Distribution Problems in Arithmetic

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy
by

Entesar Mohmed Nasr

May 2018

Abstract

In this thesis we use modern developments in ergodic theory and uniform distribution theory to investigate the the distribution of polynomials, partial quotients of convergents, random and oscillatory sequences. The original contents of the thesis are in Chapters 3-6. More specifically in Chapter 3 we study the the distribution of of sequences of the form $(a_n x)_{n\geq 1}$ modulo one. Here $(a_n)_{n\geq 1}$ is a sequence of natural numbers and x is a real number one. Usually in studies of this form, the sequence $(a_n)_{n\geq 1}$ is a specific sequence of integers and results are given for almost all x with respect to some measure usually Lebesgue measure. Here the roles are reveresed, and $(a_n)_{n\geq 1}$ is the itinerary of a random walk and x is a fixed real number with prescribed diophantine peoperties. In this context the distribution has already been studied by Michel Weber. Here we study a second order property called the pair correlation instead. The original contents of this chaper have already been publisted in a the Journal Uniform Distribution Theory

In Chapter 4 we study refinements of Weyl's famous theorem that if $P(x) = \alpha_k x^k + \dots \alpha_1 x + \alpha_0$ and one of $\alpha_1, \dots, \alpha_k$ is irrational, then $(P(n))_{n\geq 1}$ is uniformly distributed modulo one. We use a method of Furstenberg to show $(P(k_n))_{n\geq 1}$ is uniformly distributed modulo one for various broad sequences of integers, both deterministic and stochastic. This is then used to constuct new examples of sequences of integers that are Poincaré recurrent and satisfy the Glasner property. This work is to published in the high prestigue journal Ergodic Theory and Dynamical System.

Suppose $(k_n)_{n\geq 1}$ is a sequence of integers that satisfies the same conditions as in Chapter 4. In Chapter 5 we prove the following theorem. Suppose H is a finite index subgroup of a Fuchsian group Γ it self a subset of $SL(2, \mathbb{Z})$. Then for almost all x with respect to Lebesgue measure, the regular continued

fraction approximants are distributed in the cusps of H according to the relative cusp widths. That is for each cusp κ of with $w(\kappa)$,

$$\lim_{N\to\infty} \frac{1}{N} \frac{\#\{1 \leq n \leq N : p_{k_n}/q_{k_n} \in \kappa\}}{N} = \frac{w(\kappa)}{[\Gamma : H]}.$$

In the final chapter i.e. Chapter 6 we study the pair correlation of $(a_n \cos(\alpha_n x))_{n\geq}$ for deterministic $(a_n)_{n\geq 1}$. Here we calculate the Hausdorff dimension of the set where this pair correlation is a pre-determined size.

Acknowledgements

There are moments of weakness and moments of strength, failed and successful. But when you know there is a God who guarding you, then you believe you will reach your Goal. Thank you my God for being next to me at each step during my life.

I gratefully acknowledge my supervisor Dr. Radhakrishnan Nair for his valuable suggestions, constant support, encouragement and cooperation. This work would not have been possible without his guidance. Under his supervision I successfully overcame many difficulties and learned a lot during my PhD study.

Special mention goes to the staff of Mathematic Department at the University of Liverpool. My PhD has been an amazing experience; I warmly thank each member, not only for their academic support, valuable advice and providing necessary resources to accomplish my research, but also for providing me extensive personal and professional guidance and taught me a great deal about both scientific research and life in general.

This work would not have been possible without the financial support. My acknowledgment goes to the Ministry of Education- Libya and Libyan Embassy-London who are giving me this scholarship and chance to increase my knowledge and complete my degree by sponsored me during my study.

Thank you does not seem sufficient but it is said with appreciation and respect to all of My Family. I am extremely indebted to my parents. I must express my very profound gratitude to My Father My Mother My sisters Amna and Eman and all of my brothers whose love and guidance are with me in whatever I pursue and for providing me with unfailing support, continuous encouragement, and care throughout my years of study. I dedicate this thesis to them.

Words fail me to express my gratitude to the members of My small Family. I would like to thank my supportive husband, and my wonderful children; Alaa, Mohammed, Hamza and Faris, who provide unending inspiration. They are push me and motivate me to pursue this dream comes true.

I would like to extend thanks to the many people, whoever anyone gives me a support in this period of time.

Contents

Chapter 1

Introduction

The roots of our subject are very old indeed. In his book Tractatus de configurationibus qualitatum et motuum written in the 1340's Nicoli Oresme states that if two particles move on a circle with fixed velocities in "incomensurable" ratio then any arc, no matter how small, will contain both simultaneously infinitly often in the future and will have doesn so infinitely in the past. This was not proved in the sense we could understand the term. We would have to wait till the 1840's when the result was proved by L. Kroneker and in fact the result is generally refered to as Kroneker's theorem. Oresme's work took place in near complete isolation, in a time when most the tools available even during Kronecker's era, like calculus, and number theory was largely absent.

We have mentioned Kronecker but was in the late Edwardian Era, as we would describe it in Britain that real developments begun. Motivated by the problem of secular perturbations in astronomy, P. Bohl, W. Sierpinski and H. Weyl proved that if α is irrational then the sequence of $(\{n\alpha\})_{n\geq 1}$ is uniformly distributed modulo one – in the modern parlance. The concept of uniform distribution was yet to be formally defined.

For a real number x let $x = \sum_{n=1}^{\infty} \frac{c_n}{b^n}$ denote the base b expansion of x for an integer $b > 1$ and $c_n \in \{0, 1, 2, \ldots, b-1\}$ $(n = 1, 2, \ldots)$. For

$d_1, \dots, d_l \in \{0, 1, \dots, b-1\}$ let

$$N_M(x, d_1, \dots, d_l) = \#\{1 \le j \le M : x_j = d_1, \dots, x_{j-l+1} = d_l\} \qquad (M = 1, 2, \dots)$$

We say x is normal base b if

$$\lim_{M \to \infty} \frac{N_M(x, d_1, \dots, d_l)}{M} = \frac{1}{b^l},$$

for all choices of $\{d_1, \dots, d_l\}$ and $l = 1, 2, \dots$. We say x is normal if it normal base b for each b. In 1909 E. Borel proved that almost all x are normal with respect to Lebesgue measure. This was only six years after Lebesgue measure itself was invented.

A crutial development in this subject was H. Weyl's 1916 paper. This explicitly define te concept of uniform distribution, gave four equivalent characterisations, including the famous Weyl criterion, which makes the checking the the uniform distribution of a sequence much easier. Using Weyl's criterion Weyl showed that if $P(x) = \alpha_k x^k + \dots \alpha_1 x + \alpha_0$ is a polynomial with at least one of the numbers $\alpha_k, \dots, \alpha_1$ irrational then $(\{P(n)\})_{n\ge1}$ uniform distributed modulo one. To do this he developed the famous Weyl's inequality for exponential sum so crutial to modern analytic number theory. In this paper he also developed tools from harmonic analysis to show that if $(a_n)_{n\ge1}$ is any strictly increasing sequence of integers then $(\{a_n x\})_{n\ge1}$. It is easy to check that the uniform distribution modulo one of $(\{b^n x\})_{n\ge1}$ is equivalent to the nornality base b of x. This observation also forshadows the development of ergodic theory. as the unifrom distribution of $(\{b^n x\})_{n\ge1}$ could be recovered from Birkhoff's ergodic theorm. That was yet a decade and a half away.

In the late 1920's the Radamacher-Menchov method was developed in Harmonic analysis leading to a method which lead to the dyadic chaining method in probability theory. This enabled one to show estimated for the form

$$\int_\Omega |X_{M+1} + \ldots + X_{M+N}|^2 d\omega = \Phi(M,N) \qquad (M,N = 1,2,\ldots)$$

implies, given $\epsilon > 0$, that

$$X_1 + \ldots + X_{M+N} = \Phi(0,N) + O(\Phi(0,N)^{\frac{1}{2}+\epsilon}), \qquad (N = 1,2,\ldots)$$

almost everywhere. Around this time J. G. Van der Corput introduce the concept of discrepancy, as means of making uniform distribution theory qunatitative. Combined with diadic chaining eneable the refine ment of Weyl's theorem on $(\{a_n x\})_{n\geq 1}$. Important refinements were made possible by Developments like the Erdos-Turan Inequality and geometric measure theory.

In the 1930's ergodic theory came into existance, with the proof of the Von Neumann and Birkhoff ergodic Theorems. In the intervening eight and a half decades there has been a vast development of these topic , overlapping modern ergodic theory, probability theory, number theory, combinatorics and much else. In this thesis we use modern developments in ergodic theory and uniform distribution theory to investigate the the distribution of polynomials, partial quotients of convergents, random and oscillatory sequences.

Chapter 2

Background

In this chapter we give basic definitions and theorems which we are going to use in this thesis. See [18][34] for further background.

2.1 Measure Space

In this section we review background material which helps the reader understanding particularly relating to measure, ergodic theory and their applications.

Definition 2.1.1. *Let X be a non-empty set. A collection $\mathcal{B}$ of subsets of X is called a σ-algebra if it satisfies the following three conditions:*
(1) $X \in \mathcal{B}$;
(2) *If* $B \in \mathcal{B}$, *then* $X \backslash B \in \mathcal{B}$;
(3) *If* $\{B_n\}_{n=1}^{\infty}$ *is a countable collection of sets in* $\mathcal{B}$, *then* $\bigcup_{n=1}^{\infty} B_n \in \mathcal{B}$.
We call $(X, \mathcal{B})$ *a measurable space.*

Definition 2.1.2. *Let* $(X, \mathcal{B})$ *be a measurable space. A function* $\mu : \mathcal{B} \to [0, \infty)$ *is said to be a measure if it satisfies the following two conditions:*
(1) $\mu(\phi) = 0$ *for the empty set* ϕ ;

(2) *for any countable collection* $\{B_n\}_{n=1}^{\infty}$ *of pairwise disjoint sets in* $\mathcal{B}$*, we have* $\mu(\bigcup_{n=1}^{\infty} B_n) = \sum_{n=1}^{\infty} \mu(B_n)$.

Definition 2.1.3. *A collection* $\mathcal{S}$ *of subsets of* X *is called a* $semi-algebra$ *if the following three conditions hold:*
(1) $\phi \in \mathcal{S}$;
(2) *If* $A, B \in \mathcal{S}$*, then* $A \bigcap B \in \mathcal{S}$;
(3) *If* $A \in \mathcal{S}$*, then* $X \backslash A = \bigcup_{i=1}^{n} E_i \in \mathcal{S}$ *where each* $E_i \in \mathcal{S}$ *and* $E_1, ..., E_n$ *are pairwise disjoint subsets of* X .

Definition 2.1.4. *A collection* $\mathcal{A}$ *of subsets of* X *is called an algebra if the following three conditions hold:*
(1) $\phi \in \mathcal{A}$;
(2) *If* $A, B \in \mathcal{A}$*, then* $A \bigcap B \in \mathcal{A}$;
(3) *If* $A \in \mathcal{A}$*, then* $X \backslash A \in \mathcal{A}$.

Clearly every algebra is a semi-algebra and every σ-algebra is an algebra.

Definition 2.1.5. *Let* $(X, \mathcal{B}, \mu)$ *be a finite measure space, if* $\mu(X) = 1$*, then we say* $(X, \mathcal{B}, \mu)$ *is a probability space or a normalised measure space.*

Theorem 2.1.6. *(Kolmogorov Extension Theorem) Let* X *be a set, and let* $\mathcal{A}$ *be an algebra of subsets of* X*. Suppose* $\mu^* : \mathcal{A} \to [0, \infty)$ *satisfies the following three properties:*
(1) $\mu^*(\phi) = 0$;
(2) *for any countable collection* $\{E_n\}_{n=1}^{\infty}$ *of pairwise disjoint sets in* $\mathcal{A}$ *such that* $\bigcup_{n=1}^{\infty} E_n \in \mathcal{A}$*, we have* $\mu^*(\bigcup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} \mu^*(E_n)$;
(3) *there are countably many sets* $E_n \in \mathcal{A}$ *such that* $X = \bigcup_{n=1}^{\infty} E_n$ *and* $\mu^*(E_n) < \infty$*. Then there exists a unique measure* $\mu : \mathcal{B}(\mathcal{A}) \to [0, \infty)$ *which is an extension of* μ^*.

Definition 2.1.7. *(Borel Set) Let* (X, d) *be a metric space. The Borel* $\sigma-algebra$ *(*σ*-filed)* $\mathcal{B} = \mathcal{B}(X)$ *is the smallest* $\sigma - algebra$ *in* X *that contains all open subsets of* X*. The elements of* $\mathcal{B}$ *are called the Borel sets of* X.

The metric space (X, d) is called separable if it has a countable dense subset, that is, there are $x_1, x_2, ...$ in X such that $\overline{\{x_1, x_2, ...\}} = X$. ($\overline{A}$ denotes the closure of $A \subset X$).

Lemma 2.1.8. *If X is a separable metric space, then $\mathcal{B}(X)$ equals the $\sigma-$ algebra generated by the open (or closed) balls of X.*

Lemma 2.1.9. *Let (X, d) be a separable metric space. Let $\mathcal{C} \subset \mathcal{B}$ be countable. If $\mathcal{C}$ separates closed balls from points in the sense that for every closed ball B and every $x \in X \backslash B$ there exists $C \in \mathcal{C}$ such that $B \subset C$ and $x \notin C$, then the $\sigma - algebra$ generated by $\mathcal{C}$ is the Borel $\sigma - algebra$.*

Proposition 2.1.10. *Let (X, d) be a metric space. $\mathcal{B}(X)$ is the smallest $\sigma - algebra$ with respect to which all (real valued) continuous functions on X are measurable ($w.r.t.\mathcal{B}(X)$ and $\mathcal{B}(\mathbb{R})$.*

Definition 2.1.11. *(Borel probability measures) Let (X, d) be a metric space. A $finite$ Borel measure on X is a map $\mu : \mathcal{B}(X) \to [0, \infty)$ such that:*
(1) $\mu(\phi) = 0$;
(2) *If $B_1, B_2, ... \in \mathcal{B}$ are mutually disjoint, then $\mu(\bigcup_{n=1}^{\infty} B_n) = \sum_{n=1}^{\infty} \mu(B_n)$.*

We call μ a *Borel probabiliy measure* if in addition $\mu(X) = 1$.

Definition 2.1.12. *(Lebesgue measure) Take $X = [0, 1]$, and let $\mathcal{A}$ denote the algebra of all finite unions of subintervals of $[0, 1]$. For an interval $[a, b]$, define $\lambda^*([a, b]) = b - a$ and extend this to $\mathcal{A}$. Then this satisfies the hypotheses of the Kolmogorov extension theorem, and so it defines a Borel probability measure. This is the Lebesgue $measure$ on $[0, 1]$.*

In a similar way, we can define Lebesgue measure on $[0, 1]^n$. An $n-$dimensional cube is a set of the form $[a_1, b_1]...[a_n, b_n]$, where $0 \leq a_j \leq b_j \leq 1$ for each $1 \leq j \leq n$.

Let $\mathcal{A}$ denote the algebra of all finite unions of $n-$dimensional cubes. Define;

$$\lambda_n^*([a_1, b_1]...[a_n, b_n]) = \prod_{j=1}^{n}(b_j - a_j)$$

and extend this to $\mathcal{A}$. Again, this satisfies the hypotheses of the Kolmogorov extension theorem and defines the n-dimensional Lebesgue measure on $[0, 1]^n$.

2.2 Uniform Distribution mod 1

For detailed and additional background on uniform distribution modulo 1 see [56]. Let us consider the n-dimensional Euclidean space $\mathbb{R}$. Let $x, y \in \mathbb{R}^n$ and their difference $x - y$ is an integral lattice point in $\mathbb{Z}^n$. Equivalently we can say that we consider the space $\mathbb{R}^n$ *modulo* 1 or we deal with the n-dimensional torus $\mathbb{T}^n = \mathbb{R}^n \backslash \mathbb{Z}^n$.

Obviously $\mathbb{T}^n$ can be identified with the unit cube $\mathbb{U}^n = [0, 1)^n$. Formally this can be done by using the notion of fractional parts. The fractional part $\{x\}$ of a real number x is defined by $\{x\} = x - [x]$, where $[x]$ denotes the integral part of x, that is, the greatest integer $\leq x$. The fractional part $\{x\}$ and the integral part $[x]$ for $x \in \mathbb{R}^n$ are defined component-wise.

Let $I^n = [a_1, b_1) \times ... \times [a_n, b_n) \subseteq \mathbb{R}^n$ be an interval (or a product of intervals with sides parallel to the axes) in the n-dimensional space $\mathbb{R}^n$ with $0 < b_i - a_i \leq 1, i = 1, 2, ...n$, then the reduction *modulo* 1 ,$I = I^n \backslash \mathbb{Z}^n$ is called an interval (or a product of intervals with sides parallel to the axes) of the torus $\mathbb{T}^n = \mathbb{R}^n \backslash \mathbb{Z}^n$. The volume $\lambda_n(I)$ of an interval $I \subseteq \mathbb{R}^n \backslash \mathbb{Z}^n$ is given by $\prod_{i=1}^{n}(b_i - a_i)$.($\lambda_n$ denotes the n-dimensional Lebesgue measure.)

Definition 2.2.1. *The characteristic function of a probability measure μ on $\mathcal{B}(\mathbb{R})$ is the function $\phi_\mu : \mathbb{R} \to \mathbb{C}$ given by*

$$\phi_\mu(t) = \int e^{itx}\mu(dx)$$

Definition 2.2.2. *For an interval* $I \subseteq \mathbb{R}^n \backslash \mathbb{Z}^n$ *and a sequence* $(x_n)_{n\geq 1}$*, with* $x_n \in \mathbb{R}^n$*, let* $A(I, N, x_n)$ *be the number of points* $x_n, 1 \leq n \leq N$*, for which* $\{x_n\} \in I$*, i.e.*

$$A(I, N, x_n) = \sum_{n=1}^{N} \chi_I(\{x_n\}), \tag{2.1}$$

where χ_I *is the characteristic function of* I*.*

A characteristic function is simply the Fourier transform, in probabilistic language.

Definition 2.2.3. *The sequence* $\omega = (x_n), n = 1, 2, ...$ *of real numbers is said to be uniformly distributed modulo 1 (abbreviated u.d. mod 1) if for every pair* a, b *of real numbers with* $0 \leq a < b < 1$ *we have*

$$\lim_{N\to\infty} \frac{A([a, b), N, \omega)}{N} = b - a. \tag{2.2}$$

Thus, in simple terms, the sequence is *u.d. mod* 1 if every half-open subinterval of $[0, 1)$ eventually gets its "proper share" of fractional parts. Let now $c_{[a,b)}$ be the characteristic function of the interval $[a, b) \subseteq I$. Then (2.1) can be written in the form

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} c_{[a,b)}(\{x_n\}) = \int_0^1 c_{[a,b)}(x)dx. \tag{2.3}$$

This observation, together with an important approximation technique, leads to the following criterion.

Theorem 2.2.4. *The sequence* $(x_n), n = 1, 2, ...$ *of real numbers is u.d. mod 1 if and only if for every real-valued continuous function* f *defined on the closed unit interval* $\overline{I} = [0, 1]$ *we have*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(\{x_n\}) = \int_0^1 f(x)dx. \tag{2.4}$$

Corollary 2.2.5. *The sequence* (x_n) *is u.d. mod 1 if and only if for every Riemann-integrable function* f *on* $\overline{I}$ *equation (1.3) holds.*

Corollary 2.2.6. *The sequence* (x_n) *is u.d. mod 1 if and only if for every complex-valued continuous function* f *on* $\mathbb{R}$ *at with period 1 we have*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} f(x_n) = \int_0^1 f(x)dx. \tag{2.5}$$

Lemma 2.2.7. *If the sequence* (x_n), $n = 1, 2, ...$ *is u.d. mod* 1*, then the sequence* $(x_n + \alpha)$, $n = 1, 2, ...$*, where* α *is a real constant, is u.d. mod* 1*.*

Theorem 2.2.8. *If the sequence* $(x_n), n = 1, 2, ...$*, is u.d. mod 1, and if* (y_n) *is a sequence with the property* $\lim_{n\to\infty}(x_n - y_n) = \alpha$*, a real constant, then* (y_n) *is u.d. mod* 1*.*

Theorem 2.2.9. *(Kronecker's theorem) If* α *is a given irrational number, then the sequence of numbers* $\{n\alpha\}_{n\geq 1}$ *is dense in the unit interval.*

Explicitly, given any θ, $0 \leq \theta \leq 1$, and given any $\epsilon \geq 0$, there exists a positive integer k such that $\mid k\alpha - \theta \mid \leq \epsilon$.

This fact is called Kronecker's theorem who proved it in the late 1840's. However this observation was first made by Nicoli Oresme (1340s) more than half a millennium ahead of Kronecker time.

This investigation was next advanced by the introduction in 1916 by Hermann Weyl of the concept of uniform distribution.

