is measure-preserving with respect to Gauss measure γ. It is also ergodic and exact, with a Bernoulli natural extension and the measure-theoretic entropy equal to $\frac{\pi^2}{6 \log 2}$. See also [13], p. 165–177, or Chapter 4 of [27]. This point of view can be used to prove results like the following.

Suppose $F : \mathbb{R}_{\geq 0} \to \mathbb{R}$ is continuous, increasing and such that

$$\int_0^1 |F(c_1(x))| dx < \infty.$$

For each $n \in \mathbb{N}$ and arbitrary real numbers $d_1, \ldots, d_n$ we set

$$M_{F,n}(d_1, \ldots, d_n) = F^{-1} \left[\frac{F(d_1) + \ldots + F(d_n)}{n} \right].$$

Then we have

$$\lim_{n\to\infty} M_{F,n}(c_1(x), \ldots, c_n(x)) = F^{-1}\left[\int_0^1 F(c_1(x))dx\right],$$

almost everywhere with respect to Lebesgue measure on $[0,1]$, see [61]. For instance, restricting to the case $F(x) = \log x$, we recover A. Khinchin's famous result that

$$\lim_{n\to\infty} (c_1(x) \cdots c_n(x))^{\frac{1}{n}} = \prod_{k=1}^{\infty} \left(1 + \frac{1}{k^2 + 2k}\right)^{\frac{\log k}{\log 2}},$$

almost everywhere with respect to Lebesgue measure [34]. Results for means other than the geometric mean can be obtained by different choices of F. See also [27], p. 230–232 for more details.

Another well known result about the statistics of partial quotients in continued fractions is the following. Let $P_n(x,q)$ $(n = 1, 2, \ldots)$ denote the number of $c_1(x), \ldots, c_n(x)$ such that $c_i(x) = q$ where q is a positive integer. Then

$$\lim_{n\to\infty} \frac{P_n(x,q)}{n} = \frac{1}{\log 2} \log \frac{(q+1)^2}{q(q+2)},$$

almost everywhere with respect to Lebesgue measure. This result is known as the *pointwise Gauss-Kuzmin theorem* and it was first suggested by K. F. Gauss in the letters to P. S. Laplace [16]. R. O. Kuzmin gave the first proof of this problem with additional estimates in [36] and [37] and shortly after that P. Lévy developed another proof with new estimates [40]. For other results in this area see also [71]. A nice summary of the theory of the Gauss-Kuzmin distribution can be found in [33], Chapter 9.

Extensions of the above results to subsequence and moving averages appear in [42] and [32] respectively.

The purpose of this chapter is to extend this study of the ergodic and metric theory of continued fractions to the non-Archimedean settings.

2.2 Introduction of Generalised Schneider's Continued Fraction Map

Let K be a non-Archimedean local field with a normalised discrete valuation. Then $\mathcal{O} = \{x \in K : |x| \leq 1\}$ is the valuation ring, $\mathcal{M} = \{x \in K : |x| < 1\}$ is the maximal ideal in $\mathcal{O}$, $k = \mathcal{O}/\mathcal{M}$ is the residue field of K and π is the uniformiser. Our primary object of study in this chapter is the map $T_v : \mathcal{M} \to \mathcal{M}$ defined by

$$T_v(x) = \frac{\pi^{v(x)}}{x} - b(x) \tag{2.2}$$

where $v(x)$ is the valuation of x and $b(x)$ denotes the residue class to which $\frac{\pi^{v(x)}}{x}$ belongs in $k\backslash\{0\}$.

This gives rise to the continued fraction expansion of $x \in \mathcal{M}$ in the form

$$x = \cfrac{\pi^{a_1}}{b_1 + \cfrac{\pi^{a_2}}{b_2 + \cfrac{\pi^{a_3}}{b_3 + \ddots}}} \tag{2.3}$$

where $b_n \in k^* = k\backslash\{0\}, a_n \in \mathbb{N}$ for $n = 1, 2, \ldots$.

The rational approximants to $x \in \mathcal{M}$ arise in a manner similar to that in the case of the real numbers as follows. We suppose that $A_0 = 0, B_0 = 1, A_1 = \pi^{a_1}, B_1 = b_1$. Then set

$$A_n = \pi^{a_n} A_{n-2} + b_n A_{n-1} \quad \text{and} \quad B_n = \pi^{a_n} B_{n-2} + b_n B_{n-1} \tag{2.4}$$

for $n \geq 2$. The sequence of rational numbers $(\frac{A_n}{B_n})_{n=1}^{\infty}$ are the convergents to $x \in \mathcal{M}$.

Let us fix $n \in \mathbb{N}$. Then the n-th step of the expansion of $x \in \mathcal{M}$ is

$$\frac{A_n}{B_n} = \cfrac{\pi^{a_1}}{b_1 + \cfrac{\pi^{a_2}}{b_2 + \cfrac{\pi^{a_3}}{b_3 + \ddots + \cfrac{\pi^{a_n}}{b_n}}}}. \tag{2.5}$$

Now let

$$g(z) = \cfrac{\pi^{a_1}}{b_1 + \cfrac{\pi^{a_2}}{b_2 + \cfrac{\pi^{a_3}}{b_3 + \ddots + \cfrac{\pi^{a_n}}{b_n + z}}}}. \tag{2.6}$$

for some $z \in \mathcal{M}$.

Lemma 2.2.1. *We have*

$$A_{n-1}B_n - A_nB_{n-1} = (-1)^n \pi^{a_1+\cdots+a_n} \tag{2.7}$$

for $n = 1, 2, \ldots$.

Proof: We prove it by induction. First let us check that it is valid for $n = 1$. The left hand side of (2.7) is

$$A_0B_1 - A_1B_0 = -\pi^{a_1} \cdot 1 = -\pi^{a_1} = (-1)^1\pi^{a_1}$$

which is equal to the right hand side $(-1)^1\pi^{a_1}$. Now we suppose that (2.7) is valid for n and we want to prove it for $n+1$, i.e. we want to show that $A_nB_{n+1} - A_{n+1}B_n = (-1)^{n+1}\pi^{a_1+\cdots+a_{n+1}}$. Using (2.4) we have

$$\begin{aligned}
A_nB_{n+1} - A_{n+1}B_n &= A_n(\pi^{a_{n+1}}B_{n-1} + b_{n+1}B_n) - B_n(\pi^{a_{n+1}}A_{n-1} + b_{n+1}A_n) \\
&= \pi^{a_{n+1}}(A_nB_{n-1} - A_{n-1}B_n) = -\pi^{a_{n+1}}(A_{n-1}B_n - A_nB_{n-1}) \\
&= -\pi^{a_{n+1}}((-1)^n\pi^{a_1+\cdots+a_n}) = (-1)^{n+1}\pi^{a_1+\cdots+a_{n+1}}
\end{aligned}$$

which we wanted to show and thus (2.7) holds for all $n = 1, 2, \ldots$. □

Lemma 2.2.2. *For $z \in \mathcal{M}$ we have that*

$$g(z) = \frac{zA_{n-1} + A_n}{zB_{n-1} + B_n}.$$

Proof: Using the formulas (2.4) we get

$$\begin{aligned} g(z) &= \frac{\pi^{a_n} A_{n-2} + (b_n + z)A_{n-1}}{\pi^{a_n} B_{n-2} + (b_n + z)B_{n-1}} = \frac{\pi^{a_n} A_{n-2} + b_n A_{n-1} + zA_{n-1}}{\pi^{a_n} B_{n-2} + b_n B_{n-1} + zB_{n-1}} \\ &= \frac{A_n + zA_{n-1}}{B_n + zB_{n-1}}. \end{aligned}$$

□

2.3 Properties of Generalised Schneider's Continued Fraction Map

Throughout this chapter we will consider the dynamical system $(\mathcal{M}, \mathcal{B}, \mu, T_v)$ where $\mathcal{B}$ is the Haar σ-algebra restricted to $\mathcal{M}$, μ denotes Haar measure on $\mathcal{M}$ and T_v is defined by (2.2). For the Haar measure one checks that $\mu(\pi a + \pi^n \mathcal{O}) = (\#(k))^{1-n}$ where $\#(k)$ denotes the cardinality of the finite set k. So, one sees that $\mu(\pi\mathcal{O}) = \mu(\mathcal{M}) = 1$.

Lemma 2.3.1. *Let m, n be positive integers, $c_0 \in k\backslash\{0\}$ and $a \in \mathcal{O}$. We then have*

$$\frac{\pi^m}{c_0 + \pi a + \pi^n \mathcal{O}} = \frac{\pi^m}{c_0 + \pi a} + \pi^{n+m}\mathcal{O}.$$

Proof: Let $b \in \mathcal{O}$. Then $-\pi^n b \in \pi^n \mathcal{O}$. Since $c_0 \in k\backslash\{0\}$, we know c_0 cannot be equal to π, then the valuations $v(c_0 + \pi a) = 0$ and $v(c_0 + \pi a + \pi^n b) = 0$ and also $v(c_0 + \pi a) + v(c_0 + \pi a + \pi^n b) = v((c_0 + \pi a)(c_0 + \pi a + \pi^n b)) = 0$. Thus $(c_0 + \pi a)(c_0 + \pi a + \pi^n b) \in \mathcal{O}^\times$ and so

$$\frac{-\pi^n b}{(c_0 + \pi a)(c_0 + \pi a + \pi^n b)} \in \pi^n \mathcal{O}$$

which implies that

$$\frac{1}{c_0 + \pi a + \pi^n b} - \frac{1}{c_0 + \pi a} \in \pi^n \mathcal{O}.$$

From this we obtain that

$$\frac{\pi^m}{c_0 + \pi a + \pi^n b} \in \frac{\pi^m}{c_0 + \pi a} + \pi^{n+m}\mathcal{O}$$

and finally,

$$\frac{\pi^m}{c_0 + \pi a + \pi^n \mathcal{O}} \subseteq \frac{\pi^m}{c_0 + \pi a} + \pi^{n+m}\mathcal{O}.$$

Now, again let $b \in \mathcal{O}$, then $-\pi^n b \in \pi^n \mathcal{O}$. Since $v(c_0 + \pi a) = 0$, which implies that $c_0 + \pi a \in \mathcal{O}^\times$, we have $-\pi^n b(c_0 + \pi a) \in \pi^n \mathcal{O}$. Since $v(1 + \pi^n b(c_0 + \pi a)) = 0$, we get

$$\frac{-\pi^n b(c_0 + \pi a)}{1 + \pi^n b(c_0 + \pi a)} \in \pi^n \mathcal{O}.$$

This implies that

$$\frac{c_0 + \pi a}{1 + \pi^n b(c_0 + \pi a)} - c_0 - \pi a \in \pi^n \mathcal{O}$$

and

$$\frac{c_0 + \pi a}{1 + \pi^n b(c_0 + \pi a)} \in c_0 + \pi a + \pi^n \mathcal{O}.$$

Now, we have

$$\frac{1 + \pi^n b(c_0 + \pi a)}{c_0 + \pi a} \in \frac{1}{c_0 + \pi a + \pi^n \mathcal{O}}$$

leading to

$$\frac{1}{c_0 + \pi a} + \pi^n b \in \frac{1}{c_0 + \pi a + \pi^n \mathcal{O}}$$

and finally

$$\frac{\pi^m}{c_0 + \pi a} + \pi^{n+m}\mathcal{O} \subseteq \frac{\pi^m}{c_0 + \pi a + \pi^n \mathcal{O}}.$$

Thus, we proved that

$$\frac{\pi^m}{c_0 + \pi a + \pi^n \mathcal{O}} = \frac{\pi^m}{c_0 + \pi a} + \pi^{n+m}\mathcal{O}.$$

□

Proposition 2.3.2. *The map $T_v : \mathcal{M} \to \mathcal{M}$ defined by (2.2) is measure-preserving with respect to Haar measure μ, i.e. $\mu(T_v^{-1}(A)) = \mu(A)$ for all Haar measurable sets $A \in \mathcal{B}$.*

Proof: To prove that T_v preserves Haar measure on $\mathcal{M}$ we only need to check it for special sets of the form $A = \pi a + \pi^n \mathcal{O}$, where $a \in \mathcal{O}$. This is because sets of this form generate the Haar σ-algebra on $\mathcal{M}$. Suppose $c_0 \in k \backslash \{0\}$ and let m, n be positive integers. Then

$$T_v \left(\frac{\pi^m}{c_0 + \pi a + \pi^n \mathcal{O}} \right) = \pi a + \pi^n \mathcal{O}.$$

It follows that

$$T_v^{-1}(\pi a + \pi^n \mathcal{O}) = \bigcup_{c_0 \in k \backslash \{0\}} \bigcup_{m=1}^{\infty} \left(\frac{\pi^m}{c_0 + \pi a + \pi^n \mathcal{O}} \right).$$

By Lemma 2.3.1 we have that

$$\frac{\pi^m}{c_0 + \pi a + \pi^n \mathcal{O}} = \frac{\pi^m}{c_0 + \pi a} + \pi^{n+m} \mathcal{O}$$

and so

$$\mu \left(\frac{\pi^m}{c_0 + \pi a} + \pi^{n+m} \mathcal{O} \right) = (\#(k))^{1-m-n}.$$

It follows that

$$\begin{aligned} \mu(T_v^{-1}(\pi a + \pi^n \mathcal{O})) &= \sum_{c_0 \in k \backslash \{0\}} \sum_{m=1}^{\infty} \mu \left(\frac{\pi^m}{c_0 + \pi a} + \pi^{n+m} \mathcal{O} \right) \\ &= \sum_{c_0 \in k \backslash \{0\}} \sum_{m=1}^{\infty} (\#(k))^{1-n-m} \\ &= (\#(k) - 1) \frac{\frac{1}{\#(k)}}{1 - \frac{1}{\#(k)}} \cdot (\#(k))^{1-n} \\ &= (\#(k))^{1-n} = \mu(\pi a + \pi^n \mathcal{O}), \end{aligned}$$

as required. So, T_v is measure-preserving with respect to Haar measure μ. $\square$

Next, we will show that the map T_v is exact, but first we need to introduce some notation.

Let $B = k^* \times \mathbb{N}$ and let $j = (j_1, j_2, \ldots)$ be a countable sequence of elements of B. For a particular element $j^* = (b, a) \in B$ define the cylinder-set $\Delta(j^*)$ by

$$\Delta(j^*) = \left\{ x \in \mathcal{M} : v(x) = a \text{ and } \left(\frac{\pi^{v(x)}}{x} \mod \pi \right) = b \right\}.$$

Now let $\Delta^{(0)} = \mathcal{M}$ and let $\Delta_j^{(1)} = \Delta(j_1)$, where j_1 is the first element of the sequence j. Next define

$$\Delta_j^{(2)} = \Delta^{(2)}(j_1, j_2) = \{x \in \mathcal{M} : x \in \Delta(j_1) \text{ and } T_v(x) \in \Delta(j_2)\}.$$

Proceeding inductively we get

$$\begin{aligned} \Delta_j^{(n)} &= \Delta^{(n)}(j_1, \ldots, j_n) \\ &= \{x \in \mathcal{M} : x \in \Delta(j_1), T_v(x) \in \Delta(j_2), \ldots, T_v^{n-1}(x) \in \Delta(j_n)\}. \end{aligned}$$

So, $\Delta_j^{(n)}$ is the set of all $x \in \mathcal{M}$ with continued fraction expansion starting with $j_1, j_2, \ldots, j_n$. This means that $\Delta_j^{(n)}$ depends only on the first n terms of j. If $J_n = (j_1, j_2, \ldots, j_n) \in B^n$, we have

$$\mathcal{M} = \bigcup_{J_n \in B^n} \Delta_j^{(n)} \qquad \text{for all } n \geq 1$$

such that

$$\begin{aligned} &\bigcup_{j_n \in B} \Delta_j^{(n)} = \Delta^{(n-1)}(j_1, \ldots, j_{n-1}), \\ &T_v(\Delta_j^{(n)}) = \Delta^{(n-1)}(j_2, \ldots, j_n), \end{aligned}$$

and

$$T_v(\Delta_j^{(1)}) = \mathcal{M}.$$

We will also need the following lemmas.

Lemma 2.3.3. *Let $N = a_1 + \cdots + a_n$. Then $T_v^{-n}|_{\Delta_j^{(n)}}$ is bijective and*

$$d\mu(T_v^{-n}(x)) = (\#(k))^{-N} d\mu(x)$$

(on $\Delta_j^{(n)}$).

Proof: For $x \in \Delta_j^{(n)}$, its n^{th} convergent is $\frac{A_n}{B_n}$ defined by (2.5). One can see that $T_v^{-n}(x)|_{\Delta_j^{(n)}}$ is nothing else than $g(x)$ which is defined by (2.6). Using Lemma 2.2.2 we get that

$$g(x) - \frac{A_n}{B_n} = \frac{xA_{n-1} + A_n}{xB_{n-1} + B_n} - \frac{A_n}{B_n} = \frac{x(A_{n-1}B_n - A_nB_{n-1})}{B_n(xB_{n-1} + B_n)} = \frac{(-1)x\pi^N}{B_n(xB_{n-1} + B_n)}$$

and so

$$g(x) = \frac{A_n}{B_n} + \frac{(-1)^n x\pi^N}{(xB_{n-1} + B_n)B_n}.$$

As B_n is in $\mathcal{O}^\times$ and multiplication by π^N scales Haar measure by $(\#(k))^{-N}$, this lemma is proved if we show that the map $t : \mathcal{M} \to \mathcal{M}$ defined by

$$t(x) = \frac{x}{xB_{n-1} + B_n}$$

preserves Haar measure. Fix $L \in \mathbb{N}$ and $u \in \mathcal{O}$. Let $v = t(\pi u) = \frac{\pi u}{\pi u B_{n-1} + B_n}$. If $x \in \pi u + \pi^L\mathcal{O}$ then we get $t(x) \in v + \pi^L\mathcal{O}$. Conversely, if we suppose $w \in v + \pi^L\mathcal{O}$, then we get

$$t^{-1}(w) = \frac{wB_n}{1 - wB_{n-1}}$$

which belongs to the coset $t^{-1}(v) + \pi^L\mathcal{O} = \pi u + \pi^L\mathcal{O}$. So, we checked that for every u and L, t maps the coset $\pi u + \pi^L\mathcal{O}$ bijectively to the coset $t(\pi u) + \pi^L\mathcal{O}$. Cosets of this type form a basis for the open sets of $\mathcal{M}$ and have the same measure, so t is measure-preserving with respect to Haar measure. Hence our lemma is proved. □

Lemma 2.3.4. *We have that $\mu(\Delta_j^{(n)}) = \frac{1}{(\#(k))^N}$ where $N = a_1 + \cdots + a_n$.*

Proof: Since $\Delta_j^{(n)}$ is the set of all $x \in \mathcal{M}$ with continued fraction expansion starting with $j_1, j_2, \ldots, j_n$, it means that we need to compute the Haar measure of the set

$$\cfrac{\pi^{a_1}}{b_1+\cfrac{\pi^{a_2}}{b_2+\cfrac{\pi^{a_3}}{b_3+\ddots+\cfrac{\pi^{a_n}}{b_n+\pi\mathcal{O}}}}}.$$

If we apply Lemma 2.3.1 repeatedly, we obtain that

$$\cfrac{\pi^{a_1}}{b_1+\cfrac{\pi^{a_2}}{b_2+\cfrac{\pi^{a_3}}{b_3+\ddots+\cfrac{\pi^{a_n}}{b_n+\pi\mathcal{O}}}}}=\cfrac{\pi^{a_1}}{b_1+\cfrac{\pi^{a_2}}{b_2+\cfrac{\pi^{a_3}}{b_3+\ddots+\cfrac{\pi^{a_n}}{b_n}}}}+\pi^{a_1+\cdots+a_n+1}\mathcal{O},$$

which has the Haar measure equal to $(\#(k))^{-N}$ with $N=a_1+\cdots+a_n$. Hence, the lemma is proved. □

Note that the above results are generalisations of the results proved in [24].

Theorem 2.3.5. *The generalised Schneider's continued fraction map T_v is exact.*

Proof: We would like to show that $(\mathcal{M},\mathcal{B},\mu,T_v)$ is exact where $\mathcal{B}$ is the Haar σ-algebra of $\mathcal{M}$, μ is Haar measure on $\mathcal{M}$ and T_v is described by (2.2). For the proof we will use Lemma 1.2.12.

In our case $\mathcal{A}$ is the collection of all finite unions of cylinder-sets $\Delta_j^{(n)}$ forming an algebra which generates the Haar σ-algebra $\mathcal{B}$. Let $B\in\bigcap_{n=1}^{\infty}T_v^{-n}(\mathcal{B})$. It suffices to show that

$$\mu(B\cap\Delta_j^{(n)})=\mu(B)\mu(\Delta_j^{(n)})$$

for all $\Delta_j^{(n)}$.

For each $n \geq 1$ there exists a set A_n such that $\mu(A_n) = \mu(B)$ and $B = T_v^{-n}(A_n)$. By the change of variables formula if we set

$$A_j = A(j_1, j_2, \ldots, j_n) = T_v^{-n}(A) \cap \Delta^{(n)}(j_1, j_2, \ldots, j_n)$$

then we get

$$\mu(A_j) = \int_A \frac{d\mu(T_v^{-n}(x))}{d\mu(x)} d\mu(x) = \int_A d\mu(T_v^{-n}(x)).$$

By Lemma 2.3.3 we have that $d\mu(T_v^{-n}(x)) = \frac{1}{(\#(k))^N} d\mu(x)$. Thus

$$\mu(A_j) = \frac{1}{(\#(k))^N} \int_A d\mu(x) = \frac{1}{(\#(k))^N} \mu(A).$$

By Lemma 2.3.4 we know that $\frac{1}{(\#(k))^N} = \mu(\Delta_j^{(n)})$ and so we get

$$\mu(A_j) = \mu(\Delta_j^{(n)})\mu(A). \tag{2.8}$$

Applying (2.8) to the set A_n, we get

$$\mu(B \cap \Delta_j^{(n)}) = \mu(T_v^{-n}(A_n) \cap \Delta_j^{(n)}) = \mu(A_n)\mu(\Delta_j^{(n)}) = \mu(B)\mu(\Delta_j^{(n)})$$

which we wanted to show. Now from Lemma 1.2.12 it follows that $\mu(B) = 0$ or 1 so T_v is exact. $\square$

Because $(\mathcal{M}, \mathcal{B}, \mu, T_v)$ is exact, by Propositions 1.2.8 – 1.2.14, other strictly weaker properties are implied:

- T_v is strong-mixing, i.e. for all $A, B \in \mathcal{B}$ we have

  $$\lim_{n\to\infty} \mu(T_v^{-n} A \cap B) = \mu(A)\mu(B)$$

  which impies

- T_v is weak-mixing, i.e. for all $A, B \in \mathcal{B}$ we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} |\mu(T_v^{-j} A \cap B) - \mu(A)\mu(B)| = 0$$

which implies

- T_v is ergodic, i.e. $\mu(B) = 0$ or 1 for any $B \in \mathcal{B}$ with $T_v^{-1}(B) = B$.

