Results in

Metric and Analytic Number Theory

Sky Brewer
PhD

University of York
Mathematics

August 2017

Abstract

The following Thesis consists of five chapters. The first three chapters come from Metric Number theory. Chapter 1 discusses algorithms for calculating the continued fraction of algebraic numbers, as well as presenting experimental results on how well algebraic numbers fit well known conjectures on the distribution of their partial quotients. Chapter 2 discusses the Singular and Extremality theories of so called "well separated Dirichlet type systems". Chapter 3 presents an effective version of the Khintchine-Groshev theorem for simultaneously small linear forms.

The last two chapters are mostly in the area of uniform distribution. Chapter 4 proves a central limit theorem for the count of the fractional parts of imaginary parts of the zeros of the Riemann zeta function within an interval. Chapter 5 discusses the upper and lower distribution functions mod 1 of sequences of the form $(0.a_n a_{n+1} a_{n+2} \dots)_{n\in\mathbb{N}}$, where the sequence $(a_n)_{n\in\mathbb{N}}$ has polynomial growth.

Contents

List of Tables

List of Figures

Preface

The thesis consists of five chapters which are self contained and therefore stand alone.

Chapter 1: It is known that the continued fraction of quadratic irrationals have periodic continued fraction expansion, however much less is known about the continued fraction expansion of higher degree algebraic numbers. In Chapter 1 we present a discussion of conjectures on the continued fraction expansion of algebraic numbers of degree strictly larger than two. In particular whether they obey the Gauss-Kuzmin distribution, have geometric mean which tends to Khintchine's constant and whether the mean has suitable asymptotic growth.

We then overview the algorithms used to calculate the continued fraction expansion of higher degree algebraic numbers, as well as provide some experimental evidence towards the conjectures. We also investigate a link between the continued fraction expansion and the height and degree of an algebraic number, by calculating the continued fraction expansion of numbers $\sqrt[d]{k}$, for integers $d > 2$ and $k > 1$.

Chapter 2: In the recent paper [**9**], Beresnevich, Ghosh, Simmons and Velani introduced the notion of singular and extremal points associated with the limit sets of a Kleinian group. The goal of Chapter 2 is to develop a general framework of "Dirichlet systems" inspired by the ubiquity setups of [**8**] and [**60**], that naturally incorporate the Kleinian group results of [**9**]. The framework will almost certainly allow us to prove the analogous statements for rational maps and indeed general hyperbolic dynamical systems - this will be addressed in the near future. Before describing the general framework, we provide a brief overview of singular and extremal sets associated with the classical theory of Diophantine approximation. This will provide the context for the general framework.

Chapter 3: Chapter 3 is motivated by recent applications of Diophantine approximation in electronics, in particular, in the rapidly developing area of Interference Alignment found in [**71**]. Some remarkable advances in this area give substantial credit to the fundamental Khintchine-Groshev Theorem. Presented is a variant of the Khitnchine-Groshev theorem for linear forms that are simultaneously small for infinitely many integer vectors; i.e. linear forms which are close to the origin. We then present a simple example of how percentages of "Bad" simultaneously small linear forms can be calculated given a constant of approximation with a view towards the aforementioned applications.

Chapter 4: The inspiration for Chapter 4 came from the fact that the imaginary parts of the zeta zeros are uniformly distributed mod 1 (u.d. mod 1), which was first proved by Rademacher [**77**] assuming the Riemann Hypothesis (RH). Subsequently Hlawka showed that they are u.d. mod 1 without the RH [**50**]. Since then significant improvements have been made on the discrepancy of the zeros [**40**, **41**]. In this work we are interested in the more subtle question of how independent the fractional parts of the zeta zeros are. In particular, whether the count of the fractional parts of the zeta zeros obey a central limit theorem. A large influence on this work is Selberg's [**83**] study on the remainder term in the counting function $S(t)$. He proved that $S(t)$ has Gaussian moments, essentially showing that as $T \to \infty$,

$$\frac{1}{T}\int_T^{2T} \left| \frac{S(t)}{\sqrt{(\log\log T)/2\pi^2}} \right|^{2k} dt \to \frac{(2k)!}{k!2^k}.$$

Chapter 5 : Chapter 5 was influenced by Wall's result in 1949 where he showed that $x = 0.d_1d_2d_3\ldots$ is normal if and only if

$$(0.d_nd_{n+1}d_{n+2}\ldots)_{n\in\mathbb{N}}$$

is a uniformly distributed sequence. In this article, we consider sequences which are slight variants on this. In particular, we find the upper and lower distribution functions of sequences of certain normal numbers of the form

$$(0.a_na_{n+1}a_{n+2}\ldots)_{n\in\mathbb{N}},$$

where $(a_n)_{n\in\mathbb{N}}$ is a sequence of positive integers.

Motivated by a result of Davenport and Erdős [**34**], we calculate the upper and lower distribution functions of the sequence $(0.f(n)f(n+1)f(n+2)\ldots)_{n\in\mathbb{N}}$ for a non-constant integer polynomial $f(x)$.

Acknowledgments

Firstly, I would like to thank both of my supervisors Christopher Hughes and Sanju Velani for all the hard work they have put into supporting me.

I feel privileged to have been a member of the number theory study group in York during a time in which such wonderful people are working here. Special thanks goes to Faustin, who was always willing to talk about life, mathematics and academia.

I have made many friends among the PhD students in the department, thank you all for making these years so enjoyable. Especially to Demi, for being the kindest critic and Tom, my compadre in PhD life.

Thanks also goes to my parents for trying so hard to understand what I do all day.

Finally, my partner Sophie, for always being there.

Declaration

The work presented in this thesis is based on research carried out in the Department of Mathematics, University of York. This work has not previously been presented for an award at this, or any other, University.

The data in Chapter 1 was generated using the high performance computer YARCC based at the University of York. The results in Chapter 3 are joint with Jason Levesley. The results in Chapter 4 are joint with Christopher Hughes. The results in Chapter 5 are joint with Demi Allen.

CHAPTER 1

Computation of the continued fraction expansion of algebraic numbers

1.1. Background of Metric Number theory

1.1.1. Classical Diophantine Approximation. Approximation of real numbers is a classical problem leading back to the Ancient Greeks and Diophantus himself. For example, it was known since the days of Archimedes that 22/7 is a good approximation to π, and 99/70 for $\sqrt{2}$. Indeed, due the rationals being dense in the reals, we know that it is possible to arbitrarily approximate any real number by rationals. However, classical Diophantine approximation is about how "fast" we can do so. Explicitly, the object of interest is, for $x \in \mathbb{R}$ the size of

$$\left| x - \frac{p}{q} \right| \qquad \text{for } p \in \mathbb{Z} \text{ and } q \in \mathbb{N}, \tag{1.1.1}$$

as a function of q.

A simple observation is that by the spread of a fraction $1/q$, there is a bound of $1/(2q)$ on (1.1.1). However, we will see that there are multiple ways to improve on that bound. For some of the most up to date results, V. Beresnevich, F. Ramírez and S. Velani [**11**] have written a substantial overview of the area.

1.1.1.1. *Dirichlet's Theorem.* The most fundamental result of Diophantine approximation is Dirichlet's Theorem, which can be found in the overview [**11**].

THEOREM 1.1.1 (Dirichlet's Theorem 1842). *For all $x \in \mathbb{R} \setminus \mathbb{Q}$ and $N \in \mathbb{N}$ there exists $(p, q) \in \mathbb{Z} \times \mathbb{N}$ such that*

$$\left| x - \frac{p}{q} \right| \leq \frac{1}{qN} \qquad \textit{and} \qquad q < N.$$

Dirichlet first proved this result in 1842 using the "pigeonhole principle". The principle states that if n objects are placed in $n-1$ boxes then at least one box must contain at least two objects. Dirichlet's Theorem shows that the $1/(2q)$ can be significantly improved upon. In fact it also tells us about the "speed" at which irrationals can be approximated by rationals.

Corollary 1.1.2. *For all $x \in \mathbb{R} \setminus \mathbb{Q}$ there exists infinitely many (i.m.) pairs $(p, q) \in \mathbb{Z} \times \mathbb{N}$ such that*

$$\left| x - \frac{p}{q} \right| \leq \frac{1}{q^2}. \tag{1.1.2}$$

1.1.1.2. *Continued Fractions.* Given a real number x, we can find a sequence of natural numbers $a_0, a_1, a_2, \ldots$ such that

$$x = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \ldots}}} \tag{1.1.3}$$

We call this sequence *the continued fraction expansion* of x, for simplicity we write it as $x = [a_0; a_1, a_2, \ldots]$. We call a_k the *kth partial quotient of* x and the rational $p_k/q_k = [a_0; a_1, a_2, \ldots, a_k]$ is called the *kth convergent of* x. The p_k/q_k turn out to be the "best" approximations of x, in the sense that they satisfy Dirichlet's Theorem. For full details and the rest of the statements in this subsection see Hardy and Wright [**48**].

The *basic method* for calculating the continued fraction expansion of a real number is in essence the Euclidean algorithm for the computation of the greatest common divisor of two integers. If x is a real number, then the sequence of partial quotients (a_n) for x are found by the iteration

$$\alpha_0 = x, \quad a_0 = \lfloor x \rfloor, \quad \alpha_{n+1} = \frac{1}{\alpha_n - a_n}, \quad a_{n+1} = \lfloor \alpha_{n+1} \rfloor, \quad n = 0, 1, 2, \ldots. \tag{1.1.4}$$

The process terminates if $a_n = \alpha_n$, which happens if and only if x is a rational.

Some more useful properties include:

- The recursive formula

$$\begin{aligned} p_{n+1} &= a_{n+1} p_n + p_{n-1}, \\ q_{n+1} &= a_{n+1} q_n + q_{n-1}, \end{aligned} \tag{1.1.5}$$

  with initial values

$$\begin{pmatrix} p_{-1} & p_{-2} \\ q_{-1} & q_{-2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

- The difference

$$p_n q_{n-1} - p_{n-1} q_n = (-1)^{n+1}. \tag{1.1.6}$$

- The inequality

$$\left| \frac{p_n}{q_n} - \frac{p_{n-1}}{q_{n-1}} \right| > \frac{1}{q_{n-1} q_n}.$$

- Let α be a real number. If $\alpha = [a_0; a_1, a_2, \dots]$ then
$$\left|\alpha - \frac{p_n}{q_n}\right| < \frac{1}{q_n q_{n+1}} < \frac{1}{a_{n+1} q_n^2}.$$

1.1.1.3. *Bad and quadratic irrationals.* An obvious question is whether the constant 1 in Dirichlet's Theorem can be improved upon. Hurwitz [**53**] showed that it can, moreover there is a best possible constant.

THEOREM 1.1.3 (Hurwitz 1891). *Given $x \in \mathbb{R} \setminus \mathbb{Q}$ there exists infinitely many pairs $(p, q) \in \mathbb{N} \times \mathbb{Z}$, such that*
$$\left|x - \frac{p}{q}\right| \leq \frac{1}{\sqrt{5} q^2},$$
where $1/\sqrt{5}$ is optimal.

The constant is the best possible in the sense that, there exists $x \in \mathbb{R}$ such that there are finitely many solutions $(p, q) \in \mathbb{Z} \times \mathbb{N}$ to
$$\left|x - \frac{p}{q}\right| \leq \frac{\epsilon}{q^2}, \qquad \text{for } \epsilon < 1/\sqrt{5}.$$
The immediate question is then, for what numbers $x \in \mathbb{R} \setminus \mathbb{Q}$ is there a best possible constant? This leads to the following definition.

DEFINITION 1.1.1. A real number $x \in$ **Bad** if and only if there exists $c(x) > 0$ such that
$$\left|x - \frac{p}{q}\right| \geq \frac{c(x)}{q^2}, \qquad \text{for all } (p, q) \in \mathbb{Z} \times \mathbb{N}.$$

The complement set $\mathbb{R} \setminus$ **Bad** will then be called the set of *well approximable numbers.*

The **Bad** numbers in general have the interesting property that the partial quotients are bounded. The golden ratio is the most famous member of **Bad**, it has the elegant continued fraction expansion $[1; 1, 1, \dots]$. In fact all quadratic irrationals have a periodic continued fraction expansion, for example $\sqrt{2} = [1; 2, 2, 2, \dots]$. The contrapositive statement also holds true: if a real number x has a repeating continued fraction $[a_0; \overline{a_1, a_2, \dots, a_k}]$ then it is a quadratic irrational and its polynomial can be constructed. Since quadratic irrationals have periodic partial quotients, they are all members of **Bad**.

In the case of roots of higher degree polynomials there are no explicit continued fraction expansions known, and no known examples of members of **Bad**. The roots of polynomials of degree strictly greater than 2 are predicted to have unbounded partial quotients. However the following theorem of

Roth [**80**] tells us that all algebraic numbers are not far from being badly approximable.

Theorem 1.1.4 (Roth 1955). *For any irrational algebraic number x and any real $\tau > 1$ there exist only finitely many pairs $(p, q) \in \mathbb{Z} \times \mathbb{N}$ such that*

$$\left| x - \frac{p}{q} \right| \le \frac{1}{q^{\tau+1}}. \tag{1.1.7}$$

For proofs of the statements in this subsection and more recent results see Bugeaud [**22**, Appendix E].

1.1.1.4. *Very well approximable numbers.* In contrast with badly approximable numbers we can consider irrationals which are very well approximated by rational numbers. For any $\tau \ge 1$, let $W(\tau)$ be the set of real numbers $x \in (0, 1)$ for which (1.1.7) is satisfied for infinitely many pairs $(p, q) \in \mathbb{Z} \times \mathbb{N}$. We refer to $W(\tau)$ as the set of *τ-approximable numbers.* Note that from Dirichlet's theorem $W(1) = [0, 1)$.

An irrational number $x \in \mathbb{R} \backslash \mathbb{Q}$ is then said to be *very well approximable* if there exists $\tau > 1$ such that $x \in W(\tau)$. Given $\tau > 1$, it is relatively straightforward to construct numbers in $W(\tau)$ using the theory of continued fractions. However, Liouville [**66**] was first to construct explicit examples of numbers that lie in $W(\tau)$ for all $\tau > 1$, and the set of such numbers now bears his name. More precisely, we say an irrational x is a *Liouville number* if for all $\tau \in \mathbb{N}$ there exists $(p, q) \in \mathbb{Z} \times \mathbb{N}$ such that

$$\left| x - \frac{p}{q} \right| < \frac{1}{q^{\tau}}.$$

Liouville's result was that these numbers are in fact transcendental, the first establishment of the existence of transcendental numbers.

Liouville also showed the following explicit construction of such numbers.

Example 1.1.1. Let $b \ge 2$. Let $(a_k)_{k\in\mathbb{N}}$ be a sequence of integers, such that $a_k \in \{0, 1, 2, \ldots, b-1\}$ for all $k \in \mathbb{N}$, and there are infinitely many k with $a_k \ne 0$. Define the numbers

$$x = \sum_{k=1}^{\infty} \frac{a_k}{b^{k!}},$$

as Liouville numbers. In the special case of $b = 10$ and $a_k = 1$ for all k, the number x is called *Liouville's constant.*

1.1.2. Metric Diophantine Approximation. Another way of looking at possible improvements to Dirichlet's theorem is to look at "almost all" results. By *almost all* we mean that a statement is true for all but a set of Lebesgue measure zero. We will write $|A|_1$ to represent the Lebesgue measure of a measurable set $A \subseteq \mathbb{R}$. We also introduce the notation of $||\cdot||$ to represent the distance to the nearest integer, equivalently

$$||x|| := \min_{p\in\mathbb{Z}} |x-p| \text{ for all } x \in \mathbb{R}.$$

We can then change Dirichlet's theorem to, for any real number x there exists infinitely many $q \in \mathbb{N}$ satisfying the inequality

$$||qx|| \leq \frac{1}{q}.$$

We can then start to consider more general approximation results. A positive real function $\psi : \mathbb{N} \to \mathbb{R}^+$ is called an *approximation function*. We call the set of real numbers x, such that there are infinitely many solutions $q \in \mathbb{N}$, to the equation

$$||qx|| \leq \psi(q),$$

ψ-approximable numbers.