The functions f of the form $f(x) = e^{2\pi i h x}$, where h is a non-zero integer, satisfy the conditions of Corollary 2.2.6 Thus, if (x_n) is *u.d. mod* 1, the relation (2.4) will be satisfied for those f. It is one of the most important facts of the theory of *u.d. mod* 1 that these functions already suffice to determine the *u.d.mod* 1 of a sequence.

Definition 2.2.10. *A sequence of real numbers* $(x_n)_{n\geq 1}$ *is a uniformly distributed module one if given any interval* $I \subseteq [0, 1)$ *with length* $|I|$*, we have*

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

This is clearly a stronger statement than Kronecker's theorem.

Theorem 2.2.11. *(Weyl Criterion) The sequence* $(x_n), n = 1, 2, ...,$ *is u.d. mod 1 if and only if*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} e^{2\pi i h x_n} = 0. \tag{2.6}$$

for all integers $h \neq 0$

Definition 2.2.12. *The set* $\mathbf{E} \subset \mathbb{R}$ *is a null if given any* $\epsilon > 0$*, there is a countable collection of intervals* $(I_i)_{i\geq 1}$ *such that* $\mathbf{E} \subset \bigcup_i I_i$ *with* $\sum_i \mid I_i \mid \leq \epsilon$.

Theorem 2.2.13. *(Weyl's) For any strictly increasing sequence of integers* $(a_n)_n \geq 1$*, then the set of real numbers such that* $(\{a_n x\})_{n\geq 1}$ *is not u.d. mod 1 is null.*

Example 2.2.14. *Let* $a_n = b^n$ *for natural numbers* $b > 1$*. The set of such that* $(b^n x)n \geq 1$ *is not u.d. mod 1 also is null.*

Theorem 2.2.15. *The countable unions of null sets are null.*

Theorem 2.2.16. *(H. Weyl 1916) Let* $\psi(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \ldots + \alpha_k x^l$*, with at least one of the numbers* $\alpha_1, \ldots, \alpha_l$ *irrational.*

Then the sequence $(\psi(n))_{n\geq 1}$ *is uniform distribution modulo one.*

Definition 2.2.17. *Any compact abelian topological group* G *supports a unique translation invariant measure* λ *called Haar measure. We say* $(x_n)_{n\geq 1}$ *is uniformly distributed on* G *if for each continuous function* $f : G \to \mathbb{C}$ *we have*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} f(x_n) = \int_G f(x) d\lambda.$$

Remark 2.2.18. *There are two important special cases:*

1) *A sequence of real numbers* $(x_n)_{n\geq 1}$ *is uniformly distributed modulo one (u.d. mod 1) if given any interval* $I \subseteq [0, 1)$ *we have*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \chi_I(\{y_n\}) = |I|.$$

Here for a real number y we have used $\{y\}$ to denote its fractional part, χ_I to denote the characteristic function of the interval I and $|I|$ to denotes its length.

2) *We say a sequence of integers $(k_n)_{n\geq 1}$ is uniformly distributed on $\mathbb{Z}$ if for ever natural number $m \geq 2$ and every residue class a modulo m we have*

$$\lim_{N\to\infty} \frac{1}{N}\#\{n \leq N : k_n \equiv a \bmod m\} = \frac{1}{m}.$$

Definition 2.2.19. *A sequence of integers $(k_n)_{n\geq 1}$ is Hartman uniformly distributed on $\mathbb{Z}$ if for any irrational number α we have $(\{k_n\alpha\})_{n\geq 1}$ u.d.* mod 1 *and $(k_n)_{n\geq 1}$ is uniformly distributed on $\mathbb{Z}$.*

Note that if $(k_n)_{n\geq 0}$ is Hartman u.d. on $\mathbb{Z}$ and if and only if letting

$$F(N,z) := \frac{1}{N}\sum_{n=0}^{N-1} z^{k_n}, \qquad (N = 1, 2, \cdots)$$

we have $F(N,1) = 1$ for all $N \geq 1$ and if $|z| = 1$, $z \neq 1$ we have $\lim_{N\to\infty} F(N,z) = 0$.

So $(k_n)_{n\geq 1}$ is Hartman u.d. on $\mathbb{Z}$;

Definition 2.2.20. *Let x be a real number, and $b > 1$ be a natural number. Write*

$$x = [x] + \sum_{n=1}^{\infty} \frac{b_n}{b^n}$$

where $b_n \in \{0, 1, ..., b-1\}$.

We call this the base b expansion of x. In the case $b = 10$ this is the familiar decimal expansion.

If x is an irrational number, then the base b expansion of x is uniquely determined by the sequence $(b_n)_{n=1}^{\infty}$ and vice versa. Let x be an irrational number, and let b be a natural number. Suppose that $(b_n)_{n=1}^{\infty}$ be the base b expansion of x. Picking a number x at random, the chance of a particular digit taking a specific value is $\frac{1}{b}$. The chance of consecutive digits taking specific value is $\frac{1}{b^2}$. The chance of three consecutive digits taking specific a value is $\frac{1}{b^3}$ and so on.

The notion of uniform distribution can be easily generalized to compact spaces X.

Definition 2.2.21. *Let X be a compact (Hausdorff) space and μ a positive regular normalized Borel measure on X. A sequence $(x_n)_{n\geq l} \subset X$, is called uniformly distributed with respect to μ (μ-u.d.) if*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} f(x_n) = \int_X f d\mu. \tag{2.7}$$

holds for all continuous functions $f : X \to R$.

Definition 2.2.22. *A Borel set $M \subset X$ is called μ-continuity set if $\mu(\partial M) =$ 0, where ∂M denotes the boundary of M.*

Theorem 2.2.23. *[30, p. 178])A sequence $(x_n)_{n\geq 1}, x_n \in X$, is μ-u.d. if and only if*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \chi M(x_n) = \mu(M). \tag{2.8}$$

holds for all μ-continuity sets $M \subseteq X$.

Definition 2.2.24. *A collection $\mathcal{D}$ of μ-continuity sets of X is called discrepancy system if*

$$\lim_{N\to\infty} \sup_{M\in\mathcal{D}} \left| \frac{1}{N}\sum_{n=1}^{N} \chi_M(x_n) - \mu(M) \right| = 0$$

holds if and only if $(x_n)_{n\geq 1}$ is μ-u.d. For a discrepancy system $\mathcal{D}$ the supremum

$$D_N^{(\mathcal{D})}(x_n) = \sup_{M\in\mathcal{D}} \left| \frac{1}{N}\sum_{n=1}^{N} \chi_M(x_n) - \mu(M) \right|$$

is called discrepancy of $(x_n)_n \geq 1$ with respect to $\mathcal{D}$.

The discrepancy $\mathcal{D}_N$ of a sequence $(x_n)_n \geq 1$ has been introduced to quantify the convergence in (2.2).

Definition 2.2.25. *Let $x_1, ..., x_N$ be a finite sequence of points in the k-dimensional space $\mathbb{R}^k$. Then the number*

$$D_N = D_N(x_1, ..., x_N) = sup_{I \subseteq \mathbb{T}^k} \left| \frac{A(I, N, X_n)}{N} - \lambda_k(I) \right|$$

is called the discrepancy of the given sequence.

For an infinite sequence $(x_n)_n \geq 1$; $D_N(x_n)$ is also called the discrepancy of $(X_n)_{n=1}^N$ and is called discrepancy.

The essential point of the concept of discrepancy is that the notion of uniform distribution can be covered by it; i.e, the convergence in (2.2) is uniform with respect to all products of intervals $I \subseteq \mathbb{T}^k$.

Theorem 2.2.26. *A sequence $(x_n)_n \geq 1$ is u.d. mod 1 if and only if*

$$lim_{N \to \infty} D_N(X_n) = 0.$$

2.3 Ergodicity

In this section, we refer to Viana [54], Walters [63] primarily.

Definition 2.3.1. *Let $(X, \mathcal{B}, \mu)$ be a probability space, and let the map $T : X \to X$ measurable with respect to the σ-algebra $\mathcal{B}$. For $A \in \mathcal{B}$ we set $T^{-1}(A) = \{x \in X : Tx \in A\}$. The map T is said to be measure preserving with respect to μ if*

$\mu(T^{-1}A) = \mu(A)$ for all $A \in \mathcal{B}$.

Definition 2.3.2. *Let T be a measure preserving transformation on a probability space $(X, \mathcal{B}, \mu)$. The map T is said to be ergodic if for every measurable set A satisfying $T^{-1}A = A$, we have $\mu(A) = 0$ or 1.*

We say the dynamical system $(X, \mathcal{B}, \mu)$ is uniquely ergodic if μ is the only measure for which it is ergodic.

Theorem 2.3.3. *Let $(X, \mathcal{B}, \mu)$ be a probability space and $T : X \to X$ measure preserving. The following are equivalent:*
(i) T is ergodic.
(ii) If $B \in \mathcal{B}$ with $\mu(T^{-1}B \Delta B) = 0$, then $\mu(B) = 0$ or 1.
(iii) If $A \in \mathcal{B}$ with $\mu(A) > 0$, then $\mu(\bigcup_{n=1}^{\infty} T^{-n}A) = 1$.
(iv) If $A, B \in \mathcal{B}$ with $\mu(A) > 0$ and $\mu(B) > 0$, then there exists $n > 0$ such that $\mu(T^{-n}A \bigcap B) > 0$.

We denote by $L^0(X, \mathcal{F}, \mu)$ the space of all complex valued measurable functions on the probability space $(X, \mathcal{F}, \mu)$. Let

$$L^p(X, \mathcal{F}, \mu) = \{f \in L^0(X, \mathcal{F}, \mu) : \int_X |f|^p d\mu(x) < \infty\}.$$

Theorem 2.3.4. *Let $(X, \mathcal{F}, \mu)$ be a probability space, and $T : X \to X$ measure preserving. The following are equivalent:*
(i) T is ergodic.
(ii) If $f \in L^0(X, \mathcal{F}, \mu)$, with $f(Tx) = f(x)$ for all x, then f is a constant a.e.
(iii) If $f \in L^0(X, \mathcal{F}, \mu)$, with $f(Tx) = f(x)$ for a.e. x, then f is a constant a.e.
(iv) If $f \in L^2(X, \mathcal{F}, \mu)$, with $f(Tx) = f(x)$ for all x, then f is a constant a.e.
(v) If $f \in L^2(X, \mathcal{F}, \mu)$, with $f(Tx) = f(x)$ for a.e. x, then f is a constant a.e.

Theorem 2.3.5. *(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*

$$lim_{n \to \infty} \frac{1}{N} \sum_{n=1}^{N} f(T^n \alpha)$$

exists μ-almost everywhere in $\alpha \in X$. If $(X, \mathcal{B}, \mu)$ is an ergodic dynamical system, then the limit equals $\int_X f d\mu$ for μ-almost everywhere $\alpha \in X$.

Using Birkhoff's theorem and the ergodicity of the map $Tx = \{2x\}$ modulo 1 once can prove the following theorem.

Theorem 2.3.6. *(Borel's Theorem on Normal Numbers) Almost all numbers in $[0, 1)$ are normal to base 2, i.e. $x \in [0, 1)$ the frequency of $1's$ in the binary expansion of x is $\frac{1}{2}$.*

More generally we say a sequence is Hartman uniformly distributed on a locally compact abelian group G, if it is uniformly distributed on the Bohr compactification of G. This degree of generality is not relevant to our considerations in this thesis however. Hence, in the sequel, when we refer to Hartman uniform distribution, we mean Hartman uniform distribution on $\mathbb{Z}$. Note that if $(k_n)_{n\geq 0}$ is Hartman uniformly distributed, and if we set

$$F(N, z) := \frac{1}{N} \sum_{n=0}^{N-1} z^{k_n}, \qquad (N = 1, 2, \cdots)$$

we have $F(N, 1) = 1$ for all $N \geq 1$ and if $z \neq 1$ we have $\lim_{N\to\infty} F(N, z) = 0$. The converse is also true. A list of examples is given later.

Definition 2.3.7. *We say $(k_n)_{n\geq 0}$ is L^p good universal if for each dynamical system (X, β, μ, T) and for each $f \in L^p(X, \beta, \mu)$ the limit*

$$\ell_{T,f}(x) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f(T^{k_n} x),$$

exists μ almost everywhere.

One of central tool is the following lemma.

Lemma 2.3.8. *Suppose $(k_i)_{i=1}^{\infty}$ is Hartman uniformly distributed, suppose L^p-good universal for $p \in [1, 2]$ and suppose that the dynamical system (X, β, μ, T) is ergodic. Then $\ell_{T,f}(x)$ exists and equals $\int_X f d\mu$ μ almost everywhere.*

2.4 The geodesic flow

In this section we summerise what we need to know about the geodesic flow. See [17][54] for more detail.

Definition 2.4.1. *A hyperbolic surface is a smooth surface equipped with a complete Riemannian metric of constant Gaussian curvature* -1*.*

Up to isometry, there is a unique simply connected hyperbolic surface, called the hyperbolic plane, for which there are several standard models. The model we will use most frequently is the upper half-plane,

$$\mathcal{H} := \{z = x + iy \in \mathcal{C} : y > 0\}, ds^2 = \frac{dx^2 + dy^2}{y^2}$$

The other standard alternative is the unit disk model (or Poincaré disk),

$$\mathcal{B} := \{z \in \mathcal{C} : |z| < 1\}, ds^2 = 4\frac{dx^2 + dy^2}{(1 - |z|^2)^2}$$

Most calculations are simpler in $\mathcal{H}$, but $\mathcal{B}$ has the advantage that the boundary is treated uniformly.

We can identify $SL_2(\mathbb{R})/SO_2(\mathbb{R})$ with the Poincaré upper half plane $\mathcal{H}$ by identifying an element of $SL_2(\mathbb{R})$ with its action on the square root of minus one denoted by i. We can similarly identify the unit tangent bundle of $\mathcal{H}$ by $T = T_1(\mathcal{H}) := PSL_2(\mathbb{R}) = SL_2(\mathbb{R})/\pm I$. The hyperbolic metric on $\mathcal{H}$ is defined by $ds^2 = \frac{dx^2+y^2}{y^2}$. One can check that the points i and $e^t i$ are distance t apart in this metric. To a time $t > 0$, a base point in $\mathcal{H}$ and a unit vector with this base point, there s a unique geodesic passing throught the base point determined by the hyperbolic metric on $\mathcal{H}$. This determines the geodesic flow g_t on T by mapping a pair of a base point together with a unit tangent vector to another pair of a base point and a unit tangent, obtained by moving along the unique geodesic passing through both base points for time t and assigning as unit vector in the image the unique unit tangent vector to this geodesic at the image base point. This determines the geodesic flow g_t on T. Since T acts on $\mathcal{H}$ by isometries, another way to describe this is by saying that g_t takes $A \in T$ to AE_t with $E_t = \left(\begin{smallmatrix} e^{\frac{t}{2}} & 0 \\ 0 & e^{-\frac{t}{2}} \end{smallmatrix}\right)$. The Liouville measure on T is the hyperbolic area measure on $\mathcal{H}$ and the length measure on the circle of unit vectors at any point. This measure is left and right $SL_2(\mathbb{R})$ invariant and hence coincides with Haar measure on T. This measure is hence invariant under the geodesic flow. Recall that a Fuchsian group is a discrete subgroup of T. Equivalently Γ acts properly discontinuously on $\mathcal{H}$ and the quotient is a hyperbolic surface or possibly an orbifold. The unit tangent bundle of this space can be identified with $\Gamma\backslash T$. Much as the action of the geodesic flow on $\mathcal{H}$ sends $A \in T$ to AE_t, the action of the geodesic flow on $\Gamma\backslash T$ takes the coset to which A belongs to that to which AE_t belongs.

Let F be a division ring. A vector space of rank n over F can be identified with the standard space F^n (with scalars on the left) by choosing a basis. Any invertible linear transformation of V is then represented by an invertible $n \times n$ matrix, acting on F^n by right multiplication.

We let $GL(n, F)$ denote the group of all invertible $n \times n$ matrices over F, with the operation of matrix multiplication.

The group $GL(n, F)$ acts on the projective space $PG(n-1, F)$, since an invertible linear transformation maps a subspace to another subspace of the same dimension.

Definition 2.4.2. *The projective linear group and projective special linear group are the quotients of $GL(n, F)$ and $SL(n, F)$ by their centres:*

$$PGL(n, F) = GL(n, F)/Z(GL(n, F)) = GL(n, F)/F^{\times}$$

$$PSL(n, F) = SL(n, F)/Z(SL(n, F)) = SL(n, F)/\{\alpha \in F^{\times} | \alpha n = 1\}$$

These groups act on the $n-1$ dimensional *projective space* $P^{(n-1)}(F)$ (the set of one dimensional subspaces of F^n) by the following lemma.

Lemma 2.4.3. *The only elements of* $GL(n, F)$ *which stabilize every element of* $P^{(n-1)}(F)$ *are the scalar multiples of the identity matrix.*

Proposition 2.4.4. *The kernel of the action of* $GL(n, F)$ *on the set of points of* $PG(n-1, F)$ *is the subgroup*

$$\{cI : c \in (Z), c \neq 0\}$$

of central scalar matrices in F*, where* $Z(F)$ *denotes the centre of* F *.*

The projective linear group or modular group as we refer to $PSL(2, \mathbb{Z})$ as has long been known is intimately tied up with the properties of the regular continued fraction expansion. A Fuchsian group Γ is a discrete subgroup of $PSL(2, \mathbb{Z})$. From such a group one obtains a hyperbolic surface $\mathcal{H}\backslash\Gamma$. On a hyperbolic surface, an end is defined to be an equivalence class of limiting parallel rays. In the hyperbolic plane this is represented by a point on $\mathbb{R} \cup \{\infty\}$.

To a finite area hyperbolic triangle in $\mathcal{H}$ with a vertex on $\mathbb{R}\cup\{\infty\}$ and two sides identified by a generator of the stabilizer of Γ we associate a finite area triangle in $\mathcal{H}\backslash\Gamma$ one vertex of which is an end. Fix a finite index subgroup $H \subset \Gamma$. We also call this end a cusp (of H). For any cusp of H pick some parabolic fixed point p in this H orbit and the the width with respect to Γ of κ is the index $w(\kappa)$ of the H-stabilizer of p in the Γ-stabilizer of p.

We say a sequences of natural numbers $(k_n)_{n\geq 1}$ is good for the regular continued fraction if given H a finite index subgroup of a Fuchsian group $\Gamma \subseteq SL(2, \mathbb{Z})$. Then for each cusp κ of H and almost all $x \in [0, 1)$,

$$\lim_{N\to\infty} \frac{1}{N} \frac{\#\{1 \leq n \leq N : p_{k_n}/q_{k_n} \in \kappa\}}{N} = \frac{w(\kappa)}{[\Gamma : H]}.$$

2.5 Regular continued Fractions and the Gauss map

Using the Gauss map

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

we define a sequence of function of x by $c_0(x) = [x]$, $c_1(x) = [\frac{1}{x}]$ and $c_{n+1}(x) = c_n(g(x))$ $(n = 1 \dots)$, called the partial quotients of x. With these ingredienets we define the convergents

$$\frac{p_n}{q_n} = c_0 + \cfrac{1}{c_1 + \cfrac{1}{c_2 + \cfrac{}{c_3 + \ddots \cfrac{1}{+\cfrac{1}{c_n}}}}}, \qquad (n = 0, 1, \dots).$$

which is also written more compactly as $[c_0; c_1, c_2, \cdots, c_n]$. These have the property that $\frac{p_n}{q_n}$ tends to x as n tends to infinity. Recall that if we let $SL(2, \mathbb{Z})$ denote space of integer entry matricies in of determinant one and let $PSL(2, \mathbb{Z}) = SL(2, \mathbb{Z})/\pm I$ the set of equivalence classes of matricies in $SL(2, \mathbb{Z})$ identical up to multiplication by the negative 2×2 identity matrix I.

We now consider the particular ergodic properties of the Gauss map, defined on $[0, 1]$ by

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

Notice that $c_n(x) = c_{n-1}(g(x))$ $(n = 1, 2, \cdots)$. The dynamical system (X, β, μ, g) where X denotes $[0, 1]$, β is the σ-algebra of Borel sets on X, μ

is the measure on (X, β) defined for any A in β by

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

and g is the Gauss map is ergodic. See [17] for details. The ergodic properties of the dynamical system (X, β, μ, g) are not quite enough to carry out this investigation. We also need ergodic theoretic information about its natural extention. In particular we need the following lemma. See [54] for a formal definition of the natural extention.

Lemma 2.5.1. *Let* $\Omega = [0,1) \times [0,1]$. *Now let* γ *be the* σ*-algebra of Borel subsets of* Ω *and let* ω *be the probability measure on the measurable space* (Ω, β) *defined by*

$$\omega(A) \;=\; \frac{1}{(\log 2)} \int_A \frac{dxy}{(1+xy)^2}.$$

Also define the map

$$G(x,y) = \left(g(x), \frac{1}{[\frac{1}{x}] + y} \right).$$

Then the map G *preserves the measure* ω *and the dynamical system* $(\Omega, \beta, \omega, G)$ *is ergodic.*

In this section we discuss the language of suspension flows arising out of discrete transformations and the dual notion of discrete transformations on cross-sections viewed as return maps for flows. This language is a means for instance of linking the geodesic flow on a surface to the discrete continued fraction map on the boundry of the Poincaré disc. The languge is however much more general than that and useful in dynamical systems whereever flows and discrete maps interact. See [17] for further background.