The following theorem is an analogue of Khinchin's Theorem for our map T_v. It is the generalised version of the p-adic analogue of Khinchin's Theorem proved in [24].

Theorem 2.3.6. *For almost all $x \in \mathcal{M}$ the generalised Schneider's continued fraction expansion (2.3) satisfies*

$$\lim_{n\to\infty} \frac{a_1 + a_2 + \cdots + a_n}{n} = \frac{\#(k)}{\#(k) - 1}.$$

Proof: We use Ergodic Theorem 1.2.16. Since T_v is ergodic and measure-preserving transformation and we will see that $f \in L^1(\mathcal{M}, \mathcal{B}, \mu)$, then we have

$$\lim_{n\to\infty} \frac{f(x) + f(T_v x) + \cdots + f(T_v^{n-1} x)}{n} = \int_{\mathcal{M}} f d\mu$$

for almost all $x \in \mathcal{M}$. Let $f(x) = v(x)$ where $v(x)$ is the valuation. Then we have $f(T_v^i x) = a_{i+1}$ and thus we get

$$\lim_{n\to\infty} \frac{a_1 + a_2 + \cdots + a_n}{n} = \int_{x\in\mathcal{M}} v(x) d\mu(x).$$

We can write

$$\int_{\mathcal{M}} v(x) d\mu(x) = \sum_{n=1}^{\infty} n\mu(\{x : v(x) = n\})$$

where $\mu(\{x : v(x) = n\}) = \mu(\pi^n \mathcal{O} \backslash \pi^{n+1} \mathcal{O})$. Because $\pi^{n+1}\mathcal{O} \subset \pi^n \mathcal{O}$, then

$$\mu(\pi^n \mathcal{O} \backslash \pi^{n+1} \mathcal{O}) = \mu(\pi^n \mathcal{O}) - \mu(\pi^{n+1} \mathcal{O}) = \frac{1}{(\#(k))^{n-1}} - \frac{1}{(\#(k))^n} = \frac{\#(k) - 1}{(\#(k))^n}$$

for $n = 1, 2, \ldots$. Using the identity $\sum_{n=1}^{\infty} nx^n = \frac{x}{(1-x)^2}$ for $|x| < 1$ we get that

$$\begin{aligned}
\sum_{n=1}^{\infty} n\mu(\{x : v(x) = n\}) &= \sum_{n=1}^{\infty} \frac{n(\#(k)-1)}{(\#(k))^n} = (\#(k)-1) \sum_{n=1}^{\infty} \frac{n}{(\#(k))^n} \\
&= (\#(k)-1) \cdot \frac{\frac{1}{\#(k)}}{(1-\frac{1}{\#(k)})^2} = \frac{\#(k)}{\#(k)-1}.
\end{aligned}$$

So, we obtained that

$$\lim_{n\to\infty} \frac{a_1 + a_2 + \cdots + a_n}{n} = \frac{\#(k)}{\#(k)-1},$$

which we wanted to prove. Note that this also shows that $f(x) = v(x)$ belongs to $L^1(\mathcal{M}, \mathcal{B}, \mu)$. □

2.4 Entropy of Generalised Schneider's Continued Fraction Map

The main goal of this section is to compute the measure-theoretic entropy of the map T_v.

Theorem 2.4.1. *Let $\mathcal{B}$ denote the Haar σ-algebra restricted to $\mathcal{M}$ and let μ denote Haar measure on $\mathcal{M}$. Then the measure-theoretic entropy of the measure-preserving transformation $T_v : \mathcal{M} \to \mathcal{M}$ defined by (2.2) with respect to Haar measure μ is $\frac{\#(k)}{\#(k)-1}\log(\#(k))$.*

Proof: Take $j_n = (b_n, a_n), n = 1, 2, \ldots$ with $j_r \neq j_s$ if $r \neq s$ and let $\alpha = \{\Delta(j_1), \Delta(j_2), \Delta(j_3), \ldots\}$ be the partition. Notice that

$$\begin{aligned}
\Delta_j^{(n)} &= \Delta(j_1) \cap T_v^{-1}(\Delta(j_2)) \cap T_v^{-2}(\Delta(j_3)) \cap \cdots \cap T_v^{-(n-1)}(\Delta(j_n)) \\
&= \Delta_j^{(1)} \cap \bigcup_{J_1 \in B} \Delta_j^{(2)} \cap \bigcup_{J_2 \in B^2} \Delta_j^{(3)} \cap \cdots \cap \bigcup_{J_{n-1} \in B^{n-1}} \Delta_j^{(n)}.
\end{aligned}$$

To compute entropy, we first need to find the conditional information function

$I(\alpha|\bigvee_{i=1}^{n-1} T_v^{-i}\alpha)$ which is defined as

$$I(\alpha|T_v^{-1}\alpha\vee\cdots\vee T_v^{-(n-1)}\alpha) = -\sum_{\Delta(j)\in\alpha} \chi_{\Delta(j)}(x)\log\mu(\Delta(j)|T_v^{-1}\alpha\vee\cdots\vee T_v^{-(n-1)}\alpha).$$

If $x \in \Delta_j^{(n)}$, then $\chi_{\Delta(j_1)}(x) = 1$ and $\chi_{\Delta(j_i)}(x) = 0$ for all $i \geq 2$. So, we get

$$I(\alpha|T_v^{-1}\alpha \vee \cdots \vee T_v^{-(n-1)}\alpha) = -\log\mu(\Delta(j_1)|T_v^{-1}\alpha \vee \cdots \vee T_v^{-(n-1)}\alpha).$$

The conditional probability is

$$\mu(\Delta(j_1)|T_v^{-1}\alpha \vee \cdots \vee T_v^{-(n-1)}\alpha) = \sum_{C\in T_v^{-1}\alpha\vee\cdots\vee T_v^{-(n-1)}\alpha} \chi_C(x)\frac{\mu(\Delta(j_1)\cap C)}{\mu(C)}.$$

If $x \in \Delta_j^{(n)}$, we set $C_1 = T_v^{-1}(\Delta(j_2))\cap T_v^{-2}(\Delta(j_3))\cap\cdots\cap T_v^{-(n-1)}(\Delta(j_n))$. Then we can see that $\chi_{C_1}(x) = 1$ and for other

$$C_i \neq T_v^{-1}(\Delta(j_2)) \cap \cdots \cap T_v^{-(n-1)}(\Delta(j_n))$$

where $i \geq 2$ we have $\chi_{C_i}(x) = 0$. Thus, we obtain that

$$\begin{aligned}
&\mu(\Delta(j_1)|T_v^{-1}\alpha \vee \cdots \vee T_v^{-(n-1)}\alpha) \\
&\qquad = \frac{\mu(\Delta(j_1) \cap T_v^{-1}(\Delta(j_2)) \cap \cdots \cap T_v^{-(n-1)}(\Delta(j_n)))}{\mu(T_v^{-1}(\Delta(j_2)) \cap \cdots \cap T_v^{-(n-1)}(\Delta(j_n)))} \\
&\qquad = \frac{\mu(\Delta^{(n)}(j_1, j_2, \ldots, j_n))}{\mu(\Delta^{(n-1)}(j_2, j_3, \ldots, j_n))}.
\end{aligned}$$

Recall that by Lemma 2.3.4 we have $\mu(\Delta_j^{(n)}) = \frac{1}{(\#(k))^N}$ with $N = a_1 + \cdots + a_n$. Thus, we have that

$$\mu(\Delta(j_1)|T_v^{-1}\alpha \vee \cdots \vee T_v^{-(n-1)}\alpha) = \frac{1}{(\#(k))^N}\Big/\frac{1}{(\#(k))^{N-a_1}} = (\#(k))^{-a_1}$$

and the conditional information function is

$$I(\alpha|T_v^{-1}\alpha \vee \cdots \vee T_v^{-(n-1)}\alpha) = -\log((\#(k))^{-a_1}) = a_1\log(\#(k)).$$

By (1.2), we see that the entropy of T_v relative to the partition α is

$$h_\mu(T_v, \alpha) = \lim_{n\to\infty} H\left(\alpha | \bigvee_{i=1}^{n-1} T_v^{-i}\alpha\right)$$

where

$$H\left(\alpha | \bigvee_{i=1}^{n-1} T_v^{-i}\alpha\right) = \int I\left(\alpha | \bigvee_{i=1}^{n-1} T_v^{-i}\alpha\right) d\mu.$$

So, we get

$$h_\mu(T_v, \alpha) = \lim_{n\to\infty} \int a_1 \log(\#(k))\, d\mu.$$

Notice that $a_1(x) = v(x)$ and in the proof of Theorem 2.3.6 we computed that $\int_{\mathcal{M}} v(x)\, d\mu = \frac{\#(k)}{\#(k)-1}$. Hence, we have

$$h_\mu(T_v, \alpha) = \lim_{n\to\infty} \int v(x) \log(\#(k))\, d\mu = \frac{\#(k)}{\#(k)-1} \log(\#(k)).$$

We claim that α is a strong generator for T_v. This is because for almost every $x, y \in \mathcal{M}$ if $x \neq y$, the points x and y have distinct generalised Schneider's continued fraction expansions. This implies the partition α seperates almost every pair of points. Hence, by Sinai's Theorem 1.3.12, the measure-theoretic entropy of T_v with respect to the Haar measure μ is

$$h_\mu(T_v) = h_\mu(T_v, \alpha) = \frac{\#(k)}{\#(k)-1} \log(\#(k)).$$

$\square$

2.5 Bernoulli Property for Generalised Schneider's Continued Fraction Map

In this section we will prove that the natural extension of our map T_v is Bernoulli, i.e. it is isomorphic to a Bernoulli shift. To prove this, we first need to introduce some definitions which can be also found in [28] and [64].

Definition 2.5.1. Let $P = (p_1, p_2, \ldots)$ and $Q = (q_1, q_2, \ldots)$ denote two μ-measurable denumerable partitions of the same set X. Then P and Q are said to be *ε-independent* and we write $P \perp^{\varepsilon} Q$ if

$$\sum_i \sum_j |\mu(p_i \cap q_j) - \mu(p_i)\mu(q_j)| < \varepsilon.$$

Definition 2.5.2. A denumerable partition P is called *weak Bernoulli* with respect to an invertible, measure-preserving transformation T if for each $\varepsilon > 0$ there exists a positive constant $K = K(\varepsilon)$ such that for every $n \geq 0$ we have

$$\bigvee_{i=-n}^{0} T^i P \perp^{\varepsilon} \bigvee_{i=K}^{K+n} T^i P.$$

Definition 2.5.3. We say that T is a *weak Bernoulli transformation* if T has a weak Bernoulli generator.

The given definition of weak Bernoulli property is not the only way to formulate this property. As observed in [64] for a non-invertible transformation we have the following definition.

Definition 2.5.4. A denumerable partition P is *weak Bernoulli* for a non-invertible, measure-preserving transformation T if for each $\varepsilon > 0$ there exists $K = K(\varepsilon)$ such that for all $n \geq 0$ we have

$$\bigvee_{i=0}^{n} T^{-i} P \perp^{\varepsilon} \bigvee_{i=K+n}^{K+2n} T^{-i} P.$$

A natural extension of a non-invertible measure-preserving transformation T is then *weak Bernoulli* if T has a weak Bernoulli generator.

Because we need our transformation to be invertible (so the following theorem can be used), natural extensions come into consideration. From Theorem 1.4.9 we know that for any measure-preserving transformation there exists a unique natural extension and hence, our map T_v has a unique natural extension. Using Theorem 1.4.10 and results from Section 2.3 for the map T_v, we also know that the natural extension of T_v is ergodic, weak-mixing and strong-mixing and it is a K-automorphism. The measure-theoretic entropy of the natural extension of T_v is equal to $\frac{\#(k)}{\#(k)-1}\log(\#(k))$.

The existence of an isomorphism to a Bernoulli shift is then ensured by the following theorem which was proved by N.A. Friedmann and D.S. Ornstein [14].

Theorem 2.5.5. *A weak Bernoulli (invertible) transformation is isomorphic to a Bernoulli shift with the same entropy.*

Proof: The proof can be found in [14]. $\square$

Now, we will prove the following theorem.

Theorem 2.5.6. *Let $\mathcal{B}$ denote the Haar σ-algebra restricted to $\mathcal{M}$ and let μ denote Haar measure on $\mathcal{M}$. Then the transformation $T_v : \mathcal{M} \to \mathcal{M}$ defined by (2.2) has a natural extension that is Bernoulli.*

Proof: Set
$$A_j = T_v^{-n-\ell}(A) \cap \Delta_j^{(n)}.$$

Then we get
$$\mu(A_j) = \int_{T_v^{-\ell}A} \frac{d\mu(T_v^{-n}(x))}{d\mu(x)} d\mu(x) = \int_{T_v^{-\ell}A} d\mu(T_v^{-n}(x)).$$

By Lemma 2.3.3 we have that
$$\mu(A_j) = \frac{1}{(\#(k))^N} \int_{T_v^{-\ell}A} d\mu(x) = \frac{1}{(\#(k))^N}\mu(T_v^{-\ell}A) = \frac{1}{(\#(k))^N}\mu(A).$$

Recall that by Lemma 2.3.4 we know that $\frac{1}{(\#(k))^N} = \mu(\Delta_j^{(n)})$ and so we get
$$\mu(T_v^{-n-\ell}(A) \cap \Delta_j^{(n)}) = \mu(\Delta_j^{(n)})\mu(A).$$

Suppose both $\Delta_j^{(n)}$ and A belong to $\bigvee_{i=0}^{n} T_v^{-i}\alpha$ where
$$\alpha = \{\Delta(j_1), \Delta(j_2), \Delta(j_3), \ldots\}$$

is a generator for T_v. Then $\Delta = T_v^{-\ell-n} A \in \bigvee_{i=\ell+n}^{\ell+2n} T_v^{-i}\alpha$ and we get

$$\mu(\Delta \cap \Delta_j^{(n)}) - \mu(\Delta)\mu(\Delta_j^{(n)}) = 0$$

which implies

$$\sum_{\Delta_j^{(n)} \in \bigvee_{i=0}^{n} T_v^{-i}\alpha} \; \sum_{\Delta \in \bigvee_{i=\ell+n}^{\ell+2n} T_v^{-i}\alpha} \left| \mu(\Delta \cap \Delta_j^{(n)}) - \mu(\Delta)\mu(\Delta_j^{(n)}) \right| = 0 < \varepsilon.$$

Thus we proved that the generator α for T_v is weak Bernoulli and so the natural extension of T_v is weak Bernoulli which by Theorem 2.5.5 means that the natural extension of T_v is isomorphic to a Bernoulli shift with the entropy $\frac{\#(k)}{\#(k)-1}\log(\#(k))$. □

2.6 Application of Pointwise Subsequence Ergodic Theorems

In this section we state and prove a number of interesting applications of Theorem 1.6.3.

Theorem 2.6.1. *Suppose $(k_n)_{n\geq 1}$ is a L^p-good universal sequence such that $(\{k_n\gamma\})_{n\geq 1}$ is uniformly distributed modulo one for each irrational number γ and suppose $F_1 : \mathbb{R}_{\geq 0} \to \mathbb{R}$ and $F_2 : \mathbb{R}_{\geq 0} \to \mathbb{R}$ are continuous increasing functions with*

$$\int_{\mathcal{M}} |F_1(a_1(x))|^p \, d\mu < \infty \quad \textit{and} \quad \int_{\mathcal{M}} |F_2(b_1(x))|^p \, d\mu < \infty$$

For each $n \in \mathbb{N}, \ell \in \{1,2\}$ and arbitrary real numbers $d_1, \ldots, d_n$ we define

$$M_{F_\ell,n}(d_1, \ldots, d_n) = F_\ell^{-1}\left[\frac{F_\ell(d_1) + \cdots + F_\ell(d_n)}{n}\right].$$

Then we have that

$$\lim_{n\to\infty} M_{F_1,n}(a_{k_1}(x), \ldots, a_{k_n}(x)) = F_1^{-1}\left[\int_{\mathcal{M}} F_1(a_1(x)) \, d\mu\right]$$

and

$$\lim_{n\to\infty} M_{F_2,n}(b_{k_1}(x),\ldots,b_{k_n}(x)) = F_2^{-1}\left[\int_{\mathcal{M}} F_2(b_1(x))\, d\mu\right]$$

almost everywhere with respect to Haar measure on $\mathcal{M}$.

Proof: We apply Theorem 1.6.3 with $f(x) = F_1(a_1(x))$ which means we get

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} F_1(a_1(T_v^{k_n-1}x)) = \int_{\mathcal{M}} F_1(a_1(x))\, d\mu.$$

Notice that $a_1(T_v^{k_n-1}x) = a_{k_n}(x)$. Now

$$\begin{aligned}
\lim_{n\to\infty} M_{F_1,n}(a_{k_1}(x),\ldots,a_{k_n}(x)) &= \lim_{n\to\infty} F_1^{-1}\left[\frac{F_1(a_{k_1}(x)) + \cdots + F_1(a_{k_n}(x))}{n}\right] \\
&= \lim_{n\to\infty} F_1^{-1}\left[\frac{1}{n}\sum_{j=1}^{n} F_1(a_{k_j}(x))\right] = F_1^{-1}\left[\lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n} F_1(a_1(T_v^{k_j-1}x))\right] \\
&= F_1^{-1}\left[\int_{\mathcal{M}} F_1(a_1(x))\, d\mu\right]
\end{aligned}$$

which we wanted to show.

For the other part of the proof, note that $b_1(T_v^{k_n-1}x) = b_{k_n}(x)$ and as before, apply Theorem 1.6.3 with $f(x) = F_2(b_1(x))$. □

Theorem 2.6.2. *Let $(k_n)_{n\geq 1}$ be a L^p-good universal sequence $(k_n)_{n\geq 1}$ where the sequence $(\{k_n\gamma\})_{n\geq 1}$ is uniformly distributed modulo one for each irrational number γ. Let $H_1 : \mathbb{N}^m \to \mathbb{R}$ and $H_2 : \mathbb{N}^m \to \mathbb{R}$ be functions such that*

$$\int_{\mathcal{M}} |H_1(a_1(x),\ldots,a_m(x))|^p\, d\mu < \infty \text{ and } \int_{\mathcal{M}} |H_2(b_1(x),\ldots,b_m(x))|^p\, d\mu < \infty$$

where m is a positive integer. Then we have that

$$\begin{aligned}
&\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} H_1(a_{k_n}(x), a_{k_n+1}(x),\ldots,a_{k_n+m-1}(x)) \\
&= \sum_{(i_1,\ldots,i_m)\in\mathbb{N}^m} \frac{(\#(k)-1)^m}{(\#(k))^{i_1+\cdots+i_m}} H_1(i_1,\ldots,i_m)
\end{aligned} \tag{2.9}$$

and

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} H_2(b_{k_n}(x), b_{k_n+1}(x), \ldots, b_{k_n+m-1}(x)) \tag{2.10}$$
$$= \sum_{(i_1,\ldots,i_m)\in\{k\setminus\{0\}\}^m} \frac{1}{(\#(k)-1)^m} H_2(i_1,\ldots,i_m)$$

almost everywhere with respect to Haar measure on $\mathcal{M}$.

Proof: To prove (2.9), we apply Theorem 1.6.3 with $f(x) = H_1(a_1(x), \ldots, a_m(x))$. So, we obtain that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} H_1(a_1(T_v^{k_n-1}x), a_2(T_v^{k_n-1}x), \ldots, a_m(T_v^{k_n-1}x))$$
$$= \int_{\mathcal{M}} H_1(a_1(x), a_2(x), \ldots, a_m(x))\, d\mu.$$

For positive integers i_ℓ where $\ell \in \{1, 2, \ldots, m\}$ we have that

$$\int_{\mathcal{M}} H_1(a_1(x), a_2(x), \ldots, a_m(x))\, d\mu$$
$$= \sum_{(i_1,\ldots,i_m)\in\mathbb{N}^m} \frac{(\#(k)-1)^m}{(\#(k))^{i_1+\cdots+i_m}} H_1(i_1,\ldots,i_m).$$

Since we have $a_i(T_v^{k_n-1}x) = a_{k_n+i-1}(x)$ for $i \geq 1$, we get that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} H_1(a_1(T_v^{k_n-1}x), a_2(T_v^{k_n-1}x), \ldots, a_m(T_v^{k_n-1}x))$$
$$= \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} H_1(a_{k_n}(x), a_{k_n+1}(x), a_{k_n+2}(x), \ldots, a_{k_n+m-1}(x))$$
$$= \sum_{(i_1,\ldots,i_m)\in\mathbb{N}^m} \frac{(\#(k)-1)^m}{(\#(k))^{i_1+\cdots+i_m}} H_1(i_1,\ldots,i_m)$$

as required.

To prove (2.10), we apply Theorem 1.6.3 with $f(x) = H_2(b_1(x), \ldots, b_m(x))$.

So, we get that

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} H_2(b_1(T_v^{k_n-1}x), b_2(T_v^{k_n-1}x), \ldots, b_m(T_v^{k_n-1}x))$$

$$= \int_{\mathcal{M}} H_2(b_1(x), b_2(x), \ldots, b_m(x))\, d\mu.$$

For $i_\ell \in k\backslash\{0\}$ where $\ell \in \{1, 2, \ldots, m\}$ we have that

$$\int_{\mathcal{M}} H_2(b_1(x), b_2(x), \ldots, b_m(x))\, d\mu$$
$$= \sum_{(i_1,\ldots,i_m)\in\{k\backslash\{0\}\}^m} \frac{1}{(\#(k)-1)^m} H_2(i_1, \ldots, i_m).$$

Since we have $b_i(T_v^{k_n-1}x) = b_{k_n+i-1}(x)$ for $i \geq 1$, we obtain that

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} H_2(b_1(T_v^{k_n-1}x), b_2(T_v^{k_n-1}x), \ldots, b_m(T_v^{k_n-1}x))$$
$$= \lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} H_2(b_{k_n}(x), b_{k_n+1}(x), b_{k_n+2}(x), \ldots, b_{k_n+m-1}(x))$$
$$= \sum_{(i_1,\ldots,i_m)\in\{k\backslash\{0\}\}^m} \frac{1}{(\#(k)-1)^m} H_2(i_1, \ldots, i_m)$$

as required. □

Theorem 2.6.3. *For any L^p-good universal sequence $(k_n)_{n\geq 1}$ such that the sequence $(\{k_n\gamma\})_{n\geq 1}$ is uniformly distributed modulo one for each irrational number γ we have*

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} a_{k_n} = \frac{\#(k)}{\#(k)-1}, \tag{2.11}$$

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} b_{k_n} = \frac{\#(k)}{2}, \tag{2.12}$$

almost everywhere with respect to Haar measure on $\mathcal{M}$.