The set of ψ-approximable numbers are invariant under translation by integers, so to reduce the discussion on "metrical" statements we shall restrict our attentions to numbers in the unit interval $\mathbb{I} := [0, 1)$. The set of such numbers will be denoted by $W(\psi)$, explicitly

$$W(\psi) := \{x \in I : ||qx|| \leq \psi(q) \text{ for i.m. } q \in \mathbb{N}\}.$$

Not to be confused with $W(\tau) = W(r \mapsto r^{-\tau})$ for $\tau > 0$, it will be clear which set we mean by context. Note that Dirichlet's theorem implies that $W(\psi) = \mathbb{I}$ if $\psi(q) \geq \frac{1}{q}$ for all $q \in \mathbb{N}$.

A key aspect of metric number theory is then to measure the "size" of $W(\psi)$, we will be limited to Lebesgue measure statements in this work. We start with noticing that $W(\psi)$ is a lim-sup set of balls, for a fixed $q \in \mathbb{N}$ let

$$\begin{aligned} A_q(\psi) &:= \{x \in \mathbb{I} : ||qx|| < \psi(q)\}, \qquad (1.1.8) \\ &= \bigcup_{p=0}^{q} B\left(\frac{p}{q}, \frac{\psi(q)}{q}\right) \cap \mathbb{I}. \end{aligned}$$

The set $W(\psi)$ is simply the set of real numbers in $\mathbb{I}$ that lie in infinitely many sets $A_q(\psi)$ for $q \in \mathbb{N}$, equivalently

$$W(\psi) = \limsup_{q\to\infty} A_q(\psi) := \bigcap_{t=1}^{\infty} \bigcup_{q=t}^{\infty} A_q(\psi).$$

1.1.2.1. *Borel-Cantelli.* Let $(\Omega, \mathcal{A}, \mu)$ be measure space with $\mu(\Omega) < \infty$ and let E_q, for $q \in \mathbb{N}$ be a family of measurable subsets in Ω. Borel-Cantelli

[**18**, **24**] then tells us the measure of the lim sup of E_q satisfying a suitable sum condition on $\mu(E_q)$.

Lemma 1.1.5 (Borel-Cantelli Convergence)**.** *If*

$$\sum_{q=1}^{\infty}\mu(E_q)<\infty, \qquad \textit{then} \qquad \mu(\limsup_{q\to\infty}E_q)=0.$$

The contrapositive statement of: if

$$\sum_{q=1}^{\infty}\mu(E_q)=\infty \tag{1.1.9}$$

then the limsup set is of full measure. The statement is not true for all sets of E_q, if the sets "overlap" too much then the limsup set can be of measure 0.

We are interested in how much overlap can we get away with, and still get a full measure statement. If we have pairwise independence, i.e.

$$\mu(E_q\cap E_p)=\mu(E_q)\mu(E_p) \qquad \text{for all } p\neq q,$$

then under the sum condition (1.1.9) the limsup set is of full measure. However we can relax the condition to only consider quasi-independence.

A sequence of measurable sets E_q for $q\in\mathbb{N}$ are *quasi-independent on average* if there exists a constant $C>0$ such that

$$\sum_{p,q=1}^{Q}\mu(E_q\cap E_p)\leq C\left(\sum_{s=1}^{Q}\mu(E_s)\right)^2,$$

for infinitely many $Q\in\mathbb{N}$.

The full divergence statement under quasi-independence was first stated by Erdős and Chung [**29**]. A simpler proof can be found in [**21**].

Lemma 1.1.6 (Borell-Cantelli Divergence)**.** *Let* $(E_q)_{q\in\mathbb{N}}$ *be a sequence of measurable sets such that*

$$\sum_{i=1}^{\infty}\mu(E_i)=\infty,$$

then

$$\mu(\limsup_{q\to\infty}E_q)\geq\limsup_{Q\to\infty}\left(\frac{\sum_{q=1}^{Q}\mu(E_q)^2}{\sum_{p,q=1}^{Q}\mu(E_q\cap E_p)}\right).$$

1.1.2.2. *Khintchine's Theorem.* Khintchine's book [**57**] is a classic book on the area of metric number theory, it introduces measure theoretical statements relating to Diophantine approximation. The following groundbreaking theorem [**55**] is fundamental to the metrical theory of Diophantine approximation and is discussed in the book as well as many other results.

THEOREM 1.1.7 (Khintchine's Theorem 1924). *Let $\psi : \mathbb{N} \to [0,\infty)$ be an approximation function, then*

$$|(W(\psi))|_1 = \begin{cases} 1 & \text{if } \sum_{q=1}^{\infty} \psi(q) = \infty \text{ and } \psi \text{ is monotonic,} \\ \\ 0 & \text{if } \sum_{q=1}^{\infty} \psi(q) < \infty, \end{cases}$$

where $|\cdot|_1$ is the one-dimensional Lebesgue measure.

Note that Khintchine originally assumed that $q \mapsto q\psi(q)$ is monotonically decreasing and Beresnevich [**5**, **6**] removed the condition.

In view of Cassels' zero-full [**25**] law, we know that

$$|(W(\psi))|_1 = 0 \qquad \text{or} \qquad 1, \tag{1.1.10}$$

irregardless of whether or not ψ is monotonic. Therefore to prove Khintchine's theorem we use the Borel-Cantelli Divergence Lemma 1.1.6 by showing the sets A_q from 1.1.8 are quasi-independent on average.

Khintchine's Theorem implies that

$$|W(\psi)|_1 = 1 \text{ if } \psi(q) = \frac{1}{q \log q}.$$

Thus for almost all $x \in \mathbb{I}$ Dirichlet's theorem can be improved by a logarithm.

By using Khintchine's Theorem we can get a direct result on the measure of the set of badly approximable numbers.

Corollary 1.1.8.

$$|\mathbf{Bad}|_1 = 0.$$

PROOF. Consider the function $\psi(q) = 1/(q \log q)$ and observe that

$$\mathbf{Bad} \cap \mathbb{I} \subseteq \mathbb{I} \setminus W(\psi).$$

By Khintchine's Theorem, $|W(\psi)|_1 = 1$. Thus $|\mathbb{I} \setminus W(\psi)|_1 = 0$ and so $|\mathbf{Bad} \cap \mathbb{I}|_1 = 0$. □

This result show that the set **Bad** is small in the Lebesgue measure sense. A similar result exists for very well approximable numbers in that the set of $\tau > 1$ approximable numbers $W(\tau)$ and subsequently Liouville numbers are of Lebesgue measure zero.

1.2. Distribution and computation of partial quotients

1.2.1. Distribution and sum of partial quotients. There are some specific examples of real numbers where we know more about the continued fraction expansion. For instance rationals will always have a finite continued fraction expansion, and the continued fraction expansion of quadratic irrationals are periodic numbers. There are also some sets where we know the properties of the partial quotients, but not specific examples. All **Bad** numbers as mentioned earlier have their partial quotients bounded as such quadratic irrationals are a subset. Here we discuss some results on almost all numbers, it is predicted that all algebraic numbers of degree strictly greater than 2 are not in the exceptional sets of these.

Historically the following was on of the first problems in the measure theory of continued fractions. It was first posed by Gauss in a letter to Lagrange, however it was not solved until 1928 by Kuzmin [**62**]. For a discussion of the problem in detail including a proof of the result, see Khintchine's book [**57**, Chapter 3].

THEOREM 1.2.1 (Gauss-Kuzmin Distribution 1928). *For almost all* $\alpha \in \mathbb{R} \setminus \mathbb{Q}$*, with continued fraction expansion* $\alpha = [a_0; a_1, a_2, \ldots]$*, the partial quotients* a_i *have the property:*

$$\lim_{N\to\infty} \frac{\#\{a_n = k; n \leq N\}}{N} \to -\log_2 \frac{(1+k)}{k(k+1)}. \tag{1.2.1}$$

Khintchine [**57**, Chapter 3] extended Kuzmin's result to consider the average of a function on the partial quotients.

THEOREM 1.2.2 (Khintchine 1935). *Suppose that* $f : \mathbb{R} \to \mathbb{R} : r \mapsto f(r)$ *is a non-negative function and suppose that there exist positive constants* C *and* δ *such that*

$$f(r) < Cr^{1/2-\delta} \qquad r = 1, 2, 3, \ldots.$$

then, for almost all numbers in the interval $(0, 1)$*,*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} f(a_k) = \sum_{r=1}^{\infty} f(r) \frac{\log\left(1 + \frac{1}{r(r+2)}\right)}{\log 2}. \tag{1.2.2}$$

Note that if we assign $f(r) = 1$ for all $r \in \mathbb{R}$ then Theorem 1.2.2 implies Theorem 1.2.1.

Khintchine's constant K, defined by

$$K := \prod_{k=1}^{\infty} \left(1 + \frac{1}{k(k+2)}\right)^{\log k/\log 2},$$

naturally arises from Theorem 1.2.2 by substituting the function $f(r) = \log r$ into (1.2.2). Therefore, for almost all $x \in \mathbb{R}$ with continued fraction

$[a_0; a_1, a_2, \dots]$ the geometric mean tends to K, that is

$$\lim_{n\to\infty}\left(\prod_{k=1}^{n} a_k\right)^{1/n} = K.$$

As the functions $f(r)$ in Theorem 1.2.2 must be increasing slowly, the theorem does not apply to $f(r) = r$. However, Khintchine comments [**57**, Chapter 3] that the average of the partial quotients is simply unbounded,

Theorem 1.2.3. *For almost all real numbers α, with continued fraction $[0; a_1, a_2 \dots]$ there exists infinitely many n such that*

$$\sum_{j=1}^{n} a_j > n \log n. \tag{1.2.3}$$

Diamond and Vaaler [**38**] in 1986 built upon Theorem 1.2.2 to get a much more refined result on the average of the partial quotients. Define the sum $DV_n(\alpha)$ for a real number α with continued fraction expansion $[a_0, a_1, a_2, \dots]$

$$DV_n(\alpha) := \frac{\sum_{j=1}^{n} a_j}{n \log n} - \max\{a_j : j \le n\}$$

Theorem 1.2.4 (Diamond, Vaaler 1986). *For almost all real α (with respect to the Lebesgue measure) with continued fraction $[0; a_1, a_2 \dots]$*

$$\lim_{n\to\infty} DV_n(\alpha) = \frac{1}{\log 2}. \tag{1.2.4}$$

The following conjecture came from a recent work of Beresnevich, Haynes and Velani [**10**].

Conjecture 1.2.5. *For any algebraic number $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, there exists a constant $c_\alpha > 0$ such that for all $k \in \mathbb{N}$*

$$\sum_{n=1}^{k} a_n \le c_\alpha k^2. \tag{1.2.5}$$

The conjecture is in relation to investigations into the sums

$$S_N(\alpha,\gamma) := \sum_{n=1}^{N} \frac{1}{n||n\alpha - \gamma||} \quad \text{and} \quad R_N(\alpha,\gamma) := \sum_{n=1}^{N} \frac{1}{||n\alpha - \gamma||}.$$

The conjecture then implies the following bound on $S_N(\alpha, 0)$, for any algebraic $\alpha \in \mathbb{R} \setminus \mathbb{Q}$

$$S_N(\alpha, 0) \asymp (\log N)^2.$$

1.2.1.1. *Quadratic forms.* Given rationals a, b and c, define a binary quadratic form $f(x, y)$ by

$$f(x, y) = ax^2 + bxy + cy^2.$$

We call the number $D = b^2 - 4ac$ the discriminant of $f(x, y)$. Consider another binary quadratic form with rational coefficients a', b' and c' with function $g(u, v)$. If there exists rational numbers α, β, γ and δ such that $\alpha\delta - \beta\gamma = 1$ and $f(\alpha u + \beta v, \gamma u + \delta v) = g(u, v)$ then we say the two quadratic forms (a, b, c) and (a', b', c') are *equivalent*. Denote by $h(D)$ the number of quadratic forms equivalent to $f(x, y)$, and call it *the class number of the discriminant D*.

Of importance, are the numbers D such that $h(D) = 1$, Stark [**87**] showed that the only such numbers are

$$-D = 3, 4, 7, 8, 11, 19, 43, 67, 163. \tag{1.2.6}$$

Churchhouse and Muir [**30**] discussed the continued fractions of algebraic numbers associated with such discriminants. They showed that the cubic numbers $e^{\pi\sqrt{D}}$ are very close to rationals and therefore have very large partial quotients early on in the numbers continued fraction expansion. One such number in particular is $e^{\pi\sqrt{163}}$, which has a partial quotient of size 1.6×10^7 at position 121. It is predicated that the number will not have "large" partial quotients indefinitely throughout its continued fraction expansion.

1.2.2. Computational methods. There are 4 main methods for calculating the continued fraction of algebraic numbers; basic, indirect, polynomial, and the direct method. The *basic method*, as discussed earlier in §1.1.1.2, finds a good approximation using traditional numerical methods. The algorithm then follows the classical continued fraction algorithm based on Euclid's algorithm. Lehmer [**65**] improved upon this method with some explicit checking to make sure we lose as little precision as possible, this is referred to as the *indirect method*. Lang and Trotter [**64**] use the *polynomial method* which only involves integer arithmetic, avoiding any possible error encountered with floating point values. This paper uses the *direct method* from [**85**] which will be discussed later. See the work of Brent, van der Poorten and te Riele [**19**], for an extensive discussion of the methods.

1.2.2.1. *The Indirect method.* Let $\bar{\alpha}_i$ be a rational approximation of α_i with relative error bounded by δ_i. The indirect or basic method with safe error control reads as follows

for $i = 0, 1, \ldots$

$a_i = \lfloor \alpha_i \rfloor$
if $(a_i + 1)\delta_i < \bar{\alpha}_i - a_i < 1 - (a_i + 1)\delta_i$ **then**
$\bar{\alpha}_{i+1} = 1/(\bar{\alpha}_i - a_i)$
$\delta_{i+1} = \bar{\alpha}_i \bar{\alpha}_{i+1} \delta_i / (1 - \delta_i)$
else stop
endif

The problem of course with this method is that we are using floating point approximations with the $(a_i + 1)\delta_i$ and δ_i, as well as when we wish to calculate more partial quotients, we must first compute a more accurate initial approximation in binary or decimal using other methods.

1.2.2.2. *Polynomial method.* The polynomial method [**64**] is probably the simplest of them all. Assume that α is the zero of a polynomial $f(x)$ with rational coefficients c_i for $0 \le i \le d$,

$$f(x) = c_d x^d + c_{d-1} x^{d-1} + \cdots + c_0.$$

The method is to find the sequence of polynomials

$$f_h(x) = c_{h,d} x^d + c_{h,d-1} x^{d-1} + \cdots + c_{h,0}, \qquad h = 0, 1, 2 \ldots$$

with rational coefficients $c_{h,i}$ and $c_{h,0} \ge 0$ such that the complete quotients $\alpha_h = [a_h, a_{h+1}, \ldots]$ are a zero of f_h. Indeed we will have sequentially

$$f_{h+1}(x) = \pm x^d f_h(x^{-1} + a_h).$$

A core observation is that we will eventually obtain a *reduced* polynomial and the minus sign is always appropriate.

The advantages of this method is that we are only using the coefficients of the polynomial and these are integers which means any rational or floating point arithmetic is avoided. The disadvantage is that the coefficients of the polynomial increase in size rapidly, greatly increasing the computation time and memory storage required.