Definition 2.5.2. *Suppose* (X, β, μ) *is a measure space and* Φ_t *is a measure preserving flow on* X*, that is* $\Phi : X \times \mathbb{R} \to X$ *is a measureable function such that* $\Phi_t(x) = \Phi(x,t)$ *with* $\Phi_{t+s} = \Phi_t \circ \Phi_s$. *A set* $\Sigma \subset X$ *is a cross-secton for*

the flow Φ_t *if :*
(i) the flow of almost every point meets Σ*;*
(ii) for almost every point x*, the set of times* t *that* $\Phi_t(x)$ *meets* Σ *is a discrete subset of* $\mathbb{R}$ *; and*
(iii) for every $t > 0$ *the flow box* $A_{[0,\tau]} \equiv \{\Phi_t(A)|A \subseteq \Sigma; t \in [0, \tau]\}$ *is* μ *measurable.*
The set of such flow boxes defines the β_Σ *of subsets of* Σ*.*

Definition 2.5.3. *The return time function for the flow to the set* Σ *is*

$$r = r_\Sigma = \inf\{t > 0 : \Phi_t(x) \in \Sigma\}$$

Evidently a finite return time is essential for interesting dynamics. It for example allws one to define a dynamical system on the cross-section. The flow regulates the return map dynamics and vice versa.

Definition 2.5.4. *The return time map* $R : \Sigma \to \Sigma$ *is defined* $R(x) := \Phi_{r(x)}(x)$*.*

There is also a cross-sectional measure induced on the cross-section preserving the return map by the measure preserved by the ambient flow. We next describe the suspension flow, and its relationship with the discrete map it covers. This is a simple tool relating the combinatorial nature of the discrete dynamical system and the flow dynamics.

Definition 2.5.5. *The induced measure* μ_Σ *defined for flow boxes on* Σ *is defined by setting* $\mu_\Sigma(A) := \frac{1}{\tau}\mu(A_{[0,\tau]})$ *for all* $\tau \in (0, \inf_{x\in A}\{r(x)\})$*.*

Definition 2.5.6. *The flow* (X, β, μ, Φ_t) *is then naturally isomorphic to the suspension flow over the return map* $(\Sigma, \beta_\Sigma, \mu_\Sigma, r)$ *defined by setting*

$$\hat{\Sigma} := \{(x, t) : 0 \leq r(x) \leq r(x)\}/\sim,$$

where $\sim$ *is the equivalence relation* $(x, r(x)) \sim (R(x), 0)$*, equipped with the measure* $\hat{\mu}$ *which is the the product of* μ_Σ *on* Σ *and Lebesgue measure on* $\mathbb{R}$*.*

Note that the flow on Φ^R on $\hat{\Sigma}$ defined by $\Phi_t^R : (x, s) \to (z, s+t)$ preserves $\hat{\mu}$.

We now define ergodicity for a flow. We now have E. Hopf's foundational result about

Lemma 2.5.7. *A flow Φ_t is ergodic if any invariant set or its complement is a null set. (See Definition 3.1.3 for an explaination of the term null set.)*

We now have E. Hopf's foundational result about ergodicity of the geodesic flow.

Lemma 2.5.8. *(E. Hopf's result) The geodesic flow on the unit tangent bundle of any finite volume hyperbolic manifold is ergodic.*

We now define the concept of recurrence for a flow.

Definition 2.5.9. *A flow is called recurrent if the Φ orbit of almost every point meets any positive measure set infinitely often.*

There is a version of the Poincaré recurrence property for flow too.

Theorem 2.5.10. *(The Poincaré recurrence theorem) Any geodesic flow on a finite volume hyperbolic manifold is recurrent.*

The return map to a cross-section inherits many properies from the flow. One is ergodicty.

Lemma 2.5.11. *Given a cross-section the return time map $(\Sigma, \beta_\Sigma, \mu_\Sigma, r)$ is ergodic if and only if the flow is.*

Semi-conjugacy is a very general structure transfering relationship between dynamical systems.

Lemma 2.5.12. *A semi-conjugacy, homeomorphism or factor map of two measure preserving flows or maps is an almost surely onto map that preserves the measures and the dynamics.*

Lemma 2.5.13. *If a conjugacy is in additional invertible, with an inverse map also a homeomorhism.*

We now specialise to the situation considered in this chapter.

Lemma 2.5.14. *Suppose* Γ *is a finite co-volume Fuchsian subgroup of the group* $P := PSL(2,\mathbb{R})$. *Suppose* $\mathbb{A} \subseteq P$ *projects injectively onto a cross-section of the geodesic flow that is uniformised by* Γ. *Then for almost all* $A \in \mathbb{A}$, *there is a unique pair of return time and* $t = t_A$ *and* $M \in \Gamma$ *such that* $MAE_t \in \mathbb{A}$.

Let $(\overline{X}, \overline{\beta}, \overline{\mu}, \overline{\Phi}_t)$ and (X, β, μ, Φ_t) be two measure preserving flows. We say that the second is a factor of the first if there is a map $\pi : \overline{X} \to X$ called the factor map between the two, which is measure preserving, onto and such that $\pi \circ \overline{\Phi}_t = \Phi_t \circ \pi$.

Lemma 2.5.15. *Let* $(\Sigma, \mu_r.R)$ *be a cross-section for* Φ_t. *Then* $\pi^{-1}(Z)$ *is a cross section of* $\overline{\Phi}_t$ *with return function* $\overline{r} = r \circ \pi$, *return map* $\overline{R}(x) = \Phi_{\overline{r}}(\overline{x})$ *and induced measure* $\overline{\mu}$. *The two transformations are semi-conjugate via the restrition of* π *to* $\overline{\Sigma}$.

2.6 Arnoux cross-sections and Fuchsian Groups

Given $(x, y) \in \mathbb{R}^2$ set

$$A_{-1}(x,y) = \begin{pmatrix} 1 & y \\ -x & 1-xy \end{pmatrix} \text{ and } A_{+1}(x,y) = \begin{pmatrix} x & -1 \\ 1-xy & y \end{pmatrix}$$

Let

$$Z : \mathbb{R}^2 \backslash \{(x,y) : y \neq -1/x\} \to \mathbb{R}^2,$$

given by

$$(x,y) \to (x, \frac{y}{1+xy}).$$

If Ω denotes the domain of the natural extention of the regular continued fraction map, we define the following subsets of $SL_2(\mathbb{R})$, by

$$C_{-1}\{A_{-1}(x,y) : (x,y) \in Z^{-1}(\Omega)\} \text{ and } C_{+1}\{A_{+1}(x,y) : (x,y) \in Z^{-1}(\Omega)\}.$$

Let $C = C_{-1} \cup C_{+1}$. We have the following lemma of Arnoux.

Lemma 2.6.1. *$C \subseteq P$ projects injectively to give a cross section of the geodesic flow on the unit tangent bundle of the modular surface hence also on $\Gamma\backslash P = PSL(s, \mathbb{Z})\backslash PLS(s, \mathbb{R})$. Further the return map for this cross-section is given by the map*

$$A_\sigma(x,y) = MA_\sigma(x,y)E_t$$

where $t = -2\log x$ and

$$M = \begin{pmatrix} 1 & [1/x] \\ 0 & 1 \end{pmatrix} \text{ if } \sigma = -1 \text{ and } M = \begin{pmatrix} 1 & 0 \\ [1/x] & 1 \end{pmatrix} \text{ if } \sigma = 1.$$

Moreover this dynamical system by this first return map has the Gauss map as a factor, via the projection to the first coordinate.

An element of $P = PSL(2, \mathbb{R})$ is called parabolic if it fixes a unique element of $\mathbb{R} \cup \infty$. For the corresponding $\Lambda = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ this is equivalent $|tr\Lambda| = |a+d| = 2$. If Γ is a Fuchsian subgroup of P then the Γ- orbit of its parabolic fixed points consists solely of its parabolic fixed points. We call such an orbit a cusp of Γ. It is known that the Γ-stabilizer of any parabolic fixed point is infinite cyclic and thus by conjugating in P if necessary the stabilizer is a cyclic group of translations. The cusp corresponds to a finite area end, of the hyperholic manifold $\Gamma\backslash\mathcal{H}$ also called a cusp. With this normalisation this geometric cusp is the projection of a hyperbolic triangle, with one vertex at infinity, where a generator of the stabilizer, acts so as to identify two sides of the triangle.

Fix some finite index subgroup $\Gamma \subset H$. For any cusp κ of H, choose a parabolic fixed point p in this H-orbit, and define the width of κ with respect Γ is the index $w(\kappa)$ of the H-stabilizer of p in the Γ stabilizer of p.

Lemma 2.6.2. *Suppse that* Γ *is a Fuchsian group with a single cusp and that* H *is a finite index subgroup of* Γ*. Then upon the choice of a parabolic fixed point of* Γ*, there is an equivalence equivalence relation on the right cosets* $H\backslash\Gamma$*, corresponding to the cusps of* H*. The size of an equivalence class is the width of the cusp.*

Given H and Γ is as in the previous lemma, together with a parabolic fixed point p of Γ, let $[H\gamma]_p$ denote the equivalence class of the right coset $H\gamma$ under the condition of equality of image of p and $\{\kappa\}_p$ denotes the partition block of $H\backslash\Gamma$ corresponding to κ under this equivalence relation. We refer to the relationship itself to the p relation of $H\backslash\Gamma$. We will also refer to the various equivalence classes of $H\backslash\Gamma$ defined by a p relation as the cusps of H. We will have cause to switch between a group H and its conjugate by $\tau = \left(\begin{smallmatrix}0 & -1\\ 1 & 0\end{smallmatrix}\right)$. Note that $\tau = \tau^{-1}$ in $PSL(2, \mathbb{R})$.

Lemma 2.6.3. *Suppse that* Γ *is a Fuchsian group of finite co-volume and the projection map with a single cusp such that* $\tau \in \Gamma$ *and* H *is a finite index subgroup of* Γ*. If* $\{\gamma_i\}_{1\leq i\leq n}$ *is a full set of representatives for* $\tau H\tau\backslash\Gamma$*. Then the* $\tau\gamma_i$ *is a full set of representatives of* $H\backslash\Gamma$*. Furthermore the map* $p \to \tau p$ *places the* $\tau H\tau$ *cusps in correspondence with the* H *cusps. This correspondence preserves cusp widths.*

2.7 Skew Products

Lemma 2.7.1. *Suppose* Γ *is a finite co-volume Fuchsian group and that a projection map* π *maps* $C \subset P$ *to a cross section for the geodesic flow on the* $\Gamma\backslash P \equiv T_1(\Gamma\backslash\mathcal{H})$*. Let* H *denote any finite index subgroup of* Γ *and* $H\backslash\Gamma$ *the set of right cosets of* H *in* Γ*. Then the skew product transformation* $\mathcal{S} : C \times H\backslash\Gamma \to C \times H\backslash\Gamma$ *defined by* $(A, \gamma \mod H) \to (MAE_t, \gamma M^{-1} \mod H)$*, where* M *and* E_t *are as in Lemma 1.5.15 , define a dynamical system naturally isomorphic to the first return map on the lifted cross-section* *on* $H\backslash\Gamma \equiv T_1(H\backslash\mathcal{H})$. *On both* $\Gamma\backslash P$ *and* $H\backslash P$ *we have the Liouville measure. and their cross-sections* Σ *and* $\overline{\Sigma}$ *support their induced measure. The measure on* C *is pushed onto* Σ *by the bijection and on the finite measure* $H\backslash\Gamma$ *goes to counting measure on* $\overline{\Sigma}$ *, with the skew product carrying the product of both of these.*

We know $\mathcal{S}$ must be ergodic as the invariant sets of a cross-section and it flow coincide and E. Hopf has shown the geodesic flow on a finite volume hyperbolic manifold is ergodic.

Lemma 2.7.2. *If* $C \subset P$ *projects injectively to give a cross-section with a return time* r *for the geodesic flow on the surface uniformised by* Γ*, then a measureable fundamental domain for the action* Γ *on* P *is given by the set*

$$\{AE_t : A \in C, 0 \leq t \leq r\}.$$

Suppose $\Gamma = PSL(2, , \mathbb{Z})$ and H is one of its finite index subgroups. Suppose C is an Arnoux cross section for the geodesic flow $\Gamma\backslash P$. The first return map for this cross-section maps the componant $\sigma = +1$ to the one with $\sigma = -1$ and vice versa.

Lemma 2.7.3. *Let* C *is a cross-section of the geodesic flow of the unit tangent bundle of the modular surface. Then* $proj_2(\mathcal{S}^k(A, I \mod H)) \equiv \begin{pmatrix} q_k & -q_{k-1} \\ -p_k & p_{k-1} \end{pmatrix} \mod H$ *if* k *is even and* $\equiv \begin{pmatrix} q_k & -q_{k-1} \\ -p_k & p_{k-1} \end{pmatrix} \mod H$ *if* k *is odd.*

Chapter 3

Pair Correlations and Random Walks on the Integers

The contents of this chapter are published in "Pair correlations and random walks on the integers". Unif. Distrib. Theory **11**, (2016), no. 1, 159-164 [45].

In this chapter we study the sequence of fractional parts of real numbers $(\{a_n x\})_{n=1}^{\infty}$ and give the conditions to satisfy a pair correlation estimate. Here x is a fixed non-zero real number and $(a_n)_{n=1}^{\infty}$ is a random walk on the integers. Definitions and results of this chapter can be found in [18], [34], [45],[58], [63]. For ease of readership and context we recap some of our prior discussion of uniform distribution in this chapter. The paper is essentially a published paper with extra context.

3.1 Introduction

For a real number y, let $[y]$ denote the greatest integer not greater than y and let $\{y\} = y - [y]$ denote its *fractional part*, or the residue of y modulo

1. We note that the fractional part of any real number is continued in the *unit interval* $I = [0, 1)$. Suppose α is an irrational number, the sequence of numbers $(\{n\alpha\})_{n\geq 1}$ is dense in the unit interval. Obviously, given any θ, $0 \leq \theta \leq 1$, and given any $\epsilon \geq 0$, there exists a positive integer k such that $\mid k\alpha - \theta \mid \leq \epsilon$. This fact is called Kronecker's theorem who proved it in the late 1840's. However this observation was first made by Nicoli Orsme (1340s) more than half a millennium ahead of Kronecker time. This investigation was next advanced by the introduction in 1916 by Hermann Weyl of the concept of uniform distribution.

Definition 3.1.1. *A sequence of real numbers* $(x_n)_{n\geq 1}$ *is a uniformly distributed module one if given any interval* $I \subseteq [0, 1)$ *with length* $\mid I \mid$*, we have*

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

This is clearly a stronger statement than Kronecker's theorem. It tells you how often the sequence $(x_n)_{n\geq 1}$ falls into a particular interval over the long term. In the last century, many sequences and families of sequences other than $(\{n\alpha\})_{n\geq 1}$ have had their uniformity of distribution established. For a real number y let $e(y) := e^{2\pi i y}$. A crucial tool in this regard is the following famous observation. See also Theorem 2.2.11.

Theorem 3.1.2. *(Weyl Criterion)*
The sequence $(x_n), n = 1, 2, ...,$ *is u.d. mod 1 if and only if*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} e(hx_n) = 0. \tag{3.1}$$

for all integers $h \neq 0$

We now consider the sequence $x_n = \{n\alpha\}$, where $\alpha \in \mathbb{R}$. If $\alpha \in \mathbb{Q}$, then $\alpha = \frac{p}{q}$ for some coprime integers p and q with $q \neq 0$. Now the sequence

$x_n = \{n\alpha\}$ becomes $x_1 = \frac{p}{q}, x_2 = \frac{2p}{q}, x_3 = \frac{3p}{q}, ..., \frac{(q-1)p}{q}, 0, ...$, so, in this case, this sequence cannot be uniformly distributed modulo 1, as it avoids any interval that does not include some element of the set $\{\frac{p}{q}, \frac{2p}{q}, ... \frac{(q-1)p}{q}, 0, ...\}$. Now if $\alpha \in \mathbb{R}\backslash\mathbb{Q}$, then summing the geometric progressions gives

$$\left| \frac{1}{N} \sum_{n=1}^{N} e(hn\alpha) \right| = \frac{1}{N} \left| \frac{e(h\alpha N) - 1}{e(h\alpha) - 1} \right|.$$

Since $|e(h\alpha N)| = 1$ and $|-1| = 1$, we have

$$\left| \frac{1}{N} \sum_{n=1}^{N} e(hn\alpha) \right| \leq \frac{2}{N} \left| \frac{e(h\alpha N) - 1}{e(h\alpha) - 1} \right|.$$

Moreover, $\alpha \in \mathbb{R}\backslash\mathbb{Q}$, so $e(h\alpha) - 1 \neq 0$ is a constant, and hence we obtain

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} e(hn\alpha) = 0$$

That is the sequence (x_n), with $x_n = \{n\alpha\}$ for $\alpha \in \mathbb{R}\backslash\mathbb{Q}$, is uniformly distributed modulo 1.

Using his method of differences, Weyl then refined his method to prove that if $P(x) = \alpha_k x^k + ... + \alpha_1 x + \alpha$, with at least one of the coefficients $\alpha_1, ..., \alpha_k$ irrational, we have $(\{P(n)\})_{n\geq 1}$ u.d. mod. 1.

Definition 3.1.3. *The set $E \subset \mathbb{R}$ is a null if given any $\epsilon > 0$, there is a countable collection of intervals $(I_i)_{i\geq 1}$ such that $E \subset \bigcup_i I_i$ with $\bigcup_i \sum_i \mid I_i \mid \leq \epsilon$.*

Null sets should be thought of as very small. Examples include finiite sets, countable sets and some fractals. Sets which contain a positive proportion of an interval are not null and should be thought of as big. Null set also called measure zeros sets and statments that are true expect on a null set are said to be true almost everywhere.

Another result of Weyl's is that given any strictly increasing sequence of integers $(a_n)_n \geq 1$, then the set of real numbers x such that $(\{a_n x\})_{n\geq 1}$ is not u.d. mod 1 is null.

Example 3.1.4. *Let $a_n = b^n$ for natural numbers $\{b > 1\}$. The set of such that $(b^n x)n \geq 1$ is not u.d. mod 1 also is null.*

3.2 Normal Numbers

Let x be a real number, and let b be a natural number. Write

$$x = [x] + \sum_{n=1}^{\infty} \frac{b_n}{b^n},$$

where $b_n \in \{0, 1, ..., b-1\}$. We call this the base b expansion of x. In the case $b = 10$ this is the familiar decimal expansion.

If x is an irrational number, then the base b expansion of x is uniquely determined by the sequence $(b_n)_{n=1}^{\infty}$ and vice versa. Let x be an irrational number, and let b be a natural number. Suppose that $(b_n)_{n=1}^{\infty}$ be the base b expansion of x. Picking a number x at random, the chance of a specific digit taking specific value is $\frac{1}{b}$. The chance of consecutive digits taking specific value is $\frac{1}{b^2}$. The chance of three consecutive digits taking specific value is $\frac{1}{b^3}$ and so on.

Definition 3.2.1. *We define*

$$N_M(x, a_1, ..., a_l) = \#\{i \leq M : b_i = a_1, ...b_{i+l-1} = a_l\}.$$

We say x is normal base b if

$$\lim_{M\to\infty} \frac{1}{M} N_M(x, a_1, ..., a_l) = \frac{1}{b^l}.$$

We say that a real number x is a *normal number* if x is normal base b for all $b \in \mathbb{N}\backslash\{1\}$, It is routine to check a real number x is normal base $b \in \mathbb{N} \backslash \{1\}$ if and only if the sequence $\{b^n x\})_{n=1}^{\infty}$ is uniformly distribtted modulo 1. Then using Weyl's theorem on $(\{a_n x\})_n \geq 1$ we have the following theorem. See also Theorem 2.2.15

Theorem 3.2.2. *The countable unions of null sets are null.*

In 1909 Borel had stated his normal number theorem that; Almost all numbers are normal.

Remark 3.2.3. *Even though almost all numbers are normal, no single explicit example of a normal number is known. The normality of well known numbers like $\sqrt{2}$ and e are open questions.*

3.3 Discrepancy of sequences

We begin by putting our result in a general framework. We are going to study the k-dimensional Euclidean space $\mathbb{R}^k$. We will select two points $x, y \in \mathbb{R}^k$ such that $x-y$ is an integral lattice point in $\mathbb{Z}^k$. That mean our investigation is in the k-dimensional torus $\mathbb{T}^k = \mathbb{R}^k \setminus \mathbb{Z}^k$ or the space $\mathbb{R}^k$ modulo 1.