Proof: Again, we apply Theorem 1.6.3 where we take $f(x) = v(x)$ for the relation (2.11). Hence, we get that

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}v(T_v^{k_n-1}x)=\int_{\mathcal{M}}v(x)\,d\mu=\frac{\#(k)}{\#(k)-1}.$$

Since $v(T_v^{k_n-1}x)=a_{k_n}$, we get

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}v(T_v^{k_n-1}x)=\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}a_{k_n}=\frac{\#(k)}{\#(k)-1}.$$

To prove (2.12) we take $f(x)=b(x)$ and applying Theorem 1.6.3 we get that

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}b(T_v^{k_n-1}x)=\int_{\mathcal{M}}b(x)\,d\mu=\frac{\#(k)}{2}.$$

Since $b(T_v^{k_n-1}x)=b_{k_n}$, we obtain that

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}b(T_v^{k_n-1}x)=\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}b_{k_n}=\frac{\#(k)}{2}.$$

□

Next, we will investigate the Gauss-Kuzmin distribution (see Section 2.1) in the generalised Schneider's continued fraction with respect to a special sequence $(k_n)_{n\geq 1}$.

Theorem 2.6.4. *Let $(k_n)_{n\geq 1}$ be an L^p-good universal sequence such that the sequence $(\{k_n\gamma\})_{n\geq 1}$ is uniformly distributed modulo one for each irrational number γ. Then for positive integers i and j, $i<j$ we have*

$$\lim_{N\to\infty}\frac{1}{N}\#\{1\leq n\leq N: a_{k_n}=i\}=\frac{\#(k)-1}{(\#(k))^{i}},\tag{2.13}$$

$$\lim_{N\to\infty}\frac{1}{N}\#\{1\leq n\leq N: a_{k_n}\geq i\}=\frac{1}{(\#(k))^{i-1}},\tag{2.14}$$

$$\lim_{N\to\infty}\frac{1}{N}\#\{1\leq n\leq N: i\leq a_{k_n}<j\}=\frac{1}{(\#(k))^{i-1}}\left(1-\frac{1}{(\#(k))^{j-i}}\right);\tag{2.15}$$

almost everywhere with respect to Haar measure on $\mathcal{M}$.

Proof: To prove (2.13) we apply Theorem 1.6.3 with $f(x) = I_{B_1}(x)$, where I_{B_1} is the characteristic function of the set $B_1 = \{x \in \mathcal{M} : a_1(x) = i\}$. Hence, we get that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} I_{B_1}(T_v^{k_n - 1}x) = \int_{\mathcal{M}} I_{B_1}(x)\, d\mu = \mu(\{x \in \mathcal{M} : a_1(x) = i\})$$

$$\begin{aligned}
&= \mu(\{x \in \mathcal{M} : v(x) = i\}) = \mu\left(\bigcup_{b\in k\setminus\{0\}} \frac{\pi^i}{b + T_v(x)}\right) \\
&= \sum_{b\in k\setminus\{0\}} \mu\left(\frac{\pi^i}{b+T_v(x)}\right) = \sum_{b\in k\setminus\{0\}} \frac{1}{(\#(k))^i} = \frac{\#(k)-1}{(\#(k))^i}.
\end{aligned}$$

Since

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} I_{B_1}(T_v^{k_n - 1}x) = \lim_{N\to\infty} \frac{1}{N} \#\{1 \le n \le N : a_{k_n} = i\},$$

we obtain that

$$\lim_{N\to\infty} \frac{1}{N} \#\{1 \le n \le N : a_{k_n} = i\} = \frac{\#(k)-1}{(\#(k))^i}.$$

To prove (2.14) we apply Theorem 1.6.3 with $f(x) = I_{B_2}(x)$, where I_{B_2} is the characteristic function of the set $B_2 = \{x \in \mathcal{M} : a_1(x) \ge i\}$. So, we have

$$\begin{aligned}
\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} I_{B_2}(T_v^{k_n - 1}x) &= \int_{\mathcal{M}} I_{B_2}(x)\, d\mu = \mu(\{x \in \mathcal{M} : a_1(x) \ge i\}) \\
&= \mu\left(\bigcup_{l=i}^{\infty} \{x \in \mathcal{M} : a_1(x) = l\}\right) = \sum_{l=i}^{\infty} \mu(\{x \in \mathcal{M} : a_1(x) = l\}) \\
&= \sum_{l=i}^{\infty} \frac{\#(k)-1}{(\#(k))^l} = \frac{\#(k)-1}{(\#(k))^i} \cdot \left(1 + \frac{1}{\#(k)} + \frac{1}{(\#(k))^2} + \cdots\right) \\
&= \frac{\#(k)-1}{(\#(k))^i} \cdot \frac{1}{1 - \frac{1}{\#(k)}} = \frac{1}{(\#(k))^{i-1}}.
\end{aligned}$$

Because

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} I_{B_2}(T_v^{k_n-1}x) = \lim_{N\to\infty}\frac{1}{N}\#\{1\le n\le N : a_{k_n}\ge i\},$$

we get that

$$\lim_{N\to\infty}\frac{1}{N}\#\{1\le n\le N : a_{k_n}\ge i\} = \frac{1}{(\#(k))^{i-1}}.$$

To prove (2.15) we again apply Theorem 1.6.3 with $f(x) = I_{B_3}(x)$, where I_{B_3} is the characteristic function of the set $B_3 = \{x\in\mathcal{M} : i\le a_1(x) < j\}$. So,

$$\begin{aligned}
\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} I_{B_3}(T_v^{k_n-1}x) &= \int_{\mathcal{M}} I_{B_3}(x)\,d\mu = \mu(\{x\in\mathcal{M} : i\le a_1(x) < j\})\\
&= \mu\left(\bigcup_{l=i}^{j-1}\{x\in\mathcal{M} : a_1(x) = l\}\right) = \sum_{l=i}^{j-1}\mu(\{x\in\mathcal{M} : a_1(x)=l\})\\
&= \sum_{l=i}^{j-1}\frac{\#(k)-1}{(\#(k))^l} = \frac{\#(k)-1}{(\#(k))^i}\cdot\left(1+\frac{1}{\#(k)}+\frac{1}{(\#(k))^2}+\cdots+\frac{1}{(\#(k))^{j-i-1}}\right)\\
&= \frac{\#(k)-1}{(\#(k))^i}\cdot\frac{1-\frac{1}{(\#(k))^{j-i}}}{1-\frac{1}{\#(k)}} = \frac{1}{(\#(k))^{i-1}}\left(1-\frac{1}{(\#(k))^{j-i}}\right).
\end{aligned}$$

Because

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} I_{B_3}(T_v^{k_n-1}x) = \lim_{N\to\infty}\frac{1}{N}\#\{1\le n\le N : i\le a_{k_n} < j\},$$

we get that

$$\lim_{N\to\infty}\frac{1}{N}\#\{1\le n\le N : i\le a_{k_n} < j\} = \frac{1}{(\#(k))^{i-1}}\left(1-\frac{1}{(\#(k))^{j-i}}\right).$$

□

Theorem 2.6.5. *Let $(k_n)_{n\geq 1}$ be an L^p-good universal sequence such that the sequence $(\{k_n\gamma\})_{n\geq 1}$ is uniformly distributed modulo one for each irrational number γ. Then for $i \in k\backslash\{0\}$ and $j \in k\backslash\{0\}$, $i < j$ we have*

$$\lim_{N\to\infty} \frac{1}{N}\#\{1 \leq n \leq N : b_{k_n} = i\} = \frac{1}{\#(k)-1}, \tag{2.16}$$

$$\lim_{N\to\infty} \frac{1}{N}\#\{1 \leq n \leq N : b_{k_n} \geq i\} = \frac{\#(k)-i}{\#(k)-1}, \tag{2.17}$$

$$\lim_{N\to\infty} \frac{1}{N}\#\{1 \leq n \leq N : i \leq b_{k_n} < j\} = \frac{j-i}{\#(k)-1}; \tag{2.18}$$

almost everywhere with respect to Haar measure on $\mathcal{M}$.

Proof: To prove (2.16) we apply Theorem 1.6.3 with $f(x) = I_{D_1}(x)$, where I_{D_1} is the characteristic function of the set $D_1 = \{x \in \mathcal{M} : b_1(x) = i\}$. Hence, we get that

$$\begin{aligned}
\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} I_{D_1}(T_v^{k_n-1}x) &= \int_{\mathcal{M}} I_{D_1}(x)\, d\mu = \mu(\{x \in \mathcal{M} : b_1(x) = i\}) \\
&= \mu\left(\bigcup_{a_1=1}^{\infty} \frac{\pi^{a_1}}{i + T_v(x)}\right) = \sum_{a_1=1}^{\infty} \mu\left(\frac{\pi^{a_1}}{i+T_v(x)}\right) \\
&= \sum_{a_1=1}^{\infty} (\#(k))^{-a_1} = \frac{1}{\#(k)-1}.
\end{aligned}$$

Since

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} I_{D_1}(T_v^{k_n-1}x) = \lim_{N\to\infty} \frac{1}{N}\#\{1 \leq n \leq N : b_{k_n} = i\},$$

we get that

$$\lim_{N\to\infty} \frac{1}{N}\#\{1 \leq n \leq N : b_{k_n} = i\} = \frac{1}{\#(k)-1}.$$

To prove (2.17) we apply Theorem 1.6.3 with $f(x) = I_{D_2}(x)$, where I_{D_2} is the characteristic function of the set $D_2 = \{x \in \mathcal{M} : b_1(x) \geq i\}$. So, we have

$$\begin{aligned}
\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} I_{D_2}(T_v^{k_n-1}x) &= \int_{\mathcal{M}} I_{D_2}(x)\, d\mu = \mu(\{x \in \mathcal{M} : b_1(x) \geq i\}) \\
&= \mu\left(\bigcup_{l=i}^{\#(k)-1} \{x \in \mathcal{M} : b_1(x) = l\} \right) = \sum_{l=i}^{\#(k)-1} \mu(\{x \in \mathcal{M} : b_1(x) = l\}) \\
&= \sum_{l=i}^{\#(k)-1} \frac{1}{\#(k)-1} = \frac{\#(k)-i}{\#(k)-1}.
\end{aligned}$$

Because

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} I_{D_2}(T_v^{k_n-1}x) = \lim_{N\to\infty} \frac{1}{N} \#\{1 \leq n \leq N : b_{k_n} \geq i\},$$

we get that

$$\lim_{N\to\infty} \frac{1}{N} \#\{1 \leq n \leq N : b_{k_n} \geq i\} = \frac{\#(k)-i}{\#(k)-1}.$$

To prove (2.18) we again apply Theorem 1.6.3 with $f(x) = I_{D_3}(x)$, where I_{D_3} is the characteristic function of the set $D_3 = \{x \in \mathcal{M} : i \leq b_1(x) < j\}$. So,

$$\begin{aligned}
\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} I_{D_3}(T_v^{k_n-1}x) &= \int_{\mathcal{M}} I_{D_3}(x)\, d\mu = \mu(\{x \in \mathcal{M} : i \leq b_1(x) < j\}) \\
&= \mu\left(\bigcup_{l=i}^{j-1} \{x \in \mathcal{M} : b_1(x) = l\} \right) = \sum_{l=i}^{j-1} \mu(\{x \in \mathcal{M} : b_1(x) = l\}) \\
&= \sum_{l=i}^{j-1} \frac{1}{\#(k)-1} = \frac{j-i}{\#(k)-1}.
\end{aligned}$$

Since

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} I_{D_3}(T_v^{k_n-1}x) = \lim_{N\to\infty} \frac{1}{N} \#\{1 \leq n \leq N : i \leq b_{k_n} < j\},$$

we get that

$$\lim_{N\to\infty}\frac{1}{N}\#\{1\le n\le N : i\le b_{k_n}<j\}=\frac{j-i}{\#(k)-1}.$$

□

Remark 2.6.6. Note that $(n)_{n\ge1}$ is L^1-good universal sequence and $(\{n\gamma\})_{n\ge1}$ is uniformly distributed modulo 1 for any irrational number γ. So, all previous results will be the same for the case $(k_n)_{n\ge1}=(n)_{n\ge1}$.

2.7 Application of Moving Average Pointwise Ergodic Theorem

In this section we state and prove moving average variants of the results in the previous section. For proofs we will use Theorem 1.7.2 and Theorem 1.7.3.

Theorem 2.7.1. *Suppose that $(n_l,k_l)_{l\ge1}$ is Stoltz. Suppose also that we have $F_1:\mathbb{R}_{\ge0}\to\mathbb{R}$ and $F_2:\mathbb{R}_{\ge0}\to\mathbb{R}$ which are continuous increasing and such that*

$$\int_{\mathcal{M}}|F_1(a_1(x))|\,d\mu<\infty\quad \textit{and}\quad \int_{\mathcal{M}}|F_2(b_1(x))|\,d\mu<\infty.$$

Suppose $M_{F_\ell,n}(d_1,\ldots,d_n)$ for $\ell\in\{1,2\}$ is defined as in Theorem 2.6.1. Then

$$\lim_{l\to\infty}M_{F_1,k_l}(a_{n_l+1}(x),\ldots,a_{n_l+k_l}(x))=F_1^{-1}\left[\int_{\mathcal{M}}F_1(a_1(x))\,d\mu\right]$$

and

$$\lim_{l\to\infty}M_{F_2,k_l}(b_{n_l+1}(x),\ldots,b_{n_l+k_l}(x))=F_2^{-1}\left[\int_{\mathcal{M}}F_2(b_1(x))\,d\mu\right]$$

almost everywhere with respect to Haar measure on $\mathcal{M}$.

Proof: Applying Theorem 1.7.3 with $f(x)=F_1(a_1(x))$, we get that

$$\lim_{l\to\infty}\frac{1}{k_l}\sum_{i=0}^{k_l-1}F_1(a_1(T_v^{n_l+i}x))=\int_{\mathcal{M}}F_1(a_1(x))\,d\mu.$$

Note that $a_1(T_v^{n_l+i}x) = a_{n_l+i+1}(x)$. Now

$$\begin{aligned}
&\lim_{l\to\infty} M_{F_1,k_l}(a_{n_1+1}(x),\ldots,a_{n_l+k_l}(x)) \\
&\quad = \lim_{l\to\infty} F_1^{-1}\left[\frac{F_1(a_{n_1+1}(x)) + \cdots + F_1(a_{n_l+k_l}(x))}{k_l}\right] \\
&\quad = \lim_{l\to\infty} F_1^{-1}\left[\frac{1}{k_l}\sum_{i=1}^{k_l} F_1(a_{n_l+i}(x))\right] = F_1^{-1}\left[\lim_{l\to\infty}\frac{1}{k_l}\sum_{i=1}^{k_l} F_1(a_1(T_v^{n_l+i-1}x))\right] \\
&\quad = F_1^{-1}\left[\lim_{l\to\infty}\frac{1}{k_l}\sum_{i=0}^{k_l-1} F_1(a_1(T_v^{n_l+i}x))\right] = F_1^{-1}\left[\int_{\mathcal{M}} F_1(a_1(x))\, d\mu\right]
\end{aligned}$$

which we wanted to show.

For the other part of the proof, note that $b_1(T_v^{n_l+i}x) = b_{n_l+i+1}(x)$ and as before, apply Theorem 1.7.3 with $f(x) = F_2(b_1(x))$. $\square$

Theorem 2.7.2. *Suppose that $(n_l, k_l)_{l\geq 1}$ is Stoltz and $H_1 : \mathbb{N}^m \to \mathbb{R}$ and $H_2 : \mathbb{N}^m \to \mathbb{R}$ are functions such that*

$$\int_{\mathcal{M}} |H_1(a_1(x),\ldots,a_m(x))|\, d\mu < \infty \quad \textit{and} \quad \int_{\mathcal{M}} |H_2(b_1(x),\ldots,b_m(x))|\, d\mu < \infty$$

where m is a positive integer. Then we have that

$$\begin{aligned}
&\lim_{l\to\infty}\frac{1}{k_l}\sum_{j=1}^{k_l} H_1(a_{n_l+j}, a_{n_l+j+1},\ldots,a_{n_l+j+m-1})(x) \\
&\quad = \sum_{(i_1,\ldots,i_m)\in\mathbb{N}^m} \frac{(\#(k)-1)^m}{(\#(k))^{i_1+\cdots+i_m}} H_1(i_1,\ldots,i_m)
\end{aligned} \tag{2.19}$$

and

$$\begin{aligned}
&\lim_{l\to\infty}\frac{1}{k_l}\sum_{j=1}^{k_l} H_2(b_{n_l+j}, b_{n_l+j+1},\ldots,b_{n_l+j+m-1})(x) \\
&\quad = \sum_{(i_1,\ldots,i_m)\in\{k\setminus\{0\}\}^m} \frac{1}{(\#(k)-1)^m} H_2(i_1,\ldots,i_m)
\end{aligned} \tag{2.20}$$

almost everywhere with respect to Haar measure on $\mathcal{M}$.

Proof: To prove (2.19), we apply Theorem 1.7.3 with $f(x) = H_1(a_1(x), \ldots, a_m(x))$ and hence we get that

$$\begin{aligned}
&\lim_{l\to\infty} \frac{1}{k_l} \sum_{j=0}^{k_l-1} H_1(a_1(T_v^{n_l+j}x), a_2(T_v^{n_l+j}x), \ldots, a_m(T_v^{n_l+j}x)) \\
&\qquad = \int_{\mathcal{M}} H_1(a_1(x), a_2(x), \ldots, a_m(x)) \, d\mu \\
&\qquad = \sum_{(i_1,\ldots,i_m)\in\mathbb{N}^m} \frac{(\#(k)-1)^m}{(\#(k))^{i_1+\cdots+i_m}} H_1(i_1, \ldots, i_m).
\end{aligned}$$

Since $a_i(T_v^{n_l+j}x) = a_{n_l+j+i}(x)$ for $i \geq 1$, we obtain that

$$\begin{aligned}
&\lim_{l\to\infty} \frac{1}{k_l} \sum_{j=0}^{k_l-1} H_1(a_1(T_v^{n_l+j}x), a_2(T_v^{n_l+j}x), \ldots, a_m(T_v^{n_l+j}x)) \\
&\quad = \lim_{l\to\infty} \frac{1}{k_l} \sum_{j=0}^{k_l-1} H_1(a_{n_l+j+1}(x), a_{n_l+j+2}(x), \ldots, a_{n_l+j+m}(x)) \\
&\quad = \lim_{l\to\infty} \frac{1}{k_l} \sum_{j=1}^{k_l} H_1(a_{n_l+j}(x), a_{n_l+j+1}(x), a_{n_l+j+2}(x), \ldots, a_{n_l+j+m-1}(x)) \\
&\quad = \sum_{(i_1,\ldots,i_m)\in\mathbb{N}^m} \frac{(\#(k)-1)^m}{(\#(k))^{i_1+\cdots+i_m}} H_1(i_1, \ldots, i_m)
\end{aligned}$$

as required.

The proof for (2.20) is very similar – we just apply Theorem 1.7.3 with $f(x) = H_2(b_1(x), \ldots, b_m(x))$. □

Theorem 2.7.3. *Suppose $(n_l, k_l)_{l\geq 1}$ is Stoltz then we have*

$$\lim_{l\to\infty} \frac{1}{k_l} \sum_{j=1}^{k_l} a_{n_l+j} = \frac{\#(k)}{\#(k)-1}, \tag{2.21}$$

$$\lim_{l\to\infty} \frac{1}{k_l} \sum_{j=1}^{k_l} b_{n_l+j} = \frac{\#(k)}{2} \tag{2.22}$$

almost everywhere with respect to Haar measure on $\mathcal{M}$.

Proof: We apply Theorem 1.7.3 with $f(x) = v(x)$ for the relation (2.21). So, we have that

$$\lim_{l\to\infty} \frac{1}{k_l} \sum_{j=0}^{k_l-1} v(T_v^{n_l+j}x) = \int_{\mathcal{M}} v(x)\, d\mu = \frac{\#(k)}{\#(k)-1}.$$

Since $v(T_v^{n_l+j}x) = a_{n_l+j+1}$, we obtain that

$$\lim_{l\to\infty} \frac{1}{k_l} \sum_{j=0}^{k_l-1} v(T_v^{n_l+j}x) = \lim_{l\to\infty} \frac{1}{k_l} \sum_{j=0}^{k_l-1} a_{n_l+j+1} = \lim_{l\to\infty} \frac{1}{k_l} \sum_{j=1}^{k_l} a_{n_l+j} = \frac{\#(k)}{\#(k)-1}.$$

To prove (2.22) we take $f(x) = b(x)$ and using Theorem 1.7.3 we have

$$\lim_{l\to\infty} \frac{1}{k_l} \sum_{j=0}^{k_l-1} b(T_v^{n_l+j}x) = \int_{\mathcal{M}} b(x)\, d\mu = \frac{\#(k)}{2}.$$

Since $b(T_v^{n_l+j}x) = b_{n_l+j+1}$, we get that

$$\lim_{l\to\infty} \frac{1}{k_l} \sum_{j=0}^{k_l-1} b(T_v^{n_l+j}x) = \lim_{l\to\infty} \frac{1}{k_l} \sum_{j=0}^{k_l-1} b_{n_l+j+1} = \lim_{l\to\infty} \frac{1}{k_l} \sum_{j=1}^{k_l} b_{n_l+j} = \frac{\#(k)}{2}. \qquad \square$$

Theorem 2.7.4. *For Stoltz $(n_l, k_l)_{l\geq 1}$ and for positive integers i and j, $i < j$, we have that*

$$\lim_{l\to\infty} \frac{1}{k_l} \#\{1 \leq j \leq k_l : a_{n_l+j} = i\} = \frac{\#(k)-1}{(\#(k))^i}, \tag{2.23}$$

$$\lim_{l\to\infty} \frac{1}{k_l} \#\{1 \leq j \leq k_l : a_{n_l+j} \geq i\} = \frac{1}{(\#(k))^{i-1}}, \tag{2.24}$$

$$\lim_{l\to\infty} \frac{1}{k_l} \#\{1 \leq t \leq k_l : i \leq a_{n_l+t} < j\} = \frac{1}{(\#(k))^{i-1}} \left(1 - \frac{1}{(\#(k))^{j-i}}\right), \tag{2.25}$$

almost everywhere with respect to Haar measure on $\mathcal{M}$.