1.2.2.3. *Direct method.* The direct method comes from the ideas outlined in Shiu's paper [**85**]. The main aim is to find a rational approximation of the complete quotient α_{n+1} (essentially using Newton-Raphson) when the previous partial quotients $a_1, \ldots, a_n$ are known. From that approximation more partial quotients of α_{n+1} can be computed.

Step 1 First calculate a few initial partial quotients, by some other method of α say up to n.

Step 2 Check if $p_n q_{n-1} - p_{n-1} q_n \ne (-1)^{n+1}$ then stop.

Step 3 Compute the next rational approximation α' of α_{n+1} by

$$\alpha' = \frac{(-1)^{n-1}}{q_n^2}\frac{f'(p_n/q_n)}{f(p_n/q_n)} - \frac{q_{n-1}}{q_n}.$$

Step 4 Let $B = bq_n^2$ for some suitable constant $b = b(\alpha)$. While $n + m \leq N$ and $q_{n+m} < B$. Compute the next partial quotients $a_{n+1}, a_{n+2}, \ldots, a_{n+m}, \ldots$ with the basic method.

Step 5 Put $n = n + m$; if $n < N$ go back to step 3.

REMARK 1.2.1. The number $b = b(\alpha)$ is some small real number which relies only on the algebraic number α. It can be estimated by

$$b \sim \frac{|f'(\alpha)|}{|f''(\alpha)|}.$$

The b is then used to define the control value B for the iterative process.

1.2.2.4. *Our specific implementation.* The algorithm is based on the direct method, a few lines are altered to reduce the amount calls to the greatest common divisor function. Let N be the number of partial quotients to calculate and E/M an equivalent of B in the direct method.

```
find initial [a_0; a_1, a_2, ..., a_n].
while n ≤ N
    if p_n q_{n+1} − p_{n+1} q_n = (−1)^{n+1}
        halt
    let f'(p_n/q_n) = x'_n/y'_n and f(p_n/q_n) = x_n/y_n
    let x' = x'_n y_n − q_{n−1} q_n y'_n x_n  y' = q_n^2 y'_n x_n
    reduce α' = x'/y'
    let E = max(q_n^2, M(q_n + 1))
    while q_n M < E and n ≤ N
        n = n + 1
        a_n = ⌊α'⌋
        write a_n
        p_{n+1} = a_n p_n + p_{n−1}
        q_{n+1} = a_n q_n + q_{n+1}
        α' = y' / (x' − a_n y')
```

where β is between α and p_n/q_n and $\alpha_{n+1} = \alpha'$.

The algorithm was implemented in Java and computed on the YARCC machine at the University of York. The code can be found on github [**20**].

1.3. Results and Analysis

1.3.1. Analysis. Some specific example which we calculated the continued fraction expansion of are

- The cube roots

$$\sqrt[3]{2}, \quad \sqrt[3]{3}, \quad \sqrt[3]{4}, \quad \sqrt[3]{5}, \quad \sqrt[3]{7}. \tag{1.3.1}$$

- Three roots from Lang's paper [**64**]

$$\begin{aligned} f(x) &= x^3 + x^2 - 2x - 1, \quad f(2\cos(2\pi/7)) = 0 \qquad x^5 - x - 1, \\ g(x) &= x^6 - 9x^4 - 4x^3 + 27x^2 - 36x - 23, \quad g(\sqrt[3]{2} + \sqrt{3}) = 0. \end{aligned} \tag{1.3.2}$$

- The roots of the polynomials mentioned in [**30**]

$$f(x) = x^3 - 2x^2 - 2, \quad f(e^{\pi\sqrt{43}}) = 0 \qquad g(x) = x^3 - 8x - 10, \quad g(e^{\pi\sqrt{163}}) = 0. \tag{1.3.3}$$

- An example from [**19**]

$$x^4 + 6x^3 + 7x^2 - 6x - 9. \tag{1.3.4}$$

The partial quotients in these examples were calculated using the polynomial method, with 48 hours of computation time.

We then also calculated the continued fraction expansion of the numbers

- The roots of increasing height

$$\sqrt[3]{k}, \qquad 3 \leq k \leq 49.$$

- The roots of increasing degree

$$\sqrt[d]{2}, \qquad 3 \leq d \leq 12.$$

- The roots of increasing height and degree

$$\sqrt[d]{d}, \qquad 3 \leq k \leq 12.$$

The partial quotients in these examples were calculated using our specific implementation, with 24 hours of computation time.

The following tables and graphs show for a given root α of a polynomial $p(x) \in \mathbb{Z}[x]$ with continued fraction $[a_0; a_1, a_2, \dots]$ the following quantities

$$\operatorname{disc}(p(x)), \quad \max_{1 \leq i \leq n} a_i, \quad \left| DV_n(\alpha) - \frac{1}{\log 2} \right|, \quad \left| K - \left(\prod_{i=1}^{n} a_i \right)^{1/n} \right|. \tag{1.3.5}$$

Where $\operatorname{disc}(p(x))$ is the discriminant of the polynomial p. To avoid illegitimate comparisons, each of the quantities in (1.3.5) is found at the smallest

Polynomial	**Discriminant**	**Number of Partial Quotients Calculated**
$x^3 - 2$	-108	2904640
$x^3 - 3$	-243	3341326
$x^3 - 4$	-432	2895879
$x^3 - 5$	-675	2896960
$x^3 - 7$	-1323	2893707
$x^3 + x^2 - 2x - 1$	49	2895493
$x^5 - x - 1$	2869	1353924
$x^6 - 9x^4 - 4x^3 + 27x^2 - 36x - 23$	67941730271232	1068556
$x^3 - 2x^2 - 2$	-172	2894041
$x^3 - 8x - 10$	-652	2892209
$x^4 + 6x^3 + 7x^2 - 6x - 9$	14400	1826267

TABLE 1.1. Polynomials and their discriminant and the number of partial quotients of the largest root calculated in 48 hours.

n, such that n partial quotients were calculated, within each type of algebraic numbers, as classified above. The analysis was done using Jupyter notebooks with matplotlib, numpy and scipy. The notebooks can be found in the git repository [**20**].

1.3.1.1. *Some Specific examples.* First we present some analysis of the specific examples of algebraic numbers.

By looking at Table 1.1, there is an obvious link between the discriminant of the polynomial and the complexity of calculating the continued fraction of the root. Specifically for calculating the continued fraction of $\sqrt[3]{2} + \sqrt{3}$ root of polynomial with discriminant 6.7×10^{13}, less than 1.1 million partial quotients were calculated within 48 hours, whereas for the number $\sqrt[3]{3}$ root of polynomial with discriminant -243 over three million partial quotients were calculated.

Figures 1.9 and 1.10 display the averaging function

$$\lambda_N(\alpha) = \frac{1}{N} \sum_{n=1}^{N} a_n$$

as N increases. The figures show that the partial quotients obey (1.2.5) and (1.2.3), at least for those partial quotients calculated.

Table 1.7 and Table 1.8 show the ten largest partial quotients of the algebraic numbers from (1.3.2) and (1.3.3) (1.3.4) respectively. As mentioned earlier, one of the interesting outliers is the root of $x^3 - 8x - 10$ with the two large partial quotients 1501790 and 16467250 in the first 200 partial quotients. However, the maximum partial quotients after that initial large value grow much slower, as expected.

Table 1.6 shows the first ten most common partial quotients amongst different algebraic numbers. The partial quotients do not appear to deviate from the expected Gauss-Kuzmin distribution.

1.3.1.2. *Increasing height and degree.* Let the height $H(p)$ of a polynomial $p \in \mathbb{R}[x]$ be the largest coefficient in the polynomial, i.e.

$$p(x) = c_n x^n + c_{n-1} x^{n-1} + \cdots + c_0, \qquad H(p) = \max\{|c_0|, |c_1|, \ldots, |c_n|\}.$$

A curious question is what is the relation to the speed of growth of q_n in the convergents of the approximation to the root of polynomials as the height and degree of the polynomials change? Or, equivalently the maximum partial quotient a_n? As the height and degree increase, so does the discriminant, therefore the difficulty in finding the partial quotients also increases. Hence, only few partial quotients of each algebraic number of the form $\sqrt[d]{k}$, have been calculated.

Figure 1.1 shows the change in the maximum partial quotient at $n = 433664$ from the root $\sqrt[3]{k}$ as k increases. No discernible pattern can be seen, there are just large fluctuations between each value. It also plots the two quantities in (1.3.5) related to DV_n and K, both of which seem to follow the trend of the maximum partial quotient.

Figure 1.2 shows the change in the maximum partial quotient at $n = 391836$ from the root $\sqrt[d]{2}$ as d increases. The difference to K seems to be following an opposite trend as the maximum value and the difference to DV_n.

Figure 1.3 shows the change in the maximum partial quotient at $n = 418362$ of the partial quotients of $\sqrt[k]{k}$ as k increases. There appears to be a slight upwards trend of the maximum partial quotient, a possible explanation is the number $\sqrt[k]{k}$ is rapidly closer and closer to 1 as k increases and hence there are better rational approximations.

Figures 1.4, 1.5 seem to bear the same pattern as mentioned for Figure 1.3 in that the maximum partial quotient for the numbers $\sqrt[k]{k}$ seems to be increasing in k, unlike the numbers $\sqrt[d]{2}$. However the outlying maximum partial quotient of $\sqrt[11]{11}$ could be skewing the data to justify such a hypothesis.

As in Figures 1.9 and 1.10, the mean of the partial quotients in Figures 1.7 and 1.8 obey (1.2.5) and (1.2.3), at least for those calculated.

1.3.1.3. *Conclusion and Future work.* All of the data collected so far indicates that algebraic numbers are not in the exceptional sets of Theorems 1.2.1, 1.2.2 and 1.2.4. Of course this is only experimental evidence and does not prove the theorems. However, the increase in height and degree of the algebraic numbers did not make a noticeable difference in the distributions in the long run.

In terms of improving the implementation, there are several steps that could be taken to improve the speed of calculation and therefore calculate more partial quotients. While programming the implementation, it was found that the direct method applied to algebraic numbers of degree greater than 13, produced erroneous data if $b(\alpha)$ was larger than $1/1000$. However choosing $b(\alpha)$ smaller than $1/1000$ reduced the speed of calculation significantly. More work could be done to find the "best" value of $b(\alpha)$.

It would also be advantageous to find the most suitable way to calculate initial partial quotients for the direct method. In our implementation we found at least 100. Since the polynomial method is rather slow for large degree polynomials, in the case for large degree polynomials it may be more useful to use the indirect method.

In [**22**, Appendix 8], Bugeaud presents for a fixed $\epsilon > 0$ an upper bound on the number of solutions to (1.1.7) given the height and degree of a polynomial. It would be interesting to study the convergents of algebraic numbers and therefore see how "good" an upper bound it is experimentally.

1.4. Results

Poly	Discriminant	Max	DV	Khintchine
$x^3 - 3$	-243	7.384×10^5	0.0852	0.004178
$x^5 - 5$	1.953×10^6	1.993×10^6	0.02216	0.002637
$x^6 - 6$	3.628×10^8	8.292×10^6	0.2195	0.001011
$x^7 - 7$	-9.689×10^{10}	2.498×10^5	0.05741	0.001098
$x^8 - 8$	-3.518×10^{13}	1.745×10^5	0.1752	0.003253
$x^9 - 9$	1.668×10^{16}	6.418×10^5	0.03964	0.0007948
$x^{10} - 10$	1×10^{19}	2.633×10^6	0.3395	0.0007155
$x^{11} - 11$	-7.4×10^{21}	4.271×10^7	0.1305	0.001437
$x^{12} - 12$	-6.625×10^{24}	4.681×10^5	0.1201	0.004698

Table 1.2. Data of continued fraction of algebraic numbers $\sqrt[k]{k}$ for $3 \leq k \leq 12$ at $n = 441218$.

Poly	Discriminant	Max	DV	Khintchine
$x^3 - 2$	-108	5.337×10^5	0.02689	0.001714
$x^4 - 2$	-2048	9.57×10^5	0.1297	0.001122
$x^5 - 2$	5×10^4	3.391×10^6	0.5377	0.002888
$x^6 - 2$	1.493×10^6	1.117×10^6	0.03908	0.002258
$x^7 - 2$	-5.271×10^7	1.971×10^5	0.1458	0.005849
$x^8 - 2$	-2.147×10^9	6.88×10^5	0.04176	0.004013
$x^9 - 2$	9.918×10^{10}	6.725×10^5	0.2118	0.004817
$x^{10} - 2$	5.12×10^{12}	2.125×10^5	0.1904	0.002488
$x^{11} - 2$	-2.922×10^{14}	3.032×10^5	0.1247	0.0006795
$x^{12} - 2$	-1.826×10^{16}	7.527×10^5	0.01319	0.002483

Table 1.3. Data of continued fraction of algebraic numbers $\sqrt[d]{2}$ for $3 \leq d \leq 12$ at $n = 441231$

Poly	Discriminant	Max	DV	Khintchine
$x^3 - 2$	-108	4.887×10^6	0.09139	0.001202
$x^3 - 3$	-243	5.868×10^6	0.07561	0.001176
$x^3 - 4$	-432	8.313×10^6	0.1405	0.003058
$x^3 - 5$	-675	1.678×10^7	0.06549	0.002921
$x^3 - 7$	-1323	1.201×10^7	0.02145	0.0008739

TABLE 1.4. Data of continued fraction of algebraic numbers $\sqrt[3]{k}$ with $2 \leq k \leq 7$ at $n = 2893707$.

Poly	Discriminant	Max	DV	Khintchine
$x^3 + x^2 - 2x - 1$	49	5.799×10^5	0.1733	0.003367
$x^5 - x - 1$	2869	1.608×10^7	0.1356	0.002458
$x^6 - 9x^4 - 4x^3 + 27x^2 - 36x - 23$	6.794×10^{13}	1.802×10^6	0.00266	0.004648
Poly	Discriminant	Max	DV	Khintchine
$x^3 - 2x^2 - 2$	-172	4.395×10^6	0.09173	0.00253
$x^3 - 8x - 10$	-652	1.647×10^7	0.3955	0.0004621
$x^4 + 6x^3 + 7x^2 - 6x - 9$	1.44×10^4	7.296×10^6	0.1361	0.0004407

TABLE 1.5. Data of continued fraction of specific examples of algebraic numbers, the first 3 examples from Lang at $n = 1068556$, the second 3 those with class number 1 at $n = 1826267$.

$x^3 - 2$	$x^3 - 3$	$x^3 - 4$	$x^3 - 5$	$x^3 - 7$
0.41498	0.41511	0.41541	0.41543	0.41498
0.16994	0.17012	0.16984	0.16982	0.16979
0.09320	0.09294	0.09323	0.09324	0.09307
0.05880	0.05881	0.05901	0.05875	0.05885
0.04072	0.04080	0.04047	0.04079	0.04067
0.02976	0.02967	0.02971	0.02963	0.02986
0.02268	0.02261	0.02268	0.02274	0.02276
0.01798	0.01786	0.01789	0.01769	0.01805
0.01460	0.01438	0.01441	0.01449	0.01445
0.01190	0.01192	0.01192	0.01186	0.01196

TABLE 1.6. Percentage of most common partial quotients of algebraic numbers $\sqrt[3]{k}$ for $3 \leq k \leq 7$.