Let $J = [a_l, b_1) \times ... \times [a_k, b_k) \subseteq \mathbb{R}^k$ where $0 < b_i - a_i \leq 1$, $i = 1, ..., k$ is a product of intervals in the k-dimensional space $\mathbb{R}^k$; the reduction modulo 1, $I = J \setminus \mathbb{Z}^k$ is called an interval of the torus $\mathbb{T}^k = \mathbb{R}^k \backslash \mathbb{Z}^k$. The volume $\lambda_k(I)$ of an interval $I \subseteq \mathbb{R}^k \setminus \mathbb{Z}^k$ is given by $\prod_{i=1}^{k}(b_i - a_i)$. Here λ_k denotes the k-dimensional Lebesgue measure.

We can generalise notation of uniform distribution to compact space X. See also Definition 2.2.21.

Definition 3.3.1. *Let X be a compact (Hausdorff) space and μ a positive regular normalized Borel measure on X. A sequence $(X_n)_{n \geq l}, X_n \in X$, is* *called uniformly distributed with respect to* μ *(*μ*-u.d.) if*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} f(x_n) = \int_X f d\mu. \tag{3.2}$$

holds for all continuous functions $f : X \to R$.

Definition 3.3.2. *A Borel set* $M \subset X$ *is called a* μ*-continuity set if* $\mu(\partial M) = 0$*, where* ∂M *denotes the boundary of* M*.*

See also Defininitions 2.2.22, 2.2.23 and Definition 2.2.24.

Theorem 3.3.3. *A sequence* $(X_n)_{n\geq 1}, x_n \in X$*, is* μ*-u.d. if and only if*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \chi_M(x_n) = \mu(M). \tag{3.3}$$

holds for all μ*-continuity sets* $M \subseteq X$*.*

Definition 3.3.4. *A system* $\mathbb{D}$ *of* μ*-continuity sets of* X *is called discrepancy system if*

$$\lim_{N\to\infty} \sup_{M\in\mathbb{D}} \left| \frac{1}{N}\sum_{n=1}^{N} \chi_M(x_n) - \mu(M) \right| = 0$$

holds if and only if $(x_n)_{n\geq 1}$ *is* μ*-u.d. For a discrepancy system* $\mathbb{D}$ *the supremum*

$$D_N^{(\mathbb{D})}(x_n) = \sup_{M\in\mathbb{D}} \left| \frac{1}{N}\sum_{n=1}^{N} \chi_M(x_n) - \mu(M) \right|$$

is called discrepancy of $(x_n)_{n\geq 1}$ *with respect to* $\mathbb{D}$ *.*

In order to quantify the convergence in (1.2) the discrepancy $\mathbb{D}_N$ of a sequence $(x_n)_{n\geq 1}$ has been introduced.

Definition 3.3.5. *Let* $x_1, ..., x_N$ *be a finite sequence of points in the* k*-dimensional space* $\mathbb{R}^k$*. Then the number*

$$D_N = D_N(x_1, ..., x_N) = \sup_{I\subseteq\mathbb{T}^k} \left| \frac{A(I, N, x_n)}{N} - \lambda_k(I) \right|$$

is called the discrepancy of the given sequence.

The reader should compair definitions 3.3.5 and 2.2.2. For an infinite sequence $(x_n)_n \geq 1$; $D_N(x_n)$ should denote the discrepancy of $(x_n)_{n=1}^N$ and is called discrepancy,too.

The essential point of the concept of discrepancy is that the notion of uniform distribution is now quantified; i.e, the convergence in (1.2) is uniform with respect to all intervals $I \subseteq \mathbb{T}^k$. See also Theorem 2.2.26.

Theorem 3.3.6. *A sequence $(X_n)_{n\geq 1}$ is u.d. mod 1 if and only if*

$$\lim_{N\to\infty} D_N(X_n) = 0.$$

A standard problem in numerical analysis, is estimating the integral of a function, through knowledge if its value at a finite number of points $(x_n)_{n=1}^N$. This is known as Monti Carlo estimation in the case of random sequence $(x_n)_{n=1}^N$ or Quasi-Monti Carlo estimation in the case of non-random $(x_n)_{n=1}^N$. This is encapsulated in the famous Denjoy-Koksma inequality

$$\left|\frac{1}{N}\sum_{n=1}^{N} f(\{x_n\}) - \int_0^1 f(x)dx\right| \leq V(f)D(x_1,\dots,x_N),$$

where for a function f on [0,1) with variation $V(f)$ and any finite set of points $x_1,\cdots,x_N$ in [0,1) with discrepancy

$$D_N = D(x_1,\dots,x_N) = \sup_{I\subseteq[0,1)}\left|\frac{1}{N}\#\{1\leq n\leq N : x_n \in I\} - |I|\right|.$$

Here the supremum is taken over all intervals $I \subseteq [0,1)$. Evidently to estimate $\int_0^1 f(x)dx$, sufficiently precisely what is needed is a good bound for D_N and a serviceable bound for $V(f)$, which is usually straight forward. To be useful this sequence$\{x_1,\cdots,x_N\}$ must be uniformly distributed module one. The discrepancy D_n is nothing other than a quantitative measure of uniformity of distribution. In particular, the sequence $(x_n)_{n\geq 1}$ is uniformly distribution modulo one if and only if $D_N \to 0$ as $N \to \infty$. In a sense the faster D_N decays as a function of N, the better uniformly distributed the sequence $(x_n)_{n\geq 1}$ is. One of the fundamental abstractions in nature in this subject is that there is that there is a limit to how well distributed any sequence can be. This is encapsulated in the elementary inequality $D_N \geq \frac{1}{N}(N = 1, 2, \cdots)$. Playing the roll of the Weyl's criterion in this context is the Erdős-Turn inqualty, which says that

$$D_N = D(x_1, \cdots, x_N) \leq 6 \left(r^{-1} + \sum_{h=1}^{r} \frac{1}{h} \left| \frac{1}{N} \sum_{n=1}^{N} e(hx_n) \right| \right).$$

With this framework in mind, in 1981, R.C. Baker proved for Weyl's sequence $(\{a_n x\})_{n\geq 1}$ that given $\epsilon > 0$

$$D(n, x) = D(\{a_1 x\}, \cdots, \{a_n x\}) = o(N^{\frac{1}{2}}(logN)^{\frac{3}{2}+\epsilon})$$

almost everywhere. Here for two sequences $(f_n), (g_n)$ by $f_n = o(g_n)$ means that $\lim_{n\to\infty} |\frac{g_n}{f_n}| = 0$
the next line of investigation is to study the pair correlation of the sequence, i.e the properties of the discrepancy.

$$V(n, x) := D(\{a_j x\} - \{a_i x\} : \text{such that } 1 \leq i \leq j \leq n).$$

In Weyl's sequence $(\{a_n x\})_{n\geq 1}$ in essence the sequence $(a_n)_{n\geq 1}$ is fixed and the number x is random what if these roles were reversed and x was fixed and non-zero and $(\{a_n\})_{n\geq 1}$ is random walk in the integers. This is the topic of our original investigation below. A full understanding of this discussion requires some familiarity with the language of analysis and number theory.

3.4 Pair correlation and random walk

Definition 3.4.1. *Let $\{X_i\}_{i=1}^{\infty}$ be a sequence of independent, identically distributed discrete random variables. For each positive integer n, we let a_n denote the sum $X_1 + X_2 + ... + X_n$. The sequence $\{a_n\}_{n=1}^{\infty}$ is called a random walk.*

If the common range of the X_i's is $\mathbb{R}^k$, then we say that $\{a_n\}$ is a random walk in $\mathbb{R}^k$.

Let X be a $\mathbb{Z}$-valued function defined on the probability space (Ω, β, P) with characteristic function $\phi(\xi) = \mathbb{E}(e^{iX(.)\xi})$ and let $\chi = \{X_n : n \geq 1\}$ be a sequence of independent copies of X. For a positive integer $n > 0$ let $a_n = X_1 + \cdots + X_n$ and let $a_0 \equiv 0$. The sequence of integers $(a_n)_{n\geq 1}$ is the random walk which we assume to satisfy $|\phi(t) - 1| \geq C|t|$, for some $C > 0$. This last property follows, for instance, if the random walk and its absolute value have finite non-zero mean [49, p. 62]. In [57] the condition $|\phi(t) - 1| \geq C|t|$ is said to follow from the assumption that the random is aperiodic and transient—a claim the author was unable to verify. This is then used to deduce a discrepancy estimate for the sequence $(X_n(x))_{n=1}^{\infty}$. This is the case, for instance (as a consequence of the law of large numbers) if $\mathbb{E}|X| < \infty$ and $\mathbb{E}X \neq 0$ or if X is centred and $\frac{a_n}{n^{\frac{1}{\alpha}}}$ converges in distribution to F_α, a stable law of index $\alpha \in (0, 1)$. This second class of examples can be deduced using a local limit theorem of Stone [49].

For a real number x let $X_n(x) = a_n x$. For an interval I let $\chi_I(x)$ denote the characteristic function of the set I. This means that we have

$$\chi_I(x) = 1 \quad \text{if } \ x \in I \quad \text{and} \quad \chi_I(x) = 0 \quad \text{otherwise.}$$

For a real number y let $\{y\}$ denote its fractional part.

$$V_N(I)(x) = \sum_{1\leq n<m\leq N} \chi_I\left(\{X_n(x) - X_m(x)\}\right)$$

and then define

$$\Delta_N(x) = \sup_{I\subseteq \mathbb{T}} \left| V_N(I) - \frac{N(N-1)}{2}\lambda_1(I) \right|,$$

where the supremum is over all intervals I in the one dimensional torus $\mathbb{T}$. Let $||x|| = \min_{n\in\mathbb{Z}} |x - n| = \min(\{x\}, \{1 - x\})$ let η be a positive real number or infinity. The irrational number x is said to be of type η if η is the supremum of all γ for which $\liminf_{q\to\infty} q^\gamma ||qx|| = 0$.

We also say a real number x is v badly approximable if there exists a constant $c(x, v > 0$ such that

$$|x - \frac{p}{q}| \geq \frac{c(x,v)}{q^{1+v}}.$$

Using Dirichlet's theorem on diophantine approximation we can deduce for all irrational x that $\liminf_{q\to\infty} q^{\gamma}||qx|| = 0$ so $\eta \geq 1$.

On the other hand, the Thue-Siegel-Roth theorem tells us that for every irrational algebraic x and every $\epsilon > 0$ there exists a constant $c(x, \epsilon) > 0$ such that

$$\left| x - \frac{p}{q} \right| \geq \frac{c(x,\epsilon)}{q^{2+\epsilon}}$$

holds for all coprime integers $q > 0$ and p, so that algebraic η must be of type 1. Lioville numbers can easily be used to show constructively that there exist real numbers that of type strictly greater than 1.

Our theorem is the following.

Theorem 3.4.2. *Suppose* $(X_n(x))_{n=1}^{\infty}$ *is as described above that* x *has type* $\eta > 1$. *Then given* $\epsilon > 0$,

$$\Delta_N(x) = o\left(N^{2-\frac{1}{\eta}+\epsilon}\right)$$

for P *almost all* $\omega \in \Omega$.

Let $D_N(x)$ denote the N-term discrepancy of the sequence $(X_n(x))_{n\geq 1}$. See[33, p 88] for the definition. M. Weber [66, p 411] has given an estimate for almost everywhere behaviour of $D_N(x)$ as N tends to infinity in terms of the type of x and the properties of the the function ϕ. The formulation is however somewhat involved and forgone here. Results like our theorem, where $(a_n)_{n=1}^{\infty}$ is fixed and deterministic and x is random are now known. See [29] for details and further background.

We proceed by a series of lemmas. For real x let $e(x) = e^{2\pi i x}$ and let

$$\theta_N(h) = \sum_{n=1}^{N} e(ha_n x) \qquad (N = 1, 2, \cdots).$$

We need the following lemma taken for page 412 of [58].

Lemma 3.4.3. *For integers $N \geq R \geq 1$ one has*

$$\mathbb{E}|\theta_N(m) - \theta_R(m)|^2 \leq \min\left(\frac{7(N-R)}{|\phi(mx)-1|}, N-R\right).$$

Proof. An elementary computation shows for integers $N \geq P \geq 1$

$$\mathbb{E}|\theta_N(m) - \theta_P(m)|^2 =$$

$$(N-P) + (N-P+1)[\phi(mx) - \phi(-mx)] + (N-P-2)[\phi(mx)^2 + \phi(-mx)^2]$$
$$+ \cdots + [\phi(mx)^{N-P-1} + \phi(-mx)^{N-P-1}].$$

But

$$\phi(mx)^k + \phi(-mx)^k = \mathbb{E}(e^{imxa_k} + e^{-imxa_k})$$

$$= 2\mathbb{E}\cos mxa_k = 2\mathbb{R}e(\mathbb{E}(e^{imxa_k})) = 2\mathbb{R}e(\phi(mx)^k)$$

for any $k \geq 1$. Thus

$$\mathbb{E}|\theta_N(m) - \theta_P(m)|^2 =$$

$$(N-P) + 2\mathbb{R}e\{(N-P-1)\phi(mx) + (N-P-2)\phi(mx)^2 + \cdots + \phi(mx)^{N-P-1}\}.$$

For any $z \in \mathbb{C}$ and $Q \in \mathbb{N}$

$$\sum_{d=1}^{Q-1}(Q-d)z^d = Qz^{Q-1} - \frac{Q}{z-1} + \frac{z^Q - 1}{(z-1)}.$$

Therefore if $\mathbb{R}$ denotes the real part

$$\mathbb{E}|\theta_N(m) - \theta_P(m)|^2 =$$

$$(N-P) + 2\mathbb{R}\left\{(N-P)\phi(mx)^{N-P-1} - \frac{N-P}{\phi(mx)-1} + \frac{\phi(mx)^{N-P}-1}{(\phi(mx)-1)^2}\right\}$$
$$\leq \left(\frac{7(N-P)}{|\phi(mx)-1|}\right).$$

□

Let $(Y_t)_{t=1}^{\infty}$ be a sequence of measurable functions defined on a measure space Ω and then write

$$S_j = \sum_{1 \le t \le\ j} Y_t, \qquad\qquad \text{for } j{=}1, 2, \cdots$$

We can define

$$Y_{rs} = \sum_{r \le t < s} Y_t (= S_s - S_r), \quad \text{for } r{<}s,$$

and let

$$M_n = \sup_{1 \le j \le n} |S_j|.$$

We have the following elementary lemma proved in [38].

Lemma 3.4.4. *For* $K \ge 1$,

$$\int_{\Omega} M_{2^K}^2(\omega)\delta\omega \le (K+1)\left(\sum_{i=1}^{K+1}\sum_{\nu=1}^{2^i-1}\int_{\Omega} |Y_{\nu 2^{(K+1)-i},\,(\nu+1)2^{(K+1)-i}}|^2(\omega)\delta\omega\right).$$

Thus if $K = 1, 2, \cdots$

$$\begin{aligned}
\mathbb{E}|\max_{1 \le j \le 2^K} \theta_j(m)|^2 &\le (K+1)\left(\sum_{i=1}^{K+1}\sum_{\nu=1}^{2^i-1} \mathbb{E}|\theta_{\nu 2^{(K+1)-i}}(m) - \theta_{(\nu+1)2^{(K+1)-i}}(m)|^2\right)\\
&\le (K+1)\left(\sum_{i=1}^{K+1}\sum_{\nu=1}^{2^i-1}\left(\frac{7.2^{(K+1)-i}}{|\phi(mx)-1|}\right)\right)\\
&\le (K+1)\left(\sum_{i=1}^{K+1} 2^{i-1}\left(\frac{7.2^{(K+1)-i}}{|\phi(mx)-1|}\right)\right)\\
&\le \frac{7}{2}(K+1)^2\left(\frac{2^{(K+1)}}{|\phi(mx)-1|}\right).
\end{aligned}$$

Thus, using the Erdős-Turan inequality [34, p 112-4], we can show that for $L \ge 1$, there exists $C > 0$,

$$\mathbb{E}|\max_{1\le j\le 2^K}\Delta_j(x)| \le C\left(\frac{2^{2(K+1)}}{L}+\sum_{h=1}^{L}\frac{1}{h}\left(2^{K+1}+\mathbb{E}\max_{1\le j\le 2^K}|\theta_j(h)|^2\right)\right)$$

$$\le C\left(\frac{2^{2(K+1)}}{L}+\sum_{h=1}^{L}\frac{1}{h}\left(2^{K+1}+\frac{7}{2}(K+1)^2\left(\frac{2^{(K+1)}}{|\phi(hx)-1|}\right)\right)\right). \tag{3.4}$$

Let $\Lambda:\mathbb{R}_+\to\mathbb{R}_+$ be a non-decreasing and such that for any $m\in\mathbb{N}$,

$$\sum_{h=1}^{L}\frac{1}{h|\phi(hx)-1|}\le\Lambda(L).$$

Then the left hand side of (34)

$$\ll\left(\frac{2^{2(K+1)}}{L}+(\log L)2^{(K+1)}+\Lambda(L)(K+1)^2 2^{(K+1)}\right). \tag{3.5}$$

Here, of course, $\ll$ denotes Vinogradov order notation. Recall that there exists

$$C>0 \quad \text{such that} \quad |1-\phi(t)|\ge C|t|.$$

Since

$$\phi(hx)=\phi(\{hx\}),$$

we therefore have

$$\sum_{h=1}^{L}\frac{1}{h|\phi(hx)-1|}=O\left(\sum_{h=1}^{L}\frac{1}{h||hx||}\right).$$

If x is irrational of type $\eta>1$ [33, p. 123, Lemma 3.3] for any $\epsilon>0$, then

$$\sum_{h=1}^{L}\frac{1}{h||hx||}=O(L^{\eta-1+\epsilon}).$$

In consequence, we can choose $\Lambda(L)=L^{\eta-1+\epsilon}$ the right hand side of (3.5) is

$$\ll\frac{2^{2(K+1)}}{L}+(\log L)2^{(K+1)}+\Lambda(L)(K+1)^2 2^{(K+1)}. \tag{3.6}$$

Choosing L essentially optimally $2^K\approx L\Lambda(L)=L^{\eta+\epsilon}$ the right hand side of (3.5)is

$$\ll 2^{(K+1)(2-\frac{1}{\eta}+\epsilon)}(K+1)^2. \tag{3.7}$$

Now we will complete the proof of our Theorem 3.4.1.

Proof. Given $\epsilon, \epsilon_0 > 0$, we define

$$E_{\epsilon,\epsilon_0} = \{\omega \in \Omega : \Delta_N(\omega, x) > N^{2-\frac{1}{\eta}+\epsilon}(\log N)^{3+\epsilon_0} \text{ for infinitely many } N\}.$$

We now proceed to show the P measure of E_{ϵ,ϵ_0} is zero. If we denote

$$A_K = \left\{\omega \in \Omega : \max_{2^{K-1}\le m<2^K} \Delta_m(\omega, x) > 2^{K(2-\frac{1}{\eta}+\epsilon)} K^{3+\epsilon_0}\right\} \quad \text{for each} \quad K \geq 1,$$

then one easily sees that

$$E_{\epsilon,\epsilon_0} \subseteq \bigcap_{r=1}^{\infty} \bigcup_{K=r}^{\infty} A_K.$$

Using (3.7) we can bound

$$\begin{aligned}
\mathrm{P}(A_K) &\le \frac{\mathbb{E}\,|\max_{2^{K-1}\le m<2^K} \Delta_m(x)|}{2^{K(2-\frac{1}{\eta}+\epsilon)} K^{3+\epsilon_0}} \\
&\le \frac{C 2^{K(2-\frac{1}{\eta}+\epsilon)} K^2}{K^{3+\epsilon_0} 2^{K(2-\frac{1}{\eta}+\epsilon)}} \\
&\le C K^{-(1+\epsilon_0)} \quad \text{for sufficiently large } C > 0.
\end{aligned}$$

In particular, now we can observe that

$$\sum_{K=1}^{\infty} \mathrm{P}(A_K) \le \sum_{K=1}^{\infty} K^{-(1+\epsilon_0)} < +\infty.$$

It follows from the Borel-Cantelli lemma that $\mathrm{P}(E_{\epsilon,\epsilon_0}) = 0$. □

We hope the result considered in this chapter can be extended the more general frameworks described in the start of this chapter.

Chapter 4

On Uniform Distribution of Polynomials and Good Universality

The contents of this chapter have been accepted for publication in Ergodic Theory and Dynamical Systems. For ease of readership and completeness we recap ideas for ergodic theory and uniform distribution theory

4.1 Introduction

Since the proof of Weyl's theorem [64] now over a century ago, innumerable examples of uniformly distributed sequences have been studied. See the standard references [18] and [34] for more background. Dispite its age, Weyl's theorem remains of interest, in part because of its applictions. Most proofs of Weyl's theorem rely on exponential sums. An ergodic theoretic proof of Weyl's theorem, making no use of exponential sums, was given by the ergodic theorist H. Furstenberg [17]. This is the point of view taken in this chapter. That said, exponential sums will also play a role in this chapter.

Inspired by K.F. Roth's proof that a set of integers of positive density contains a three term arithmetic progression [Ch. 10, 53], A. Sárközy [53] showed that any set of integers of positive density contains two terms whose difference is a square. Later, in stages, he also considered more general polynomials than squares and also a primes minus one [54][55].