Proof: For the proof of (2.23) we apply Theorem 1.7.3 with $f(x) = I_{B_1}(x)$, where I_{B_1} is the characteristic function of the set $B_1 = \{x \in \mathcal{M} : a_1(x) = i\}$.

So, we obtain that

$$\lim_{l\to\infty}\frac{1}{k_l}\sum_{j=0}^{k_l-1} I_{B_1}(T_v^{n_l+j}x) = \int_{\mathcal{M}} I_{B_1}(x)\,d\mu = \mu(\{x\in\mathcal{M} : a_1(x)=i\}) = \frac{\#(k)-1}{(\#(k))^i}.$$

Because

$$\lim_{l\to\infty}\frac{1}{k_l}\sum_{j=0}^{k_l-1} I_{B_1}(T_v^{n_l+j}x) = \lim_{l\to\infty}\frac{1}{k_l}\#\{0\le j\le k_l-1 : a_{n_l+j+1}=i\} = \lim_{l\to\infty}\frac{1}{k_l}\#\{1\le j\le k_l : a_{n_l+j}=i\},$$

we get that

$$\lim_{l\to\infty}\frac{1}{k_l}\#\{1\le j\le k_l : a_{n_l+j}=i\} = \frac{\#(k)-1}{(\#(k))^i}.$$

To prove (2.24) we apply Theorem 1.7.3 with $f(x)=I_{B_2}(x)$, where I_{B_2} is the characteristic function of the set $B_2=\{x\in\mathcal{M} : a_1(x)\ge i\}$. So, we get

$$\lim_{l\to\infty}\frac{1}{k_l}\sum_{j=0}^{k_l-1} I_{B_2}(T_v^{n_l+j}x) = \int_{\mathcal{M}} I_{B_2}(x)\,d\mu = \mu(\{x\in\mathcal{M} : a_1(x)\ge i\}) = \frac{1}{(\#(k))^{i-1}}.$$

Since

$$\lim_{l\to\infty}\frac{1}{k_l}\sum_{j=0}^{k_l-1} I_{B_2}(T_v^{n_l+j}x) = \lim_{N\to\infty}\frac{1}{k_l}\#\{0\le j\le k_l-1 : a_{n_l+j+1}\ge i\} = \lim_{l\to\infty}\frac{1}{k_l}\#\{1\le j\le k_l : a_{n_l+j}\ge i\},$$

we obtain that

$$\lim_{l\to\infty}\frac{1}{k_l}\#\{1\le j\le k_l : a_{n_l+j}\ge i\} = \frac{1}{(\#(k))^{i-1}}.$$

To prove (2.25) we apply again Theorem 1.7.3 with $f(x) = I_{B_3}(x)$, where I_{B_3} is the characteristic function of the set $B_3 = \{x \in \mathcal{M} : i \le a_1(x) < j\}$ and so we have that

$$\begin{aligned} \lim_{l\to\infty} \frac{1}{k_l} \sum_{t=0}^{k_l-1} I_{B_3}(T_v^{n_l+t}x) &= \int_{\mathcal{M}} I_{B_3}(x)\, d\mu = \mu(\{x \in \mathcal{M} : i \le a_1(x) < j\}) \\ &= \frac{1}{(\#(k))^{i-1}} \left(1 - \frac{1}{(\#(k))^{j-i}}\right). \end{aligned}$$

Because

$$\begin{aligned} \lim_{l\to\infty} \frac{1}{k_l} \sum_{t=0}^{k_l-1} I_{B_3}(T_v^{n_l+t}x) &= \lim_{l\to\infty} \frac{1}{k_l} \#\{0 \le t \le k_l - 1 : i \le a_{n_l+t+1} < j\} \\ &= \lim_{l\to\infty} \frac{1}{k_l} \#\{1 \le t \le k_l : i \le a_{n_l+t} < j\}, \end{aligned}$$

we get that

$$\lim_{l\to\infty} \frac{1}{k_l} \#\{1 \le t \le k_l : i \le a_{n_l+t} < j\} = \frac{1}{(\#(k))^{i-1}} \left(1 - \frac{1}{(\#(k))^{j-i}}\right). \qquad \square$$

Theorem 2.7.5. *For Stoltz $(n_l, k_l)_{l\ge1}$ and for $i, j \in k\backslash\{0\}$, $i < j$, we have*

$$\lim_{l\to\infty} \frac{1}{k_l} \#\{1 \le j \le k_l : b_{n_l+j} = i\} = \frac{1}{\#(k) - 1}, \tag{2.26}$$

$$\lim_{l\to\infty} \frac{1}{k_l} \#\{1 \le j \le k_l : b_{n_l+j} \ge i\} = \frac{\#(k) - i}{\#(k) - 1}, \tag{2.27}$$

$$\lim_{l\to\infty} \frac{1}{k_l} \#\{1 \le t \le k_l : i \le b_{n_l+t} < j\} = \frac{j - i}{\#(k) - 1}, \tag{2.28}$$

almost everywhere with respect to Haar measure on $\mathcal{M}$.

Proof: For the proof of (2.26) we apply Theorem 1.7.3 with $f(x) = I_{D_1}(x)$, where I_{D_1} is the characteristic function of the set $D_1 = \{x \in \mathcal{M} : b_1(x) = i\}$.

So, we obtain that

$$\begin{aligned}\lim_{l\to\infty}\frac{1}{k_l}\sum_{j=0}^{k_l-1} I_{D_1}(T_v^{n_l+j}x) &= \int_{\mathcal{M}} I_{D_1}(x)\, d\mu = \mu(\{x\in\mathcal{M} : b_1(x)=i\}) \\ &= \frac{1}{\#(k)-1}.\end{aligned}$$

Since

$$\begin{aligned}\lim_{l\to\infty}\frac{1}{k_l}\sum_{j=0}^{k_l-1} I_{D_1}(T_v^{n_l+j}x) &= \lim_{l\to\infty}\frac{1}{k_l}\#\{0\le j\le k_l-1 : b_{n_l+j+1}=i\} \\ &= \lim_{l\to\infty}\frac{1}{k_l}\#\{1\le j\le k_l : b_{n_l+j}=i\},\end{aligned}$$

we get that

$$\lim_{l\to\infty}\frac{1}{k_l}\#\{1\le j\le k_l : b_{n_l+j}=i\} = \frac{1}{\#(k)-1}.$$

To prove (2.27) we apply Theorem 1.7.3 with $f(x)=I_{D_2}(x)$, where I_{D_2} is the characteristic function of the set $D_2=\{x\in\mathcal{M} : b_1(x)\ge i\}$. So, we get

$$\begin{aligned}\lim_{l\to\infty}\frac{1}{k_l}\sum_{j=0}^{k_l-1} I_{D_2}(T_v^{n_l+j}x) &= \int_{\mathcal{M}} I_{D_2}(x)\, d\mu = \mu(\{x\in\mathcal{M} : b_1(x)\ge i\}) \\ &= \frac{\#(k)-i}{\#(k)-1}.\end{aligned}$$

Since

$$\begin{aligned}\lim_{l\to\infty}\frac{1}{k_l}\sum_{j=0}^{k_l-1} I_{D_2}(T_v^{n_l+j}x) &= \lim_{N\to\infty}\frac{1}{k_l}\#\{0\le j\le k_l-1 : b_{n_l+j+1}\ge i\} \\ &= \lim_{l\to\infty}\frac{1}{k_l}\#\{1\le j\le k_l : b_{n_l+j}\ge i\},\end{aligned}$$

we obtain that

$$\lim_{l\to\infty}\frac{1}{k_l}\#\{1\le j\le k_l : b_{n_l+j}\ge i\} = \frac{\#(k)-i}{\#(k)-1}.$$

To prove (2.28) we apply again Theorem 1.7.3 with $f(x) = I_{D_3}(x)$, where I_{D_3} is the characteristic function of the set $D_3 = \{x \in \mathcal{M} : i \le b_1(x) < j\}$ and so we have that

$$\begin{aligned}\lim_{l\to\infty} \frac{1}{k_l} \sum_{t=0}^{k_l-1} I_{D_3}(T_v^{n_l+t}x) &= \int_{\mathcal{M}} I_{D_3}(x)\, d\mu = \mu(\{x \in \mathcal{M} : i \le b_1(x) < j\}) \\ &= \frac{j-i}{\#(k)-1}.\end{aligned}$$

Since

$$\begin{aligned}\lim_{l\to\infty} \frac{1}{k_l} \sum_{t=0}^{k_l-1} I_{D_3}(T_v^{n_l+t}x) &= \lim_{l\to\infty} \frac{1}{k_l} \#\{0 \le t \le k_l - 1 : i \le b_{n_l+t+1} < j\} \\ &= \lim_{l\to\infty} \frac{1}{k_l} \#\{1 \le t \le k_l : i \le b_{n_l+t} < j\},\end{aligned}$$

we get that

$$\lim_{l\to\infty} \frac{1}{k_l} \#\{1 \le t \le k_l : i \le b_{n_l+t} < j\} = \frac{j-i}{\#(k)-1}.$$

$\square$

2.8 Special Cases

2.8.1 Schneider's p-adic Continued Fraction Map

In the case where $K = \mathbb{Q}_p$, the map T_v becomes the original Schneider's continued fraction map T_p, which motivates the previous investigation of the generalised map T_v. Recall that in the case when $K = \mathbb{Q}_p$, the valuation ring is $\mathcal{O} = \mathbb{Z}_p$, the maximal ideal in $\mathcal{O}$ is $\mathcal{M} = p\mathbb{Z}_p$, the uniformiser is $\pi = p$ and the residue field of K is $k = \mathbb{Z}_p/p\mathbb{Z}_p$. The map T_p was introduced by T. Schneider in [63] and is defined as follows. For $x \in p\mathbb{Z}_p$ define the map $T_p : p\mathbb{Z}_p \to p\mathbb{Z}_p$ by

$$T_p(x) = \frac{p^{v(x)}}{x} - \left(\frac{p^{v(x)}}{x} \bmod p\right) = \frac{p^{a(x)}}{x} - b(x) \tag{2.29}$$

where $v(x)$ is the p-adic valuation of x, $a(x) \in \mathbb{N}$ and $b(x) \in \{1, 2, \ldots, p-1\}$. Then using the continued fraction algorithm for x, we get the expansion

$$x = \cfrac{p^{a_1}}{b_1 + \cfrac{p^{a_2}}{b_2 + \cfrac{p^{a_3}}{b_3 + \ddots}}} \tag{2.30}$$

where $b_n \in \{1, 2, \ldots, p-1\}, a_n \in \mathbb{N}$ for $n = 1, 2, \ldots$.

As in the generalised case, the map T_p defined by (2.29) gives rise to an integer recurrence relationship. This is as follows. Suppose that $A_0 = 0, B_0 = 1, A_1 = p^{a_1}, B_1 = b_1$. Then set

$$A_n = p^{a_n} A_{n-2} + b_n A_{n-1} \quad \text{and} \quad B_n = p^{a_n} B_{n-2} + b_n B_{n-1} \tag{2.31}$$

for $n \geq 2$. As in the proof of Lemma 2.2.1 an inductive argument gives for $n = 1, 2, \ldots$

$$A_{n-1} B_n - A_n B_{n-1} = (-1)^n p^{a_1 + \cdots + a_n}. \tag{2.32}$$

Because p does not divide B_n we deduce that the integers A_n and B_n are coprime. The sequence of rational numbers $(\frac{A_n}{B_n})_{n=1}^{\infty}$ are the convergents to x in $p\mathbb{Z}_p$ arising from (2.30).

Since the map T_p is a special case of the map T_v, the p-adic continued fraction map T_p has all properties which were proved for T_v in previous sections. We will therefore consider the dynamical system $(p\mathbb{Z}_p, \mathcal{C}, \rho, T_p)$ where $\mathcal{C}$ is the Haar σ-algebra on $p\mathbb{Z}_p$, ρ is the Haar measure on $p\mathbb{Z}_p$ and T_p is defined by (2.29). For the Haar measure we have $\rho(pa + p^m \mathbb{Z}_p) = p^{1-m}$.

The metric theory of the p-adic continued fraction map is initiated in the paper by J. Hirsh and L. C. Washington [24] where the following was proved:

- T_p is measure-preserving with respect to ρ, i.e. $\rho(T_p^{-1}(A)) = \rho(A)$ for all $A \in \mathcal{C}$.

- T_p is ergodic, i.e. $\rho(B) = 0$ or 1 for any $B \in \mathcal{C}$ with $T_p^{-1}(B) = B$.

- The p-adic analogue of Khinchin's Theorem: For almost all $x \in p\mathbb{Z}_p$ the p-adic continued fraction expansion (2.30) satisfies

$$\lim_{n\to\infty} \frac{a_1 + a_2 + \cdots + a_n}{n} = \frac{p}{p-1}.$$

Other properties of this map are developed in the paper by J. Hančl, A. Jaššová, P. Lerchoosakul and R. Nair [21] where exactness and results about various averages and moving averages of partial quotients using subsequence pointwise and moving average ergodic theorems are proved. A summary of theorems follows. The proofs are special cases of the proofs for the generalised continued fraction map T_v and are therefore omitted. Moreover, we can consider the theorems as corollaries of the theorems in Sections 2.3, 2.6 and 2.7 where we have $k = \mathbb{Z}_p/p\mathbb{Z}_p \cong \mathbb{Z}/p\mathbb{Z}$ and so $\#(k) = p$.

Theorem 2.8.1. *The p-adic continued fraction map T_p is exact. This means that $\bigcap_{n=0}^{\infty} T_p^{-n}\mathcal{C} = \mathcal{N}$ where $\mathcal{N} = \{B \in \mathcal{C} \mid B = \emptyset$ a.e. or $B = X$ a.e.$\}$.*

Again, since $(p\mathbb{Z}_p, \mathcal{C}, \rho, T_p)$ is exact, we have other strictly weaker properties:

- T_p is strong-mixing, i.e. for all $A, B \in \mathcal{C}$ we have

  $$\lim_{n\to\infty} \rho(T_p^{-n} A \cap B) = \rho(A)\rho(B),$$

  which implies

- T_p is weak-mixing, i.e. for all $A, B \in \mathcal{C}$ we have

  $$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} |\rho(T_p^{-j} A \cap B) - \rho(A)\rho(B)| = 0,$$

  which implies

- T_p is ergodic, i.e. $\rho(B) = 0$ or 1 for any $B \in \mathcal{C}$ with $T_p^{-1}(B) = B$.

Theorem 2.8.2. *Suppose $(k_n)_{n\geq 1}$ is a L^p-good universal sequence such that $(\{k_n\gamma\})_{n\geq 1}$ is uniformly distributed modulo one for each irrational number γ.*

Suppose $F_1 : \mathbb{R}_{\geq 0} \to \mathbb{R}$ and $F_2 : \mathbb{R}_{\geq 0} \to \mathbb{R}$ are continuous increasing functions with

$$\int_{\mathcal{M}} |F_1(a_1(x))|^p \, d\rho < \infty \quad \text{and} \quad \int_{\mathcal{M}} |F_2(b_1(x))|^p \, d\rho < \infty$$

For each $n \in \mathbb{N}$, $\ell \in \{1, 2\}$ and arbitrary real numbers $d_1, \ldots, d_n$ we define

$$M_{F_\ell,n}(d_1, \ldots, d_n) = F_\ell^{-1} \left[\frac{F_\ell(d_1) + \cdots + F_\ell(d_n)}{n} \right].$$

Then we have

$$\lim_{n\to\infty} M_{F_1,n}(a_{k_1}(x), \ldots, a_{k_n}(x)) = F_1^{-1} \left[\int_{p\mathbb{Z}_p} F_1(a_1(x)) \, d\rho \right]$$

and

$$\lim_{n\to\infty} M_{F_2,n}(b_{k_1}(x), \ldots, b_{k_n}(x)) = F_2^{-1} \left[\int_{p\mathbb{Z}_p} F_2(b_1(x)) \, d\rho \right]$$

almost everywhere with respect to Haar measure on $p\mathbb{Z}_p$. □

Theorem 2.8.3. *Let $(k_n)_{n\geq 1}$ be a L^p-good universal sequence $(k_n)_{n\geq 1}$ where the sequence $(\{k_n\gamma\})_{n\geq 1}$ is uniformly distributed modulo one for each irrational number γ. Let $H_1 : \mathbb{N}^m \to \mathbb{R}$ and $H_2 : \mathbb{N}^m \to \mathbb{R}$ be functions such that*

$$\int_{p\mathbb{Z}_p} |H_1(a_1(x), \ldots, a_m(x))|^p \, d\rho < \infty \text{ and } \int_{p\mathbb{Z}_p} |H_2(b_1(x), \ldots, b_m(x))|^p \, d\rho < \infty$$

where m is a positive integer. Then we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} H_1(a_{k_n}(x), a_{k_n+1}(x), \ldots, a_{k_n+m-1}(x))$$
$$= \sum_{(i_1,\ldots,i_m)\in\mathbb{N}^m} \frac{(p-1)^m}{p^{i_1+\cdots+i_m}} H_1(i_1, \ldots, i_m)$$

and

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} H_2(b_{k_n}(x), b_{k_n+1}(x), \dots, b_{k_n+m-1}(x))$$
$$= \sum_{(i_1,\dots,i_m)\in\{1,\dots,p-1\}^m} \frac{1}{(p-1)^m} H_2(i_1,\dots,i_m)$$

almost everywhere with respect to Haar measure on $p\mathbb{Z}_p$. □

Theorem 2.8.4. *For any L^p-good universal sequence $(k_n)_{n\geq 1}$ such that the sequence $(\{k_n\gamma\})_{n\geq 1}$ is uniformly distributed modulo one for each irrational number γ we have that*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} a_{k_n} = \frac{p}{p-1},$$

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} b_{k_n} = \frac{p}{2},$$

almost everywhere with respect to Haar measure on $p\mathbb{Z}_p$. □

Note that in the case $k_n = n$ where $n = 1, 2, \dots$, the first part of this result is the p-adic analogue of Khinchin's Theorem [24]. Next, we will investigate the Gauss-Kuzmin distribution (see Section 2.1) in Schneider's p-adic continued fraction with respect to a special sequence $(k_n)_{n\geq 1}$.

Theorem 2.8.5. *Let $(k_n)_{n\geq 1}$ be a L^p-good universal sequence such that the sequence $(\{k_n\gamma\})_{n\geq 1}$ is uniformly distributed modulo one for each irrational number γ. Then for positive integers i and j, $i < j$, we have that*

$$\lim_{N\to\infty} \frac{1}{N} \#\{1 \leq n \leq N : a_{k_n} = i\} = \frac{p-1}{p^i},$$

$$\lim_{N\to\infty} \frac{1}{N} \#\{1 \leq n \leq N : a_{k_n} \geq i\} = \frac{1}{p^{i-1}},$$

$$\lim_{N\to\infty} \frac{1}{N} \#\{1 \leq n \leq N : i \leq a_{k_n} < j\} = \frac{1}{p^{i-1}} \left(1 - \frac{1}{p^{j-i}}\right);$$

almost everywhere with respect to Haar measure on $p\mathbb{Z}_p$. □

Theorem 2.8.6. *Let $(k_n)_{n\geq 1}$ be an L^p-good universal sequence such that the sequence $(\{k_n\gamma\})_{n\geq 1}$ is uniformly distributed modulo one for each irrational number γ. Then for $i,j \in \{1,\ldots,p-1\}$, $i<j$, we have that*

$$\lim_{N\to\infty}\frac{1}{N}\#\{1\leq n\leq N : b_{k_n}=i\}=\frac{1}{p-1},$$

$$\lim_{N\to\infty}\frac{1}{N}\#\{1\leq n\leq N : b_{k_n}\geq i\}=\frac{p-i}{p-1},$$

$$\lim_{N\to\infty}\frac{1}{N}\#\{1\leq n\leq N : i\leq b_{k_n}< j\}=\frac{j-i}{p-1};$$

almost everywhere with respect to Haar measure on $p\mathbb{Z}_p$. □

The moving average variants of previous theorems follow.