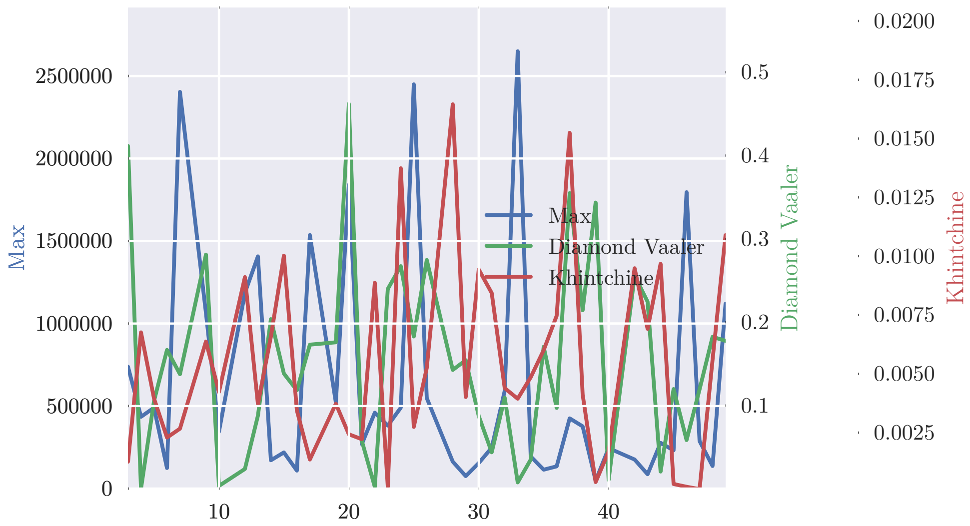

FIGURE 1.1. Plot of the change of the maximum partial quotient, difference to Khintchine's constant, and difference to $1/\log 2$ of Diamond-Vaaler sum of algebraic numbers $\sqrt[3]{k}$ for $3 \leq k \leq 49$ at $n = 433664$.

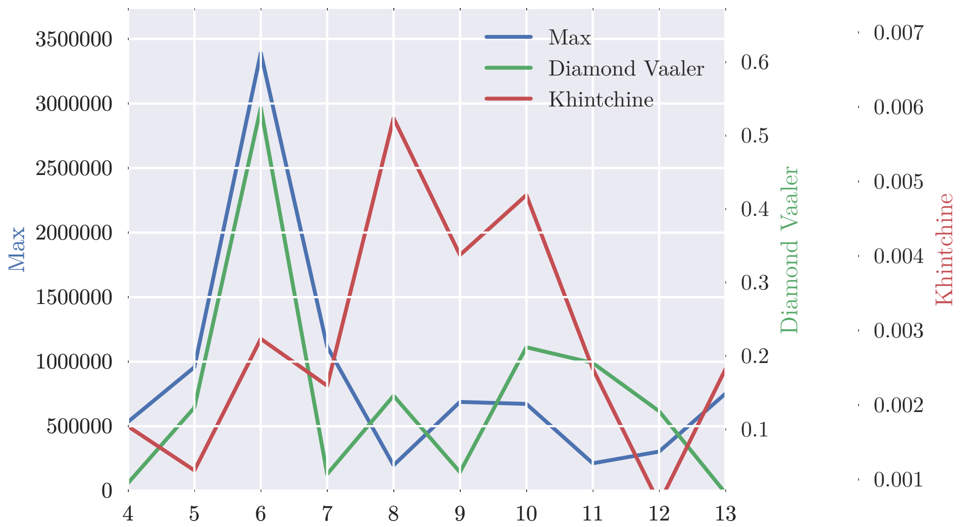

FIGURE 1.2. Plot of the change of the maximum partial quotient, difference to Khintchine's constant, and difference to $1/\log 2$ of Diamond-Vaaler sum of algebraic numbers $\sqrt[d]{2}$ for $3 \leq d \leq 12$ at $n = 391836$.

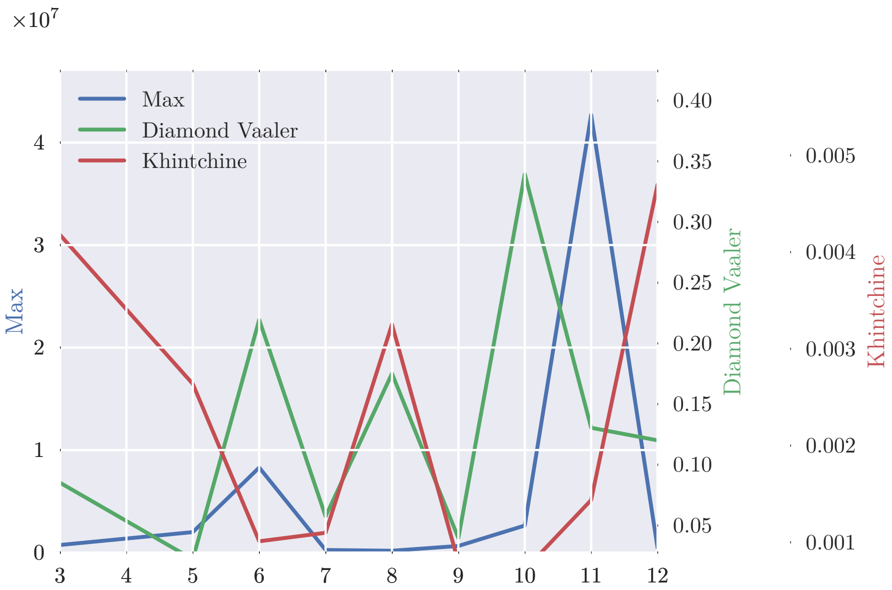

FIGURE 1.3. Plot of the change of the maximum partial quotient, difference to Khintchine's constant, and difference to $1/\log 2$ of Diamond-Vaaler sum of algebraic numbers $\sqrt[k]{k}$ for $3 \leq k \leq 12$ at $n = 418362$.

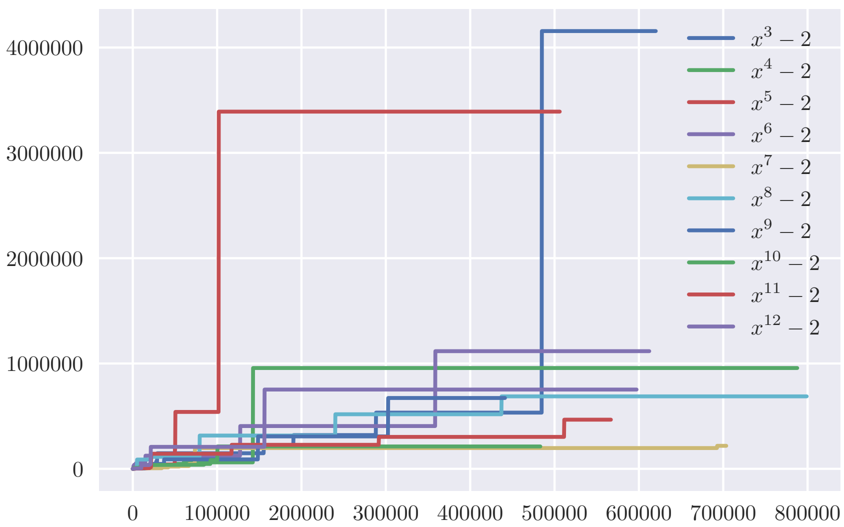

FIGURE 1.4. Plot of the change of the maximum partial quotient of algebraic numbers $\sqrt[d]{2}$ for $3 \leq d \leq 12$.

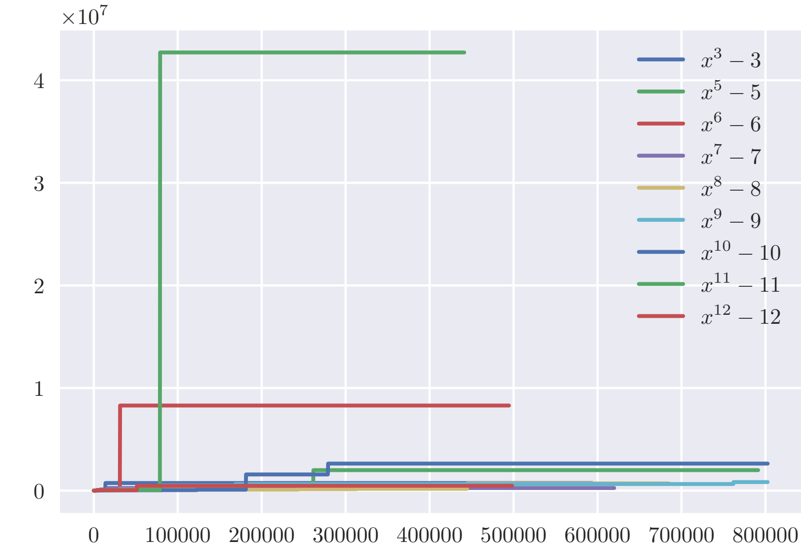

FIGURE 1.5. Plot of the change of the maximum partial quotient of algebraic numbers $\sqrt[k]{k}$ for $3 \leq k \leq 12$.

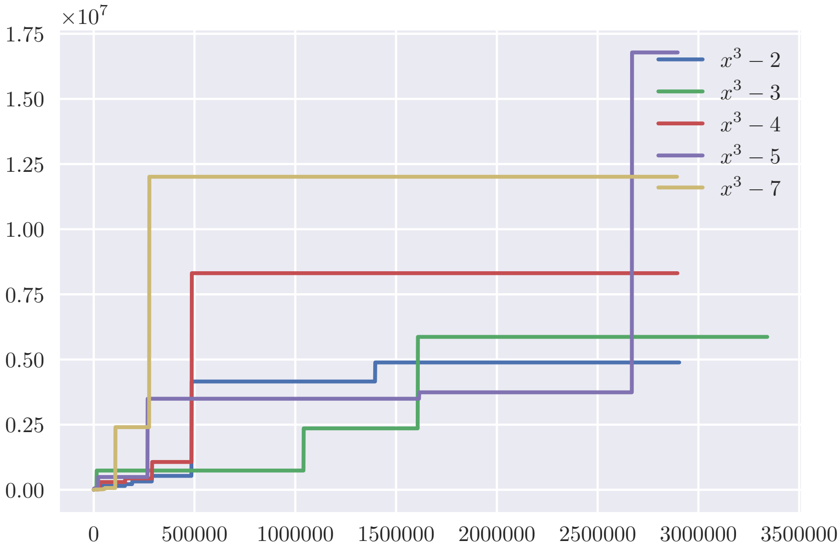

FIGURE 1.6. Plot of the change of the maximum partial quotient of algebraic numbers $\sqrt[3]{k}$ for $2 \leq k \leq 7$.

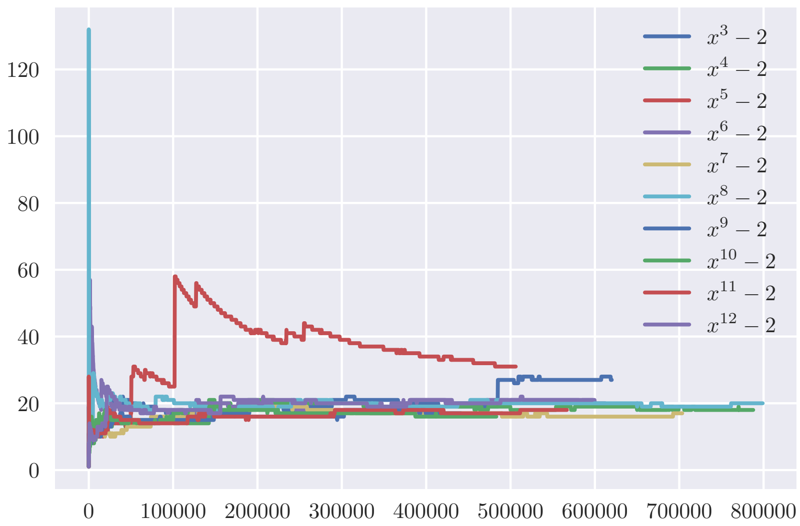

FIGURE 1.7. Plot of the mean of the partial quotients of algebraic numbers $\sqrt[d]{2}$ for $3 \leq d \leq 12$.

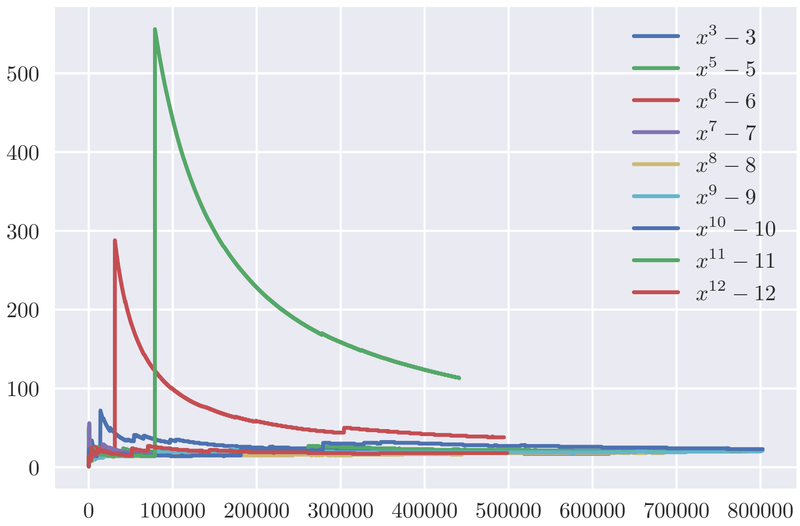

FIGURE 1.8. Plot of the mean of the partial quotients of algebraic numbers $\sqrt[k]{k}$ for $3 \leq k \leq 12$.

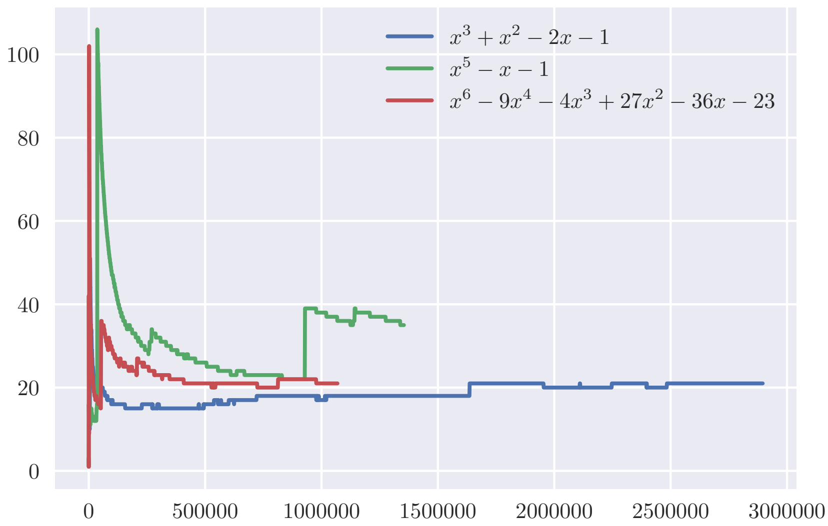

FIGURE 1.9. Plot of the mean of the partial quotients of algebraic numbers $2\cos\frac{2\pi}{7}$, root of $x^5 - x - 1$ and $\sqrt[3]{2} + \sqrt{3}$.

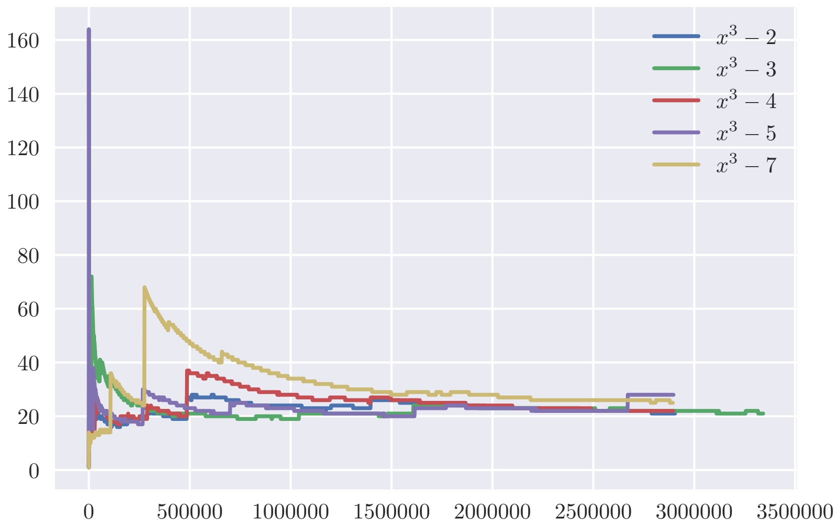

FIGURE 1.10. Plot of the mean of the partial quotients of algebraic numbers $\sqrt[3]{k}$ for $3 \leq k \leq 7$.