Soon thereafter H. Furstenberg used Weyl's theorem and an ergodic method to prove that if $P(x)$ is a non-constant polynomial with integer coefficiants such that $P(0) = 0$ then $(P(n))_{n\geq 1}$ is a set of recurrence [22].

Let $(p_n)_{n\geq 1}$ denote the sequence of prime numbers. In [38] polynomials such that $(P(p_n))_{n\geq 1}$ and $(P(n))_{n\geq 1}$ are sets of strong recurrence are fully classified. The primary tools used to prove the first authors single recurrence results are a method of Furstenberg, together with the uniform distribution of the sequences $(\psi(p_n))_{n\geq 1}$ and $(\psi(n))_{n\geq 1}$. The extension of these results to $\mathbb{Z}^r$ for $r \in \mathbb{N}$ rather than just $\mathbb{Z}$ is routine and for this reason not discussed in either [38] or [40]. The modifications amount to working with characters of $\mathbb{Z}^r$ rather than $\mathbb{Z}$ which work very similarly. In [9] it is shown that if η is a polynomial with integer coefficients mapping the natural numbers to themselves, then $(\eta(p_n - 1))_{n\geq 1}$ strongly Poincaré recurrent. Evidently this is just a special case of Theorem 1 of [36]. There is now of course a literature on the multiple recurrence of these sequences, too extensive to conveniently discuss here.

A sequence in a second countable locally compact group is called Hartman uniformly distributed if it is uniformly distributed on the Bohr compactification of the locally compact group. In [40] it is shown that any such sequence is strongly recurrent. In the special cases of $\mathbb{Z}$ and $\mathbb{R}$ this observation is anticipated in [1] and [14]. As is shown in [35], one can use Weyl's criterion on the Bohr compactification to constuct examples. In the particular case of $\mathbb{Z}$, the methods of analytic number theory are used to constuct a number of families of strongly Poincaré recurrent sequences. See Theorem

4 in [40] for details of this. Let $(\alpha_1, \ldots, \alpha_r)$ be an r-tuple of non-integer real numbers all greater than 1. Then $([{p_n}^{\alpha_1}], \ldots, [{p_n}^{\alpha_r}])_{n\geq 1}$ is shown in [7] to be strongly Poincaré recurrent. In fact, while not discussed in [38], the sequence $([{p_n}^{\alpha_1}], \ldots, [{p_n}^{\alpha_r}])_{n\geq 1}$ is Hartman uniformly distributed on $\mathbb{Z}^r$, so the fact that it is strongly recurrent follows which follows from Theorem 1 in [40].

In [33] we adapt the ideas of H. Furstenberg to study the Poincaré recurrence and intersectivity phenomenon in positive characteristic. In this context the analogue of the integers is the ring of polynomials $\mathbb{F}_q[x]$ over the finite field $\mathbb{F}_q$ in characteristic q. We show in [33] that the set of irreducibles in $\mathbb{F}_q[x]$, which is the analogue of the primes in $\mathbb{Z}$, once shifted by one, is both a set of strong recurrence and a set of intersectivity. Poincaré recurrent sets are intersective sets in this context. Whether the converse is true is unknown to the authors. In [33] it is shown that if $q = p^r$ for some natural number r. and exceed the degree of a polynomial P over $\mathbb{F}_q$ then $\{P(x) : x \in \mathbb{F}_q\}$ is also both a set of strong recurrence and a set of intersectivity. Later in [10] this restriction on degree is removed, by proving the analogue of Weyl's Theorem in positive characteristic.

In [2] it is shown that if P is a non-constant polynomial mapping the natural numbers to themselves, then the sets $\{P(n) : n \in \mathbb{N}\}$ and $\{p_n : n \in \mathbb{N}\}$, are Glasner. In [2] there is also an emphersis on the quantitative forms of results. The methods in [2] are Fourier analytic and this is what allows these more quantitative forms of the result. See also [7] for earlier results. In [45] it is shown that $\{P(p_n) : n \in \mathbb{N}\}$ is Glasner and in [27] it is shown that any sequence that is Hartman uniformly distributed on $\mathbb{Z}$ is again Glasner. Here again, the methods are Fourier analytic and the theorems quantitative.

4.2 Good Universal

Let $(X, \mathcal{B}, \mu)$ be a probability space and let $T : X \to X$ be a measurable map, that is also measure-preserving. That is, given $A \in \mathcal{B}$, we have $\mu(T^{-1}A) = \mu(A)$, where $T^{-1}A$ denotes the set $\{x \in X : Tx \in A\}$. We call $(X, \mathcal{B}, \mu, T)$ a dynamical system. We say the dynamical system is ergodic if $T^{-1}A = A$ for $A \in \mathcal{B}$ means that either $\mu(A)$ or $\mu(X \backslash A)$ is 0. For instance; $Tx = \{x + \alpha\}$ on $[0, 1)$ is ergodic for irrational α. The dynamical system (X, β, μ, T) is said to be uniquely ergodic if μ is the only measure for which it is ergodic.

We say $(k_n)_{n\geq 0} \in \mathbb{N}$ is L^p good universal if for each dynamical system $(X, \mathcal{B}, \mu, T)$ and for each $f \in L^p(X, \beta, \mu)$ the limit

$$\ell_{T,f}(x) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f(T^{k_n}x),$$

exists μ almost everywhere.

In the sequel we will call $(k_n)_{n\geq 1}$ good if it is L^∞-good universal and we have $\ell_{T,f}(Tx) = \ell_{T,f}(x)$ μ almost everywhere.

Recall we say an infinite sequence sequence $x_1, \ldots, x_N, \ldots$ is uniformly distributed modulo one if

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

for every interval $I \subseteq [0, 1)$. For a real number y we have used $\{y\}$ to denote its fractional part and let $[y] = y - \{y\}$ denote its integer part.

In this section we prove the following theorem,

Theorem 4.2.1. *Let $\psi(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \ldots + \alpha_k x^k$, with at least one of the numbers $\alpha_1, \ldots, \alpha_l$ irrational. Then the sequence $(\psi(k_n))_{n\geq 1}$ is uniform distribution modulo one if $(k_n)_{n\geq 1}$ is good.*

The theorem is proved using H. Furstenberg's skew product method [17]. This seems to be the first new family of uniformly distributed sequences provided by this method.

We say a sequence of natural numbers $(k_n)_{n\geq 1}$ is Hartman uniform distributed (on $\mathbb{Z}$) if it is uniformly distributed in residue classes modulo m, for each natural number $m > 1$, and for each irrational number α, the sequence $(\{k_n\alpha\})_{n\geq 1}$ is uniformly distributed modulo one.

This condition coincides with $(k_n)_{n\geq 1}$ being uniformly distributed on the Bohr compactification of $\mathbb{Z}$. Note that if $(k_n)_{n\geq 1}$ is Hartman uniformly distributed on $\mathbb{Z}$, and if we set

$$F(N,z) := \frac{1}{N}\sum_{n=0}^{N-1} z^{k_n}, \qquad (N = 1, 2, \cdots)$$

then

$$\lim_{N\to\infty} F(N,z) = \begin{cases} 1, & \text{if } z = 1 \\ 0, & \text{if } z \neq 1. \end{cases}$$

The converse is also true. See Example 5.11 on page 296 in [33]. In the sequel when we say a sequence is Hartman uniformly distributed we mean it is Hartman uniformly distributed on $\mathbb{Z}$.

The following lemma is a Corollary 3 from [13].

Lemma 4.2.2. *The following are equivalent.*

(i) $(k_n)_{n\geq 1}$ *is Hartman uniformly distributed,*

and

(ii) for any dynamical system $(X, \mathcal{B}, \mu, T)$ *and* $f \in L^2(X, \mathcal{B}, \mu)$*, if* $P_T f$ *denotes the projection of* f *onto the* T *invariant subspace of* $L^2(X, \mathcal{B}, \mu)$ *for some* $p \geq 1$*, we have*

$$\lim_{N\to\infty} \left\| \frac{1}{N}\sum_{n=0}^{N-1} f(T^{k_n}x) - P_T f \right\|_2 = 0.$$

The following lemma is a special case of a theorem due to S. Sawyer [56].

Lemma 4.2.3. *For a dynamical system (X, β, μ, T), suppose that $f \in L^2(X, \beta, \mu)$. Set*

$$Mf(x) = \sup_{N \geq 1} \left| \frac{1}{N} \sum_{n=1}^{N} f(T^{k_n} x) \right|. \qquad (N = 1, 2, \dots)$$

If $(k_n)_{n\geq1}$ is L^q-good universal for some $q > 1$, then there exists $C > 0$ such that $||Mf||_q \leq C||f||_q$.

A central tool of ours is the following lemma.

Lemma 4.2.4. *Suppose $(k_i)_{i=1}^{\infty}$ is L^p-good universal for $p \in (1, 2]$. Then $(k_i)_{i=1}^{\infty}$ is Hartman uniformly distributed if and only if $\ell_{T,f}(Tx) = \ell_{T,f}(x)$ μ almost every where for any dynamical system (X, β, μ, T). Further $\ell_{T,f}(x) = \int_X f d\mu$ μ almost everywhere if (X, β, μ, T) is ergodic .*

The sequence $k_n = n^2$ $(n = 1, 2, \dots)$ is L^p good universal for all $p > 1$ [4] but not Hartman uniform distributed. To see this note that a square integer is never congruent to 3 modulo 4. If we take $X = \mathbb{Z}_p$ – the p adic integers, set $\mathcal{B}$ to be the Haar σ-algebra on $\mathbb{Z}_p$, set μ to be Haar measure on $\mathbb{Z}_p$ and define $T : \mathbb{Z}_p \to \mathbb{Z}_p$ by $Tx = x + 1$. The dynamical system we obtain is the p-adic adding machine, which is ergodic. For $f \in L^p(\mathbb{Z}_p)$ it is possible using Fourier analysis to calculate

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(x + n^2)$$

almost everywhere, which we know must exist because $k_n = n^2$ $(n = 1, 2, \dots)$ is L^p good universal. The limit is however not the Haar integral of f on $\mathbb{Z}_p$ as you might expect but rather a more complicated expression involving Fourier multipliers and Gauss sums. See [4] for these details.. This means that for the limit $\ell_{T,f}(x)$ for squares to be the integral of f we need more than ergodicty. The ergodicty of all powers of T is a sufficient condition [58].

Proof of Lemma 4.2.4: The assertion that $\ell(x)$ is invariant for ergodic (X, β, μ, T) immediately implies $\ell(x) = \int_X f d\mu$.

Because $\left|\frac{1}{N}\sum_{n=1}^{N} f(T^{a_n}x)\right| \leq Mf(x)$ $(N = 1, 2, \dots)$ 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. Lemma 4.2.2 tells us that $g(x)$ is T invariant if and only if $(k_n)_{n\geq 1}$ is Hartman uniformly distributed.

All we have to do now is show the pointwise limit is the same as the norm limit, i.e. that $\ell_{T,f}(x) = g(x)$ μ almost everywhere. We consider a sequence of natural numbers $(N_t)_{t\geq 1}$ such that

$$\left\|\frac{1}{N_t}\sum_{n=1}^{N_t} f(T^{k_n}x) - g(x)\right\|_q \leq \frac{1}{t}.$$

Thus

$$\sum_{t=1}^{\infty}\int_X \left|\frac{1}{N_t}\sum_{n=1}^{N_t} f(T^{k_n}x) - g(x)\right|^q d\mu < \infty.$$

Fatou's lemma tells us that

$$\int_X \left(\sum_{t=1}^{\infty}\left|\frac{1}{N_t}\sum_{n=1}^{N_t} f(T^{k_n}x) - g(x)\right|^q\right) d\mu < \infty,$$

which implies that

$$\sum_{t=1}^{\infty}\left|\frac{1}{N_t}\sum_{n=1}^{N_t} f(T^{k_n}x) - g(x)\right|^q < \infty,$$

almost everywhere. This means that

$$\left|\frac{1}{N_t}\sum_{n=1}^{N_t} f(T^{k_n}x) - g(x)\right| = o(1),$$

μ almost everywhere. As $(k_n)_{n\geq 1}$ is L^q-good universal we must have $\ell_{T,f}(x) = g(x)$ μ almost everywhere. □

We also use the following lemma from [30], which in the case $k_n = n(n = 1, 2, \dots)$ is classical and due to J. C. Oxtoby [50].

Lemma 4.2.5. *Suppose $(k_n)_{n\geq 1}$ is Hartman uniformly distributed and L^p-good universal. Let T be a continuous map of a compact metrizable space X. Also let μ denote a measure defined on a σ-algebra $\mathcal{B}$ of subsets of X. The following statements are equivalent:*

a) the transformation $(X, \mathcal{B}, \mu, T)$ is uniquely ergodic;

b) whenever f is a continuous function on X we have

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

pointwise on X, i.e. for all $x \in X$.

Let δ_a denote the Kronecker delta function at a. We say $(x_n)_{n\geq 1} \subseteq X$ for compact metric space X is asymptotically distributed with respect to a measure m if

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \delta_{x_n} = m,$$

in the weak star limit. For $x \in X$ a sequence of natural numbers $\kappa = (k_n)_{n\geq 1}$ and a measure preserving transformation $T : X \to X$ we say x is κ -generic with respect to m if $x_n = T^{k_n}x$ $(n = 1, 2 \dots)$ is asymptotically distributed with respect to m.

In light of Lemmas 4.2.2 and 4.2.4 we see that if κ is both Hartman uniformly distributed and L^∞ good-universal then μ almost all x is κ-generic with respect to μ. For the following lemma see [p2–7, 34].

Lemma 4.2.6 (Weyl's criteria). *The following are equivalent.*

(i) $(x_n)_{n\geq 1}$ being uniformly distributed modulo one;

(ii) we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} e(hx_n) = 0,$$

for $h \in \mathbb{Z}\backslash\{0\}$.

and

(iii) for each continuous function $f : [0,1) \to \mathbb{R}$ we have

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

We have the following lemma of H. Furstenberg [17].

Lemma 4.2.7. *Let $T : X \to X$ be a be a homeomorphism of a compact metric space that preserves a measure μ such that $(X, \mathcal{B}, \mu, T)$ is uniquely ergodic. Let G be a be a compact group with Haar measure m_G and let $c : X \to G$ be a continuous map. Define the skew product map $\mathcal{S}$ on $Y = X \times G$ by $\mathcal{S}(x, g) = (T(x), c(x)g)$. If $\mathcal{S}$ is ergodic with respect to $\mu \times m_G$ then it is uniquely ergodic.*

We now specialise to the case $G = \mathbb{T}^l$. For irrational α and $\mathbf{x} = (x_1, \dots, x_l)$ set $\mathcal{S}(\mathbf{x}) = (x_1+\alpha, x_2+x_1, \dots, x_l+x_{l-1})$. A standard Fourier series argument shows this map is ergodic. See p.175 in [54] for details. Furstenberg's lemma now shows it must be uniquely ergodic. Suppose $\phi_l(x) := \phi(x)$ is of degree l and let $\phi_{i-1}(x) := \phi_i(x+1)-\phi_i(x)$ $(i = 2, \dots, l)$. So $\phi_1(x) = \alpha x+\beta$ for some real number β. Observe that $\mathcal{S}^n((\phi_1(0), \dots, \phi_l(0))) = (\phi_1(n), \dots, \phi_l(n))$ $(n = 1, 2, \dots)$. Using Lemma 4.2.3 all points in $\mathbb{T}^k$ are k-generic if k is good, so given a continuous function $f : \mathbb{T}^l \to \mathbb{C}$ we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(\phi_1(k_n), \dots, \phi_l(k_n)) = \int_{\mathbb{T}^l} f(t_1, \dots, t_l)dt_1 \dots dt_l.$$

Setting $g(t_l) = f(t_1, \dots, t_l)$, we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} g(\phi_l(k_n)) = \int_{\mathbb{T}} g(t_l)dt_l,$$

for any continuous $g : \mathbb{T} \to \mathbb{C}$ as required, this proving Theorem 4.2.1 by Weyl's criteria on uniform distribution. □

Theorem 4.2.1 has a number of aplications. We are going to intrduce them.

4.3 Application I: Norm Convergence

For $f \in L^p(X, \mathcal{B}, \mu)$, with $p \geq 1$ set $\|f\|_q = \left(\int_X |f|^p d\mu\right)^{\frac{1}{q}}$; we have the following corollary:

Corollary 4.3.1. *Suppose that ϕ is a non-constant polynomial mapping the natural numbers to themselves, and that $(k_n)_{n\geq 1}$ is good. Then for any dynamical system $(X, \mathcal{B}, \mu, T)$ and any $f \in L^2(X, \mathcal{B}, \mu)$, there exists $\overline{f} \in L^q(X, \mathcal{B}, \mu)$ such that*

$$\lim_{N\to\infty} \left\| \frac{1}{N} \sum_{n=0}^{N-1} f(T^{\phi(k_n)}x) - \overline{f} \right\|_2 = 0.$$

It is natural to ask if Corollary 4.3.1 is true almost everywhere.

Proof. For coprime integers p, q with q non-zero let

$$G(p, q, \phi) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} e\left(\frac{p}{q}\phi(n)\right),$$

which must exist as $(k_j)_{j\geq 1}$ is uniformly distributed on $\mathbb{Z}$. Using Theorem 4.2.1 we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} e(\phi(k_n)\theta) = \begin{cases} 0, & \text{if } \theta \notin \mathbb{Q} \\ G(p, q, \phi), & \text{if } \theta = \frac{p}{q}. \end{cases}$$

Let $Uf(x) = f(Tx)$. This is a unitary operator on L^2 as T is measure preserving. Also let U^{-1} denote the L^2 adjoint of U. Recall that we say any sequence $(c_n)_{n\in\mathbb{Z}}$ is 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 $\sum_{n,m\in\mathbb{Z}} c_{n-m} z_n \overline{z_m} \geq 0$. Here $\overline{z}$ is the conjugate of the complex number z. Let $\langle f, g\rangle = \int_X f\overline{g} d\mu$ (i.e. the standard inner product on L^2). Notice that

$(\langle U^n f, f\rangle)_{n\in\mathbb{Z}}$ is positive definite. Recall that the Bochner-Herglotz theorem [27] says that there is a measure ω_f on $\mathbb{T}$ such that

$$\langle U^n f, f\rangle = \int_{\mathbb{T}} z^n d\omega_f(z). \qquad (n \in \mathbb{Z})$$

Here by T^{-n} for positive integer n we mean $(T^*)^n$, were T^* denotes the adjoint of T. From this we see that

$$\left\| \frac{1}{N}\sum_{n=0}^{N-1} f(T^{\phi(k_n)}x) - \frac{1}{M}\sum_{n=0}^{M-1} f(T^{\phi(k_n)}x) \right\|_2^2$$

$$= \int_{\mathbb{T}} \left| \frac{1}{N}\sum_{n=0}^{N-1} e(\phi(k_n)\theta) - \frac{1}{M}\sum_{n=0}^{M-1} e(\phi(k_n)\theta) \right|^2 d\omega(\theta).$$

Using Theorem 4.2.1 and Weyl's criteria this can be as small as you like by choosing M and N large enough. This means $(\frac{1}{N}\sum_{n=0}^{N-1} f(T^{\phi(k_n)}x))_{N\geq 1}$ is a Cauchy sequence in L^2, proving Corollary 1.5 in L^2. □

4.4 Application II: Intersectivity and Recurrence

We say a set of natural numbers S has positive Banach density $B(S)$ if there exists a sequence of intervals $(I_p)_{p=1}^{\infty}$ in $\mathbb{N}$ with $I_p = [a_p, b_p] \cap \mathbb{N}$ and $|I_p| = b_p - a_p$ tending to infinity as p tends to infinity such that

$$\lim_{p\to\infty} \frac{|S \cap I_p|}{|I_p|} = B(S)$$

and for any other sequence of intervals $(I'_p)_{p=1}^{\infty}$ in $\mathbb{N}$ such that $|I'_p|$ tends to infinity as p tends to infinity we have

$$\limsup_{p\to\infty} \frac{|\, S \cap I'_p \,|}{|\, I'_p \,|} \leq B(S).$$

We say a sequence $\kappa = (k_n)_{n=1}^{\infty}$ of positive integers is a set of intersectivity if given any set of natural numbers S of positive Banach density $B(S)$ there exists an integer k in κ such that we can find s_1 and s_2 both in S satisfying

$$k = s_1 - s_2.$$

Following [18] we say a sequence of natural numbers $\kappa = (k_n)_{n=1}^{\infty}$ is Poincaré recurrent if given any dynamical system on probability space (X, β, μ, T) and any set A in β of positive measure there exists an element k of κ such that

$$\mu(A \cap T^{-k}A) > 0.$$

We say a sequence of natural numbers $\kappa = (k_n)_{n=1}^{\infty}$ is strongly recurrent if given any dynamical system on a probability space (X, β, μ, T) and any set A in β of positive measure there exists $\gamma_{\kappa,A} > 0$ and an element k of κ such that

$$\mu(A \cap T^{-k}A) \geq \gamma_{\kappa,A}.$$

Corollary 4.4.1. *Suppose ϕ is a non-constant polynomial mapping the natural numbers to themselves such that $\phi(0) = 0$. Then the sequence $(\phi(k_n))_{n\geq1}$ is strongly recurrent if $(k_n)_{n\geq1}$ is good.*

Poincaré recurrent sequences are the same as intersective sequences [12], hence we have the following corollary.