Theorem 2.8.7. *Suppose that $(n_l,k_l)_{l\geq 1}$ is Stoltz and suppose that we have $F_1:\mathbb{R}_{\geq 0}\to\mathbb{R}$ and $F_2:\mathbb{R}_{\geq 0}\to\mathbb{R}$ which are continuous increasing and such that*

$$\int_{p\mathbb{Z}_p}|F_1(a_1(x))|\,d\rho<\infty \quad \textit{and} \quad \int_{p\mathbb{Z}_p}|F_2(b_1(x))|\,d\rho<\infty.$$

Suppose $M_{F_\ell,n}(d_1,\ldots,d_n)$ for $\ell\in\{1,2\}$ is defined as in Theorem 2.8.2. Then

$$\lim_{l\to\infty}M_{F_1,n_l}(a_{n_l+1}(x),\ldots,a_{n_l+k_l}(x))=F_1^{-1}\left[\int_{p\mathbb{Z}_p}F_1(a_1(x))\,d\rho\right]$$

and

$$\lim_{l\to\infty}M_{F_2,n_l}(b_{n_l+1}(x),\ldots,b_{n_l+k_l}(x))=F_2^{-1}\left[\int_{p\mathbb{Z}_p}F_2(b_1(x))\,d\rho\right]$$

almost everywhere with respect to Haar measure on $p\mathbb{Z}_p$. □

Theorem 2.8.8. *Suppose that $(n_l,k_l)_{l\geq 1}$ is Stoltz and $H_1:\mathbb{N}^m\to\mathbb{R}$ and $H_2:\mathbb{N}^m\to\mathbb{R}$ are functions such that*

$$\int_{p\mathbb{Z}_p}|H_1(a_1(x),\ldots,a_m(x))|\,d\rho<\infty \quad \textit{and} \quad \int_{p\mathbb{Z}_p}|H_2(b_1(x),\ldots,b_m(x))|\,d\rho<\infty$$

where m is a positive integer. Then we have

$$\begin{aligned}\lim_{l\to\infty}\frac{1}{n_l}\sum_{j=1}^{n_l}&H_1(a_{k_l+j},a_{k_l+j+1},\ldots,a_{k_l+j+m-1})(x)\\&=\sum_{(i_1,\ldots,i_m)\in\mathbb{N}^m}\frac{(p-1)^m}{p^{i_1+\cdots+i_m}}H_1(i_1,\ldots,i_m)\end{aligned}$$

and

$$\begin{aligned}\lim_{l\to\infty}\frac{1}{n_l}\sum_{j=1}^{n_l}&H_2(b_{k_l+j},b_{k_l+j+1},\ldots,b_{k_l+j+m-1})(x)\\&=\sum_{(i_1,\ldots,i_m)\in\{1,\ldots,p-1\}^m}\frac{1}{(p-1)^m}H_2(i_1,\ldots,i_m)\end{aligned}$$

almost everywhere with respect to Haar measure on $p\mathbb{Z}_p$. □

Theorem 2.8.9. *Suppose $(n_l,k_l)_{n\geq 1}$ is Stoltz. Then we have that*

$$\lim_{l\to\infty}\frac{1}{k_l}\sum_{j=1}^{k_l}a_{n_l+j}=\frac{p}{p-1},$$

$$\lim_{l\to\infty}\frac{1}{k_l}\sum_{j=1}^{k_l}b_{n_l+j}=\frac{p}{2}$$

almost everywhere with respect to Haar measure on $p\mathbb{Z}_p$. □

Theorem 2.8.10. *For Stoltz $(n_l,k_l)_{l\geq 1}$ and positive integers i and j, $i<j$, we have that*

$$\lim_{l\to\infty}\frac{1}{k_l}\#\{1\leq j\leq k_l : a_{n_l+j}=i\}=\frac{p-1}{p^i},$$

$$\lim_{l\to\infty}\frac{1}{k_l}\#\{1\leq j\leq k_l : a_{n_l+j}\geq i\}=\frac{1}{p^{i-1}},$$

$$\lim_{l\to\infty}\frac{1}{k_l}\#\{1\leq t\leq k_l : i\leq a_{n_l+t}<j\}=\frac{1}{p^{i-1}}\left(1-\frac{1}{p^{j-i}}\right),$$

almost everywhere with respect to Haar measure on $p\mathbb{Z}_p$. □

Theorem 2.8.11. *For Stoltz $(n_l, k_l)_{l\geq 1}$ and $i, j \in \{1, \dots, p-1\}$, $i < j$, we have that*

$$\lim_{l\to\infty} \frac{1}{k_l} \#\{1 \leq j \leq k_l : b_{n_l+j} = i\} = \frac{1}{p-1},$$

$$\lim_{l\to\infty} \frac{1}{k_l} \#\{1 \leq j \leq k_l : b_{n_l+j} \geq i\} = \frac{p-i}{p-1},$$

$$\lim_{l\to\infty} \frac{1}{k_l} \#\{1 \leq t \leq k_l : i \leq b_{n_l+t} < j\} = \frac{j-i}{p-1},$$

almost everywhere with respect to Haar measure on $p\mathbb{Z}_p$. □

In Section 2.4 we also computed the measure-theoretic entropy of the map T_v. Because the map T_p is the special case we can also immediately conclude that the measure-preserving transformation $(p\mathbb{Z}_p, \mathcal{C}, \rho, T_p)$ has the measure-theoretic entropy with respect to the Haar measure ρ equal to $\frac{p}{p-1}\log p$.

Further, in Section 2.5 we showed that the map T_v has a natural extension that is Bernoulli. Hence, the natural extension of the p-adic continued fraction map T_p is isomorphic to a Bernoulli shift with the entropy $\frac{p}{p-1}\log p$. Since we know that two dynamical systems with different entropies cannot be isomorphic, this means that for each prime p the corresponding Schneider's continued fraction maps are mutually non-isomorphic. Each of them is however isomorphic to the analogue of Schneider's map on the field of formal Laurent series over a finite field of p elements.

Now, we show some examples of p-adic continued fraction expansions.

Example 2.8.12. Let $x = \frac{325}{289}$ and $p = 5$. We want to write x as a p-adic continued fraction expansion in the form (2.30).

We use the continued fraction algorithm using the map T_p (2.29). So, first we find $v(\frac{325}{289}) = v(\frac{5^2\cdot 13}{17^2}) = 2$, that is $a_1 = 2$. Now, we have

$$T_p(x) = \frac{5^2}{\frac{325}{289}} - \left(\frac{5^2}{\frac{325}{289}} \mod 5\right) = \frac{289}{13} - \left(\frac{289}{13} \mod 5\right). \tag{2.33}$$

Hence, $b_1 \equiv \frac{289}{13} \mod 5$ where $\frac{1}{13}$ is treated as the inverse of 13 modulo 5, i.e.

$(13)^{-1} \mod 5$. So, we get

$$b_1 \equiv 4 \cdot (3)^{-1} \mod 5 \equiv 4 \cdot 2 \mod 5 \equiv 3 \mod 5$$

and hence $b_1 = 3$. Getting back to (2.33), we have $T_p(x) = \frac{289}{13} - 3 = \frac{250}{13}$ which we denote x_1.

We perform another step of the algorithm by applying the map T_p again to x_1. So, we have $v(\frac{250}{13}) = v(\frac{5^3 \cdot 2}{13}) = 3 = a_2$. Now,

$$T_p(x_1) = \frac{5^3}{\frac{250}{13}} - \left(\frac{5^3}{\frac{250}{13}} \mod 5\right) = \frac{13}{2} - \left(\frac{13}{2} \mod 5\right). \tag{2.34}$$

Hence, we get

$$b_2 \equiv 13 \cdot (2)^{-1} \mod 5 \equiv 3 \cdot 3 \mod 5 \equiv 4 \mod 5$$

and so $b_2 = 4$. Getting back to (2.34), we obtain $T_p(x_1) = \frac{13}{2} - 4 = \frac{5}{2}$ which we denote x_2. Again, we apply T_p to x_2. We have $v(\frac{5}{2}) = 1 = a_3$ and

$$T_p(x_2) = \frac{5^1}{\frac{5}{2}} - \left(\frac{5^1}{\frac{5}{2}} \mod 5\right) = 2 - (2 \mod 5). \tag{2.35}$$

So, we get $b_3 \equiv 2 \mod 5$ and hence $b_3 = 2$. Getting back to (2.35), we have $T_p(x_2) = 2 - 2 = 0$ which means that the algorithm terminates and we can write the 5-adic continued fraction expansion of $x = \frac{325}{289}$ as

$$x = \cfrac{5^2}{3 + \cfrac{5^3}{4 + \cfrac{5^1}{2}}}.$$

Example 2.8.13. Let $x = \frac{234}{49}$ and $p = 3$. Again, we want to write x as a p-adic continued fraction expansion in the form (2.30).

As in the previous example, we use the continued fraction algorithm using the map T_p (2.29). So, we have $v(\frac{234}{49}) = v(\frac{2 \cdot 3^2 \cdot 13}{7^2}) = 2$, that is $a_1 = 2$. Now, we have

$$T_p(x) = \frac{3^2}{\frac{234}{49}} - \left(\frac{3^2}{\frac{234}{49}} \mod 3\right) = \frac{49}{26} - \left(\frac{49}{26} \mod 3\right). \tag{2.36}$$

Hence, $b_1 \equiv \frac{49}{26} \mod 3 \equiv 49{\cdot}(26)^{-1} \mod 3 \equiv 1{\cdot}(2)^{-1} \mod 3 \equiv 1{\cdot}2 \mod 3 \equiv 2 \mod 3$ and so $b_1 = 2$. Getting back to (2.36), we have $T_p(x) = \frac{49}{26} - 2 = -\frac{3}{26}$ which we denote x_1.

Now, we apply the map T_p again to x_1. So, we have $v(-\frac{3}{26}) = 1 = a_2$. Now,

$$T_p(x_1) = \frac{3^1}{-\frac{3}{26}} - \left(\frac{3^1}{-\frac{3}{26}} \mod 3\right) = -26 - (-26 \mod 3). \tag{2.37}$$

Hence, we obtain

$$b_2 \equiv -26 \mod 3 \equiv 1 \mod 3$$

and so $b_2 = 1$. Getting back to (2.37), we obtain that $T_p(x_1) = -26 - 1 = -27$ which we denote by x_2. Again, we apply T_p to x_2. We now have $v(-27) = v(-3^3) = 3 = a_3$ and

$$T_p(x_2) = \frac{3^3}{-27} - \left(\frac{3^3}{-27} \mod 3\right) = -1 - (-1 \mod 3). \tag{2.38}$$

Hence, we get $b_3 \equiv -1 \mod 3 \equiv 2 \mod 3$ and hence $b_3 = 2$. Getting back to (2.38), we obtain $T_p(x_2) = -1 - 2 = -3$ and denote it x_3.

Again, $v(-3) = 1 = a_4$ and

$$T_p(x_3) = \frac{3^1}{-3} - \left(\frac{3^1}{-3} \mod 3\right) = -1 - (-1 \mod 3). \tag{2.39}$$

So, $b_4 \equiv -1 \mod 3 \equiv 2 \mod 3$ and hence $b_4 = 2$. Getting back to (2.39), we obtain $T_p(x_3) = -1 - 2 = -3$. We can see that from this point the partial quotients will still be the same, so $a_i = 1$ and $b_i = 2$ for all $i \geq 4$. Finally, we can write the 3-adic continued fraction expansion of $x = \frac{234}{49}$ as

$$x = \cfrac{3^2}{2 + \cfrac{3^1}{1 + \cfrac{3^3}{2 + \cfrac{3^1}{2 + \cfrac{3^1}{2 + \ddots}}}}}.$$

One will notice that in both examples we were working with rational numbers, however in the first example the p-adic continued fraction expansion is terminating while in the second example the expansion is non-terminating. So, unlike the regular continued fraction expansion in the real case where the continued fraction expansions of rational numbers always terminate, in the p-adic case, the continued fraction expansions of rational numbers can be both terminating and non-terminating. When a negative number occurs in the p-adic continued fraction expansion, this expansion cannot terminate. Moreover, in [10] the following was proved. If the p-adic continued fraction expansion of a rational number does not terminate, then the tail of the expansion is of the form

$$p - 1 + \cfrac{p}{p - 1 + \cfrac{p}{p - 1 + \ddots}}.$$

In [24] the authors gave data indicating that rational numbers with non-terminating p-adic continued fraction expansions seem to be more common than rationals with terminating expansions.

2.8.2 Schneider's Continued Fraction Map in Positive Characteristic

We now consider Schneider's continued fraction map in the case where K is the field of formal Laurent series, i.e. $K = \mathbb{F}_q((X^{-1}))$. Recall that in this case the valuation ring is $\mathcal{O} = \mathbb{L} = \{x \in \mathbb{F}_q((X^{-1})) : |x| \leq 1\}$, the maximal ideal in $\mathcal{O}$ is $\mathcal{M} = X^{-1}\mathbb{L} = \{x \in \mathbb{F}_q((X^{-1})) : |x| < 1\}$, the uniformiser is $\pi = X^{-1}$ as $v(X^{-1}) = 1$ and the residue field of K is $k = \mathbb{L}/X^{-1}\mathbb{L} = \mathbb{F}_q$. The map T_v is then defined on $X^{-1}\mathbb{L}$. Let us denote Schneider's continued fraction map T_v on $X^{-1}\mathbb{L}$ by T_q.

For $x \in X^{-1}\mathbb{L}$ define the map $T_q : X^{-1}\mathbb{L} \to X^{-1}\mathbb{L}$ by

$$T_q(x) = \frac{X^{-v(x)}}{x} - b(x) \tag{2.40}$$

where $v(x)$ is the valuation of x and $b(x)$ denotes the residue class to which $\frac{X^{-v(x)}}{x}$ belongs in $\mathbb{F}_q\backslash\{0\}$.

This gives rise to the continued fraction expansion of $x \in X^{-1}\mathbb{L}$ in the form

$$x = \cfrac{X^{-a_1}}{b_1 + \cfrac{X^{-a_2}}{b_2 + \cfrac{X^{-a_3}}{b_3 + \ddots}}} \tag{2.41}$$

where $b_n \in \mathbb{F}_q\backslash\{0\}, a_n \in \mathbb{N}$ for $n = 1, 2, \ldots$.

The properties and results in Sections 2.3 – 2.7 introduced for the generalised Schneider's continued fraction map T_v are the same for the map T_q with $\#(k) = q$. We will not further discuss the details here.

However, we would like to give an example of a continued fraction expansion of $x \in X^{-1}\mathbb{L}$ in the form of (2.41).

Let $\mathbb{F}_q$ be a finite field with 3 elements. So, we have $\mathbb{F}_3 = \mathbb{Z}\backslash 3\mathbb{Z} = \{-1, 0, 1\}$. We want to find a continued fraction expansion (2.41) of $\alpha \in X^{-1}\mathbb{L}$ given by $X^{-1} + X^{-2}$.

We apply the continued fraction algorithm using the map (2.40). So, we have $v(X^{-1}+X^{-2}) = -(-1) = 1 = a_1$. Now,

$$T_q(\alpha) = \frac{X^{-1}}{X^{-1}+X^{-2}} - b_1$$

where b_1 is the equivalence class to which $\frac{X^{-1}}{X^{-1}+X^{-2}}$ belongs in $\{-1,1\}$. The next step is to write $\frac{X^{-1}}{X^{-1}+X^{-2}}$ in the form of formal Laurent series. We write

$$\frac{X^{-1}}{X^{-1}+X^{-2}} = a_nX^n + \cdots + a_0 + a_{-1}X^{-1} + a_{-2}X^{-2} + a_{-3}X^{-3} + \cdots .$$

So, we get

$$\begin{aligned} X^{-1} &= (a_nX^n + \cdots + a_0 + a_{-1}X^{-1} + a_{-2}X^{-2} + \cdots)(X^{-1}+X^{-2}) \\ &= a_nX^{n-1} + \cdots + a_2X + a_1 + a_0X^{-1} + a_{-1}X^{-2} + a_{-2}X^{-3} + \cdots \\ &\quad + a_nX^{n-2} + \cdots + a_3X + a_2 + a_1X^{-1} + a_0X^{-2} + a_{-1}X^{-3} + \cdots \end{aligned}$$

and we compare coefficients. Thus, we obtain that $a_i = 0$ for all $i \geq 1$ and $a_0 = 1, a_{-1} = -1, a_{-2} = 1, a_{-3} = -1, \ldots$, so $a_{-2j} = 1$ and $a_{-2j-1} = -1$ for $j \geq 0$. Hence, we have

$$\frac{X^{-1}}{X^{-1}+X^{-2}} = 1 - X^{-1} + X^{-2} - X^{-3} + X^{-4} - X^{-5} + \cdots .$$

Now, we get that $b_1 = 1$ and so

$$T_q(\alpha) = 1 - X^{-1} + X^{-2} - X^{-3} + \cdots - 1 = -X^{-1} + X^{-2} - X^{-3} + \cdots .$$

Let us denote $-X^{-1} + X^{-2} - X^{-3} + \cdots$ by α' and again apply the continued fraction algorithm using the map (2.40). We have $v(\alpha') = 1 = a_2$ and

$$T_q(\alpha') = \frac{X^{-1}}{-X^{-1} + X^{-2} - X^{-3} + \cdots} - b_2$$

where b_2 is the equivalence class to which $\frac{X^{-1}}{-X^{-1}+X^{-2}-X^{-3}+\cdots}$ belongs in $\{-1,1\}$. Similarly, as before, we can write $\frac{X^{-1}}{-X^{-1}+X^{-2}-X^{-3}+\cdots}$ in the form of formal Laurent series as

$$\frac{X^{-1}}{-X^{-1}+X^{-2}-X^{-3}+\cdots}=-1-X^{-1}.$$

So, we have that $b_2=-1$ and

$$T_q(\alpha')=-1-X^{-1}+1=-X^{-1}.$$

Let us denote $\alpha''=-X^{-1}$ and again using the continued fraction algorithm, we get

$$T_q(\alpha'')=\frac{X^{-1}}{-X^{-1}}-b_3=-1-b_3$$

where $v(\alpha'')=1=a_3$ and $b_3=-1$. Thus, $T_q(\alpha'')=0$ which concludes the algorithm. Hence, the continued fraction expansion in the form (2.41) of $X^{-1}+X^{-2}$ is

$$X^{-1}+X^{-2}=\cfrac{X^{-1}}{1+\cfrac{X^{-1}}{-1+\cfrac{X^{-1}}{-1}}}.$$

Chapter 3

Uniform Distribution of β-adic Halton Subsequences

In this chapter we will give conditions which will be used to construct variants of Halton sequences that are uniformly distributed on $[0,1)^s$.

3.1 Motivation

As suggested earlier in the introduction, the central role in Quasi-Monte Carlo integration is played by low-discrepancy sequences. One of the most famous examples of such a sequence is the so called van der Corput sequence. This is described as follows. Let $\mathbb{Z}_b$ denote b-adic integers and let $\tau : \mathbb{Z}_b \to \mathbb{Z}_b$ be defined by $\tau(x) = x + 1$. The dynamical system $(\mathbb{Z}_b, \tau)$ is uniquely ergodic. Every $x \in \mathbb{Z}_b$ has a unique expansion of the form

$$x = \sum_{n=0}^{\infty} a_n b^n$$

for an integer $b \geq 2$ and with $a_n \in \{0, 1, \dots, b-1\}$ for all $n \in \mathbb{N}_0$. For $x \in \mathbb{Z}_b$ we define a map $\varphi_b : \mathbb{Z}_b \to [0,1)$ to be

$$\varphi_b \left(\sum_{n=0}^{\infty} a_n b^n \right) = \sum_{n=0}^{\infty} a_n b^{-n-1}$$

and this map is called the *b-adic Monna map.* The sequence $(\varphi_b(n))_{n\geq 0}$ is then the *van der Corput sequence in base b.* This sequence is also uniformly distributed modulo 1 which can be shown using an isomorphism and unique ergodicity of $(\mathbb{Z}_b, \tau)$, see [18].

Given pairwise coprime integers $b_i, i = 1, \ldots, s$, all greater than 1, we define the $\mathbf{b}$-*adic Halton sequence* as

$$(\varphi_{\mathbf{b}}(n))_{n\geq 0} = (\varphi_{b_1}(n), \ldots, \varphi_{b_s}(n))_{n\geq 0}.$$

The $\mathbf{b}$-adic Halton sequence is uniformly distributed in $[0,1)^s$ and it is also a low-discrepancy sequence in $[0,1)^s$.

In recent years there has been an interest in analogues of the van der Corput and Halton sequences where the role of the numbers $\{b_1, \ldots, b_s\}$ is taken by real numbers $\{\beta_1, \ldots, \beta_s\}$ with $\beta_i > 1$ for all $i = 1, 2, \ldots, s$. A dynamical approach to the distribution of these sequences is based on the ergodic theory of the Parry-Renyi beta transformations $T_i(x) = \{\beta_i x\}$ where $i = 1, 2, \ldots, s$. In this chapter we study the distribution of subsequences of $\boldsymbol{\beta}$-adic Halton sequences, primarily using ergodic but also some other analytic methods. The results in this chapter are suggested by work of M. Hofer, M. R. Iaco and R. Tichy in [26] where they investigate the distribution behaviour of Halton sequences with special bases for the special case $k_j = j, j = 0, 1, 2, \ldots$.

3.2 Hartman Uniform Distribution and Unique Ergodicity

In this section we introduce the definition of Hartman uniform distribution of a sequence of integers and state and prove results using this concept together with some ergodic properties. This will be one of the important conditions used later in Section 3.4. Definitions and results in this section also appear in [21] and [30].

Definition 3.2.1. A sequence of integers $(a_n)_{n\geq 1}$ is *uniformly distributed on* $\mathbb{Z}$ if for each $m \in \mathbb{N} \setminus \{1\}$ and $j \in [0, m-1] \cap \mathbb{N}$ we have

$$\lim_{M\to\infty} \frac{1}{M} \#\{n : n \leq M, a_n \equiv j \mod m\} = \frac{1}{m}.$$

Definition 3.2.2. A sequence of integers $(a_n)_{n=1}^{\infty}$ is called *Hartman uniformly distributed on* $\mathbb{Z}$ if $(\{a_n\gamma\})_{n=1}^{\infty}$ is uniformly distributed modulo one for each irrational γ and $(a_n)_{n=1}^{\infty}$ is uniformly distributed on $\mathbb{Z}$.

Remark 3.2.3. Note that a sequence of integers $(a_n)_{n=1}^{\infty}$ is Hartman uniformly distributed if and only if

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} e^{2\pi i t a_n} = 0$$

for all $t \in \mathbb{R}\backslash\mathbb{Z}$. This is observed in [35] on page 296.

Remark 3.2.4. Throughout this chapter, we will assume that a sequence of integers $(a_n)_{n=1}^{\infty}$ is L^p-good universal and Hartman uniformly distributed. Under these conditions the relation

$$\overline{f}(x) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(T^{a_n - 1}x)$$

is T-invariant where T is an ergodic transformation. That is,

$$\overline{f}(Tx) = \overline{f}(x) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(T^{a_n - 1}x) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(T^{a_n}x)$$

almost everywhere with respect to a corresponding measure. This follows from the proof of Theorem 3.2.6.

Let us consider the space $L^2(X, \mathcal{B}, \mu)$ of complex-valued square-integrable functions where $(X, \mathcal{B}, \mu)$ is a measure space, i.e. the set of all measurable functions $f : X \to \mathbb{C}$ such that

$$\|f\|_2 = \left(\int_X |f|^2 \, d\mu\right)^{\frac{1}{2}} < \infty.$$

It is know that this space is a Hilbert space with respect to the inner product

$$\langle f, g\rangle = \int_X f\overline{g}\, d\mu.$$

Notice that for any $f \in L^2(X, \mathcal{B}, \mu)$ we have

$$\langle f, f\rangle = \int_X f\overline{f}\, d\mu = \int_X |f|^2 \; d\mu = \|f\|_2^2 .$$

Any measurable transformation T of a measure space $(X, \mathcal{B}, \mu)$ generates a linear operator U on $L^2(X, \mathcal{B}, \mu)$ by defining $Uf(x) = f(Tx)$. Sometimes, operators can be useful since ergodic-theoretic properties of T can often be related with spectral properties of U.