$x^3-8x-10$		x^3-2x^2-2		x^3-2x^2-2x-2	
n	a_n	n	a_n	n	a_n
133001	1017397	33	1501790	142839	7295890
343360	439547	121	16467250	200537	273863
529540	4394508	832979	2699885	216860	333158
710600	781932	968939	1758896	412227	400318
1360767	425531	970223	723895	951717	256244
1516087	630718	1290101	2319808	1125342	277275
1558423	1255067	1516898	7005770	1128424	2098560
1912564	645034	1775920	596178	1292228	481573
2306580	935373	2001985	1689240	1703812	410318
2845433	1627737	2216288	2084037	1722508	499576

TABLE 1.7. Largest 10 partial quotients of examples (1.3.3) and (1.3.4) at $n = 1826267$.

$2\cos(2\pi/7)$		x^5-x-1		$\sqrt[3]{2}+\sqrt{3}$	
n	a_n	n	a_n	n	a_n
720371	579913	36377	3297074	52062	1075748
1563150	910774	169519	238865	84267	282368
1635477	3758763	260254	659182	179369	224193
1704847	609719	268941	559520	206496	716296
2108691	812891	418443	284809	315752	242099
2246320	1245916	634794	423878	551657	358103
2482959	588120	928061	16084222	602486	305699
2484071	952134	948168	518861	812259	1801591
2699190	880465	1141464	3508145	830638	224052
2817914	640703	1171445	420754	892765	229232

TABLE 1.8. Largest 10 partial quotients of examples (1.3.2) at $n = 1068556$

CHAPTER 2

Singular and extremal sets associated with Dirichlet systems

2.1. Introduction

In the recent paper [**9**], Beresnevich, Ghosh, Simmons and Velani introduced the notion of singular and extremal points associated with the limit sets of a Kleinian group. The goal of this work is to develop a general framework of "Dirichlet systems" inspired by the ubiquity setups of [**8**] and [**60**], that naturally incorporate the Kleinian group results of [**9**]. The framework will almost certainly allow us to prove the analogous statements for rational maps and indeed general hyperbolic dynamical systems - This will be addressed in the near future. Before describing the general framework, we provide a brief overview of singular and extremal sets associated with the classical theory of Diophantine approximation. This will provide the context for the general framework.

2.1.1. The classical singular theory. We start by recalling Dirichlet's fundamental theorem in the theory of simultaneous Diophantine approximation.

THEOREM 2.1.1 (Dirichlet's Theorem). *For any* $\mathbf{x} = (x_1, x_2, \ldots, x_m) \in \mathbb{R}^m$ *and* $N \in \mathbb{N}$*, there exists* $\mathbf{p} = (p_1, \ldots, p_m) \in \mathbb{Z}^m$*,* $q \in \mathbb{N}$ *such that*

$$\max_{1\le i\le m}\left|x_i - \frac{p_i}{q}\right| \le \frac{1}{qN^{1/m}} \qquad \text{and} \qquad q < N.$$

An obvious question to ask is: can Dirichlet's theorem be improved? In particular, does there exist vectors $\mathbf{x} \in \mathbb{R}^m$ such that the constant 1 can be arbitrarily improved? With this is mind, a vector $\mathbf{x} \in \mathbb{R}^m$ is said to be *singular* if for every $\epsilon > 0$ there exists N_0 with the following property: for each $N \ge N_0$, there exist $\mathbf{p} \in \mathbb{Z}^m$, $q \in \mathbb{N}$ so that

$$\max_{1\le i\le m}\left|x_i - \frac{p_i}{q}\right| < \frac{\epsilon}{qN^{\frac{1}{m}}} \qquad \text{and} \qquad q < N\,. \tag{2.1.1}$$

In short, $\mathbf{x}$ is singular if Dirichlet's Theorem can be "improved" by an arbitrarily small constant factor $\epsilon > 0$. It is not difficult to see that the set $\mathrm{Sing}(m)$ of singular vectors contains every rational hyperplane in $\mathbb{R}^m$ and thus its Hausdorff dimension is between $m-1$ and m. In the case $m=1$, a nifty argument (which we shall utilise) due to Khintchine [**56**] shows that a real number is singular if and only if it is rational; that is, $\mathrm{Sing}(1)=\mathbb{Q}$. Davenport & Schmidt [**36**] in the seventies showed that $\mathrm{Sing}(m)$ is a set of m-dimensional Lebesgue measure zero. Recently, Cheung & Chevallier [**28**], building on the pioneering $m=2$ work of Cheung [**27**], have shown that $\mathrm{Sing}(m)$ has Hausdorff dimension $\frac{m^2}{m+1}$.

2.1.2. The classical extremal theory and beyond . Theorem 2.1.1 implies the following well known statement which is also referred to as Dirichlet's Theorem.

THEOREM 2.1.2 (Dirichlet's Theorem). *For any* $\mathbf{x}\in\mathbb{R}^m$ *and* $N\in\mathbb{N}$, *there exist infinity many* $(\mathbf{p},q)\in\mathbb{Z}^m\times\mathbb{N}$ *such that*

$$\max_{1\le i\le m}\left|x_i-\frac{p_i}{q}\right|\le q^{-\frac{m+1}{m}}\ . \tag{2.1.2}$$

This statements describes to what extent irrational points in $\mathbb{R}^m$ may be approximated by rational points; namely every irrational point $\mathbf{x}$ can be approximated by rational points $(p_1/q,\dots,p_m/q)$ with "rate" of approximation given by $q^{-(m+1)/m}$ – the right-hand side of (2.1.2) determines the "rate" or "error" of approximation. It is natural to broaden the discussion to include general *approximating functions* ψ. More precisely, let $\psi:\mathbb{R}^+\to\mathbb{R}^+$ be a decreasing function and let

$$W(m,\psi)=\left\{x\in\mathbb{I}^m:\max_{1\le i\le m}\left|x_i-\frac{p_i}{q}\right|\le\psi(q)\text{ for i.m. }(\mathbf{p},q)\in\mathbb{Z}^m\times\mathbb{N}\right\}$$

where 'i.m.' means 'infinitely many'. This is the classical set of ψ-well approximable points in the theory of Diophantine approximation. The fact that we have restricted our attention to points $\mathbf{x}$ in the unit cube $\mathbb{I}^m:=[0,1]^m$ is purely for convenience – it makes statements less ambiguous and easier to write. Indeed, it is readily verified that the set of ψ-approximable points in $\mathbb{R}^m$ is invariant under translations by integer vectors.

In the case $\psi:q\to q^{-\tau}$ for some $\tau>0$, we write $W(m,\tau)$ for $W(m,\psi)$. The set $W(m,\tau)$ is usually referred to as the set of τ*-well approximable points* and note that in view of Theorem 2.1.2 we have that

$$W(m,\tau)=\mathbb{I}^m\qquad\text{for}\qquad\tau\le\frac{m+1}{m}\,.$$

On the other hand, a simple consequence of the Borel-Cantelli Lemma from probability theory is that

$$|W(m,\tau)|_m=0\qquad\text{for}\qquad\tau>\frac{m+1}{m}\,.$$

Here and throughout, $|X|_m$ denotes the m-dimensional Lebesgue measure of a subset $X \subseteq \mathbb{R}^m$. Concerning the set of ψ-well approximable points, we have the following fundamental statement that provides an elegant criterion for the size of $W(m, \psi)$ expressed in terms of Lebesgue measure.

THEOREM 2.1.3 (Khintchine). *Let ψ be an approximating function. Then*

$$|W(m,\psi)|_m = \begin{cases} 0 & \text{if} \quad \sum_{r=1}^{\infty} (\psi(r)\, r)^m < \infty \,, \\ \\ 1 & \text{if} \quad \sum_{r=1}^{\infty} (\psi(r)\, r)^m = \infty \,. \end{cases}$$

We now turn our attention to the manifold theory. In short, Diophantine approximation on manifolds is the study of the Diophantine properties of points in $\mathbb{R}^m$ whose coordinates are constrained by (differentiable) functional relations, or equivalently points which are known to be members of a submanifold $\mathcal{M} \subseteq \mathbb{R}^m$. Actually, there is no harm in restricting our attention to submanifolds $\mathcal{M} \subseteq \mathbb{I}^m$, and the specific aspect of the manifold theory that we will be concerned with is that of describing the measure of $\mathcal{M} \cap W(m, \psi)$ (with respect to the Lebesgue measure on $\mathcal{M}$). The fact that the points of interest $\mathbf{x} \in \mathbb{I}^m$ are constrained by functional relations, or in other words that they are required to be members of a fixed manifold $\mathcal{M}$, introduces major difficulties in attempting to analyse the measure-theoretic structure of $\mathcal{M} \cap W(m, \psi)$. This is true even for seemingly simple curves such as the unit circle or the parabola.

The goal is to obtain a Khintchine-type theorem that describes the Lebesgue measure of the set of ψ–approximable points lying on any given manifold. Notice that if the dimension k of the manifold $\mathcal{M}$ is strictly less than m then $|\mathcal{M} \cap W(m, \psi)|_m = 0$ irrespective of the approximating function ψ. Thus, in attempting to develop a general Lebesgue theory for $\mathcal{M} \cap W(m, \psi)$ it is natural to use the normalised k-dimensional Lebesgue measure on $\mathcal{M}$. This will be denoted by $|\cdot|_{\mathcal{M}}$. In order to make any reasonable progress with developing a general theory, we insist that the manifolds $\mathcal{M}$ under consideration are *nondegenerate manifolds*. Essentially, these are smooth submanifolds of $\mathbb{R}^m$ which are sufficiently curved so as to deviate from any hyperplane. For a formal definition and indeed a more in-depth overview of the manifold theory, we refer the reader to [**11**, Section 6] and the references within. In terms of examples, any connected analytic manifold not contained in any hyperplane of $\mathbb{R}^m$ is nondegenerate. Also, a planar curve $\mathcal{C}$ is nondegenerate if the set of points on $\mathcal{C}$ at which the curvature vanishes is a set of one-dimensional Lebesgue measure zero.

The claim is that the notion of nondegeneracy is the right criterion for a manifold $\mathcal{M}$ to be "sufficiently" curved in order to obtain a Khintchine-type theorem (both convergence and divergence cases) for $\mathcal{M} \cap W(m, \psi)$.

Conjecture 2.1.4 (The Dream Theorem). *Let $\mathcal{M}$ be a nondegenerate submanifold of $\mathbb{R}^m$. Then*

$$|\mathcal{M} \cap W(m,\psi)|_{\mathcal{M}} = \begin{cases} 0 & \text{if} \quad \sum_{r=1}^{\infty} (\psi(r)\,r)^m < \infty \,, \\ \\ 1 & \text{if} \quad \sum_{r=1}^{\infty} (\psi(r)\,r)^m = \infty \,. \end{cases} \tag{2.1.3}$$

We now describe various "general" contributions towards the Dream Theorem.

- *Extremal manifolds.* A submanifold $\mathcal{M}$ of $\mathbb{R}^m$ is called *extremal* if

  $$|\mathcal{M} \cap W(m,\tau)|_{\mathcal{M}} = 0 \qquad \forall \;\; \tau > \tfrac{m+1}{m} \,.$$

  Note that Dirichlet's theorem implies that $W(m, \frac{m+1}{m}) = \mathbb{I}^m$ and so it trivially follows that $\mathcal{M} \cap W(m, \frac{m+1}{m}) = \mathcal{M}$. In their pioneering work [**59**], Kleinbock & Margulis proved that any nondegenerate submanifold $\mathcal{M}$ of $\mathbb{R}^m$ is extremal. It is easy to see that this implies the convergence case of the Dream Theorem for functions $\psi : r \mapsto r^{-\tau}$. It is worth mentioning that Kleinbock & Margulis established a stronger (multiplicative) form of extremality that settled the Baker–Sprindžuk Conjecture from the eighties. Establishing extremality has essentially been the catalyst for subsequent work described below.
- *Planar curves.* The Dream Theorem is true when $m = 2$; that is, when $\mathcal{M}$ is a nondegenerate planar curve. The convergence case of (2.1.3) for planar curves was established in [**93**] and subsequently strengthened in [**15**]. The divergence case of (2.1.3) for planar curves was established in [**8**].
- *Beyond planar curves.* The divergence case of the Dream Theorem is true for analytic nondegenerate submanifolds of $\mathbb{R}^m$ [**7**]. In current work [**13**] being written up, the divergence case of (2.1.3) will be shown to be true for nondegenerate curves, as well as manifolds that can be "fibred" into such curves [**13**]. The latter includes C^∞ nondegenerate submanifolds of $\mathbb{R}^m$ which are not necessarily analytic. The convergence case of the Dream Theorem is true for a large class of nondegenerate submanifolds of $\mathbb{R}^m$ with dimension k satisfying $k(k+3)/2 > m$, and this class includes "most" manifolds when $k(k+1)/2 \geq m$ [**86**]. The work in [**86**] builds upon

the approach taken in [**12**] in which the convergence case is shown to be true for a large subclass of nondegenerate submanifolds with $k > (m+1)/2$.

The upshot of the above is that the Dream Theorem actually holds for a fairly generic class of nondegenerate submanifolds $\mathcal{M}$ of $\mathbb{R}^m$ apart from the case of convergence when $m \geq 3$ and $k(k+1)/2 < m$.

REMARK 2.1.1. In [**58**], Kleinbock, Lindenstrauss, & Weiss made a serious generalisation of the "extremal" work of [**59**] to subsets K of $\mathbb{R}^m$ supporting so-called *friendly* measures. Within the context of this work, it suffices to say that friendly measures form a large and natural class of measures on $\mathbb{R}^m$ which includes Riemannian measures supported on nondegenerate manifolds, fractal measures supported on self-similar sets satisfying the open set condition (e.g. regular Cantor sets, the Koch snowflake, the Sierpiński gasket), and conformal (Patterson) measures supported on the limit sets of geometrically finite Kleinian groups, as long as they are not contained in any hyperplane. These facts are proven in [**58**, Theorem 2.3] and [**33**, Theorem 1.9], respectively. Recently, the concept of friendly measures has been generalised even further to the notion of *quasi-decaying* measures, see [**32**, **33**].