Corollary 4.4.2. *Suppose ϕ is a non-constant polynomial mapping the natural numbers to themselves such that $\phi(0) = 0$. Then the sequence $(\phi(k_n))_{n\geq1}$ is intersective if $(k_n)_{n\geq1}$ is good.*

It is natural to ask if the sequences from Corollaries 4.4.1 and 4.4.2 are also multiply recurrent and multiply intersective.

Proof. Corollaries 4.4.1 and 4.4.2

Now for each natural number N, the complex function $\frac{1}{N}\sum_{n=1}^{N} z^n$ is equal to 1 for all N if $z = 1$, and it tends to 0 for all other z such that z is of absolute value 1 as N tends to infinity. This means that if

$$A_N f(x) = \frac{1}{N}\sum_{n=1}^{N} f(T^n x), \qquad (N = 1, 2, \cdots)$$

then $< A_N f, f >= \int_{\mathbb{T}} (\frac{1}{N}\sum_{n=1}^{N} e(n\theta)) d\omega_f(\theta)$ tends to $w_f(\{0\})$ as N tends to infinity. By the mean ergodic theorem however, for f in L^2, if $P_T f$ is the projection of f onto the T invariant subspace of L^2 then $A_N f$ tends to $P_T f$ in L^2 norm as N tends to infinity. This means that $< P_T f, f >= w_f(\{0\})$ and so, by Cauchy's inequality,

$$\omega_f(\{0\}) \ = < P_T f, f > = < P_T f, P_T f > \geq \left| \int P_T f d\mu \right|^2 \ = \ \left| \int_X f d\mu \right|^2.$$

Let $L_{s,n} = \{Q(s^* k) : k \in \kappa \cap [1, n] \ (n = 1, 2 \cdots)\}$ where s^* is the least common multiple of the first s integers let $F_k = \{\frac{a}{q} : 1 \leq a < q \leq k; (a, q) = 1\}$ and let $\omega_f = \omega_i + \omega_r$ where ω_r are the atoms of ω_f on the rationals. Then for arbitrary k_0 we have

$$\frac{1}{|L_{k_0,n}|}\sum_{v \in L_{k_0,n}} < U^v f, f >= \omega_f(\{0\}) + \sum_{\theta \in F_{k_0}} \omega_r(\{e(\theta)\})$$

$$+ \sum_{\theta \in F_{k_0}^c} \omega_r(\{e(\theta)\}) \left(\frac{1}{|L_{k_0,n}|}\sum_{v \in L_{k_0,n}} e(v\theta) \right)$$

$$+ \int_{\mathbb{T}} \left(\frac{1}{|L_{k_0,n}|}\sum_{v \in L_{k_0,n}} e(v\alpha) \right) d\omega_i(e(\alpha)).$$

The second term on the right hand side of this identity is non-negative, the third become $\leq \epsilon$ for large enough k_0, and the the final term tends to zero as n tends to infinity by Theorem 4.2.1 and Weyl's criteria, so choosing $f = \chi_A$ with $\mu(A) > 0$ we have

$$\limsup_{n\to\infty} \frac{1}{|L_{k_0,n}|}\sum_{v \in L_{k_0,n}} \mu(A \cap T^{-v}A) \geq \mu^2(A) - \epsilon,$$

as required.

$\square$

4.5 Application III: Glasner Property

Following [7] we say a set S contained in $\mathbb{Z}$ is Glasner if for every infinite set Y contained in $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ and $\epsilon > 0$, some dilation $nY = \{ny : y \in Y\}$ is ϵ dense (that is nY intersects every open interval of length ϵ). This definition is motivated by the 1979 result of S. Glasner [22] in which he showed that given an infinite set $Y \subset \mathbb{T}$ there exists a natural number n such that nY is ϵ dense in $\mathbb{T}$.

We have the following theorem, which in the case $k_n = n$ $(n = 1, 2, \dots)$, appears in [2].

Corollary 4.5.1. *Let ϕ be a polynomial of degree $l \geq 1$ mapping the natural numbers to themselves. Suppose $\delta > 0$. Then there exists $\epsilon(\phi, \delta) > 0$ such that if $0 < \epsilon < \epsilon(\phi, \delta)$ for any set Y contained in $\mathbb{T}$ of cardinality s where*

$$s > \left(\frac{1}{\epsilon}\right)^{2l+\delta}$$

we have an ϵ dense dilation of the form $\phi(k_n)Y$, for some natural number n.

To prove this corollary we begin with some technical results we will need in the sequel. The first is abstracted from special cases dealt with in [2].

Lemma 4.5.2. *Given $\epsilon > 0$ if $X = \{x_1, \cdots, x_s\}$ is any set of finitely many points contained in $\mathbb{T}$ such that for every natural number n indexing the sequence $(a_n)_{n\geq 1}$ the dilation a_nX is not ϵ dense. Then there is an absolute constant $C > 0$ such that if*

$$M = \left[\left(\frac{1}{\epsilon}\right)\log^2\left(\frac{1}{\epsilon}\right)\right]$$

we have

$$s^2 \leq \left(\frac{C}{\epsilon}\right) \sum_{m=1}^{M} \sum_{j=1}^{s} \sum_{l=1}^{s} \frac{1}{N} \sum_{n=1}^{N} e_m(a_n(x_j - x_l)),$$

where $e_m(t) = e^{2\pi imt}$.

We next need the following estimate for rational exponential sums which is due to L.K. Hua [24].

Lemma 4.5.3. *Let θ denote a polynomial of degree L mapping the natural numbers to themselves. Then if $\epsilon_1 > 0$, there is a constant $C > 0$ such that*

$$\left| \frac{1}{b} \sum_{a=1}^{b} e_1\left(\frac{\theta(a)}{b}\right) \right| \leq \frac{C}{b^{\frac{1}{L} - \epsilon_1}}.$$

The following is also taken from [2].

Lemma 4.5.4. *Let $\{x_1, \cdots, x_s\}$ be an arbitrary set of s distinct points in the unit interval $[0, 1)$. Denote by h_m the number of pairs (i, j) with $1 \leq i < j \leq s$, such that $m(x_i - x_j)$ is an integer. Suppose $\beta > 0$. Then if s is sufficiently large for any $m \geq 1$, the partial sum*

$$H_m = \sum_{l=1}^{m} h_l$$

satisfies the inequality

$$H_m \leq (sm)^{\beta+1}.$$

The trivial bound here is $H_m \leq sm^2$.

Proof. Following Lemma 4.5.2, given $\epsilon > 0$ and a finite set $X = \{x_1, \cdots, x_s\}$ contained in $\mathbb{T}$ the dilation $\phi(k_n)X$ is not ϵ dense, setting

$$M = \left[\left(\frac{1}{\epsilon}\right) \log^2 \left(\frac{1}{\epsilon}\right)\right]$$

we have

$$s^2 \leq \left(\frac{C}{\epsilon}\right) \sum_{m=1}^{M} \sum_{j=1}^{s} \sum_{l=1}^{s} \frac{1}{N} \sum_{n=1}^{N} e_m(\phi(k_n)(x_j \; - \; x_l)). \qquad (4.1)$$

As a consequence of Theorem 4.2.1 and Weyl's criterion, as N tends to infinity, when for a particular j and l, the difference $x_j \; - \; x_k$ is irrational the average further most to the right in (4.5.1) tends to zero. This means that in estimating the right hand side of (4.5.1) we need only consider the contribution of the terms in the double sum in j and l for which the corresponding $x_j \; - \; x_l$ is rational.

Now for rational $\frac{a}{b}$, in reduced form we may clearly write

$$m\phi(n)\frac{a}{b} \; = \; \frac{(a'_l n^l \; + \; \cdots \; + \; a'_1 n)}{b'} \; = \; \frac{\theta(n)}{b'}$$

where the highest common factor of the integers $a'_L, \cdots, a_1$, and b' is one, b' denotes $\frac{b}{(m,b)}$, and (m, b) denotes the highest common factor of m and b. Because $(k_n)_{n=1}^{\infty}$ is uniformly distributed amongst the residue classes modulo b' and further because of the fact that the value of $m\phi(k_n)\frac{a}{b}$ is $\frac{\theta(c)}{b'}$ when $k_n \; \equiv \; c \; (mod \; b')$, if $x_j \; - \; x_k \; = \; \frac{a}{b}$, we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} e_m(\phi(k_n)(x_j \; - \; x_l)) = \frac{1}{b'} \sum_{a=1}^{b'} e_1\left(\frac{\theta(a)}{b'}\right).$$

From Lemma 4.5.3 we have that

$$\left|\frac{1}{b'} \sum_{a=1}^{b'} e_1\left(\frac{\theta(a)}{b'}\right)\right| < \frac{C}{(b')^{\frac{1}{l}-\epsilon_1}} \leq C\left(\frac{(m,b)}{b}\right)^{\frac{1}{l}-\epsilon_1}.$$

Also because there are at most $\frac{M}{r}$ multiples of r that divide b less than M we know that

$$\sum_{m=1}^{M} (m,b)^{\frac{1}{l}-\epsilon_1} \leq \sum_{\substack{r|b \\ r\leq M}} \left(\frac{M}{r}\right) r^{\frac{1}{l}-\epsilon_1}.$$

We also note that

$$\sum_{\substack{r|b \\ r\leq M}} \left(\frac{M}{r}\right) r^{\frac{1}{l}-\epsilon} \leq Md(b), \qquad (4.2)$$

where $d(n)$ denotes the number of integers between one and n inclusive there are that divide n. Note that given $\epsilon_2 > 0$ we have $d(n) = o(n^{\epsilon_2})$. From this (3.5.2) tells us that

$$\sum_{m=1}^{M} \left| \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} e_m(\phi(k_n)(x_j - x_l)) \right| \leq CMb^{-\frac{1}{l}+\epsilon_1 + \epsilon_2}.$$

Let

$$g_b = \#\{(j,k) : 1 \leq j < l \leq s : x_j - x_l = \frac{a}{b} \text{ for some } a \; ; \; (a,b) = 1\}$$

and let

$$G_b = \sum_{i=1}^{b} g_i.$$

Then setting $\epsilon = \epsilon_1 + \epsilon_2$ and using partial summation we have

$$s^2 \leq C\left(\frac{M}{\epsilon}\right)\left(s + \sum_{b\geq 2} g_b b^{-\frac{1}{l}+\epsilon}\right)$$

$$= C\left(\frac{M}{\epsilon}\right)\left(s + \sum_{b\geq 2} G_b(b^{-\frac{1}{l}+\epsilon} - (b+1)^{-\frac{1}{l}+\epsilon})\right).$$

From the trivial estimate in Lemma 4.5.4 we see that we have $G_b \leq s^2$ for $b \geq s$. If $b < s$ we know $G_b \leq H_b \leq sb^{1+\beta}$. Therefore the above expression is majorised by

$$= C\left(\frac{M}{\epsilon}\right)\left(s + \sum_{b\geq 2}(sb)^{1+\beta}b^{-\frac{1}{l}+\epsilon} + s^2 s^{-\frac{1}{l}+\epsilon}\right)$$

$$\leq C\left(\frac{M}{\epsilon}\right)s^{1+\beta}\left(\sum_{b=1}^{s} b^{2\beta-\frac{1}{l}} + s^{1-\frac{1}{l}}\right)$$

$$\leq C\left(\frac{M}{\epsilon}\right)s^{2+3\beta-\frac{1}{l}}(1+s^{-2\beta}).$$

which tells us, on noting our know value for M that

$$s \leq \left(\frac{1}{\epsilon}\right)^{2+\delta}$$

as required.

□

Chapter 5

On Convergents and Cusps of Fuchsian Groups

5.1 Introduction

For a real number x let $[x]$ denote the greatest integer not greater than x and let $\{x\} = x - [x]$ denote its fractional part. Using the Gauss map

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

we define a sequence of function of x by $c_0(x) = [x]$, $c_1(x) = [\frac{1}{x}]$ and $c_{n+1}(x) = c_n(g(x))$ $(n = 1 \dots)$, called the partial quotients of x. With these ingredients we define the convergents

$$\frac{p_n}{q_n} = c_0 + \cfrac{1}{c_1 + \cfrac{1}{c_2 + \cfrac{}{c_3 + \ddots \cfrac{1}{+\cfrac{1}{c_n}}}}}, \qquad (n = 0, 1, \dots).$$

which are also written more compactly as $[c_0; c_1, c_2, \cdots, c_n]$. These convergents have the property that $\frac{p_n}{q_n}$ tends to x as n tends to infinity.

Let $SL(2,\mathbb{Z})$ denote space of integer entry matricies of determinant one and let $PSL(2,\mathbb{Z}) = SL(2,\mathbb{Z})/\pm I$ the set of equivalence classes of matricies in $SL(2,\mathbb{Z})$ identical up to multiplication by the negative 2×2 identity matrix I. The projective linear group or modular group as we refer to $PSL(2,\mathbb{Z})$ as has long been known is intimately tied up with the properties of the regular continued fraction expansion. A Fuchsian group Γ is a discrete subgroup of $PSL(2,\mathbb{Z})$. From such a group one obtains a hyperbolic surface $\mathbb{H}\backslash\Gamma$. On a hyperbolic surface, an end is defined to be an equivalence class of limiting parallel rays. In the hyperbolic plane this is represented by a point on $\mathbb{R} \cup \{\infty\}$.

To a finite area hyperbolic triangle in $\mathcal{H}$ with a vertex on $\mathbb{R} \cup \{\infty\}$ and two sides identified by a generator of the stabilizer of Γ we associate a finite area triangle in $\mathcal{H}\backslash\Gamma$ one vertex of which is an end. We also call this end a cusp. Fix a finite index subgroup $H \subset \Gamma$. For any cusp of H pick some parabolic fixed point p in this H orbit and the the width with respect to Γ of κ is the index $w(\kappa)$ of the H-stabilizer of p in the Γ-stabilizer of p.

We say a sequences of natural numbers $(k_n)_{n\geq 1}$ is good for the regular continued fraction if given H, a finite index subgroup of a Fuchsian group $\Gamma \subseteq SL(2,\mathbb{Z})$, for each cusp κ of H and almost all $x \in [0,1)$,

$$\lim_{N\to\infty} \frac{1}{N} \frac{\#\{1 \leq n \leq N : p_{k_n}/q_{k_n} \in \kappa\}}{N} = \frac{w(\kappa)}{[\Gamma : H]}.$$

5.2 Example of Regular Continued Fractions of Sequences

A.M. Fisher and A. Schmidt showed that $(k_n)_{n\geq 1}$ is good for regular continued fractions if $k_n = n$ $(n = 1, 2, \dots)$. Consider the following families of sequences.

A. Denote by $[y]$ the integer part of real number y. Set $k_n = [g(n)]$ $(n = 1, \dots)$ where $g : [1, \infty) \to [1, \infty)$ is a differentiable function whose derivative increases with its argument. Let A_n denote the cardinality of the set $\{n : k_n \leq n\}$, and suppose for some function $a : [1, \infty) \to [1, \infty)$ increasing to infinity as its argument does, we set

$$b_M = \sup_{\{z\} \in [\frac{1}{a(M)}, \frac{1}{2})} \left| \sum_{n: a_n \leq M} e(zk_n) \right|.$$

(Here $e(x) = e^{2\pi i x}$ for real x.) Suppose also for some decreasing function $c : [1, \infty) \to [1, \infty)$ and some positive constant $C > 0$ that

$$\frac{b_M + A_{[a(M)]} + \frac{M}{a(M)}}{A_M} \leq C.c(M).$$

Then, if we have

$$\sum_{s=1}^{\infty} c(\theta^s) < \infty$$

for $\theta > 1$ we say that $\underline{k} = (k_n)_{n=1}^{\infty}$ satisfies condition H [20]. Condition H sequences are all L^p good universal for all $p > 1$. Condition H arises out of doing this using harmonic analysis.

Specific sequences of integers that satisfy conditions H include $k_n = [g(n)]$ $(n = 1, 2, \dots)$ where:

I. $g(n) = n^{\omega}$ if $\omega > 1$ and $\omega \notin \mathbb{N}$. In fact there is a interval $\omega \in [1, a]$ for very small a where $([n^{\omega}])_{n\geq 1}$ is L^1 good universal [28].

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 multiples of the same real number.

IV. Hardy Fields: By a Hardy Field we mean a closed subfield (under differentiation), of the ring of germs at $+\infty$ of continuous real valued functions with addition and multiplication taken to be pointwise.

Let L denote the union of all Hardy fields. If $(k_n)_{n=1}^{\infty} = ([a(n)])_{n=1}^{\infty}$, where a satisfies the following conditions:

$$a \in L;$$

for some $k \in \mathbb{Z}$, $k \geq 2$

$$\lim_{x\to\infty} \frac{a(x)}{x^{k-1}} = \infty$$

and

$$\lim_{x\to\infty} \frac{a(x)}{x^k} = 0;$$

then $(k_n)_{n=1}^{\infty}$ satisfies condition H. This example is observed in [3].

B. A random L^p example for $p > 1$: (i) Suppose $S = (n_k)_{n=1}^{\infty} \subseteq \mathbb{N}$ 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 may 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 $X_n = \{0,1\}$, and 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 β is that 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 $\alpha_{i_1}, \ldots, \alpha_{i_r}$. Let $(k_n)_{n=1}^{\infty}$ be almost any point in Λ with respect to the measure π [4].

C. Block L^1 good universal sequences: If $(k_n)_{n\geq 1} = \cup_{k=1}^{\infty}[d_k, d_k + e_k]$ ordered by absolute value for disjoint $([d_k, d_k + e_k])_{k\geq 1}$ with $d_{k-1} = O(e_k)$ as k tends to infinity. Note that if $d_{k-1} = o(e_k)$ the sequence $(k_n)_{n=}^{\infty}$ is zero density.

D. Random perturbation of L^2 good sequences: Suppose $(k_n)_{n\geq 1}$ is a L^2-good universal sequence of integers that is also Hartman uniformly distributed. Suppose $\theta = \{\theta_n, n \geq 1\}$ 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-field $\mathcal{A}$. We assume that there exist $0 < \alpha < 1$ and $B > 1/\alpha$, such that

$$k_n = O(e^{n^\alpha}),$$

and if $\mathbb{E}$ denotes expectation with respect to the basic probability space $(\Omega, \mathcal{A}, \mathcal{P})$ we have

$$\mathbb{E} \log_+^B |\theta_1| < \infty.$$

Then $(k_n + \theta_n(\omega))_{n\geq 1}$ is L^2-good universal and Hartman uniformly distributed [23].

Recall that we say a sequence $(k_n)_{n\geq 1}$ is good if it s L^p - good universal, for some $p \geq 1$ and also Hartman uniformly distributed.

Theorem 5.2.1. *Suppose $(k_n)_{n\geq 1}$ is an element of one of the families of sequences A, B, C and D is good for the regular continued fraction expansion.*

We also have a moving average version of this property. Let Z be a collection of points in $\mathbb{Z} \times \mathbb{N}$ and let

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

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

and

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

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 5.2.2. *We say a sequence of pairs of natural numbers* $(n_l, m_l)_{l=1}^\infty$ *is 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 $(n_l, m_l)_{l=t}^\infty \in Z^{h(t)}$ and there exist h_0, α_0 and $A > 0$ such that for all integers $\lambda > 0$ we have

$$|Z_{\alpha_0}^{h_0}(\lambda)| \leq A\lambda.$$

We have the following moving average result.

Theorem 5.2.3. *Suppose* $(n_l, k_l)_{l\geq 1}$ *is Stoltz and that* H *is a finite index subgroup of a Fuchsian group* $\Gamma \subseteq SL(2,\mathbb{Z})$*. Then for each cusp* κ *of* H *and almost all* $x \in [0,1)$*,*

$$\lim_{l\to\infty} \frac{1}{m_l} \#\{n_l \leq n \leq n_l + m_l - 1 : p_n/q_n \in \kappa\} = \frac{w(\kappa)}{[\Gamma : H]}.$$

.

In the case where $\Gamma = SL(2,\mathbb{Z})$ with $k_n = n$ $(n = 1, 2, \dots)$ the Theorem 4.2.1 is due T. Nakanishi who strengthening earlier work of R. Moeckel who proved this result subject to additional technical restrictions that were not necessary. Their proof relies on the properties of the geodesic flow on the unit tangent manifold of a finite cover of modular surface. An extension of this result congruence subgroups of the modular group using a different method relying on a skew product over the continued fraction map rather than the geodesic flow method.