Proposition 3.2.5. *The transformation T is measure-preserving if and only if the operator U is an isometry, i.e. $\langle Uf, Ug\rangle = \langle f, g\rangle$ for all f, g in $L^2(X, \mathcal{B}, \mu)$. Furthermore, if T is invertible and measure-preserving then U is unitary.*

Proof: The proof can be found in [68] on page 25. □

Let T be an invertible, measure-preserving transformation of the measure space $(X, \mathcal{B}, \mu)$. For any element $f \in L^2(X, \mathcal{B}, \mu)$ one can consider the sequence $(\langle U^n f, f\rangle)_{n\in\mathbb{Z}}$. Now, let us recall that any sequence $(c_n)_{n\in\mathbb{Z}}$ is called *positive definite* if, given a bi-sequence of complex numbers $(z_n)_{n\in\mathbb{Z}}$, only finitely many of whose terms are non-zero, we have $\sum_{n,m\in\mathbb{Z}} c_{n-m} z_n \overline{z_m} \geq 0$ where $\overline{z}$ is the conjugate of the complex number z. The sequence is $(\langle U^n f, f\rangle)_{n\in\mathbb{Z}}$ is positive definite since U is an isometry and so for any family of complex numbers $\psi_0, \psi_1, \ldots, \psi_m$ we have

$$0 \leq \left\| \sum_{k=0}^{m} \psi_k U^k f \right\|_2^2 = \left(\sum_{i=0}^{m} \psi_i U^i f, \sum_{j=0}^{m} \psi_j U^j f \right) = \sum_{i,j=0}^{m} \left\langle U^{i-j} f, f \right\rangle \psi_i \overline{\psi}_j. \quad (3.1)$$

Then by the Bochner-Herglotz Theorem we have that there is a finite measure ω_f on the unit circle $\mathbb{T}$ such that

$$\langle U^n f, f\rangle = \int_{\mathbb{T}} z^n \, d\omega_f(z) \quad (3.2)$$

with $n \in \mathbb{Z}$.

In the case when T is non-invertible, the sequence $(\langle U^n f, f\rangle)_{n\in\mathbb{Z}}$ is defined only for $n \geq 0$. We put $(\langle U^n f, f\rangle)_n = (\overline{\langle U^n f, f\rangle})_{-n}$ when n is negative. Then the relations (3.1) and (3.2) remain valid.

For Hartman uniform distributed sequences of integers, it is possible to prove a version of Theorem 1.6.3 using only ergodicity. The following theorem shows that for Hartman uniformly distributed sequences, we need far more elementary mixing information about a transformation than in the general setting. See also [21], [30] and [45].

Theorem 3.2.6. *Suppose that $(X, \mathcal{B}, \mu, T)$ is an ergodic dynamical system. Further suppose that $(a_n)_{n=1}^{\infty}$ is Hartman uniformly distributed on $\mathbb{Z}$ and also L^p-good universal for $p \in (1, 2]$. Then $\overline{f}(x)$ exists and $\overline{f}(x) = \int_X f d\mu$ almost everywhere with respect to μ.*

To prove this theorem we need the following lemma which is a special case of a theorem due to S. Sawyer [62].

Lemma 3.2.7. *Let $(X, \mathcal{B}, \mu, T)$ be a dynamical system and $f \in L^2(X, \mathcal{B}, \mu)$ where $||f||_2 = (\int_X |f|^2 d\mu)^{\frac{1}{2}}$. Set*

$$Mf(x) = \sup_{N\geq 1} \left| \frac{1}{N} \sum_{n=1}^{N} f(T^{a_n} x) \right|. \qquad (N = 1, 2, \dots)$$

If $(a_n)_{n\geq 1}$ is L^p-good universal for $p > 1$, then there exists $C > 0$ such that $||Mf||_2 \leq C||f||_2$.

Proof of Theorem 3.2.6: Let $f \in L^2(X, \mathcal{B}, \mu)$. Since by Lemma 3.2.7 we have that

$$\left| \frac{1}{N} \sum_{n=1}^{N} f(T^{a_n} x) \right| \leq Mf(x) \qquad (N = 1, 2, \dots)$$

and $Mf^2 \in L^1$, the Dominated Convergence Theorem implies

$$g(x) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(T^{a_n} x)$$

exists in L^2 norm. Our next step is to show that $g(Tx) = g(x)$. Let U be an operator associated to the transformation T, that is $Uf(x) = f(Tx)$. Since T is measure-preserving, U is an isometric operator on L^2 and hence the relations (3.1) and (3.2) hold. So, we obtain that

$$
\begin{aligned}
&\left\|\frac{1}{N}\sum_{n=1}^{N} f(T^{a_n+1}x) - \frac{1}{N}\sum_{n=1}^{N} f(T^{a_n}x)\right\|_2^2 \\
&= \int \left|\frac{1}{N}\sum_{n=1}^{N} f(T^{a_n+1}x) - \frac{1}{N}\sum_{n=1}^{N} f(T^{a_n}x)\right|^2 d\mu \\
&= \int \left(\frac{1}{N}\sum_{n=1}^{N} f(T^{a_n+1}x) - \frac{1}{N}\sum_{n=1}^{N} f(T^{a_n}x)\right) \\
&\qquad\qquad \cdot \overline{\left(\frac{1}{N}\sum_{n=1}^{N} f(T^{a_n+1}x) - \frac{1}{N}\sum_{n=1}^{N} f(T^{a_n}x)\right)}\, d\mu \\
&= \int \left(\frac{1}{N}\sum_{n=1}^{N} f(T^{a_n+1}x) - \frac{1}{N}\sum_{n=1}^{N} f(T^{a_n}x)\right) \\
&\qquad\qquad \cdot \left(\frac{1}{N}\sum_{n=1}^{N} \overline{f}(T^{a_n+1}x) - \frac{1}{N}\sum_{n=1}^{N} \overline{f}(T^{a_n}x)\right) d\mu \\
&= \int \Bigg(\frac{1}{N}\sum_{n=1}^{N} f(T^{a_n+1}x)\cdot\frac{1}{N}\sum_{n=1}^{N} \overline{f}(T^{a_n+1}x) - \frac{1}{N}\sum_{n=1}^{N} f(T^{a_n+1}x)\cdot\frac{1}{N}\sum_{n=1}^{N} \overline{f}(T^{a_n}x) \\
&\quad -\frac{1}{N}\sum_{n=1}^{N} f(T^{a_n}x)\cdot\frac{1}{N}\sum_{n=1}^{N} \overline{f}(T^{a_n+1}x) + \frac{1}{N}\sum_{n=1}^{N} f(T^{a_n}x)\cdot\frac{1}{N}\sum_{n=1}^{N} \overline{f}(T^{a_n}x)\Bigg) d\mu \\
&= \frac{1}{N^2}\left\langle \sum_{n=1}^{N} U^{a_n+1}f, \sum_{m=1}^{N} U^{a_m+1}f \right\rangle - \frac{1}{N^2}\left\langle \sum_{n=1}^{N} U^{a_n+1}f, \sum_{m=1}^{N} U^{a_m}f \right\rangle \\
&\qquad - \frac{1}{N^2}\left\langle \sum_{n=1}^{N} U^{a_n}f, \sum_{m=1}^{N} U^{a_m+1}f \right\rangle + \frac{1}{N^2}\left\langle \sum_{n=1}^{N} U^{a_n}f, \sum_{m=1}^{N} U^{a_m}f \right\rangle.
\end{aligned}
$$

Now using (3.1), we get that

$$= \frac{1}{N^2} \sum_{1 \le n,m \le N} \left\langle U^{a_n - a_m} f, f \right\rangle - \frac{1}{N^2} \sum_{1 \le n,m \le N} \left\langle U^{a_n + 1 - a_m} f, f \right\rangle$$
$$- \frac{1}{N^2} \sum_{1 \le n,m \le N} \left\langle U^{a_n - a_m - 1} f, f \right\rangle + \frac{1}{N^2} \sum_{1 \le n,m \le N} \left\langle U^{a_n - a_m} f, f \right\rangle .$$

Using (3.2), we obtain that

$$= 2 \cdot \frac{1}{N^2} \sum_{1 \le n,m \le N} \int_{\mathbb{T}} z^{a_n - a_m} \, d\omega_f(z) - \frac{1}{N^2} \sum_{1 \le n,m \le N} \int_{\mathbb{T}} z^{a_n - a_m + 1} \, d\omega_f(z)$$
$$- \frac{1}{N^2} \sum_{1 \le n,m \le N} \int_{\mathbb{T}} z^{a_n - a_m - 1} \, d\omega_f(z)$$
$$= \int_{\mathbb{T}} 2 \left(\frac{1}{N} \sum_{n=1}^{N} z^{a_n} \cdot \frac{1}{N} \sum_{m=1}^{N} z^{-a_m} \right) \, d\omega_f(z)$$
$$- \int_{\mathbb{T}} z \left(\frac{1}{N} \sum_{n=1}^{N} z^{a_n} \cdot \frac{1}{N} \sum_{m=1}^{N} z^{-a_m} \right) \, d\omega_f(z)$$
$$- \int_{\mathbb{T}} z^{-1} \left(\frac{1}{N} \sum_{n=1}^{N} z^{a_n} \cdot \frac{1}{N} \sum_{m=1}^{N} z^{-a_m} \right) \, d\omega_f(z)$$
$$= \int_{\mathbb{T}} (2 - z - z^{-1}) \left| \frac{1}{N} \sum_{n=1}^{N} z^{a_n} \right|^2 \, d\omega_f(z).$$

Now, we use the parametrization $z = e^{2\pi i \theta}$ for $\theta \in [0, 1)$ and we obtain that

$$= 4 \int_{\mathbb{T}} \sin^2 (\theta\pi) \left| \frac{1}{N} \sum_{n=1}^{N} z^{a_n} \right|^2 d\omega_f(z).$$

Using the fact that $\sin(\theta\pi) = 0$ if $\theta = 0$, we conclude that the inner integrand is zero. For $\theta \neq 0$ we use the fact that $(a_n)_{n \ge 1}$ is Hartman uniformly distributed and hence, by Remark 3.2.3, the integrand tends to zero as $N \to \infty$. Thus, we see that $g(Tx) = g(x)$ almost everywhere. We also have that $\int_X g(x) d\mu = \int_X f(x) d\mu$. The same observation extends to L^p since L^2 is dense in L^p for $p \in (1, 2]$.

By Theorem 1.2.6 we know that if T is ergodic and $g(Tx) = g(x)$ almost everywhere, then $g(x)$ is constant almost everywhere, say C_f.

Now we want to show that the pointwise limit is the same as the norm limit, i.e. that $\overline{f}(x) = g(x) = C_f$. We consider a sequence of natural numbers $(N_t)_{t\geq 1}$ such that

$$\left\| \frac{1}{N_t} \sum_{n=1}^{N_t} f(T^{a_n}x) - C_f \right\|_p \leq \frac{1}{t}.$$

Thus

$$\sum_{t=1}^{\infty} \int_X \left| \frac{1}{N_t} \sum_{n=1}^{N_t} f(T^{a_n}x) - C_f \right|^p d\mu < \infty.$$

Fatou's Lemma tells us that

$$\int_X \left(\sum_{t=1}^{\infty} \left| \frac{1}{N_t} \sum_{n=1}^{N_t} f(T^{a_n}x) - C_f \right|^p \right) d\mu < \infty,$$

which implies that

$$\sum_{t=1}^{\infty} \left| \frac{1}{N_t} \sum_{n=1}^{N_t} f(T^{a_n}x) - C_f \right|^p < \infty$$

almost everywhere. This means that

$$\lim_{t\to\infty} \left| \frac{1}{N_t} \sum_{n=1}^{N_t} f(T^{a_n}x) - C_f \right|^p = 0$$

μ almost everywhere and thus we obtain that

$$\lim_{t\to\infty} \frac{1}{N_t} \sum_{n=1}^{N_t} f(T^{a_n}x) = C_f$$

almost everywhere. As $(a_n)_{n\geq 1}$ is L^p-good universal, we must have $\overline{f}(x) = C_f$ almost everywhere with respect to μ.

Now suppose that $0 < M < \infty$ and $|f| \leq M$ almost everywhere. Then using the Bounded Convergence Theorem, we get that

$$\lim_{N\to\infty}\int_X\left(\frac{1}{N}\sum_{n=1}^{N}f(T^{a_n}x)\right)d\mu=\int_X C_f\;d\mu=C_f.$$

Since we have that

$$\int_X\left(\frac{1}{N}\sum_{n=1}^{N}f(T^{a_n}x)\right)d\mu=\int_X f(x)\;d\mu,$$

we conclude that $C_f=\int_X f d\mu$. For a general function f, set $f_M=\min(f,M)$. Then $\lim_{M\to\infty}f_M=f$ almost everywhere. Evidently, we also have that $C_{f_M}=\int_X f_M d\mu$. Letting $M\to\infty$, we get that $C_f=\int_X f d\mu$. Thus we proved that $\overline{f}(x)=\int_X f d\mu$ almost everywhere with respect to μ. □

We have the following theorem which is a stronger version of Theorem 3.2.6 and restricted to the case $k_n=n$ $(n=0,1,2,\ldots)$, it reduces to Theorem 1.8.7. See also [30] and the partial proof can be also found in [45].

Theorem 3.2.8. *Suppose $(k_n)_{n\geq 0}$ is Hartman uniformly distributed and L^2-good universal. Let T be a continuous map of a compact metrisable space X. Also let μ denote a measure defined on a σ-algebra $\mathcal{B}$ of subsets of X. The following statements are equivalent:*

1. *the transformation $(X,\mathcal{B},\mu,T)$ is uniquely ergodic;*

2. *for each function f in $C(X)$ (the space of continuous functions on X), there is a constant C_f independent of x such that*

   $$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}f(T^{k_n}x)=C_f$$

   pointwise on X;

3. *for each function f in $C(X)$, there is a constant C_f independent of x such that*

   $$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}f(T^{k_n}x)=C_f,$$

uniformly on X;

4. *there is $\mu \in M(X,T)$ such that for all f in $C(X)$ and all $x \in X$,*

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

Proof: It is evident that 3) implies 2). We next consider the proof that 4) implies 1). Let

$$S_N f(x) = \frac{1}{N} \sum_{n=0}^{N-1} f(T^{k_n}x). \qquad (N = 1, 2, \ldots)$$

Suppose that μ and ν are in $M(X,T)$. By the assumption we get that

$$\int_X \lim_{N\to\infty} S_N f(x)\, d\nu = \int_X \left(\int_X f\, d\mu \right) d\nu = \int_X f\, d\mu$$

and by the Dominated Convergence Theorem we have

$$\begin{aligned} \int_X \lim_{N\to\infty} S_N f(x)\, d\nu &= \lim_{N\to\infty} \int_X S_N f(x)\, d\nu(x) = \lim_{N\to\infty} \frac{1}{N} \left(N \int_X f\, d\nu \right) \\ &= \int_X f\, d\nu = \int_X f\, d\mu. \end{aligned}$$

This holds for all f in $C(X)$ and hence by the Riesz Representation Theorem 1.8.1 we have $\nu = \mu$, as required. We now prove 2) implies 4). Set

$$k(f) = \lim_{N\to\infty} \frac{1}{N} \sum_{j=0}^{N-1} f(T^{k_j}x).$$

Observe that k is a linear operator and is continuous since

$$\left| \frac{1}{N} \sum_{j=0}^{N-1} f(T^{k_j}x) \right| \le \frac{1}{N} \sum_{j=0}^{N-1} \left| f(T^{k_j}x) \right| \le \frac{1}{N} \cdot N \sup_{x\in X} |f(x)| = \|f\|.$$

Also as $k(1) = 1$ and $k(f) \ge 0$ if $f \ge 0$ we have $k(f) \ge 0$. Thus by the

Riesz Representation Theorem 1.8.1 $k(f) = \int_X f d\mu$ with respect to a Borel probability measure μ. Also note $k(f \circ T) = k(f)$ which one can see from the proof of Theorem 3.2.6. So $\int_X f \circ T \, d\mu = \int_X f \, d\mu$. Thus $\mu \in M(X, T)$.

We now show how 1) implies 3). Let us assume that $M(X, T) = \{\mu\}$. Then by the Dominated Convergence Theorem we have that $C_f = \int_X f \, d\mu$ because

$$\begin{aligned} C_f &= \int_X C_f \, d\mu = \int_X \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f(T^{k_n}x) \, d\mu = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \int_X f(T^{k_n}x) \, d\mu \\ &= \int_X f \, d\mu. \end{aligned}$$

Now suppose 3) does not hold. Then there exists an $\epsilon > 0$, a function g in $C(X)$ and a sequence $(x_n)_{n=1}^{\infty}$ in X such that

$$|S_n g(x_n) - \int_X g \, d\mu| \geq \epsilon.$$

Let

$$\mu_n = \frac{1}{n} \sum_{i=0}^{n-1} \delta_{T^{k_i}(x_n)} = \frac{1}{n} \sum_{i=0}^{n-1} T_*^{k_i} \delta_{x_n}, \qquad (n = 1, 2, \ldots)$$

where δ_y denotes the Dirac measure at y. Using Lemma 1.8.2 and the fact that $\int_X f \, d\delta_x = f(x)$ we get

$$\begin{aligned} \int_X g \, d\mu_n &= \int_X g \, d\left(\frac{1}{n} \sum_{i=0}^{n-1} T_*^{k_i} \delta_{x_n}\right) \\ &= \frac{1}{n} \int_X g \, d(T_*^{k_0} \delta_{x_n}) + \frac{1}{n} \int_X g \, d(T_*^{k_1} \delta_{x_n}) + \cdots + \frac{1}{n} \int_X g \, d(T_*^{k_{n-1}} \delta_{x_n}) \\ &= \frac{1}{n} \int_X g \circ T^{k_0} \, d\delta_{x_n} + \frac{1}{n} \int_X g \circ T^{k_1} \, d\delta_{x_n} + \cdots + \frac{1}{n} \int_X g \circ T^{k_{n-1}} \, d\delta_{x_n} \\ &= \frac{1}{n} \sum_{i=0}^{n-1} g(T^{k_i} x_n) = S_n g(x_n). \end{aligned}$$

So, we have

$$\left| \int_X g \, d\mu_n - \int_X g \, d\mu \right| \geq \epsilon.$$

As the set of measures $M(X)$ is compact, we can choose a subsequence (μ_{n_j}) convergent to μ_∞. The next step is to prove that μ_∞ is T-invariant which will lead to conclusion that $\mu_\infty \neq \mu$ and so 1) would imply 3) as required. We have

$$\begin{aligned}
\int_X g\, d\mu_\infty - \int_X g \circ T\, d\mu_\infty &= \int_X (g - g\circ T)\, d\mu_\infty \\
&= \lim_{j\to\infty} \int_X (g - g(T))\, d\mu_{n_j} \\
&= \lim_{j\to\infty} \int_X (g - g(T))\, d\left(\frac{1}{n_j}\sum_{i=0}^{n_j-1} T_*^{k_i}\delta_{x_{n_j}}\right), \\
&= \lim_{j\to\infty} \int_X \left(\frac{1}{n_j}\sum_{i=0}^{n_j-1} g(T^{k_i}x_{n_j}) - \frac{1}{n_j}\sum_{i=0}^{n_j-1} g(T^{k_i+1}x_{n_j})\right)\, d\delta_{x_{n_j}}.
\end{aligned}$$

This means that

$$\left|\int_X (g - g(T))\, d\mu_\infty\right| \leq \lim_{j\to\infty} \int_X \left|\frac{1}{n_j}\sum_{i=0}^{n_j-1} g(T^{k_i}x_{n_j}) - \frac{1}{n_j}\sum_{i=0}^{n_j-1} g(T^{k_i+1}x_{n_j})\right| d\delta_{x_{n_j}}.$$

Since $\int_X f\, d\delta_x = f(x)$, we get that

$$\left|\int_X (g - g(T))\, d\mu_\infty\right| \leq \lim_{j\to\infty} \left|\frac{1}{n_j}\sum_{i=0}^{n_j-1} g(T^{k_i}x_{n_j}) - \frac{1}{n_j}\sum_{i=0}^{n_j-1} g(T^{k_i+1}x_{n_j})\right|.$$

Integrating both sides of this inequality with respect to μ and noting the left hand side is a constant, we have that

$$\left|\int_X (g - g(T))\, d\mu_\infty\right| \leq \int_X \left(\lim_{j\to\infty} \left|\frac{1}{n_j}\sum_{i=0}^{n_j-1} g(T^{k_i}x_{n_j}) - \frac{1}{n_j}\sum_{i=0}^{n_j-1} g(T^{k_i+1}x_{n_j})\right|\right) d\mu.$$

Using the Dominated Convergence Theorem, this gives

$$\left|\int_X (g - g(T))\, d\mu_\infty\right| \leq \lim_{j\to\infty} \int_X \left|\frac{1}{n_j}\sum_{i=0}^{n_j-1} g(T^{k_i}x_{n_j}) - \frac{1}{n_j}\sum_{i=0}^{n_j-1} g(T^{k_i+1}x_{n_j})\right| d\mu.$$

Using the Cauchy-Schwarz Inequality (for L^2-functions), we get

$$\left|\int_X (g - g(T))\, d\mu_\infty\right| \leq \lim_{j\to\infty} \left\| \frac{1}{n_j}\sum_{i=0}^{n_j-1} g(T^{k_i}x_{n_j}) - \frac{1}{n_j}\sum_{i=0}^{n_j-1} g(T^{k_i+1}x_{n_j}) \right\|_2 .$$

Using the Bochner-Herglotz Theorem, there is a spectral measure ω_g attached to the function g and implicitly to the map T. Performing the same calculation as in the proof of Theorem 3.2.6, we obtain

$$\left|\int_X (g - g(T))\, d\mu_\infty\right| \leq \lim_{j\to\infty} \left(\int_{\mathbb{T}} (2 - z - z^2) \left| \frac{1}{n_j}\sum_{i=0}^{n_j-1} z^{k_i} \right|^2 d\omega_g \right)^{\frac{1}{2}} .$$

If we write $z = e^{2\pi i\theta}$ on $\mathbb{T}$, then this is

$$\left|\int_X (g - g(T))\, d\mu_\infty\right| \leq 4 \lim_{j\to\infty} \left(\int_{\mathbb{T}} \sin^2(\theta\pi) \left| \frac{1}{n_j}\sum_{i=0}^{n_j-1} z^{k_i} \right|^2 d\omega_g \right)^{\frac{1}{2}} .$$

When $\theta = 0$, the inner integrand is zero. When $\theta \neq 0$, we use the assumption that $(k_n)_{n\geq 0}$ is Hartman uniformly distributed and so, by Remark 3.2.3, the integrand tends to zero as $j \to \infty$. This implies that $\int_X g\, d\mu_\infty = \int_X g \circ T\, d\mu_\infty$ and thus μ_∞ is in $M(X, T)$. Because we have that $\left|\int_X g\, d\mu_\infty - \int_X g\, d\mu\right| \geq \epsilon$ and so, $\mu_\infty \neq \mu$. This contradicts the unique ergodicity of T. Thus 1) implies 3) as required. □

Remark 3.2.9. We can see from Theorems 1.5.2 and 3.2.8 that if a dynamical system $(X, \mathcal{B}, T, \mu)$ is uniquely ergodic and if $(k_n)_{n\in\mathbb{N}}$ is Hartman uniformly distributed and L^2-good universal, then the sequence $(x_{k_n})_{n\in\mathbb{N}} = (T^{k_n}x)_{n\in\mathbb{N}}$ is uniformly distributed modulo 1.