2.2. The general framework: Dirichlet systems

Let (Ω, d) be a compact metric space equipped with a non-atomic, probability measure ν. Let $\mathcal{R} = \{R_\alpha \subseteq \Omega : \alpha \in J\}$ be a family of points R_α of Ω indexed by an infinite, countable set J. The points R_α will be referred to as *resonant points* for reasons which will become apparent later. Next, let $\beta : J \to \mathbb{R}^+ : \alpha \mapsto \beta_\alpha$ be a positive function on J. Thus, the function β attaches a 'weight' β_α to the resonant point R_α. To avoid pathological situations, we shall assume that the number of α in J with β_α bounded above is always finite; i.e. for any $T \geq 1$

$$\#\{\alpha \in J : \beta_\alpha \leq T\} < \infty \,. \tag{2.2.1}$$

Given a decreasing function $\psi : \mathbb{R}^+ \to \mathbb{R}^+$ let

$$\Lambda(\psi) = \{x \in \Omega : d(x, R_\alpha) \leq \psi(\beta_\alpha)) \text{ for infinitely many } \alpha \in J\} \,.$$

The set $\Lambda(\psi)$ is a 'lim sup' set; it consists of points x in Ω which lie in infinitely many of the balls $B(R_\alpha, \psi(\beta_\alpha))$ centred at resonant points. In general, $B(c, r) := \{x \in \Omega : d(x, c) \leq r\}$ is the ball in Ω centred at c and radius r. points. As in the classical setting, it is natural to refer to the function ψ as the approximating function. It governs the 'rate' at which points in Ω must be approximated by resonant sets in order to lie in $\Lambda(\psi)$. Let us rewrite $\Lambda(\psi)$ in a fashion which brings its 'lim sup' nature to the forefront – we shall make use of this later. For $t \in \mathbb{N}$, let

$$\Delta(\psi, t) := \bigcup_{\alpha \in J \,:\, k^t < \beta_\alpha \leq k^{t+1}} B(R_\alpha, \psi(\beta_\alpha)) \quad \text{where } k > 1 \text{ is fixed.} \tag{2.2.2}$$

By assumption, the number of α in J with $k^{t-1} < \beta_\alpha \leq k^t$ is finite regardless of the value of k. Thus, $\Lambda(\psi)$ is precisely the set of points in Ω which lie in infinitely many $\Delta(\psi, t)$; that is

$$\Lambda(\psi) = \limsup_{t \to \infty} \Delta(\psi, t) := \bigcap_{s=1}^{\infty} \bigcup_{t=s}^{\infty} \Delta(\psi, t) \ . \tag{2.2.3}$$

Observe that the classical set $W(\psi) = W(\psi, 1)$ of ψ-well approximable numbers can be expressed in the form $\Lambda(\psi)$ with

$$\Omega := [0,1] \ , \quad J := \{(p,q) \in \mathbb{Z} \times \mathbb{N} : 0 \leq p \leq q\} \ , \quad \alpha := (p,q) \in J \ ,$$

$$\beta_\alpha := q \ , \quad R_\alpha := p/q \quad \text{and} \quad \Delta(R_\alpha, \psi(\beta_\alpha)) := B(p/q, \psi(q)) \ .$$

The metric d is of course the standard Euclidean metric; $d(x,y) := |x - y|$. Thus in this **basic example**, the resonant points R_α are simply rational points p/q. Furthermore,

$$\Delta(\psi, t) := \bigcup_{k^t < q \leq k^{t+1}} \ \bigcup_{p=0}^{q} B(p/q, \psi(q)) \qquad \text{and} \qquad W(\psi) = \limsup_{t \to \infty} \Delta(\psi, t) \ .$$

With respect to the general setup, given a pair $(\mathcal{R}, \beta)$ the goal is to develop a singular and extremal theory akin to the classical theory. Clearly, in order to this we need the existence of a Dirichlet type theorem. This is precisely the purpose of the following condition.

Condition D. *There exist (i) decreasing functions* $g : \mathbb{R}^+ \to \mathbb{R}^+$ *and* $h : \mathbb{R}^+ \to \mathbb{R}^+$ *with* $h(r) \to 0$ *as* $r \to \infty$ *and (ii) constants* $k > 1$, $C_1 > 0$ *and* $t_0 > 0$ *such that: for all* $x \in \Omega$ *and integers* $t > t_0$ *there exists* $\alpha \in J$ *such that*

$$d(x, R_\alpha) \leq C_1 g(\beta_\alpha) h(k^t) \qquad \textit{and} \qquad \beta_\alpha \leq k^t.$$

Whenever Condition D is satisfied, we say that the pair $(\mathcal{R}, \beta)$ is a *Dirichlet system relative to* (g, h, k). Observe that with respect to the basic example, Condition D with g and h given by

$$g(r) = h(r) = r^{-1} \tag{2.2.4}$$

naturally coincides with Dirichlet's classical one dimensional statement; namely Theorem 2.1.1 with $m = 1$.

2.2.1. Dirichlet systems and Singular Points. Let (Ω, d) be a compact metric space and suppose that $(\mathcal{R}, \beta)$ is a Dirichlet system relative to (g, h, k). Motivated by the classical "singular" theory we say that a point x in Ω is singular if Condition D can be "improved" by an arbitrary small constant. We now state this formally.

DEFINITION 2.2.1. Let (Ω, d) be a compact metric space and suppose that $(\mathcal{R}, \beta)$ is a Dirichlet system relative to (g, h, k). A point $x \in \Omega$ is said to be *singular* if for all $\epsilon > 0$ there exists $t_0 > 0$ with the following property: for each integer $t > t_0$ there exists $\alpha \in J$ so that

$$d(x, R_\alpha) < \epsilon g(\beta_\alpha) h(k^t) \qquad \text{and} \qquad \beta_\alpha < k^t.$$

Clearly, with respect to the basic example and with g and h given by (2.2.4), the above definition reduces to the classical one dimensional definition of singular numbers.

In order for us to say anything sensible regarding the set of singular points in Ω we impose the following separation condition on the set $\mathcal{R}$ of resonant points. For convenience, given the functions h and g associated with a Dirichlet system, let

$$f : \mathbb{R}^+ \to \mathbb{R}^+ : x \to f(x) := g(x)h(x). \tag{2.2.5}$$

Observe that $f(r) \to 0$ as $r \to \infty$ and that this together with (2.2.1) implies that for any $x \in \Omega \setminus \mathcal{R}$, there exists infinitely many $\alpha \in J$ such that

$$d(x, R_\alpha) \leq C_1 f(\beta_\alpha)\,.$$

CONDITION S. *There exist constants $t_0 > 0$ and $C_2 > 0$ with the following property: for any $\alpha, \alpha' \in J$ with $R_\alpha \neq R_{\alpha'}$ and $\beta_\alpha, \beta_{\alpha'} > t_0$ we have that*

$$d(R_\alpha, R_{\alpha'}) > C_2 \min\{f(\beta_\alpha), f(\beta_{\alpha'})\}. \tag{2.2.6}$$

Note that, with respect to the basic example and with h and g given by (2.2.4), Condition S is trivially satisfied since for any two distinct rationals $p_1/q_1, p_2/q_2$ we have that

$$\left|\frac{p_1}{q_1} - \frac{p_2}{q_2}\right| \geq \frac{1}{q_1 q_2} > \min\left\{\frac{1}{q_1^2}, \frac{1}{q_2^2}\right\}\ .$$

However, it is worth pointing out that Condition S does not hold in higher dimensions for the classical setup. Indeed, the higher dimension analogue of the above rational separation property is given by the simplex lemma – see [**60**, Lemma 4].

Whenever both Conditions D and S are satisfied, we say that the pair $(\mathcal{R},\beta)$ is a *well separated Dirichlet system relative to* (g,h,k). For such systems we are able to prove the following statement for singular points.

THEOREM 2.2.1. *Let* (Ω,d) *be a compact metric space. Suppose that* $(\mathcal{R},\beta)$ *is a well separated Dirichlet system relative to* (g,h,k)*. Furthermore, suppose there exist constants* $\lambda=\lambda(k)>1$ *and* $t_0>0$ *such that*

$$\frac{g(x)}{g(y)} \leq \frac{h(x)}{h(y)} \qquad \forall \quad y\geq x>t_0 \tag{2.2.7}$$

and

$$h(k^t) < \lambda\, h(k^{t+1}) \qquad \forall \quad t>t_0\,. \tag{2.2.8}$$

Then a point $x\in\Omega$ *is singular if and only if* $x\in\mathcal{R}$*.*

A decreasing function that satisfies (2.2.8) is usually said to be *k-regular*. As indicated the associated constant λ is independent of t but may depend on k. Clearly the functions h and g given by (2.2.4) satisfy conditions (2.2.7) and (2.2.8) and thus the theorem implies the one dimensional classical singular statement; i.e. $Sing(1)=\mathbb{Q}$.

PROOF OF THEOREM 2.2.1. Trivially, if $x=R_\alpha$ for some $\alpha\in J$, then x is singular. To prove the opposite implication, assume x is singular. Then by definition, for any $\epsilon>0$ and $t>t_0$ there exists $\alpha\in J$ so that

$$d(x,R_\alpha) < \epsilon g(\beta_\alpha)h(k^t) \qquad \text{and} \qquad \beta_\alpha < k^t, \tag{2.2.9}$$

and $\alpha'\in J$ so that

$$d(x,R_{\alpha'}) < \epsilon g(\beta_{\alpha'})h(k^{t+1}) \qquad \text{and} \qquad \beta_{\alpha'} < k^{t+1}. \tag{2.2.10}$$

If $R_\alpha = R_{\alpha'}$ for all $t>t_0$ then since the right hand side of (2.2.9) tends to zero as $t\to\infty$, we have that $x=R_\alpha$. In other words, $x\in\mathcal{R}$ and we are done. Thus, assume that $R_\alpha\neq R_{\alpha'}$. Furthermore, without loss of generality, assume that t is large enough so that (2.2.6), (2.2.7) and (2.2.8) are all valid and that

$$\beta_\alpha, \beta_{\alpha'} > k^{t_0}$$

in (2.2.9) and (2.2.10). Note that if the latter was not the case, then since the right hand side of (2.2.9) tends to zero as $t\to\infty$ we have that $x\in\mathcal{R}$ and again we are done.

It now follows by (2.2.9), (2.2.10) and the triangle inequality that

$$\begin{aligned} d(R_\alpha, R_{\alpha'}) &\leq d(x, R_\alpha) + d(x, R_{\alpha'}), \\ &< \epsilon g(\beta_\alpha) h(k^t) + \epsilon g(\beta_{\alpha'}) h(k^{t+1}) \\ &< \epsilon g(\beta_\alpha) h(\beta_\alpha) + \epsilon g(\beta_{\alpha'}) h(\beta_{\alpha'}). \end{aligned} \tag{2.2.11}$$

The last inequality makes use of the fact that h is decreasing. Assume for the moment that

$$\beta_\alpha \leq \beta_{\alpha'}. \tag{2.2.12}$$

Then, in view of (2.2.7) it follows that

$$g(\beta_\alpha) \leq g(\beta_{\alpha'}) \frac{h(\beta_\alpha)}{h(\beta_{\alpha'})} \tag{2.2.13}$$

which together with (2.2.11) implies that

$$d(R_\alpha, R_{\alpha'}) < \epsilon \left(g(\beta_{\alpha'}) \frac{h^2(\beta_\alpha)}{h(\beta_{\alpha'})} + f(\beta_{\alpha'}) \right). \tag{2.2.14}$$

Now, on using (2.2.8) we find that

$$h(\beta_\alpha) < h(k^{t^*}) < \lambda^{t+1-t^*} h(k^{t+1}) < c h(\beta_{\alpha'})$$

where $t^* \geq t_0$ is the largest integer such that $k^{t^*} < \beta_\alpha$ and $c := \lambda^{t+1+t^*}$. This together with (2.2.14) implies that

$$d(R_\alpha, R_{\alpha'}) < \epsilon f(\beta_{\alpha'})(c^2 + 1). \tag{2.2.15}$$

On the other hand, in view of (2.2.6) and (2.2.12), we have that

$$d(R_\alpha, R_{\alpha'}) > C_2 \min\{f(\beta_\alpha), f(\beta_{\alpha'})\} = C_2 f(\beta_{\alpha'}) \ . \tag{2.2.16}$$

Thus, for any $\epsilon > 0$, on combining (2.2.15) and (2.2.16) we have that

$$C_2 f(\beta_{\alpha'}) < d(R_\alpha, R_{\alpha'}) < \epsilon f(\beta_{\alpha'})(c^2 + 1) \ .$$

The upshot is that we obtain a contradiction by setting

$$\epsilon < \frac{C_2}{c^2 + 1}.$$

This completes the proof of Theorem 2.2.1 in the case (2.2.12). It is easily verified, that an analogous argument works when $\beta_\alpha > \beta_{\alpha'}$. □

2.2.2. Dirichlet systems and Extremal Points. Let $(\mathcal{R}, \beta)$ be a Dirichlet system relative to (g, h, k). Then, as already mentioned, with f given by (2.2.5) we have that for any $x \in \Omega \setminus \mathcal{R}$, there exists infinitely many $\alpha \in J$ such that

$$d(x, R_\alpha) \leq C_1 f(\beta_\alpha) \, . \tag{2.2.17}$$

In other words,

$$\Lambda(\psi) = \Omega \setminus \mathcal{R} \qquad \text{with} \qquad \psi(r) = C_1 f(r)\,. \tag{2.2.18}$$

Clearly, with respect to the basic example with g and h given by (2.2.4) so that $f(r) = r^{-2}$, the above statement naturally coincides with the classical statement given by Theorem 2.1.2 with $m = 1$ and $x \in \mathbb{R} \setminus \mathbb{Q}$.

Now let K be a subset of the Ω which supports a nonatomic probability measure μ. Then motivated by the classical extremal theory described in §2.1.2, K will play the role of the manifold and μ the role of the Lebesgue measure on the manifold. Indeed, in view of the statement associated with (2.2.18) it is natural to introduce the following notion of extremality within the framework of Dirichlet system. In keeping with the classical theory, in the case $\psi : r \to C_1 f(r)^\tau$ for some $\tau > 0$, we write $\Lambda(\tau)$ for $\Lambda(\psi)$. Note that in view of (2.2.18) and the fact that $\mathcal{R}$ is countable, we trivially have that

$$\mu(K \cap \Lambda(\tau)) = 1 \qquad \text{if} \qquad \tau \leq 1\,.$$

Definition 2.2.2. Let (Ω, d) be a compact metric space and suppose that $(\mathcal{R}, \beta)$ is a Dirichlet system relative to (g, h, k). Let K be a subset of the Ω equipped with a nonatomic probability measure μ. Then K is said to be *μ-extremal* if

$$\mu(K \cap \Lambda(\tau)) = 0 \qquad \forall\ \tau > 1\,.$$

To have any hope of developing a general extremal theory for the subsets $K \subseteq \Omega$ we impose the following "decaying" condition on the measure μ. Given $\alpha > 0$, the measure μ supported on K is said to be *weakly absolutely α-decaying* if there exist strictly positive constants C, r_0 such that for all $\epsilon > 0$ we have

$$\mu(B(x, \epsilon r)) \ \leq\ C\,\epsilon^\alpha \mu(B(x, r)) \qquad \forall\, x \in K \qquad \forall\, r < r_0\,.$$

For sets supporting such measures, we are able to prove the following "extremality" result.

Theorem 2.2.2. *Let (Ω, d) be a compact metric space. Suppose that $(\mathcal{R}, \beta)$ is a well separated Dirichlet system relative to (g, h, k) and that K is a subset of Ω equipped with a weakly absolutely α-decaying measure μ. Furthermore, suppose there exist constants $\lambda = \lambda(k) > 1$ and $t_0 > 0$ such that*

$$f(k^t) < \lambda\, f(k^{t+1}) \qquad \forall \quad t > t_0\,. \tag{2.2.19}$$

Then

$$\mu(K \cap \Lambda(\psi)) = 0 \quad \text{if} \quad \sum_{t=1}^{\infty} \Big(\frac{\psi(k^t)}{f(k^t)}\Big)^{\alpha} \; < \; \infty \, . \tag{2.2.20}$$

REMARK 2.2.1. Note that since f is k-regular (i.e. satisfies (2.2.19)) and that ψ is monotonic, the convergence/divergence property of the sum appearing in (2.2.20) is equivalent to that of the sum

$$\sum_{r=1}^{\infty} \frac{1}{r} \Big(\frac{\psi(r)}{f(r)}\Big)^{\alpha} .$$

The following statement is a trivial consequence of Theorem 2.2.2 and the definition of μ-extremal.

Corollary 2.2.3. *Let (Ω, d) be a compact metric space. Suppose that $(\mathcal{R}, \beta)$ is a well separated Dirichlet system relative to (g, h, k) and that K is a subset of Ω equipped with a weakly absolutely α-decaying measure μ. Furthermore, suppose that f satisfies* (2.2.19) *and that*

$$\sum_{t=1}^{\infty} f(k^t)^{\alpha(\tau-1)} \; < \; \infty \qquad \forall \, \tau > 1 \, .$$

Then K is μ-extremal.