5.3 Preliminary results

Theorem 5.3.1. *Fix a cusp* κ *a suppose* $\Psi : C \times H\backslash\Gamma \to \mathbb{R}$ *is defined by*

$$(A, \gamma \mod H) \to \mathbb{I}_{A_{-1}\times\kappa.0} + \mathbb{I}_{A_{+1}\times\kappa.\infty}.$$

Suppose $(k_n)_{n\geq 1}$ is good. Then for almost all $(A, \gamma \mod H)$ in $C \times H\backslash\Gamma$, we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \Psi(S^{k_n}((A, \gamma \mod H)) = \frac{w(\kappa)}{[\Gamma : H]}.$$

Proof. The function Ψ is the sum of two function each with an integral of $\frac{w(\kappa)}{2[\Gamma:H]}$. The result now follows from the goodness of the sequence $(k_n)_{n\geq 1}$. □

Corollary 5.3.2. *For each cusp $\kappa \in H$ and almost all $x \in [0, 1)$, if $(k_n)_{n\geq 1}$ is good, we have*

$$\lim_{N\to\infty} \frac{1}{N} \frac{\#\{1 \leq n \leq N : p_{k_n}/q_{k_n} \in \kappa\}}{N} = \frac{w(\kappa)}{[\Gamma : H]}.$$

Proof. The subset $\mathbb{A}_{-1} \times I \mod H \subset \mathbb{C} \times H\backslash\Gamma$ has positive measure. hence the limit in Theorem 2 holds for almost all x. This imples the map defined by

$$\overline{\Psi}(\Lambda_\sigma(x, y), \gamma \mod H) = \begin{cases} |H\gamma|_{.\infty} & \text{if } \sigma = -1, \\ |H\gamma|_{.0} & \text{if } \sigma = +1. \end{cases}$$

sends the corresponding sequence $S^{k_n}(\Lambda_{-1}(x, y), I \mod H)$ into the cusp κ with limiting frequency $\frac{w(\kappa)}{[\Gamma:H]}$. From Lemma 1.7.3, it follows for almost every $x \in [0, 1)$, that $(-\frac{p_{k_n}}{q_{k_n}})_{n\geq 1}$ has a limiting frequency. We now use the above argument with the roll of H replace by that of iHi we see that $(-\frac{p_{k_n}}{q_{k_n}})_{n\geq 1}$ has this limiting frequency in any cusp of iHi. Lemma 1.6.3 tells us that that this means the sequence $((\frac{p_{k_n}}{q_{k_n}}))_{n\geq 1} = (i(-\frac{p_{k_n}}{q_{k_n}}))_{n\geq 1}$ has the same limiting frequency for the corresponding cusps in H. By Lemma 1.6.3 the cusp widths are the same for the two groups. Also conjugate groups have the same index and so we are done. □

Lemma 5.3.3. *Each of the elements of the families of sequences (A), (B), (C), (D) and (E) are both Hartman uniformly distributed and L^∞ good universal.*

We now consider moving averages.

Lemma 5.3.4. *Let (X, β, μ, T) denote a dynamical system, with set X, a σ-algebra of its subsets β, a measure μ defined on the measurable space (X, β) such that $\mu(X) = 1$ and a measurable, measure preserving map T from X to itself. Suppose f is in $L^1(X, \beta, \mu)$ and that the sequence of pairs on natural numbers $(n_l, k_l)_{l=1}^{\infty}$ is Stoltz then if (X, β, μ, T) is ergodic,*

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

exists almost everywhere with respect to Lebesgue measure.

Note that if

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

then

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

This means that $m_f(Tx) = m_f(x)$ μ almost everywhere.

Using the same argument used to produce Theorem 2 we have the following.

Theorem 5.3.5. *Fix a cusp κ a suppose $\Psi : C \times H\backslash\Gamma \to \mathbb{R}$ is defined by*

$$(A, \gamma \mod H) \to \mathbb{I}_{A_{-1} \times \kappa.0} + \mathbb{I}_{A_{+1} \times \kappa.\infty}.$$

Suppose $(n_l, m_l)_{l \geq 1}$ is Stoltz. Then for almost all $(A, \gamma \mod H)$ in $C \times H\backslash\Gamma$, we have

$$\lim_{l \to \infty} \frac{1}{m_l} \sum_{n=1}^{m_l} \Psi(S^{n+n_l-1}((A, \gamma \mod H)) = \frac{w(\kappa)}{[\Gamma : H]}.$$

5.4 Two variants of the regular continued fraction expansions

(a) **Nakada's α-continued fractions :** For $\alpha \in (0,1]$ set $\mathbb{I}_\alpha = [\alpha - 1, \alpha)$ and Nakada's map is defined by

$$T_\alpha(x) = \begin{cases} \left|\frac{1}{x}\right| - \left[\left|\frac{1}{x}\right| + 1 - \alpha\right] & \text{if} \quad x \neq 0 \\ 0, & \text{if} \quad x = 0 \end{cases}$$

For $x \in \mathbb{I}_\alpha$ let

$$\epsilon(x) = \operatorname{sign}(x).$$

and

$$d_\alpha(x) = \left[\left|\frac{1}{x}\right| + 1 - \alpha\right].$$

We now set $\epsilon_n = \epsilon(T_\alpha^{-n-1}(x))$ and $d_n = d_\alpha(T_\alpha^{-n-1}(x))$. This gives rise to the sequence

$$\frac{p_n}{q_n} = d_0 + \cfrac{\epsilon_1}{d_1 + \cfrac{\epsilon_2}{d_2 + \cfrac{}{d_3 + \ddots + \cfrac{\epsilon_3}{+\frac{\epsilon_n}{d_n}}}}}, \qquad (n = 0, 1, \dots),$$

which converges to x as n tends to infinity. Here $d_0 \in \mathbb{Z}$ and $x - d_0 \in \mathbb{I}_\alpha$. Analogously to the case $\alpha = 1$ one can define

$$\mathcal{C}_{\alpha,-1} = \{A_{-1}(x,y) : (x,) \in Z^{-1}(\Omega_\alpha)\},$$

and

$$\mathcal{C}_{\alpha,+1} = \{A_{+1}(x,y) : (x,) \in Z^{-1}(\Omega_\alpha)\}.$$

where Ω_α is the closure of $\{\mathcal{T}_\alpha^n(x,0) : x \in [\alpha-1,\alpha] \text{ for } n \geq 0\}$ with

$$\mathcal{T}_\alpha(x,y) = \left(T_\alpha^n(x), \frac{1}{d_\alpha(x)+\epsilon(x)y}\right).$$

Also set $\mathcal{C}_\alpha = \mathcal{C}_{\alpha,-1} \cup \mathcal{C}_{\alpha,+1}$.

Arnoux and Schmidt [17] show the following.

Lemma 5.4.1. *Up to measure zero, the set $\mathcal{C}_\alpha$ projects to a cross section of the geodesic flow on the unit tangent bundle of the modular surface. Furthermore this cross section is given by*

$$A_\sigma(x,y) \to MA_\sigma(x,y)E_t$$

with $t = -2\log|x|$ and

$$M = \begin{pmatrix} 0 & 1 \\ -1 & d \end{pmatrix}, \begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} d & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ d & 1 \end{pmatrix},$$

where $d = d_\alpha(x)$ as $(\sigma, \epsilon(x)) = (-1,-1), (-1,+1), (+1,-1)$ or $(+1,+1)$ respectively. Moreover the dynamical system defined by the first return map has T_α as a factor.

Using this lemma Fisher and Schmidt prove the following.

Lemma 5.4.2. *Let $\mathcal{C}_\alpha$ be the cross section of the the geodesic flow on the modular flow in the previous lemma. Suppose $A_{-1}(x,y) \in \mathcal{C}_\alpha$. Then the second composite of the k^{th} iterate of the skew product transformation $\mathcal{S}$ with itself appplied to $(A, I \mod H)$ denoted $proj_2(\mathcal{S}^k(A, I \mod H))$, takes the value*

$$proj_2(\mathcal{S}^k(A, I \mod H)) = \begin{pmatrix} q_k & -q_{k-1} \\ -p_k & p_{k-1} \end{pmatrix} \mod H$$

if $proj_1(\mathcal{S}^k(A, I \mod H)) \in \mathcal{C}_{\alpha-1}$ and

$$proj_2(\mathcal{S}^k(A, I \mod H)) = \begin{pmatrix} q_{k-1} & -q_k \\ -p_{k-1} & p_k \end{pmatrix} \mod H$$

otherwise.

We can now deduce the following.

Proposition 5.4.3. *Suppose* $(k_n)_{n\geq 1}$ *is good and that* H *is a finite index subgroup of a Fuchsian group* $\Gamma \subseteq SL(2,\mathbb{Z})$. *Then for each cusp* κ *of* H *and almost all* $x \in \mathbb{R}$, *if* $(\frac{p_n}{q_n})_{n\geq 1}$ *denotes the* α*-Nakada convergents to* x *we have*

$$\lim_{N\to\infty} \frac{1}{N} \frac{\#\{1 \leq n \leq N : p_{k_n}/q_{k_n} \in \kappa\}}{N} = \frac{w(\kappa)}{[\Gamma : H]}.$$

Proposition 5.4.4. *Suppose* $(n_l, k_l)_{l\geq 1}$ *is Stoltz and that* H *is a finite index subgroup of a Fuchsian group* $\Gamma \subseteq SL(2,\mathbb{Z})$. *Then for each cusp* κ *of* H *and almost all* $x \in \mathbb{R}$, *if* $\left(\frac{p_n}{q_n}\right)_{n\geq 1}$ *denotes the* α*-Nakada convergents to* x

$$\lim_{l\to\infty} \frac{1}{m_l} \#\{n_l \leq n \leq n_l + m_l - 1 : p_n/q_n \in \kappa\} = \frac{w(\kappa)}{[\Gamma : H]}.$$

(b)**Rosen continued fractions:** Let $\lambda = \lambda_m = \cos(\frac{\pi}{m})$. Then Γ_m is the Hecke group of index m generated by i and $\gamma_m = \begin{pmatrix} 1 & \lambda_m \\ 1 & 0 \end{pmatrix}$. The group Γ_3 is the the modular group $PSL(2,\mathbb{Z})$. Let $\mathbb{I}_m = [-\frac{\lambda}{2}, \frac{\lambda}{2})$ for $m \geq 3$. For a fixed $m \geq 3$

$$T(x) = T_m(x) = \begin{cases} \left|\frac{1}{x}\right| - \lambda\left[|1\lambda x| + \frac{1}{2}\right] & \text{if } x \neq= 0 \\ 0 & \text{if } x = 0 \end{cases}$$

for $x \in \mathbb{I}_m$. Set

$$\epsilon(x) = \text{sign}(x) \text{ and } r(x) = \left[\left|\frac{1}{\lambda x}\right| + \frac{1}{2}\right]$$

and set

$$\epsilon_n(x) = \epsilon(T^{n-1}(x)) \text{ and } r_n(x) = r(T^{n-1}(x)).$$

$$x = c_0 + \cfrac{\epsilon_1}{c_1 + \cfrac{\epsilon_2}{c_2 + \cfrac{\epsilon_3}{c_3 + \ddots}}}.$$

The convergents $\left(\frac{p_n}{q_n}\right)_{n\geq 1}$ are defined via the relations

$$\begin{pmatrix} p_{-1} & p_0 \\ q_{-1} & q_0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and

$$\begin{pmatrix} p_{n-1} & p_n \\ q_{n-1} & q_n \end{pmatrix} = \begin{pmatrix} 0 & \epsilon_1 \\ 1 & \lambda r_1 \end{pmatrix}\begin{pmatrix} 0 & \epsilon_2 \\ 1 & \lambda r_2 \end{pmatrix}\cdots\begin{pmatrix} 0 & \epsilon_n \\ 1 & \lambda r_n \end{pmatrix}.$$

Recall that we call a sequence of natural numbers good if it is L^∞ good universal and Hartman uniform distribution. We can now deduce the following.

Proposition 5.4.5. *Suppose $(k_n)_{n\geq 1}$ is good and that H is a finite index subgroup of H of the Hecke group Γ_m. Then for each cusp κ of H and almost all $x \in \mathbb{R}$, if $(\frac{p_n}{q_n})_{n\geq 1}$ denotes the Rosen convergents to x we have*

$$\lim_{N\to\infty} \frac{1}{N}\frac{\#\{1 \leq n \leq N : p_{k_n}/q_{k_n} \in \kappa\}}{N} = \frac{w(\kappa)}{[\Gamma_m : H]}.$$

Proposition 5.4.6. *Suppose $(n_l, k_l)_{l\geq 1}$ is Stoltz and that H is a finite index subgroup of H of the Hecke group Γ_m. Then for each cusp κ of H and almost all $x \in \mathbb{R}$, if $\left(\frac{p_n}{q_n}\right)_{n\geq 1}$ denotes the Rosen convergents to x then*

$$\lim_{l\to\infty} \frac{1}{m_l}\#\{n_l \leq n \leq n_l + m_l - 1 : p_n/q_n \in \kappa\} = \frac{w(\kappa)}{[\Gamma_m : H]}.$$

Chapter 6

Pair Correlation and Oscillating Sequences

The contents of this chapter are joint with Lgiia Loretta Cristea of the Karl Franzens-Universität in Graz, Austria.

6.1 Introduction

Let $(X_n(x))_{n\geq 1}$ denote a sequence of real numbers. For an interval I let $\chi_I(x)$ denote the characteristic function of the set I. This means that we have $\chi_I(x) = 1$ if $x \in I$ and $\chi_I(x) = 0$ otherwise. For a real number y let $\{y\}$ denote its fractional part. Recall that

$$V_N(I)(x) = \sum_{1\leq n<m\leq N} \chi_I(\{X_n(x) - X_m(x)\})$$

and that we then define

$$\Delta_N(x) = \sup_{I\subseteq\mathbb{T}} \left| V_N(I) - \frac{N(N-1)}{2}\lambda_1(I) \right|$$

where the supremum is over all intervals I in the one dimensional torus $\mathbb{T}$. See also Section 3.4. Suppose $X_n(x) = a_n \cos(a_n x)$ $(n = 1, 2 \ldots)$ for a sequence of distinct integers $(a_n)_{n\geq 1}$. In Theorem 4 in [40], it is shown that given $\epsilon > 0$ we have $\Delta_N(x) = o(N^{\frac{3}{2}}(\log N)^{4+\epsilon})$ almost everywhere with respect to Lebesgue measure.

For a set $A \subseteq [0,1)$ let diam $A = \ sup_{x,y\in A}|y-x|$. Also for $\delta > 0$ and a set $E \subseteq [0,1)$, we say a countable collection of intervals $I = (I_i)_i$ is a δ-cover if $E \subseteq \cup_i I_i$. Let

$$\mathcal{H}^s_\delta(E) = \sup_{All\ \delta-covers\ I} \sum_i |I_i|^s.$$

Also let

$$\mathcal{H}^s(E) = \mathcal{H}^s_\delta(E),$$

which always exists as $\mathcal{H}^s_\delta(E)$ is a monotone function of δ. There is a specific value of s denoted $s_0 \in [0,\infty]$ (say) called the Hausdorff dimension of E. The concept of Hausdorff dimension can be defined for any set E contained in any metric space. In our context $s_0 \in [0,1]$. If a set has positive Lebesgue measure in $[0,1]$, then its Hausdorff dimension is 1. Thus Hausdorff dimension provides a means of comparing the size of sets whose Lebesgue measure is zero.

6.2 Main Theorem

For a set E let $\dim E$ denote its Hausdorff dimension. In this note we show the following.

Theorem 6.2.1. *Suppose $(a_n)_{n\geq 1}$ is a sequence of distinct integers such that there exist $p > 1$ such that $a_n = O(n^p)$ and suppose $q \in (0, \frac{3}{2})$. Let*

$$E_q = \{x \in \mathbb{T} : \limsup_{N\to\infty} N^{2-q}\Delta_N(x) > 0\}.$$

Then

$$\dim E_q \leq 1 - \frac{(1-q)}{2p+q}.$$

The uniform distribution of the sequence $(a_n \cos(a_n x))_{n\geq 1}$, for almost all x with respect to Lebesgue measure, was first proved by W. J. Leveque. Let

$$D(x,N) = \sup_{I\subset[0,1)} \frac{1}{N}\sum_{n=1}^{N} \chi_I(\{a_n \cos(\{a_n x\}) - |I||. \qquad (N = 1,2,\dots)$$

Here the supremum is taken over all intervals $I \subseteq [0,1)$. Clearly the uniform distribution of $(a_n \cos(a_n x))_{n\geq 1}$ is equivalent to $D(x,N) = o(1)$ almost everywhere. It is shown in [33] that given $\epsilon > 0$ we have

$$D(x,N) = o(N^{\frac{1}{4}}(\log N)^{\frac{2}{3}+\epsilon}),$$

almost everywhere with respect to Lebesgue. Suppose $a_n = O(n^p)$ $(n = 1,2,\dots)$ with $p > 1$ and that $q \in (0,\frac{1}{4})$. We define

$$F_q = \{x \in [0,1) : \limsup_{N\to\infty} N^q D(x,N) > o\}.$$

In the author's supervisors thesis, it is shown that

$$\dim F_q \leq 1 - \frac{1-2q}{p+q}.$$

Theorem 6.2.1 is a second order version of this phenomenon.

6.3 Proof of Theorem 6.2.1:

We start with some lemmas, the first proved in [37], the second from [44] and the third a straightforward consequence of the famous Erdős-Turán inequality [34 p. 112-14] .

Lemma 6.3.1. *Let μ denote a Borel measure on $\mathbb{T}$ such that for each arc $[x,y]$ we have $\mu([x,y]) \leq (y-x)^\nu$. Also let*

$$S_h(M,N) = \sum_h(M,N,x) = \sum_{n=M+1}^{M+N} e^{2\pi i h a_n \cos(a_n)}.$$
$$(M = 1,2,\dots,N = 1,2,\dots)$$

Then

$$\int_{\mathbb{T}} |S_h(M,N,x)|^2 d\mu \ll (M+N)^{2p(1-\nu)} N^{(\frac{7-\nu}{8})} h^{(\frac{1-\nu}{2})}.$$

Let $(Y_t)_{t=1}^{\infty}$ be a sequence of measurable functions defined on a measure space Ω and then write

$$S_j \ = \sum_{1 \leq t \leq \ j} Y_t, \text{ for } j \ = \ 1, 2, \cdots.$$

We can define

$$Y_{rs} \ = \sum_{r<t\leq s} Y_t \quad (= S_s - S_r), \text{ for } r < s,$$

and let $M_n = \sup_{1\leq j\leq n} |S_j|$. We have the following elementary lemmas proved in [NP].

Lemma 6.3.2. *For $K \geq 1$,*

$$\int_{\Omega} M_{2^K}^2(\omega) d\omega \leq (K+1) \left(\sum_{i=1}^{K+1} \sum_{\nu=1}^{2^i-1} \int_{\Omega} |Y_{\nu 2^{(K+1)-i}, (\nu+1)2^{(K+1)-i}}|^2(\omega) d\omega \right).$$

See [30] for the following lemma.

Lemma 6.3.3. *For $L \geq 1$ we have*

$$\Delta_N \ll \left(\frac{N^2}{L} + \sum_{h=1}^{L} \frac{1}{h} \left(|\sum_{n=1}^{N} e^{2\pi i h X_n}|^2 + N \right) \right).$$

We also need a lemma of O. Frostman [20].

Lemma 6.3.4. *Suppose $C \subseteq \mathbb{T}$ is compact and has Hausdorff dimension greater tha ν then there exists a Borel measure supported on C such that $\mu([x,y)) \leq (y-x)^{\nu}$.*

Let $||f||_r = (\int_{\mathbb{T}} |f|^r d\mu)^{\frac{1}{r}}$. We have the following Lemma.