3.3 Examples of Hartman Uniformly Distributed Sequences

In this section, a list of constructions of Hartman uniformly distributed sequences is given. The first five are also examples of L^p-good universal sequences for some $p \geq 1$. The examples 6–13 also appear in [43]. The list of examples can be also found in [21] and [30].

1. *The positive integers:* The sequence $(n)_{n=1}^{\infty}$ is L^1-good universal and Hartman uniformly distributed.

2. *Sequences satisfying condition H:* Let $g : [1, \infty) \to [1, \infty)$ be a differentiable function whose derivative increases with its argument. Set $k_n = [g(n)]$ $(n = 1, 2, \ldots)$ where $[y]$ denotes the integer part of a real number y. Let A_M denote the cardinality of the set $\{n : k_n \leq M\}$. Suppose for some function $a : [1, \infty) \to [1, \infty)$ increasing to infinity as its argument does, that we set

$$b(M) = \sup_{\{z\} \in \left[\frac{1}{a(M)}, \frac{1}{2}\right)} \left| \sum_{n:\, k_n \leq M} e^{2\pi i z k_n} \right|.$$

For some decreasing function $c : [1, \infty) \to [1, \infty)$ and some positive constant $K > 0$, suppose that

$$\frac{b(M) + A_{[a(M)]} + \frac{M}{a(M)}}{A_M} \leq K c(M).$$

Then if we have

$$\sum_{s=1}^{\infty} c(\theta^s) < \infty$$

for $\theta > 1$, we say that $(k_n)_{n=1}^{\infty}$ satisfies condition H [44].

Specific examples of sequences of integers satisfying condition H are $k_n = [g(n)]$ $(n = 1, 2, \ldots)$ where

I. $g(n) = n^{\omega}$ if $\omega > 1$ and $\omega \notin \mathbb{N}$.

II. $g(n) = e^{\log^{\gamma} n}$ for $\gamma \in (1, \frac{3}{2})$.

III. $g(n) = P(n) = b_k n^k + \ldots + b_1 n + b_0$ for $b_k, \ldots, b_1$ not all rational multiplies of the same real number.

IV. *Hardy fields:* Hardy field is a closed field (under differentiation) of germs at $+\infty$ of continuous real-valued functions with addition and multiplication taken to be pointwise. Let $\mathcal{H}$ denote the union of all Hardy fields. Let $a \in \mathcal{H}$ and suppose that for some $k \in \mathbb{Z}$, $k \geq 2$ we have that

$$\lim_{x\to\infty} \frac{a(x)}{x^{k-1}} = \infty \quad \text{and} \quad \lim_{x\to\infty} \frac{a(x)}{x^k} = 0;$$

then $(k_n)_{n=1}^{\infty} = ([a(n)])_{n=1}^{\infty}$ satisfies condition H. This example is observed in [6].

3. *A random example:* Suppose $S = (c_n)_{n=1}^{\infty} \subseteq \mathbb{N}$ is a strictly increasing sequence of natural numbers. By identifying S with its characteristic function I_S, we may view it as a point in $\Lambda = \{0,1\}^{\mathbb{N}}$, the set of maps from $\mathbb{N}$ to $\{0,1\}$. We endow Λ with a probability measure by viewing it as a Cartesian product $\Lambda = \prod_{n=1}^{\infty} X_n$ where for each natural number n we have $X_n = \{0,1\}$. We specify the probability π_n on X_n by $\pi_n(\{1\}) = q_n$ with $0 \leq q_n \leq 1$ and $\pi_n(\{0\}) = 1 - q_n$ such that $\lim_{n\to\infty} q_n n = \infty$. The desired probability measure on Λ is the corresponding product measure $\pi = \prod_{n=1}^{\infty} \pi_n$. The underlying σ-algebra $\mathcal{A}$ is generated by the 'cylinders'

$$\{\lambda = (\lambda_n)_{n=1}^{\infty} \in \Lambda : \lambda_{i_1} = \alpha_{i_1}, \ldots, \lambda_{i_r} = \alpha_{i_r}\}$$

for all possible choices of $i_1, \ldots, i_r$ and $\alpha_{i_1}, \ldots, \alpha_{i_r}$. Then if $(k_n)_{n=1}^{\infty}$ is almost every point in Λ with respect to the measure π, it is Hartman uniformly distributed [7].

4. *Block sequences:* These are sequences of the form $(k_n)_{n\geq 1} = \bigcup_{n=1}^{\infty}[d_n, e_n]$ ordered by absolute value for disjoint $([d_n, e_n])_{n\geq 1}$ with $d_{n-1} = O(e_n)$ as n tends to infinity. Note that this allows the possibility that $(k_n)_{n\geq 1}$ is zero density. This example is an immediate consequence of Templeman's semigroup ergodic theorem [67].

5. *Random perturbation of good sequences:* Let $(k_n)_{n\geq 1}$ be an L^p-good universal sequence of integers that is also Hartman uniformly distributed. Let us suppose that $\theta = \{\theta_n, n \geq 1\}$ denotes a sequence of $\mathbb{N}$-valued independent, identically distributed random variables with basic probability space $(\Omega, \mathcal{A}, \mathcal{P})$, and a $\mathcal{P}$-complete σ-algebra $\mathcal{A}$. Assume that there exist $0 < \alpha < 1$ and $B > 1/\alpha$ such that

$$k_n = O(e^{n^\alpha}) \quad \text{and} \quad \mathbb{E} \log_+^B |\theta_1| < \infty.$$

Here $\mathbb{E}$ denotes expectation with respect to the probability space $(\Omega, \mathcal{A}, \mathcal{P})$. Then $(k_n + \theta_n(\omega))_{n\geq 1}$ is L^p-good universal and Hartman uniformly distributed, see [46].

6. Sequences $k_n = [P(n)]$ $(n = 1, 2, \ldots)$ where $P(x) = a_k x^k + \cdots + a_1 x + a_0$ such that the numbers $a_k, \ldots, a_1$ are not all rational multiples of the same real number, see [43].

7. Sequences $k_n = [P(p_n)]$ $(n = 1, 2, \ldots)$ where $(p_n)_{n=1}^\infty$ denotes the sequence of rational primes and $P(z)$ is as in 6), see [43].

8. Sequences $k_n = [f(n)]$ $(n = 1, 2, \ldots)$ where $f(z)$ denotes a non-polynomial entire function which is real on the real numbers and such that $|f(z)| \ll e^{(\log z)^\alpha}$ with $\alpha < \frac{4}{3}$, see [43].

9. Sequences $k_n = [f(p_n)]$ $(n = 1, 2, \ldots)$ where $f(z)$ is as in 8) and p_n denotes the n^{th} rational prime, see [43].

10. Sequences $k_n = [a_n \cos(a_n x)]$ $(n = 1, 2, \ldots)$ for a strictly increasing sequence of integers $(a_n)_{n=1}^\infty$ and almost all x with respect to Lebesgue measure, see [43].

11. Sequences $k_n = [a_n \cos(a_n x)]$ $(n = 1, 2, \ldots)$ for a strictly increasing sequence of integers $(a_n)_{n=1}^\infty$ such that $a_n \ll n^p$ and $p > 1$ and all x outside a set of Hausdorff dimension not greater than $1 - \frac{1}{4p+\frac{1}{2}}$, see [43].

12. Sequences $k_n = [g_n(x)]$ $(n = 1, 2, \ldots)$ for almost all x with respect to Lebesgue measure in $[a, b]$ where $(g_n(x))_{n=1}^\infty$ is a sequence of continuously differentiable functions defined on $[a, b]$ satisfying the following hypothesis. For each pair of distinct natural numbers m and n we have:

(a) $g_n'(x) - g_m'(x)$ is monotonic on $[a, b]$,

(b) there is an absolute constant λ such that $|g'_n(x) - g'_m(x)| \geq \lambda > 0$. See [43].

13. Sequences $k_n = [g_n(x)]$ $(n = 1, 2, \ldots)$ for all x lying outside a set of Hausdorff dimension at most $1 - \frac{1}{p}$ in $[a, b]$ where $(g_n(x))_{n=1}^{\infty}$ is a sequence of continuously differentiable functions defined on $[a, b]$ which satisfies the hypothesis (a), (b) of 12) and additional two conditions:

(c) for all x in $[a, b]$ we have

$$\sup_{x \in [a,b]} |g'_n(x)| \ll n^p$$

for some $p > 1$ and with an implied constant independent of x,

(d) for each pair of distinct positive integers m and n the function

$$\frac{g'_n(x) g'_m(x)}{g'_m(x) - g'_n(x)}$$

is monotonic on $[a, b]$. See [43].

3.4 Uniform Distribution of $\boldsymbol{\beta}$-adic Halton Subsequences

To introduce the concept of the $\boldsymbol{\beta}$-adic Halton sequence, first we need to set the relevant background which can be also found in [19] and [26].

Let $(G_n)_{n \geq 0}$ be an increasing sequence of positive integers with $G_0 = 1$. Then every non-negative integer n can be written as

$$n = \sum_{k=0}^{\infty} g_k(n) G_k, \tag{3.3}$$

where $g_k(n) \in \{0, \ldots, [G_{k+1}/G_k]\}$ and $[x]$ denotes the integral part of x. This expansion is called the *G-expansion*. Provided that for every finite $K > 0$ such that

$$n = \sum_{k=0}^{K} g_k(n) G_k < G_{K+1}, \tag{3.4}$$

the expansion is unique and finite. We call g_k the *k-digit* of the G-expansion. The digits $(g_k)_{k\geq 0}$ can be calculated using the greedy algorithm and $G = (G_n)_{n\geq 0}$ is called a *numeration system*.

Let $E_j = \{k \in \mathbb{N}_0 : 0 \leq k \leq [G_{j+1}/G_j]\}$. There is a natural injective map from $\mathbb{N}$ to the infinite product space $E = \prod_{j\geq 0} E_j$ given by $n \to \overline{n}$ where

$$\overline{n} = g_0(n)\dots g_K(n)0^\infty$$

is determined by the G-expansion of $n = g_0(n)G_0 + \cdots + g_K(n)G_K$. The closure of the image $\overline{\mathbb{N}}$ in E is the subset of sequences satisfying (3.4) and is denoted by $\mathcal{K}_G$, that is

$$\mathcal{K}_G = \{x = (x_0x_1x_2\dots) \in E : \forall j \geq 0, x_0G_0 + \cdots + x_jG_j < G_{j+1}\}.$$

The elements of $\mathcal{K}_G$ are said to be *G-admissible.* A finite sequence $x_0 \dots x_n$ is then G-admissible if $x_0 \dots x_n 0^\infty$ is G-admissible. To extend the addition-by-1 map from $\mathbb{N}$ to $\mathcal{K}_G$ we introduce the set

$$\mathcal{K}_G^0 = \{x \in \mathcal{K}_G : \exists M_x, \forall j \geq M_x, x_0G_0 + \cdots + x_jG_j < G_{j+1} - 1\} \subseteq \mathcal{K}_G. \quad (3.5)$$

Put $x(j) = \sum_{k=0}^{j} x_kG_k$ and set

$$\tau(x) = (g_0(x(j)+1)\dots g_j(x(j)+1))x_{j+1}x_{j+2}\dots, \quad (3.6)$$

for every $x \in \mathcal{K}_G^0$ and $j \geq M_x$. This definition does not depend on the choice of $j \geq M_x$ and can be extended to x in $\mathcal{K}_G \backslash \mathcal{K}_G^0$ by setting $\tau(x) = 0 = (0)^\infty$. Thus, we have defined the map τ on $\mathcal{K}_G$ and we call τ the *G-odometer* or *G-adding machine.*

In the sequel of this chapter we restrict attention to numeration systems where $G = (G_n)_{n\geq 0}$ is a linear recurrence, i.e. we require that $G_0 = 1$ and $G_k = a_0G_{k-1} + \cdots + a_{k-1}G_0 + 1$ for $k < d$. Then for each $n \geq d$, G_n is given by a recurrence of order $d \geq 1$ which is

$$G_{n+d} = a_0G_{n+d-1} + \cdots + a_{d-1}G_n. \quad (3.7)$$

To this linear recurrence we can associate the characteristic equation

$$x^d = a_0 x^{d-1} + \cdots + a_{d-1}. \tag{3.8}$$

We further confine attention to numeration systems with a characteristic equation (3.8) having a Pisot-Vijayragahavan number (PV-number), say β, as a root. In [8] it was shown that this is always the case when

$$a_0 \geq a_1 \geq \ldots \geq a_{d-1} \geq 1.$$

Under this assumption, W. Parry [50] showed that the β-expansion of β is finite, that is

$$\beta = a_0 + \frac{a_1}{\beta} + \cdots + \frac{a_{d-1}}{\beta^{d-1}}, \tag{3.9}$$

where $a_0 = [\beta]$.

To numeration systems, whose characteristic root β is a PV-number satisfying (3.9), a sum $\sum_{k=0}^{M} g_k G_k$ for finite M is the expansion of an integer if and only if for the digits g_k of the G-expansion we have

$$(g_k, g_{k-1}, \ldots, g_0, 0^\infty) \prec (a_0, a_1, \ldots, a_{d-1})^\infty,$$

for each k with $\prec$ denoting the lexicographic order [50]. Representations $(g_k, \ldots, g_0)$ satisfying this condition are said to be *admissible representations* and thus belong to $\mathcal{K}_G$.

Let Z denote a cylinder of length K with digits $g_0, \ldots, g_{K-1}$ for the dynamical system $(\mathcal{K}_G, \tau)$ and let $F_{K,r} = \#\{n < G_{K+r} : (g_0(n), g_1(n), \ldots) \in Z\}$. We can define the measure μ on $\mathcal{K}_G$ by

$$\mu(Z) = \frac{F_{K,0}\beta^{d-1} + (F_{K,1} - a_0 F_{K,0})\beta^{d-2} + \cdots + (F_{K,d-1} - a_0 F_{K,d-2} - \cdots - a_{d-2}F_{K,0})}{\beta^K(\beta^{d-1} + \beta^{d-2} + \cdots + 1)}.$$

In [19] it was showed that $(\mathcal{K}_G, \tau)$, i.e. the odometer on an admissible numeration system G, is uniquely ergodic with respect to the measure μ.

In [26] the following theorem was proved.

Theorem 3.4.1. *Let $G^1, \ldots, G^s$ be numeration systems defined by linear recurrences (3.7) where the coefficients are given by $a_j^i = b_i$ where $j = 0, \ldots, (d_i - 1)$ and $i = 1, \ldots, s$ with pairwise coprime positive integers b_i, $i = 1, \ldots, s$. Further, suppose $\frac{\beta_i^k}{\beta_j^l} \notin \mathbb{Q}$ for all positive integers k, l with $\beta_1, \ldots, \beta_s$ being the roots of the characteristic equations (3.8). Then the dynamical system*

$$((\mathcal{K}_{G^1}, \tau_1) \times (\mathcal{K}_{G^2}, \tau_2) \times \cdots \times (\mathcal{K}_{G^s}, \tau_s)),$$

that is s-dimensional Cartesian product of the corresponding odometers, is uniquely ergodic.

We now define the Monna map ϕ_β for irrational bases $\beta > 1$ as follows.

Definition 3.4.2. Let $n = \sum_{j \geq 0} g_j(n) G_j$ be the G-expansion of a non-negative integer n. Then the transformation $\phi_\beta : \mathcal{K}_G \to \mathbb{R}_0^+$ defined by

$$\phi_\beta(n) = \phi_\beta \left(\sum_{j \geq 0} g_j(n) G_j \right) = \sum_{j \geq 0} g_j(n) \beta^{-j-1}$$

is called the *β-adic Monna map.*

The restriction of ϕ_β to $\mathcal{K}_G^0$ has a well defined inverse which is called *pseudo-inverse* $\phi_\beta^+ : \mathbb{R}_0^+ \to \mathcal{K}_G^0$ defined by

$$\phi_\beta^+ \left(\sum_{j \geq 0} g_j(n) \beta^{-j-1} \right) = \sum_{j \geq 0} g_j(n) G_j.$$

In this context, we have the following definition of the $\boldsymbol{\beta}$-adic Halton sequence.

Definition 3.4.3. The *$\boldsymbol{\beta}$-adic Halton sequence* is defined by

$$(\phi_{\boldsymbol{\beta}}(n))_{n \geq 0} = (\phi_{\beta_1}(n), \ldots, \phi_{\beta_s}(n))_{n \geq 0},$$

where $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_s)$, and β_i is the solution of the corresponding characteristic equation of a numeration system G^i.

Note that it is not evident that the image of $\mathcal{K}_G^0$ under ϕ_β is contained in $[0,1)$ or dense in $[0,1)$ even if the solution of (3.8) is a PV-number. The following proposition gives conditions for this to hold [26].

Proposition 3.4.4. *Let* $\mathbf{a} = (a_0, \ldots, a_{d-1})$ *where* $a_0, \ldots, a_{d-1} \geq 0$ *are the coefficients defining the numeration system* G *and suppose that the corresponding characteristic root* β *satisfying (3.8) is a PV-number. Then* $\phi_\beta(\mathbb{N}) \subset [0,1)$ *and* $\phi_\beta(\mathbb{N})$ *is not contained in* $[0,x)$ *for all* $x \in (0,1)$ *if and only if* $\mathbf{a}$ *can be written either as* $\mathbf{a} = (a_0, \ldots, a_0)$ *or as* $\mathbf{a} = (a_0, a_0 - 1 \ldots, a_0 - 1, a_0)$ *where* $a_0 > 0$.

Proof: The proof can be found in [26]. □

Proposition 3.4.5. *Let* G *be a numeration system given by (3.7) where the coefficients of linear recurrences are of the form* $a_j = a$ *where* a *is a positive integer and* $j = 0, 1, \ldots, (d-1)$. *Let* β *be a solution of the corresponding characteristic equation. Then* $\mu(Z) = \lambda(\phi_\beta(Z))$ *for every cylinder set* Z.

Proof: The proof can be found in [26]. □

Next, we will be interested in the more general case of $\boldsymbol{\beta}$-adic Halton sequences and that is the sequence $(\phi_{\boldsymbol{\beta}}(k_j))_{j\geq 1}$ where $(k_j)_{j\geq 1}$ is a sequence of non-negative integers. In the sequel, we will also assume that the sequence $(k_j)_{j\geq 1}$ is Hartman uniformly distributed and L^2-good universal. The following result can be also found in [30].

Theorem 3.4.6. *Let* $G^1, \ldots, G^s$ *be numeration systems defined by linear recurrences (3.7) where the coefficients are given by* $a_j^i = b_i$ *where* $j = 0, \ldots, (d_i - 1)$ *and* $i = 1, \ldots, s$ *with pairwise coprime positive integers* $b_i, i = 1, \ldots, s$. *Further, suppose* $\frac{\beta_i^k}{\beta_j^l} \notin \mathbb{Q}$ *for all positive integers* k, l *with* $\beta_1, \ldots, \beta_s$ *being the roots of the characteristic equations (3.8). Then if* $(k_j)_{j\geq 1}$ *is Hartman uniformly distributed and* L^2*-good universal, the sequence* $(\phi_{\boldsymbol{\beta}}(k_j))_{j\geq 1}$ *is uniformly distributed on* $[0,1)^s$.

Proof: Using Proposition 3.4.5 and the definition of the Monna map, we get an isomorphism between the dynamical systems $((\mathcal{K}_{G^1}, \tau_1) \times \cdots \times (\mathcal{K}_{G^s}, \tau_s))$ and $(([0,1), T_1) \times \cdots \times ([0,1), T_s))$ with $T_i : [0,1) \to [0,1)$ given by

$$T_i(x) = \phi_{\beta_i} \circ \tau_i \circ \phi_{\beta_i}^{+}(x).$$

We can see this in the diagram below illustrating the isomorphism.

$$\begin{array}{ccc} \mathcal{K}_{G^i} & \xrightarrow{\tau_i} & \mathcal{K}_{G^i} \\ {\scriptstyle \phi_{\beta_i}}\Big\downarrow & & {\scriptstyle \phi_{\beta_i}}\Big\downarrow \\ [0,1) & \xrightarrow{T_i} & [0,1) \end{array}$$

Let $\mathbf{Tx} = (T_1x_1, \ldots, T_sx_s)$ where $\mathbf{x} = (x_1, \ldots, x_s)$ is in $[0,1)^s$. Since by Theorem 3.4.1, the dynamical system $((\mathcal{K}_{G^1}, \tau_1) \times \cdots \times (\mathcal{K}_{G^s}, \tau_s))$ is uniquely ergodic and also it is isomorphic to $(([0,1), T_1) \times \cdots \times ([0,1), T_s))$, then also $(([0,1), T_1) \times \cdots \times ([0,1), T_s))$ or equivalently $\mathbf{T}$ is uniquely ergodic. Moreover, since $(k_j)_{j\geq 1}$ is Hartman uniformly distributed and L^2-good universal, by Remark 3.2.9, we obtain that $(\mathbf{T}^{k_j}\mathbf{x})_{j\geq 1}$ is uniformly distributed in $[0,1)^s$ for all $\mathbf{x} \in [0,1)^s$. In particular, $(\mathbf{T}^{k_j}\mathbf{0})_{j\geq 1}$ is uniformly distributed in $[0,1)^s$.