PROOF OF THEOREM 2.2.2. For $t \in \mathbb{N}$, let

$$\Delta^+(\psi, t) := \bigcup_{\alpha \in J \,:\, k^t < \beta_\alpha \le k^{t+1}} B(R_\alpha, \psi(k^t)). \tag{2.2.21}$$

Then by definition (see (2.2.2)) and the fact that ψ is decreasing, it follows that

$$\Delta(\psi, t) \subseteq \Delta^+(\psi, t)$$

and thus

$$\Lambda(\psi) \subseteq \Lambda^+(\psi) := \limsup_{t \to \infty} \Delta^+(\psi, t) \; . \tag{2.2.22}$$

The upshot of this is that

$$\mu(K \cap \Lambda^+(\psi)) = 0 \quad \Longrightarrow \quad \mu(K \cap \Lambda(\psi)) = 0 \, .$$

and by the Borel-Cantelli Lemma

$$\mu(K \cap \Lambda^+(\psi)) = 0 \qquad \text{if} \qquad \sum_{n=1}^{\infty} \mu(\Delta^+(\psi, t)) < \infty \, . \tag{2.2.23}$$

Thus, to complete the proof of the theorem we need to show that the above sum converges. With this in mind, we start making two simple observations.

Firstly, for any $\epsilon > 0$, we can assume that

$$\psi(k^t) < \epsilon f(k^t) \quad \text{for all } t \text{ sufficiently large.} \tag{2.2.24}$$

If this was not the case then $\psi(k^t)/f(k^t) > \epsilon$ for infinitely many t and thus violate the convergent sum condition associated with (2.2.20).

Secondly, the separation Condition S together with (2.2.19) implies that for t sufficiently large: for any $\alpha, \alpha' \in J$ with $R_\alpha \neq R_{\alpha'}$ and $\beta_\alpha, \beta_{\alpha'} < k^{t+1}$ we have that

$$d(R_\alpha, R_{\alpha'}) > C_2 f(k^{t+1}) > \frac{C_2}{\lambda} f(k^t) \ .$$

In turn this implies that

$$B\big(R_\alpha, C_3 f(k^t)\big) \cap B\big(R_{\alpha'}, C_3 f(k^t)\big) = \emptyset \tag{2.2.25}$$

where

$$C_3 := \frac{C_2}{\lambda} \ .$$

In view of (2.2.24), for t sufficiently large we can assume that

$$\psi(k^t) < \frac{C_3}{4} \, f(k^t)$$

and together with (2.2.25) this implies that the union of balls associated with $\Delta^+(\psi, t)$ is a disjoint union. Thus, for t sufficiently large

$$\mu(\Delta^+(\psi, t)) := \sum_{\alpha \in J \ : \ k^t < \beta_\alpha \leq k^{t+1}} \mu(B(R_\alpha, \psi(k^t))). \tag{2.2.26}$$

The measure μ is supported on K and so the only balls that can potentially make a positive contribution to the above sum are those that intersect K. With this in mind, take such a ball $B\big(R_\alpha, \psi(k^t)\big)$ and choose a point

$$\widehat{R_\alpha} \in K \cap B\big(R_\alpha, \psi(k^t)\big) \ .$$

It is easily verified that

$$B\Big(R_\alpha, \psi(k^t))\Big) \subseteq B\Big(\widehat{R_\alpha}, 2\psi(k^t)\Big) \subseteq B\Big(\widehat{R_\alpha}, \tfrac{C_3}{2} f(k^t)\Big) \subseteq B\Big(R_\alpha, C_3 f(k^t)\Big) \ .$$

Since μ is weakly absolutely α-decaying, it follows that for t sufficiently large

$$\begin{aligned}
\mu(\Delta^+(\psi,t)) &\leq \sum_{\substack{\alpha\in J:\\ k^t<\beta_\alpha\leq k^{t+1}}} \mu\Big(B\big(\widehat{R_\alpha},\ 2\,\psi(k^t)\big)\Big) \\
&= \sum_{\substack{\alpha\in J:\\ k^t<\beta_\alpha\leq k^{t+1}}} \mu\Big(B\Big(\widetilde{R_\alpha},\ 2\,\psi(k^t)\,\frac{2C_3 f(k^t)}{2C_3 f(k^t)}\Big)\Big) \\
&\leq \sum_{\substack{\alpha\in J:\\ k^t<\beta_\alpha\leq k^{t+1}}} C\Big(\frac{4\psi(k^t)}{C_3 f(k^t)}\Big)^\alpha \mu\Big(B\Big(\widehat{R_\alpha},\ \tfrac{C_3}{2} f(k^t)\Big)\Big) \\
&\leq C\Big(\frac{4\psi(k^t)}{C_3 f(k^t)}\Big)^\alpha \sum_{\substack{\alpha\in J:\\ k^t<\beta_\alpha\leq k^{t+1}}} \mu\Big(B\big(R_\alpha, C_3 f(k^t)\big)\Big). \qquad (2.2.27)
\end{aligned}$$

The measure μ is a probability measure and by (2.2.25) the balls associated with the above sum are disjoint. Hence

$$\sum_{\substack{\alpha\in J:\\ k^t<\beta_\alpha\leq k^{t+1}}} \mu\Big(B\big(R_\alpha, C_3 f(k^t)\big)\Big) \ \leq\ 1$$

which together with (2.2.27) and the convergent sum hypothesis implies that

$$\sum_{t=1}^{\infty} \mu(\Delta^+(\psi,t)) \ \ll\ \sum_{t=1}^{\infty} \Big(\frac{\psi(k^t)}{f(k^t)}\Big)^\alpha \ <\ \infty\ .$$

This establishes (2.2.23) and thereby completes the proof of Theorem 2.2.2.

□

REMARK 2.2.2. Given an increasing function $\theta:\mathbb{R}\to\mathbb{R}$, the measure μ is called *θ-decaying* if there exists strictly positive constants C, r_0 such that for all $\epsilon > 0$

$$\mu(B(x,\epsilon r)) \leq C\theta(\epsilon)\mu(B(x,r)) \quad \forall x\in K \quad \forall r<r_0\ .$$

We can generalise Theorem 2.2.2 further to subsets $K\subseteq\Omega$ supporting a θ-decaying measure μ.

2.3. Application to Kleinian groups

As an application of the Dirichlet systems framework introduced in the §2.2 we consider the theory of Diophantine approximation on limit sets of Kleinian groups. In particular, we obtain the singular and extremal results obtained in [**9**] as a simple consequence of the general theorems associated with well separated Dirichlet systems. Naturally, we start by describing the Kleinian group setup and the main results from [**9**].

The classical results of Diophantine approximation, in particular those from the one-dimensional theory, have natural counterparts and extensions in the hyperbolic space setting. In this setting, instead of approximating real numbers by rationals, one approximates the limit points of a fixed Kleinian group G by points in the orbit (under the group) of a distinguished limit point y. Beardon & Maskit [**4**] have shown that the geometry of the group is reflected in the approximation properties of points in the limit set.

Throughout, G is a nonelementary, geometrically finite Kleinian group acting on the unit ball model (B^{m+1}, ρ) of $(m+1)$–dimensional hyperbolic space with metric ρ derived from the differential $d\rho = |d\mathbf{x}|/(1-|\mathbf{x}|^2)$. Thus, G is a discrete subgroup of $\text{Möb}(B^{m+1})$, the group of orientation-preserving Möbius transformations of the unit ball B^{m+1}. By assumption, there is some finite-sided convex fundamental polyhedron for the action of G on B^{m+1}. Since G is nonelementary, the limit set $L(G)$ of G (the set of limit points in the unit sphere S^m of any orbit of G in B^{m+1}) is uncountable. The group G is said to be *of the first kind*[1] if $L(G) = S^m$ and *of the second kind* otherwise. Let δ denote the Hausdorff dimension of $L(G)$. Trivially, if G is of the first kind then we have $\delta := \dim L(G) = m$. In general, it is well known that δ is equal to the exponent of convergence of the group [**74**, **92**].

For each element $g \in G$ we shall use the notation $L_g := |g'(0)|^{-1}$, where $|g'(0)| = 1 - |g(0)|^2$ is the (Euclidean) conformal dilation of g at the origin. Note that by definition, $L_g = e^{\rho(0,g(0))}$. With this setup and notation in mind, we are in the position to state two fundamental results originating from Patterson's pioneering paper [**73**]. In short, they represent natural generalisations to the hyperbolic space setting of the classical theorems of Dirichlet and Khintchine in the theory of Diophantine approximation.

2.3.1. A Dirichlet-type statement and singular sets. The following two Dirichlet-type theorems were first established by Patterson [**73**, Section 7: Theorems 1 & 2] for finitely generated Fuchsian groups, i.e. Kleinian groups acting on the unit disc model of 2–dimensional hyperbolic space. Recall that in this $m = 1$ case, the class of finitely generated groups coincides with the class of geometrically finite groups.

[1] A geometrically finite group of the first kind is also called a *lattice*.

THEOREM 2.3.1. *Let G be a nonelementary, geometrically finite Kleinian group containing parabolic elements and let P be a complete set of inequivalent parabolic fixed points of G. Then there is a constant $c > 0$ with the following property: for each $\xi \in L(G)$, $N > 1$, there exist $p \in P$, $g \in G$ so that*

$$|\xi - g(p)| \leq \frac{c}{\sqrt{L_g N}} \qquad \textit{and} \qquad L_g \leq N\,.$$

As pointed out in [**94**], the $m = 1$ proof of Patterson can be easily generalised to higher dimensions when the ranks[2] of the parabolic fixed points are all maximal; i.e. when $\operatorname{rank}(p) = m$ for all $p \in P$. Without this rank assumption, the theorem is proved in [**90**, Theorem 1]. We now consider the case where the geometrically finite group G has no parabolic elements; i.e. where G is convex cocompact.

THEOREM 2.3.2. *Let G be a nonelementary, geometrically finite Kleinian group without parabolic elements and let $\{\eta, \eta'\}$ be the pair of fixed points of a hyperbolic element of G. Then there is a constant $c > 0$ with the following property: for all $\xi \in L(G)$, $N > 1$, there exist $y \in \{\eta, \eta'\}$, $g \in G$ so that*

$$|\xi - g(y)| \leq \frac{c}{N} \qquad \textit{and} \qquad L_g \leq N\,.$$

Patterson's $m = 1$ proof of the above theorem easily generalises to higher dimensions.

When interpreted on the upper half-plane $\mathbb{H}^2$ and applied to the modular group $SL(2, \mathbb{Z})$, it is easily verified that Theorem 2.3.1 reduces to the $m = 1$ case of Dirichlet's Theorem.

Motivated by the classical theory, the above Dirichlet-type theorems naturally lead to the notion of singular limit points. Let G be a Kleinian group and let Y be a complete set P of inequivalent parabolic fixed points of G if the group has parabolic elements; otherwise let Y be the pair $\{\eta, \eta'\}$ of fixed points of a hyperbolic element of G. A point $\xi \in L(G)$ is said to be *singular* if for every $\epsilon > 0$ there exists N_0 with the following property: for each $N \geq N_0$, there exist $y \in Y$, $g \in G$ so that

$$|\xi - g(y)| \ < \begin{cases} \frac{\epsilon}{\sqrt{L_g N}} & \text{if} \quad Y = P \\ \\ \frac{\epsilon}{N} & \text{if} \quad Y = \{\eta, \eta'\} \end{cases} \qquad \text{and} \qquad L_g < N\,. \quad (2.3.1)$$

[2]The stabiliser $G_p = \{g \in G : g(p) = p\}$ of a parabolic fixed point p is an infinite group which contains a free abelian subgroup of finite index. The *rank* of p is defined to be the number $k \in [1, m]$ such that this subgroup is isomorphic to $\mathbb{Z}^k$.

The following statement is established in [**9**, Theorem 1].

THEOREM 2.3.3. *Let G be a nonelementary, geometrically finite Kleinian group, and let Y be as above. Then a point $\xi \in L(G)$ is singular if and only if $\xi \in G(Y) := \{g(y) : g \in G, y \in Y\}$.*

REMARK 2.3.1. In the case where G is convex cocompact, the set of singular limit points is dependent on the choice of Y; i.e. on the chosen pair $\{\eta, \eta'\}$ of hyperbolic fixed points of G. If G has parabolic elements, the set of singular limit points is precisely the set of parabolic fixed points of G. Dynamically, the set corresponds to geodesics on the associated hyperbolic manifold $\mathcal{H} = B^{m+1}/G$ that travel straight into the cuspidal end.

REMARK 2.3.2. The parabolic fixed points of the modular group are the rationals together with the point at infinity. Thus, Theorem 2.3.3 when interpreted on $\mathbb{H}$ and applied to $SL(2, \mathbb{Z})$ precisely coincides with the $m = 1$ classical results; namely that $\mathrm{Sing}(1) = \mathbb{Q}$.

2.3.2. A Khintchine-type theorem and extremal sets. Let G be a nonelementary, geometrically finite Kleinian group G and let y be a parabolic fixed point of G if the group has parabolic elements and a hyperbolic fixed point otherwise. The Dirichlet-type theorems of §2.3.1 together with natural "decoupling" results (see for example [**90**, Proposition 2.3] and [**94**, Proposition 2]) imply the following statement for any nonelementary, geometrically finite Kleinian group G: *for each point $\xi \in L(G)$ which is not a parabolic fixed point there exist infinitely many $g \in G$ such that*

$$|\xi - g(y)| < \frac{c}{L_g}\,. \tag{2.3.2}$$

Here, c is a positive group constant. It is easy to see that if G has only one equivalence class of parabolic fixed points then we can take ξ to be any limit point. In any case, the statement describes to what extent any (non-parabolic) limit point ξ may be approximated by the orbit of the distinguished point y; namely that every non-parabolic limit point can be approximated by orbit points $g(y)$ with "rate" of approximation given by c/L_g – the right-hand side of inequality (2.3.2) determines the "rate" or "error" of approximation. It is natural to broaden the discussion to include general approximating functions. More precisely, let $\psi : \mathbb{R}^+ \to \mathbb{R}^+$ be a decreasing function and let

$$W_y(\psi) := \{\xi \in L(G) : |\xi - g(y)| < \psi(L_g) \quad \text{for i.m. } g \in G\}.$$

This is the set of points in the limit set which are "close" to infinitely many ("i.m.") images of the "distinguished" point y. As above, y is taken to be a parabolic fixed point of G if the group has parabolic elements and a hyperbolic fixed point of G otherwise. A natural problem is to determine the "size" of the set $W_y(\psi)$ in terms of the Patterson measure ν – a nonatomic, δ-conformal probability measure supported on $L(G)$. For groups of the first kind, since $\delta := \dim L(G) = m$, the Patterson measure is simply normalised m-dimensional Lebesgue measure on the unit sphere S^m. The following Khintchine-type theorem was first established by Patterson [**73**, Section 9] for finitely generated Fuchsian groups of the first kind. For convenience, let

$$w(y) := \begin{cases} 2\delta - \text{rank}(y) & \text{if} \quad y \text{ is parabolic,} \\ \\ \delta & \text{if} \quad y \text{ is hyperbolic.} \end{cases}$$

THEOREM 2.3.4. *Let G be a nonelementary, geometrically finite Kleinian group and let y be a parabolic fixed point of G, if there are any, and a hyperbolic fixed point otherwise. Then*

$$\nu(W_y(\psi)) = \begin{cases} 0 & \text{if} \quad \sum_{r=1}^{\infty} \ \psi\,(r)^{w(y)} \quad r^{w(y)-1} < \infty \ , \\ \\ 1 & \text{if} \quad \sum_{r=1}^{\infty} \ \psi\,(r)^{w(y)} \quad r^{w(y)-1} = \infty \ . \end{cases}$$

REMARK 2.3.3. In terms of this note, there are two special cases of the above theorem that are of particular interest to us.