Lemma 6.3.5. *For $(X_n(x))_{n\geq 1}$ with $X_n = a_n \cos(a_n x)$ $(n = 1, 2, \dots)$ we have*

$$|| \max_{1\leq j\leq 2^K} \Delta_j |||_1 \ll 2^{K(2-\frac{2p(\nu-1)-1}{\nu-1})}.$$

Proof of Lemma 6.3.5: Using Lemma 6.3.3 one readily checks that

$$\left\| \max_{1\leq j\leq 2^K} \Delta_j \right\|_1 \ll \left(\frac{2^{2K}}{L} + \sum_{h=1}^{L} \frac{1}{h} \left(\left\| \max_{1\leq n\leq j} \left| \sum_{n=1}^{j} e^{2\pi i h a_n \cos(a_n x)} \right| \right\|_2^2 + 2^K \right) \right).$$

Using Lemma 6.3.2

$$|| \max_{1\leq n\leq j} | \sum_{n=1}^{j} e^{2\pi i h a_n \cos(a_n x)} |||_2^2$$

$$\ll (K+1) \left(\sum_{i=1}^{K+1} \sum_{\eta=1}^{2^i-1} ||S_h(\eta 2^{(K+1)-i}, (\eta+1)2^{(K+1)-i}||_2^2 \right).$$

Using Lemma 6.3.1, this is

$$\ll (K+1) \sum_{i=1}^{K+1} \sum_{\eta=1}^{2^i-1} (\eta+1)^{2p(1-\nu)} (2^{(K+1)-i)2p(1-\nu)}).2^{((K+1)-i)(\frac{7-\nu}{4})} h^{(1-\nu)}.$$

$$\ll (K+1)h^{(1-\nu)} \sum_{i=1}^{K+1} \sum_{\eta=1}^{2^i-1} (\eta+1)^{2p(1-\nu)} (2^{(K+1)-i}))^{2p(1-\nu)}.2^{((K+1)-i)(\frac{7-\nu}{4})}.$$

$$\ll (K+1)h^{(1-\nu)} \sum_{i=1}^{K+1} \left(2^{K+1-i} \right)^{\{2p(1-\nu)+(\frac{7-\nu}{4})\}} \left(\sum_{\eta=1}^{2^i-1} (\eta+1)^{2p(1-\nu)} \right).$$

$$\ll (K+1)h^{(1-\nu)} \left(\sum_{i=1}^{K+1} (2^{K+1-i})^{\{2p(1-\nu)+(\frac{7-\nu}{4})\}} 2^{i(2p(1-\nu)+1)} \right).$$

$$\ll (K+1)h^{(1-\nu)} (2^{K+1})^{\{2p(1-\nu)+(\frac{7-\nu}{4})\}} \sum_{i=1}^{K+1} (2^{-i})^{\{2p(1-\nu)+(\frac{7-\nu}{4})\}} (2^i)^{2p(1-\nu)+1}).$$

$$\ll (K+1)h^{(1-\nu)} (2^{K+1})^{\{2p(1-\nu)+(\frac{7-\nu}{4})\}} (2^{K+1})^{(1-(\frac{7-\nu}{4}))}.$$

$$\ll (K+1)h^{(1-\nu)} (2^{K+1})^{\{2p(1-\nu)+1\}}.$$

This means that

$$\left\| \max_{1\leq j\leq 2^K} \Delta_j \right\|_1 \ll \left(\frac{2^{2K}}{L} + \sum_{h=1}^{L} \frac{1}{h} \left((h^{(1-\nu)} (2^{K+1})^{\{2p(1-\nu)+1\}} \right) + 2^K \right) \right)$$

$$\ll \frac{2^{2K}}{L} + (2^{K+1})^{\{2p(1-\nu)}L^{1-\nu} + (\log L)2^K.$$

Take $L = 2^{K\alpha}$. Then this is

$$\ll 2^{2K-\alpha K} + 2^{(K+1)(2p(1-\nu)+1)}2^{K\alpha(1-\nu)}.$$

Choosing α optimally $(2-\alpha) = (1-\nu)(2p-\alpha)+1$ or $\alpha = \frac{2p(\nu-1)-1}{\nu-1}$. Hence

$$|| \max_{1\le j\le 2^K} \Delta_j|||_1 \ll 2^{K(2-\frac{2p(\nu-1)-1}{\nu-1})},$$

as required. □

We now complete the proof of the theorem. So suppose for the sake of contradiction that there exist ν such that

$$\dim E_q > \nu > 1 - \frac{(2-q)}{2p}.$$

By Lemma 6.3.4, there is a Borel measure μ supported on a compact subset of E_q such that

$$\mu([x,y]) \le (y-x)^\nu.$$

Lemma 6.3.5 now tells us that

$$|| \max_{1\le j\le 2^K} \Delta_j|||_1 \ll 2^{K(2p(1-\nu))}.$$

One checks easily that

$$E_q \subseteq \bigcap_{r=1}^{\infty}\bigcup_{k=r}^{\infty} A_k,$$

where for given $\epsilon_0 > 0$

$$A_l = \{x \in \mathbb{T} : |\max_{2^{l-1}\le n\le 2^l} \Delta_n(x)| > 2^{l(2p(1-\nu)}l^{1+\epsilon_0}\}.$$

Notice that

$$|A_l| \le \frac{\int_{\mathbb{T}} |\max_{1\le j\le 2^K} \Delta_j|(x)d\mu}{2^{l(2p(1-\nu)}l^{1+\epsilon_0}} \ll \frac{1}{l^{1+\epsilon_0}}.$$

The Borel-Cantelli lemma now tells us that E_q has μ measure zero. Thus $\Delta_N(x) = o(N^{2p(1-\nu)})$ which for any $\epsilon_1 > 0$ is $o(N^{-q+\epsilon_1})$ as required. So Theorem 6.2.1 is proved. □

Appendix

List of known good universal sequences

We give some examples of L^p-good universal sequences for some $p \geq 1$. All examples save 2 example below are also Hartman uniformly distributed.

1. *The natural numbers:* The sequence $(n)_{n=1}^{\infty}$ is L^1-good universal. This is Birkhoff's pointwise ergodic theorem.

2. *Polynomial like sequences:* If $\phi(x)$ is a polynomial such that $\phi(\mathbb{N}) \subseteq \mathbb{N}$ and $p > 1$, then $(\phi(n))_{n=1}^{\infty}$ and $(\phi(p_n))_{n=1}^{\infty}$, where p_n is the n^{th} prime, are L^p-good universal sequences. See [13], and [26], respectively.

3. *Condition* H*:* Sequences $(k_n)_{n=1}^{\infty}$ that are both L^p-good universal and Hartman uniformly distributed can be constructed as follows. Set $k_n = [\tau(n)]$ $(n = 1, 2, \dots)$, where $\tau : [1, \infty) \to [1, \infty)$ is a differentiable function whose derivative increases with its argument. Let Ω_m denote the cardinality of the set $\{n \colon a_n \leq m\}$, and suppose, for some function $\varphi : [1, \infty) \to [1, \infty)$ increasing to infinity as its argument does, that we set

$$\varrho(m) = \sup_{\{z\} \in \left[\frac{1}{\varphi(m)}, \frac{1}{2}\right)} \left| \sum_{n \colon k_n \leq m} e(zk_n) \right|.$$

Suppose also, for some decreasing function $\rho : [1, \infty) \to [1, \infty)$ and some positive constant $\omega > 0$, that

$$\frac{\varrho(m) + \Omega_{[\varphi(m)]} + \frac{m}{\varphi(m)}}{\Omega_m} \leq \omega\rho(m).$$

Then if we have

$$\sum_{n=1}^{\infty} \rho(\theta^n) < \infty$$

for all $\theta > 0$, we say that $(a_n)_{n=1}^{\infty}$ satisfies condition H, see [26].

Sequences satisfying condition H are known to be both Hartman uniformly distributed and L^p-good universal. Specific sequences of integers that satisfy condition H include $a_n = [\tau(n)]$ $(n = 1, 2, \dots)$ where:

I. $\tau(n) = n^{\gamma}$ if $\gamma > 1$ and $\gamma \notin \mathbb{N}$.

II. $\tau(n) = e^{\log^{\gamma} n}$ for $\gamma \in (1, \frac{3}{2})$.

III. $\tau(n) = b_k n^k + \cdots + b_1 n + b_0$ for $b_k, \dots, b_1$ not all rational multiplies of the same real number.

IV. *Hardy fields:* By a Hardy field, we mean a closed subfield (under differentiation) of the ring 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. Conditions for $(a_n)_{n=1}^{\infty} = ([\psi(n)])_{n=1}^{\infty}$, where $\psi \in \mathcal{H}$ to satisfy condition H are given by the hypotheses of Theorems 3.4, 3.5 and 3.8. in [12]. Note the term ergodic is used in this paper in place of the older term Hartman uniformly distributed.

4. *A random example:* Suppose that $S = (k_n)_{n=1}^{\infty}$ is a strictly increasing sequence of natural numbers. By identifying S with its characteristic function χ_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 may 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 $X_n = \{0,1\}$ and specify the probability ν_n on X_n by $\nu_n(\{1\}) = \omega_n$ with $0 \leq \omega_n \leq 1$ and $\nu_n(\{0\}) = 1 - \omega_n$ such that $\lim_{n\to\infty} \omega_n n = \infty$. The desired probability measure on Λ is the corresponding product measure $\nu = \prod_{n=1}^{\infty} \nu_n$. The underlying σ-algebra $\mathcal{A}$ is that generated by the cylinders

$$\{(\Delta_n)_{n=1}^{\infty} \in \Lambda \colon \Delta_{n_1} = \alpha_{n_1}, \dots, \Delta_{n_k} = \alpha_{n_k}\}$$

for all possible choices of $n_1, \dots, n_k$ and $\alpha_{n_1}, \dots, \alpha_{n_k}$. Then almost every point $(a_n)_{n=1}^{\infty}$ in Λ, with respect to the measure ν, is Hartman uniformly distributed. See Proposition 8.2 (i) in [13] for the details of this. Again Hartman uniformly distributed sequences are called ergodic sequences in this paper.

5. *Block sequences:* Suppose that $(a_n)_{n=1}^{\infty} = \bigcup_{n=1}^{\infty}[d_n, e_n]$ is ordered by absolute value for disjoint $([d_n, e_n])_{n=1}^{\infty}$ with $d_{n-1} = O(e_n)$ as n tends to infinity. Note that this allows the possibility that $(a_n)_{n=1}^{\infty}$ is zero density. This example is an immediate consequence of Tempelman's semigroup ergodic theorem. See page 218 of [4]. Being a group average ergodic theorem this pointwise limit must be invariant. which ensures that the block sequence must be Hartman uniformly distributed. The proof of this, which we don't need in this paper and is hence forgone is a simple exercise in spectral theory.

6. *Random perturbation of good sequences:* Suppose that $(a_n)_{n=1}^{\infty}$ is an L^p-good universal sequence which is also Hartman uniformly distributed. Let $\theta = (\theta_n)_{n=1}^{\infty}$ be a sequence of $\mathbb{N}$-valued independent, identically distributed random variables with basic probability space $(Y, \mathcal{A}, \mathcal{P})$, and a $\mathcal{P}$-complete σ-field $\mathcal{A}$. Let $\mathbb{E}$ denote expectation with respect to the basic probability space $(Y, \mathcal{A}, \mathcal{P})$. Assume that there exist $0 < \alpha < 1$

and $\beta > 1/\alpha$ such that

$$a_n = O(e^{n^\alpha}) \qquad \text{and} \qquad \mathbb{E}\log_+^\beta |\theta_1| < \infty.$$

Then $(k_n+\theta_n(\omega))_{n=1}^\infty$ is both L^p-good universal and Hartman uniformly distributed [30].

Bibliography

[1] M. Ajtai, I. Havas, J. Komlós: *Every group admits a bad topology*, Studies in pure mathematics, 21-34, Birkhuser, Basel, (1983).

[2] N. Alon, Y. Peres: *Uniform Dilations*, Geometric and Functional Analysis, **2**, No. 1, (1992), 1–28.

[3] P. Arnoux: *Le codage du flot géodésique sur la surface modulaire*, Enseign. Math. **40** (1994), 29-48.

[4] N. H. Asmar, R. Nair: *Certain averages on the a-adic numbers*, Proc. Amer. Math. Soc. **114** (1992), no. 1, 21-28.

[5] A. Bellow, V. Losert: *On sequences of zero density in ergodic theory*, Conference in modern analysis and probability (New Haven, Conn. 1982), 49–60, Contemp. Math. **26**, Amer. Math. Soc. Providence. RI. 1984.

[6] A. Bellow, V. Losert: *The weighted pointwise ergodic theorem and the individual ergodic theorem along subsequences*, Trans. Amer. Math. Soc. **288** (1985), no. 1, 307–345.

[7] D. Berend, Y. Peres: *Asymptotically dense dilations of sets on the circle*, J. London Math. Soc. (2) **47**, (1993), no. 1, 1-17.

[8] V. Bergelson, G. Kolesnik, M. Madritsch, Y. Son, R. Tichy: *Uniform distribution of prime powers and sets of recurrence and van der Corput sets in $\mathbb{Z}^k$*, Israel J. Math. **201** (2014), no. 2, 729-760.

[9] V. Bergelson, E. Lesigne: *Van der Corput sets in $\mathbb{Z}^d$.*, Colloq. Math. **110** (2008), no. 1, 1–49.

[10] V. Bergelson, A. Leibman: *A Weyl-type equidistribution theorem in finite characteristic*, Adv. Math. **289** (2016), 928-950.

[11] I. Berkes, W. Philipp and R. Tichy: *Pair correlation and U-statistics for independent and weakly dependent random variables*, Illinois Journal of Mathematics 45 (2001), 559-580.

[12] A. Bertrand-Mathis: *Ensembles intersectifs et récurrence de Poincaré*, Israel J. Math. **55** (1986), no. 2, 184–198.

[13] J. Blum, B.Eisenberg: *Generalized summing sequences and the mean ergodic theorem*, Proc. Amer. Math. Soc. **42**, (1974), 423–429.

[14] M. Boshernitzan: *Homogeneously distributed sequences and Poincaré sequences of integers of sublacunary growth*, Monatsh. Math. **96** (1983), no. 3, 173–181.

[15] M. Boshernitzan, G. Kolesnik, A. Quas, M. Wierdl: *Ergodic averaging sequences*, J. Anal. Math. **95** (2005), 63–103.

[16] J. Bourgain: *On the maximal ergodic theorem for certain subsets of the integers*, Israel J. Math. Vol. **61** no. 1, 39–72 (1988).

[17] J. W. S. Cassels: *Some metrical theorems on diophantine approximation II*, Journal of the London Mathematical Society 25 (1950), 180-184.

[18] M. Drmota, R.F. Tichy: *Sequences, Discrepancies and Applications*, Springer Verlag, New York, 19 **45** (1983), 281–299.

[19] K. J. Falconer: *The Geometry of Fractal Sets*, Cambridge University Press, Cambridge, 1986.

[20] A. M. Fisher, T. A. Schmidt: *Distribution of approximants and geodesic flows*, Ergodic Theory Dynam. Systems **34** (2014), no. 6, 1832–1848.

[21] H. Furstenberg: *Strict ergodicity and transformation of the torus*, Amer. J. Math., **83**, 573–601.

[22] H. Furstenberg: *Poincaré recurrence and number theory*, Bull. Amer. Math. Soc. (N.S.) **5** (1981), no. 3, 211-234.

[23] O. Frostman *Thesis*, Lund 1935.

[24] S. Glasner: *Almost periodic sets and measures on the Torus*, Israel J. Math. **32** (1979) 161–172.

[25] A. Haddley, R. Nair, E. Nasr, P. Lertchoosakul: *On Convergents and Cusps of Fuchsian Groups* Preprint (2017).

[26] J. Hancl, A. Jassová, P. Lertchoosakul, R. Nair: *On the metric theory of p-adic continued fractions*, Indag. Math. (N.S.) 24 (2013), no. 1, 42–56.

[27] G. Harman: *Metric Number Theory*, London Mathematical Society Monographs, New Series, 18, The Clarendon Press, Oxford University Press, New York, 1998.

[28] L.K. Hua: *Sur une somme exceptionelle*, C.R. Acad. Sci., **200**, 1940, 520–523.

[29] H. Jager, P. Liardet: *Distributions arithmtiques des dnominateurs de convergents de fractions continues*, Indag. Math. **50** (2) (1988), 181-197.

[30] A. Jassová, P. Lertchoosakul, R. Nair: *On variants of the Halton sequence*, Monatsh. Math. **180** (2016), no. 4, 743-764.

[31] H. Kamarul Haili, R. Nair: *On certain Glasner sets*, Proc. Roy. Soc. Edinburgh Sect. A **133** (2003), no. 4, 849–853.

[32] Y. Katznelson: *An introduction to Harmonic Analysis'*, Dover (1976).

[33] S. Kristensen, A. Jaššová, P. Lertchoosakul, R. Nair: *On recurrence in positive characteristic*, Indag. Math. (N.S.) **26** (2015), no. 2, 346-354.

[34] L. Kuipers, H. Niederreiter: *Uniform Distribution of Sequences*, John Wiley & Sons, New York 1974.

[35] A. Manning: *Dynamics of geodesic and horocycle flows on surfaces of constant negative curvature*, Ergodic theory, symbolic dynamics, and hyperbolic spaces (Trieste, 1989), 71-91, Oxford Sci. Publ., Oxford Univ. Press, New York, 1991.

[36] R. Moeckel: *Geodesics on modular surfaces and continued fractions*, Ergodic Theory. Dynam. Systems. **2** (1982), 69–83.

[37] R. Nair: *On LeVeque's theorem about the uniform distribution (mod 1) of* $(a_j cos a_j x)_{j=1}$*"*. Israel J. Math. **65** (1989), no. 1, 96112.

[38] R. Nair: *On certain solutions of the Diophantine equation x-y=p(z)*, Acta Arith. **62** (1992), no. 1, 61-71.

[39] R. Nair: *On Pair Correlations and Hausdorff Dimension*, Israel Journal of Mathematics 171(2009), 197-219.

[40] R. Nair: *On polynomials in primes and J. Bourgain's circle method approach to ergodic theorems II*, Studia Math., **05** (3), (1993), 207–233.

[41] R. Nair: *On uniformly distributed sequences of integers and Poincaré recurrence*, Indag. Math., N. S. vol **9** (1998), no. 1, 55–63.

[42] R. Nair: *On uniformly distributed sequences of integers and Poincaré recurrence II*, Indag. Math., N. S. vol **9** (1998), no. 3, 405–415.

[43] R. Nair: *Some theorems on metric uniform distribution using L2 methods*, Journal of Number Theory 35 (1990), 18-52.

[44] R. Nair, E. Nasr: *On Uniform Distribution of Polynomials and Good Universality* Preprint (2017).

[45] R. Nair, E. Nasr: *Pair correlations and random walks on the integers* Unif. Distrib. Theory **11** (2016), no.1, 159-164.

[46] R. Nair, M. Pollicott: *On pair correlations of sequences in higher dimensions*, Israel J. Math. **157** (2007), no. 1, 219–238.

[47] R. Nair, S.L. Velani: *Glasner sets and polynomials in primes*, Proc. Amer. Math. Soc. **126** (1998), no. 10, 2835–2840.

[48] R. Nair, M. Weber: *On random perturbation of some intersective set*, Indag. Math. N.S. **15** (2004), no. 3, 373–381.

[49] T. Nakanishi: *A remark on R. Moeckel's paper: Geodesics on modular surfaces and continued fractions.* Ergodic Theory Dynam. Systems. **9**(3) (1989), 511-514.

[50] J.C. Oxtoby: *Ergodic Sets*, Bull. Amer. Math. Soc. **58** (1952), 116–136.

[51] C. A. Roger: *Hausdorff Measures*, Cambridge University Press, London New York, 1970.

[52] Z. Rudnick and P. Sarnak: *The pair correlation functions of fractional parts of polynomials*, Communications in Mathematical Physics 194 (1998), 61-70.

[53] A. Sárkőzy: *On difference sets of sequences of integers. I*, Acta Math. Acad. Sci. Hungar. **31** (1978), no. 12, 125-149.

[54] A. Sárkőzy: *On difference sets of sequences of integers. II.* Ann. Univ. Sci. Budapest. Eőtvős Sect. Math. **21** (1978), 45-53 (1979).

[55] A. Sárkőzy: *On difference sets of sequences of integers. III*, Acta Math. Acad. Sci. Hungar. **31** (1978), no. 3-4, 355–386.

[56] S. Sawyer: *Maximal inequalities of weak type*, Ann. Math. (2) **84**, 157–174 (1966).

[57] F. L. Spitzer: *Principles of Random Walks*, Second ed. Graduate Texts in Mathematics, Vol. 34. Springer-Verlag, New York-Heidelberg, 1976.

[58] C. T. Stone: *On local and ratio limit theorems*, in: Proc. of the Fifth Berkeley Symp. on Math. Statist. and Probab. Vol. 2 : Contributions to Probability Theory, Part 2, (L. M. Le Cam and J. Neyman, Eds.), University of California Press, Berkeley, California, 1967, pp. 217–224.

[59] J.P. Thouvenot: *La convergence presque sûre des moyennes ergodiques suivant certaines sous-suites d'entiers (d'aprés Jean Bourgain)*, Séminaire Bourbaki, Vol. 1989/90. Astérisque No. **189-190** (1990), Exp. No. 719, 133–153.

[60] R.Urban, J.Zienkiewicz: *Weak type (1,1) estimates for a class of discrete rough maximal functions*, Math. Res. Lett. **14** (2007), no. 2, 227–237.

[61] R.C. Vaughan: *The Hardy-Littlewood method*, Second edition. Cambridge Tracts in Mathematics, **125**. Cambridge University Press, Cambridge, 1997.

[62] M.Viana, K. Oliveira: *Foundations of Ergodic Theory*, Cambridge Studies in Advanced Mathematics,**151**, 2016.

[63] P. Walters: *An Introduction to Ergodic Theory*, Graduate Texts in Mathematics **79**, Springer-Verlag, New York, 1982.

[64] H. Weyl: *Über die Gleichverteiling von Zahlen mod. Eins.*, Mathematische Annalen **77**(1), 313-352, (1916).

[65] M. Weber: *Discrepancy of randomly sampled sequences of reals*, Math. Nachr. **271** (2004), 105–110.

[66] M. Weber: *Dynamical Systems and Processes*, IRMA Lectures in Mathematics and Theor. Physics Vol 14, European Mathematical Society (EMS), Zürich, 2009.