Now, we want to show that $(\mathbf{T}^{k_j}\mathbf{0})_{j\geq 1} = (\phi_{\boldsymbol{\beta}}(k_j))_{j\geq 1}$. For any $i = 1, \ldots, s$ we have

$$T_i(0) = \phi_{\beta_i} \circ \tau_i \circ \phi^+_{\beta_i}(0) = \phi_{\beta_i} \circ \tau_i(0^\infty)$$

where $\tau_i(0^\infty) = (g_0(1)\ldots g_j(1))0^\infty$. So, we obtain

$$T_i(0) = \phi_{\beta_i}((g_0(1)\ldots g_j(1))0^\infty).$$

Since we can write $(g_0(1)\ldots g_j(1))0^\infty)$ as $g_0(1)G_0 + \cdots + g_j(1)G_j = 1$, we get

$$T_i(0) = \phi_{\beta_i}(1).$$

Now, we have

$$T_i^2(0) = \phi_{\beta_i} \circ \tau_i \circ \phi^+_{\beta_i}(\phi_{\beta_i}(1)) = \phi_{\beta_i} \circ \tau_i(1) = \phi_{\beta_i} \circ \tau_i((g_0(1)\ldots g_j(1))0^\infty).$$

We use again the fact that one can write $(g_0(1)\ldots g_j(1))0^\infty)$ as $g_0(1)G_0 + \cdots + g_j(1)G_j = 1$ and we get

$$\tau_i((g_0(1)\ldots g_j(1))0^\infty) = (g_0(1+1)\ldots g_j(1+1))0^\infty = (g_0(2)\ldots g_j(2))0^\infty$$

and so

$$T_i^2(0) = \phi_{\beta_i}((g_0(2)\dots g_j(2))0^\infty) = \phi_{\beta_i}(2).$$

Proceeding inductively, we get $T_i^n(0) = \phi_{\beta_i}(n)$ or $T_i^{k_j}(0) = \phi_{\beta_i}(k_j)$ for all $k_j \in \mathbb{N}_0, j = 1, 2, \dots$. So, we have

$$(\mathbf{T}^{k_j}\mathbf{0})_{j\geq 1} = (T_1^{k_j}0, \dots, T_s^{k_j}0)_{j\geq 1} = (\phi_{\beta_1}(k_j), \dots, \phi_{\beta_s}(k_j))_{j\geq 1} = (\phi_{\boldsymbol{\beta}}(k_j))_{j\geq 1}.$$

Thus we can conclude that $(\phi_{\boldsymbol{\beta}}(k_j))_{j\geq 1}$ is uniformly distributed in $[0,1)^s$. $\square$

Example 3.4.7. Let us have numerations systems G^1 and G^2 defined as follows.

$$G^1 : G_0 = 1, G_1 = G_0 + 1 = 2 \text{ and } G_{n+2} = G_{n+1} + G_n$$
$$G^2 : G_0 = 1, G_1 = 2G_0 + 1 = 3 \text{ and } G_{n+2} = 2G_{n+1} + 2G_n$$

Solving characteristic equations of corresponding numerations systems, we get solutions which are PV-numbers and these are $\beta_1 = \frac{1+\sqrt{5}}{2}$ and $\beta_2 = 1 + \sqrt{3}$.

Let $(k_j)_{j\geq 1}$ be given by

$$([n^{\frac{3}{2}}])_{n=1}^\infty = 1, 2, 5, 8, 11, 14, 18, 22, 27, 31, 36, 41, 46, 52, 58, 64, 70, 76, \dots.$$

Then the corresponding $\boldsymbol{\beta}$-adic Halton sequence is of the form

$$(\phi_{\beta_1}([n^{3/2}]), \phi_{\beta_2}([n^{3/2}]))_{n=1}^\infty.$$

The values for terms of this $\boldsymbol{\beta}$-adic Halton sequence were computed using algorithm in Maple which can be found in the Appendix. The algorithm is based on definitions introduced in this section. First few values of the $\boldsymbol{\beta}$-adic Halton sequence are (rounded to 4 decimal places):

$$0.6180, 0.3660, 0.3820, 0.7321, 0.1459, 0.8660, 0.0902, 0.0490, 0.3262, 0.1830,$$
$$0.6738, 0.3170, 0.2016, 0.8301, 0.6525, 0.0179, 0.7984, 0.8840, 0.5066, 0.4330, \dots$$

Since we know that $([n^{\frac{3}{2}}])_{n=1}^\infty$ is Hartman uniformly distributed and L^p-good universal and moreover, the conditions for the coefficients of linear recurrences and for β_1, β_2 are satisfied, by Theorem 3.4.6 $(\phi_{\beta_1}([n^{3/2}]), \phi_{\beta_2}([n^{3/2}]))_{n=1}^{\infty}$ is uniformly distributed in $[0, 1)^s$.

Chapter 4

Conclusion and Future Research

The first main outcome of this thesis is proving ergodic and metric properties of Schneider's continued fraction map in non-Archimedean settings. It was shown that the natural extension of this map is isomorphic to a Bernoulli shift with the entropy $\frac{\#(k)}{\#(k)-1}\log(\#(k))$ which was also calculated. Since the natural extension of the map is Bernoulli, this implies a number of strictly weaker properties which are exactness, strong-mixing, weak-mixing and ergodicity. Further, interesting results about various averages and moving averages of partial quotients of the generalised Schneider's continued fraction expansion are proved applying subsequence pointwise ergodic theorems and the moving average ergodic theorem.

There are many other results for the regular continued fraction on the real numbers which one can try to recover for Schneider's continued fraction expansion in non-Archimedean settings. For example, an adaptation of the original Gauss-Kuzmin theorem ([16], [36], [37]) to our continued fraction expansion in non-Archimedean settings could yield an interesting result. Another thing one might try to do is to prove quantitative versions of metrical theorems about averages and moving averages of partial quotients of the generalised Schneider's continued fraction expansions. Specifically, we would like to find error terms in the theorems in Sections 2.6 and 2.7. To determine the error terms one can use I. S. Gál and J. F. Koksma's method [15] or to get slightly better results, one can try to adapt the method introduced in [23].

Another interesting problem in non-Archimedean settings in general is to recover the p-adic analogue of Rudolph's Theorem [60] which is the partial result to a well known problem due to H. Furstenberg who asks whether Lebesgue measure is the only non-atomic measure preserved by multiplication by both a and b where a, b are coprime natural numbers. Rudolph used symbolic dynamics to give a proof of this problem under the assumptions that a, b are coprime, the joint action is ergodic and that one of the multiplications has positive entropy. A. Johnson then proved Rudolph's Theorem assuming only that a divides no power of b instead of requiring coprimality of a, b [31]. The analogue of Rudolph's result in positive characteristic was proved in [20] and the proof of Johnson's result in positive characteristic was given in [22]. However, to prove these results in the p-adic case is not so straightforward as there are complications with the assumption of ergodicity of the joint action which need to be addressed.

Another main result proved in this thesis is that under certain assumptions a subsequence of $\boldsymbol{\beta}$-adic Halton sequences is uniformly distributed in $[0,1)^s$. Specifically, if a sequence of non-negative integers $(k_j)_{j=1}^{\infty}$ is Hartman uniformly distributed and L^2-good universal and the bases $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_s)$ are special PV-numbers then $(\phi_{\boldsymbol{\beta}}(k_j))_{j\geq 1}$ is uniformly distributed in $[0,1)^s$.

For one-dimensional Halton sequences, that is for van der Corput sequences, one might try to extend the result proved in [39] to β-adic van der Corput sequences. Specifically, we would like to prove the following. If $(k_j)_{j=1}^{\infty}$ is Hartman uniformly distributed, $(m_1, \ldots, m_s)$ is an s-tuple of non-negative integers and if $\beta > 1$ is a PV-number then the asymptotic distribution function of the sequence

$$(\phi_\beta(k_j + m_1), \ldots, \phi_\beta(k_j + m_s))_{j\geq 1}$$

exists and is a copula.

Appendix

Maple Code of Program for Calculating Terms in β-adic Halton sequence

```
### Setup
 ## Calculate up to this term of the sequence
 highest_order:=120:
 ## Start output at this term of sequence
 min_out:=1:
 ## End output at this term of sequence, must be
    less than highest_order
 max_out:=40:
 ## Number of digits in approximation
 decimal_approx:=6:

### Solve Characteristic Equation
 i:='i':
 ## Input coefficients of equations
 a:=[1,1]:
 d:=nops(a):
 eq:=x^d=sum(a[i]*x^(d-i),i=1..d):
 sols:={solve(eq)}:

### Get real solution greater than 1
 solarray:={}:
```

```
 i:='i':
 for i from 1 to nops(a) do
   if evalf(sols[i]-1)>0 then
     solarray:=solarray union {sols[i]}:
   fi:
 od:
 ## Visual check that there is only one element
 solarray;
 ## Set solution
 X:=solarray[1]:

### Initialise sequence
i:='i':
g:=[1,2]: ### Input first terms
for i from 1 to nops(g) do
 G[i-1]:=g[i]:
od:

### Generate G sequence
 i:='i':
 for i from 1 to highest_order-d+1 do
   G[i+d-1]:=sum(a[j]*G[i+d-j-1],j=1..nops(a)):
 od:

### Linear equation with coefficients from G sequence
 i:='i':
 lin_eq:=sum(x[i]*G[i],i=0..10):

### Output terms in Halton sequence using greedy algorithm
 i:='i':
   for n from 1 to highest_order do
     max_i:=0:
     while(n>=G[max_i]) do
```

```
      max_i:=max_i+1:
    od:
    max_i:=max_i-1:
    r:=n:
    for i from 0 to max_i do
      j:=max_i-i:
      y[n,j]:=min(floor(r/G[j]),floor(G[j+1]/G[j])):
      r:=r-y[n,j]*G[j]:
    od:
  j:='j':
  ans:=simplify(expand(rationalize(sum(y[n,j]*X^(-j-1),
   j=0..max_i)))):
## Output term, including position and approximation
  print(n,ans,evalf[decimal_approx+1](ans));
od:
```

Bibliography

[1] G. Barat, P. J. Grabner: *Distribution properties of G-additive functions.* J. Number Theory **60**, no. 1, 1996, 103–123.

[2] P. G. Becker: *Periodizitätseigenschaften p-adischer Kettenbrüche.* [Periodicity properties of p-adic continued fractions.] Elem. Math. **45**, no. 1, 1990, 1–8.

[3] E. Bedocchi: *A note on p-adic continued fractions.* Ann. Mat. Pura Appl. **152**, no. 4, 1988, 197–207.

[4] E. Bedocchi: *Sur le développement de* $\sqrt{m}$ *en fraction continue p-adique.* [On the expansion of $\sqrt{m}$ into a p-adic continued fraction.] Manuscripta Math. **67**, no. 2, 1990, 187–195.

[5] A. Bellow, R. Jones, J. Rosenblatt: *Convergence of Moving Averages.* Ergodic Theory and Dynam. Systems **10**, no. 1, 1990, 43–62.

[6] M. Boshernitzan, G. Kolesnik, A. Quas, M. Wierdl: *Ergodic averaging sequences.* J. Anal. Math. **95**, 2005, 63–103.

[7] J. Bourgain: *On the maximal ergodic theorem for certain subsets of the integers.* Israel J. Math. **61**, no. 1, 1988, 39–72.

[8] A. Brauer: *On algebraic equations with all but one root in the interior of the unit circle.* Mathematische Nachrichten **4**, 1951, 250–257.

[9] Y. Bugeaud, T. Pejkovíc: *Quadratic Approximation in* $\mathbb{Q}_p$. Preprint, 2014.

[10] P. Bundschuh: *p-adische Kettenbrüche und Irrationalität p-adischer Zahlen.* Elem. Math. **32**, no. 2, 1977, 36–40.

[11] P. Bundschuh, R. Wallisser: *Algebraische Unabhängigkeit p-adischer Zahlen.* Math. Ann. **221**, no. 3, 1976, 243–249.

[12] J. W. S. Cassels, A. Frohlich: *Algebraic Number Theory.* Academic Press, London, 1967.

[13] I. P. Cornfeld, S. V. Formin, Ya. G. Sinai: *Ergodic Theory.* Springer-Verlag, New York, 1982.

[14] N. A. Friedman, D. S. Ornstein: *On Isomorphism of Weak Bernoulli Transformations.* Advances in Mathematics **5**, 1970, 365–394.

[15] I. S. Gál, J. F. Koksma: *Sur l'ordre de grandeur des fonctions sommables.* Indag. Math. **12**, 1950, 638–653.

[16] C. F. Gauss: *Recherches Arithmétiques.* Blanchard, Paris, 1807.

[17] F. Q. Gouvea: *p-adic Numbers: An Introduction.* Springer-Verlag, Berlin, Heidelberg, 1993.

[18] P. Grabner, P. Hellekalek, P. Liardet: *The dynamical point of view of low-discrepancy sequences.* Uniform Distribution Theory **7**, no. 1, 2012, 11–70.

[19] P. Grabner, P. Liardet, R. Tichy: *Odometers and systems of numeration.* Acta Arithmetica **70**, 1995, 103–123.

[20] J. Hančl, A. Jaššová, P. Lertchoosakul, R. Nair: *Polynomial Actions in Positive Characteristic.* Proc. Steklov Inst. Math. **280**, no. 2, 2013, 37–42.

[21] J. Hančl, A. Jaššová, P. Lertchoosakul, R. Nair: *On the metric theory of p-adic continued fractions.* Indag. Math. (N.S.) **24**, no. 1, 2013, 42–56.

[22] J. Hančl, A. Jaššová, P. Lertchoosakul, R. Nair: *Polynomial Actions in Positive Characteristic II.* Preprint, 2014.

[23] J. Hančl, A. Jaššová, P. Lertchoosakul, R. Nair: *On the Quantitative Metric Theory of Continued Fractions.* Preprint, 2014.

[24] J. Hirsh and L. C. Washington: *P-adic continued fractions.* Ramanujan J. Math., 2011, 389–403.

[25] E. Hlawka: *Funktionen von beschränkter Variation in der Theorie der Gleichverteilung.* Ann. Math. Pura Appl. **54**, no. 4, 1961, 325–333.

[26] M. Hofer, M. R. Iaco, R. Tichy: *Ergodic properties of β-adic Halton sequences.* Ergodic Theory and Dynam. Systems, 2014 (to appear).

[27] M. Iosifescu, C. Kraaikamp: *Metrical Theory of Continued Fractions.* In: Mathematics and its Applications, vol. **547**, Kluwer Academic Publishers, Dordrecht, 2002.

[28] S. Ito, H. Murata, H. Totoki: *Remarks on the isomorphism theorems for weak Bernoulli transformations in general case.* Publ. Res. Inst. Math. Sci. **7**, 1971/1972, 541–580.

[29] A. Jaššová, R. Nair: *On Schneider's Continued Fraction Map on a Complete Non-Archimedean Field.* Preprint, 2014.

[30] A. Jaššová, P. Lertchoosakul, R. Nair: *On Variants of the Halton Sequences.* Preprint, 2014.

[31] A. S. A. Johnson: *Measures on the circle invariant under multiplication by a nonlacunary subsemigroup of the integers.* Israel J. Math. **77**, no. 1–2, 1992, 211–240.

[32] H. Kamarul-Haili, R. Nair: *On moving averages and continued fractions.* Unif. Distrib. Theory **6**, no. 1, 2011, 65–78.

[33] O. Karpenkov: *Geometry of Continued Fractions.* Algorithms and Computation in Mathematics **26**, Springer-Verlag, Berlin, Heidelberg, 2013.

[34] A. Ya. Khinchin: *Continued Fractions.* University of Chicago Press, Chicago, 1964.

[35] L. Kuipers, H. Niederreiter: *Uniform distribution of sequences.* Dover Publications, Inc., Mineola, New York, 2006.

[36] R. O. Kuzmin: *On one problem of Gauss.* Dokl. Akad. Nauk SSSR, Ser. A, 1928, 375–380.

[37] R. O. Kuzmin: *On a problem of Gauss.* Atti del Congresso Internazionale dei Matematici, Bologna **6**, 1932, 83–89.

[38] V. Laohakosol, P. Ubolsri: *p-adic continued fractions of Liouville type.* Proc. Amer. Math. Soc. **101**, no. 3, 1987, 403–410.

[39] P. Lertchoosakul, R. Nair: *Distribution functions for subsequences of the van der Corput sequence.* Indag. Math. **24**, no. 3, 2013, 593–601.

[40] P. Lévy: *Sur les lois de probabilité dont dependent les quotients complets et incomplets d'une fraction continue.* Bull. Soc. Math. France **57**, 1929, 178–194.

[41] R. Nair: *On Polynomials in Primes and J. Bourgain's Circle Method Approach to Ergodic II.* Studia Math. **105**, no. 3, 1993, 207–233.

[42] R. Nair: *On the Metrical Theory of Continued Fractions.* Proc. Amer. Math. Soc. **120**, no. 4, 1994, 1041–1046.

[43] R. Nair: *On asymptotic distribution of the a-adic integers.* Proc. Indian Acad. Sci. Math. Sci. **107**, no. 4, 1997, 363–376.

[44] R. Nair: *On uniformly distributed sequences of integers and Poincaré recurrence II.* Indag. Math. (N.S.) **9**, no. 3, 1998, 405–415.

[45] R. Nair: *On Hartman uniform distribution and measures on compact spaces.* Harmonic Analysis and Hypergroups, Delhi, 1995, Trends Math., 59–75, Birkhäuser Boston, Boston, MA, 1998.

[46] R. Nair, M. Weber: *On random perturbation of some intersective sets.* Indag. Math. (N.S.) **15**, no. 3, 2004, 373–381.

[47] J. Neukirch: *Algebraic Number Theory.* Springer-Verlag, Berlin, Heidelberg, 1999.

[48] S. Ninomiya: *Constructing a new class of low-discrepancy sequences by using the β-adic transformation.* IMACS Seminar on Monte Carlo Methods, Brussels, 1997, Math. Comput. Simulation **47**, no. 2–5, 1998, 403–418.

[49] D. Ornstein: *Bernoulli shifts with the same entropy are isomorphic.* Advances in Math. **4**, 1970, 337–352.

[50] W. Parry: *On the β-expansions of real numbers.* Acta Math. Acad. Sci. Hungar. **11**, 1960, 401–416.

[51] W. Parry: *Entropy and Generators in Ergodic Theory.* W. A. Benjamin, Inc., New York, 1969.

[52] W. Parry: *Topics in ergodic theory.* Cambridge University Press, Cambridge, 1981.

[53] K. R. Parthasarathy: *Introduction to Probability and Measure.* MacMillan, London, 1977.

[54] K. Petersen: *Ergodic Theory.* Cambridge University Press, Cambridge, 1983.

[55] M. Pollicott, M. Yuri: *Dynamical Systems and Ergodic Theory.* Cambridge University Press, Cambridge, 1998.

[56] D. Ramakrishnan, R. J. Valenza: *Fourier Analysis on Number Fields.* Graduate Texts in Mathematics **186**, Springer-Verlag, New York, 1999.

[57] A. M. Robert: *A Course in p-adic Analysis.* Graduate Texts in Mathematics **198**, Springer-Verlag, New York, 2000.

[58] V. A. Rokhlin: *Exact endomorphisms of a Lebesgue space.* Izv. Akad. Nauk SSSR Ser. Mat. **25**, no. 4, 1961, 499–530.

[59] A. A. Ruban: *Some metric properties of p-adic numbers.* Sib. Math. J. **11**, 1970, 176–180.

[60] D. J. Rudolph: $\times 2$ *and* $\times 3$ *invariant measures and entropy.* Ergodic Theory and Dynam. Systems **10**, no. 2, 1990, 395–406.

[61] C. Ryll-Nardzewski: *On the ergodic theorems II. Ergodic theory of continued fractions.* Studia Math. **12**, 1951, 74–79.

[62] S. Sawyer: *Maximal inequalities of weak type.* Ann. of Math. **84**, no. 2, 1966, 157–174.

[63] T. Schneider: *Über p-adische Kettenbrüche.* In: Symposia Mathematica, Vol. **IV**, INDAM, Rome, 1968/69, 181–189. Academic Press, London, 1970.

[64] M. Smorodinsky: *β-automorphisms are Bernoulli shifts.* Acta Math. Acad. Sci. Hungar. **24**, 1973, 273–278.

[65] V. G. Sprindžuk: *Mahler's problem in metric number theory.* Translated from the Russian by B. Volkmann. Translations of Mathematical Monographs **25** American Mathematical Society, Providence, R.I., 1969.

[66] W. Steiner: *Regularities of the distribution of β-adic van der Corput sequences.* Monatsh. Math. **149**, no. 1, 2006, 67–81.

[67] A. Tempelman: *Ergodic Theorems for Group Actions.* Mathematics and Its Applications **78**, 1992.

[68] P. Walters: *An Introduction to Ergodic Theory.* Graduate Texts in Mathematics **79**, Springer-Verlag, New York, 1982.

[69] B. M. M. de Weger: *Periodicity of p-adic continued fractions.* Elem. Math. **43**, no. 4, 1988, 112–116.

[70] A. Weil: *Basic Number Theory.* Springer-Verlag, New York, 1974.

[71] E. Wirsing: *On the theorem of Gauss-Kusmin-Lévy and a Frobenius-type theorem for function spaces.* Acta Arith. **24**, 1973/1974, 507–528. Collection of articles dedicated to Carl Ludwig Siegel on the occasion of his seventy-fifth birthday.

List of the Author's Publications

1. A. Jaššová, S. Kristensen, P. Lertchoosakul, R. Nair: *On Recurrence in Positive Characteristic.* Submitted to Indag. Math., 2014.

2. A. Jaššová, P. Lertchoosakul, R. Nair: *On Variants of the Halton Sequences.* Preprint, 2014.

3. A. Jaššová, R. Nair: *On Schneider's Continued Fraction Map on a Complete Non-Archimedean Field.* Preprint, 2014.

4. J. Hančl, A. Jaššová, P. Lertchoosakul, R. Nair: *Polynomial Actions in Positive Characteristic II.* Preprint, 2014.

5. J. Hančl, A. Jaššová, P. Lertchoosakul, R. Nair: *On the Quantitative Metric Theory of Continued Fractions.* Preprint, 2014.

6. J. Hančl, A. Jaššová, P. Lertchoosakul, R. Nair: *On the metric theory of p-adic continued fractions.* Indag. Math. (N.S.) **24**, no. 1, 2013, 42–56.

7. J. Hančl, A. Jaššová, P. Lertchoosakul, R. Nair: *Polynomial Actions in Positive Characteristic.* Proc. Steklov Inst. Math. **280**, no. 2, 2013, 37–42.

8. J. Hančl, A. Jaššová, J. Šustek: *Lebesgue measure and Hausdorff dimension of special sets of continued fractions.* Ramanujan J. **28**, no. 1, 2012, 15–23.