(i) For $\epsilon \geq 0$, let $\psi_\epsilon : r \mapsto r^{-1}(\log r)^{-\frac{1+\epsilon}{w(y)}}$. Then it follows that

$$\nu(W_y(\psi_\epsilon)) = \begin{cases} 0 & \text{if} \quad \epsilon > 0 \ , \\ 1 & \text{if} \quad \epsilon = 0 \ . \end{cases}$$

This statement has a well-known dynamical interpretation in terms of the "rate" of excursions by geodesics into a cuspidal end of the associated hyperbolic manifold $\mathcal{H} = B^{m+1}/G$; namely Sullivan's logarithm law for geodesics [**16, 90, 91**].

(ii) For $\tau \geq 1$, consider the function $\psi : r \mapsto r^{-\tau}$ and write $W_y(\tau)$ for $W_y(\psi)$. Then it follows that

$$\nu(W_y(\tau)) = 0 \qquad \text{if} \quad \tau > 1 \ .$$

The fact that $\nu(W_y(\tau)) = 1$ for $\tau = 1$ can be easily deduced from the statement associated with inequality (2.3.2) and the fact the

$\nu(W_y(c\psi)) = \nu(W_y(\psi))$ for any constant $c > 0$ [**90**, Lemma 4.6] - we do not need the full power of the divergence case of Theorem 2.3.4.

Without assuming that G is of the first kind, Theorem 2.3.4 is essentially established in [**89**] if y is a hyperbolic fixed point of G and in [**90**] if y is a parabolic fixed point of G. We say essentially, since in both [**89**] and [**90**] an extra regularity condition on the approximating function ψ is assumed. The theorem as stated above, without any regularity condition on ψ beyond monotonicity, is established in [**8**, Section 10.3: Theorems 5 & 9] and is the perfect Kleinian group analogue of Khintchine's Theorem in the classical theory of metric Diophantine approximation. Indeed, when interpreted on the upper half-plane $\mathbb{H}^2$ and applied to the modular group $SL(2, \mathbb{Z})$, it is easily verified that Theorem 2.3.4 reduces to the $m = 1$ case of Khintchine's Theorem.

In view of the recent progress within the classical manifold theory, it would be highly desirable to obtain an analogous theory within the hyperbolic space setup. With this in mind as the ultimate goal, let K be a subset of the limit set $L(G)$ which supports a nonatomic probability measure μ. In view of (2.3.2) and the comments made in Remark 2.3.3(ii), it is natural to say that a subset $K \subseteq L(G)$ is μ-*extremal* if

$$\mu(K \cap W_y(\tau)) = 0 \qquad \forall \ \tau > 1 \ .$$

Note that $L(G)$ is ν-extremal where ν is the Patterson measure — see Remark 2.1.1. To have any hope of developing a general extremal theory for the subsets K we impose the condition that the measure μ supported on K is weakly absolutely α-decaying. For sets supporting such measures, the following statement is established in [**9**, Theorem 2].

THEOREM 2.3.5. *Let G be a nonelementary, geometrically finite Kleinian group and let y be a parabolic fixed point of G, if there are any, and a hyperbolic fixed point otherwise. Fix $\alpha > 0$, and let K be a compact subset of $L(G)$ equipped with a weakly absolutely α-decaying measure μ. Then*

$$\mu(K \cap W_y(\psi)) = 0 \quad \text{if} \quad \sum_{r=1}^{\infty} r^{\alpha-1}\psi(r)^{\alpha} \ < \ \infty \ . \tag{2.3.3}$$

REMARK 2.3.4. It is easily verified that if a measure μ is absolutely α-decaying as defined in [**75**] then it is weakly absolutely α-decaying. Also it is worth pointing out that although the Lebesgue measure $| \ . \ |_{\mathcal{M}}$ on a nondegenerate manifold $\mathcal{M}$ is not necessarily absolutely α-decaying, it is weakly absolutely α-decaying.

Observe that if we write $\psi_\tau(r) = r^{-\tau}$, then

$$\sum_{r=1}^{\infty} r^{\alpha-1}\psi_\tau(r)^\alpha = \sum_{r=1}^{\infty} r^{\alpha(1-\tau)-1} \ < \ \infty \qquad \forall\, \tau > 1 \ \ \forall\, \alpha > 0.$$

Hence the following statement is a trivial consequence of Theorem 2.3.5.

Corollary 2.3.6. *Let G be a nonelementary, geometrically finite Kleinian group and let y be a parabolic fixed point of G, if there are any, and a hyperbolic fixed point otherwise. Let K be a compact subset of $L(G)$ equipped with a weakly absolutely decaying measure μ. Then K is μ-extremal.*

2.3.3. Dirichlet systems: Proof of Theorems 2.3.3 & 2.3.5. Let G be a nonelementary, geometrically finite Kleinian group and let Y be a complete set P of inequivalent parabolic fixed points of G if the group has parabolic elements; otherwise let Y be the pair $\{\eta, \eta'\}$ of fixed points of a hyperbolic element of G. We show that the Dirichlet systems framework of §2.2 naturally incorporates the singular and extremal theory associated with Kleinian groups as described in §2.3.1 and §2.3.2 . With this in mind, given $y \in Y$ let

$$\Omega := L(G)\ , \quad J := \{g : g \in G\}\ , \quad \alpha := g \in J\ , \quad \beta_\alpha := L_g\ ,$$

$$R_\alpha := g(y) \quad \text{and} \quad \Delta(R_\alpha, \psi(\beta_\alpha)) := B(g(y), \psi(L_g))\ .$$

Thus, the family $\mathcal{R}$ of resonant sets R_α consists of orbit points $g(y)$ with $g \in G$. Furthermore,

$$\Delta(\psi, t) := \bigcup_{\substack{g \in G\,: \\ k^t < L_g \leq k^{t+1}}} B\left(g(y), \psi(L_g)\right)$$

and

$$W_y(\psi) = \Lambda(\psi) := \limsup_{t\to\infty} \Delta(\psi, t)\ .$$

The metric d is of course the standard Euclidean metric in $\mathbb{R}^m$ and the non-atomic, probability measure ν supported on Ω is the Patterson measure supported on $L(G)$.

In view of Theorems 2.3.1 & 2.3.2, it follows that $(\mathcal{R},\beta)$ is a Dirichlet system relative to (g,h,k) for any $k>1$ and with

$$g(r)=h(r)=r^{-\frac{1}{2}} \qquad \text{if} \qquad Y=P \tag{2.3.4}$$

$$g(r)=1 \;\; \text{and} \;\; h(r)=r^{-1} \qquad \text{if} \qquad Y=\{\eta,\eta'\}\,. \tag{2.3.5}$$

Now observe that in both cases (i.e. when $Y=P$ and $Y=\{\eta,\eta'\}$) the above functions g and f satisfy conditions (2.2.7) and (2.2.8) associated with Theorem 2.2.1 and the corresponding function $f:r\to f(r):=g(r)h(r)$ satisfies condition (2.2.19) associated with Theorem 2.2.2. Moreover, in both cases, condition (2.2.6) associated with Condition S is satisfied courtesy of the following two well know statements [**9**, Lemmas 1 & 2].

Lemma 2.3.7. *Let $Y=P$. Then there is a constant $c_1>0$ depending only on G and P with the following property: for all $p,q\in P$ and $g,h\in G$ such that $g(p)\neq h(q)$, we have*

$$|g(p)-h(q)|>\frac{c_1}{\sqrt{L_gL_h}}\,. \tag{2.3.6}$$

Lemma 2.3.8. *Let $Y=\{\eta,\eta'\}$. Then there is a constant $c_1>0$ depending only on G and η,η' with the following property: for all $u,v\in Y$ and $g,h\in G$ such that $g(u)\neq h(v)$, we have*

$$|g(u)-h(v)|>c_1\,\min\{L_g^{-1},L_h^{-1}\}\,.$$

Trivially, concerning the right hand side of (2.3.6), we have that

$$\frac{c_1}{\sqrt{L_gL_h}}\;\geq\;c_1\,\min\{L_g^{-1},L_h^{-1}\}\;=\;c_1\,\min\{f(L_g),f(L_h)\}\,.$$

The upshot of the above discussion is that in both cases, $(\mathcal{R},\beta)$ is a well separated Dirichlet system relative to (g,h,k) for any $k>1$. Thus, Theorem 2.3.3 and Theorem 2.3.5 follow at once from Theorem 2.2.1 and Theorem 2.2.2 respectively.

CHAPTER 3

Effective approximation of simultaneously small linear forms

3.1. Introduction

Various publications utilise the theory of metric Diophantine approximation to develop new approaches in interference alignment, a concept in the field of wireless communication. The main tool is the fundamental Khintchine-Groshev Theorem and its variations. Adiceam, Beresnevich, Levesley, Velani and Zorin [**2**] presented a quantitative version of Khintchine-Groshev on sub-manifolds, for use in interference alignment. This statement and others are discussed in a recent survey by Nazer and Ordentlich [**71**] of results and applications from the conference in “Workshop on interactions between number theory and wireless communication”.

In this work we consider the size of linear forms under the Euclidean distance $|\cdot|$ to 0, so called small linear forms. In contrast to the distance to linear forms under the nearest integer $||\cdot||$ called well approximable linear forms as in classical Diophantine approximation. M. Hussain and J. Levesley’s paper [**54**] finds a Hausdorf measure variant of Khintchine-Groshev for simultaneously small linear forms. Here we find a quantitative variation on the corresponding convergence part of the probability measure statement. The result is based on methods in [**2**], however we use an explicit formula for the volume of slabs through hypercubes from [**68**] to prove the main result.

Firstly we present some notation and background results and describe a non-effective version of the main result. The following section is on the volume of slabs through hypercubes. Next is the proof of the non-effective version of the result, where various constants and terms are defined. We can then prove the effective statement, after which we give a Corollary for use in applications. We then apply the Corollary to an explicit example where the measure has density of the normal distribution.

3.1.1. Metric Number theory. Let $m, n \in \mathbb{N}$, then define $M_{m,n}(\mathbb{R})$ to be the set of $m \times n$ matrices with real entries. Let ψ be a real positive decreasing function such that $\psi(r) \to 0$ as $r \to \infty$. Such a function will be referred to as an *approximation* function. Given a subset S of $M_{m,n}(\mathbb{R})$, we will write $|S|_{mn}$ for its mn-dimensional Lebesgue measure. Given a function $\Psi : \mathbb{Z}^n \to \mathbb{R}^+$, a matrix $\mathbf{X} = (x_{i,j}) \in M_{m,n}(\mathbb{R})$ is Ψ-*approximable* if

$$||\mathbf{X}\mathbf{a}|| := \max_{1\leq i\leq m} ||x_{i,1}a_1 + x_{i,2}a_2 + \cdots + x_{i,n}a_n|| \leq \Psi(\mathbf{a})$$

has infinitely many (i.m.) solutions $\mathbf{a} = (a_i)_{i=1}^n \in \mathbb{Z}^n \setminus \{\mathbf{0}\}$, where $||x||$ is the distance to the nearest integer of x. In classical Diophantine approximation you study set of Ψ-approximable points $W(m, n; \Psi)$

$$W(m, n; \Psi) = \{\mathbf{X} \in M_{m,n}(\mathbb{R}) : ||\mathbf{X}\mathbf{a}|| \leq \Psi(\mathbf{a}) \text{ for i.m. } \mathbf{a} \in \mathbb{Z}^n \setminus \{\mathbf{0}\}\}.$$

One of the main results in classical Diophantine approximation $M_{m,n}(\mathbb{R})$ is the Khintchine-Groshev Theorem, which tells us about the Lebesgue measure of $W(m, n; \Psi)$ when we replace $\Psi(\mathbf{a})$ by an approximation function $\psi(|\mathbf{a}|)$.

THEOREM 3.1.1 (Khintchine-Groshev and others). *Let $m, n \in \mathbb{N}$ with $nm > 1$, $\psi : \mathbb{N} \to \mathbb{R}^+$ be an approximating function, let $\Psi : \mathbb{Z}^n \to \mathbb{R}^+$ be given by $\Psi(\mathbf{a}) := \psi(|\mathbf{a}|)$ for $\mathbf{a} \in \mathbb{Z}^n \setminus \{\mathbf{0}\}$, where $|\mathbf{a}| = \max_{1\leq i\leq n} |a_i|$. Then*

$$|W(m, n; \Psi)|_{mn} = \begin{cases} 0 & \text{if } \sum_{q=1}^{\infty} q^{n-1}\psi(q)^m < \infty, \\ \\ \text{Full} & \text{if } \sum_{q=1}^{\infty} q^{n-1}\psi(q)^m = \infty. \end{cases}$$

Theorem 3.1.1 was first obtained by Groshev under the assumption that $q^n\psi(q)^m$ is monotonic in the divergence case. The redundancy of the monotonicity condition for $n \geq 3$ follows from Schmidt's paper [**81**] and for $n = 1$ from Gallagher's paper [**45**]. The final case $n = 2$ of the statement was finally shown by Beresnevich and Velani [**14**].

We now shift our attention to the notion of badly approximable points. In particular, it does make sense to consider badly approximable points in linear forms. The set of *badly approximable linear forms in m variables* is defined by

$$\mathbf{Bad}(n, m) = \{\mathbf{x} \in \mathbb{I}^{mn} : \inf_{\mathbf{a}\in\mathbb{Z}^n\setminus\{\mathbf{0}\}} |\mathbf{a}|^{m/n}||\mathbf{X}\mathbf{a}|| > 0\}.$$

When $m = n = 1$ it is easily seen that $\mathbf{Bad}(m, n)$ reduces to $\mathbf{Bad}$. An immeditate consequence of Groshev' thoerem is that $\mathbf{Bad}(m, n)$ is of mn-dimensional Lebesgue measure zero.

An immediate consequence of Theorem 3.1.1 is that for almost every $\mathbf{X} \in M_{m,n}(\mathbb{R})$ there exists a constant $\kappa(\mathbf{X}) > 0$ dependent on $\mathbf{X}$ such that

$$||\mathbf{X}\mathbf{a}|| > \kappa(\mathbf{X})\psi(|\mathbf{a}|) \qquad \text{for all } \mathbf{a} \in \mathbb{Z}^n \setminus \{\mathbf{0}\}.$$

Estimating the value of $\kappa(\mathbf{X})$, independent of $\mathbf{X}$ is an important ingredient in studying the achievable number of degrees of freedom in schemes of Interference Alignment, for an example see [**70**].

The complement of $W(m, n; \Psi)$ is called the set of Ψ-badly approximable numbers $\mathbf{Bad}(m, n; \Psi)$. We define the more refined set $\mathbf{Bad}(m, n; \Psi, \kappa)$ of (Ψ, κ)-*badly approximable numbers* for $\kappa > 0$ by

$$\mathbf{Bad}(m, n; \Psi, \kappa) = \{\mathbf{X} \in M_{m,n}(\mathbb{R}) : ||\mathbf{X}\mathbf{a}|| > \kappa\Psi(\mathbf{a}) \text{ for all } \mathbf{a} \in \mathbb{Z}^n \setminus \{\mathbf{0}\}\}.$$

We are interested in finding the measure of $\mathbf{Bad}(m, n; \Psi, \kappa)$ for a large set of measures. In particular, we are considering measures μ on $M_{m,n}(\mathbb{R})$ which are linked to the Lebesgue measure. A measure μ on $M_{m,n}(\mathbb{R})$ is *absolutely continuous with respect to Lebesgue measure*, if there exists a Lebesgue integrable function $f : M_{m,n}(\mathbb{R}) \to \mathbb{R}^+$ such that for every measurable subset $A \subseteq M_{m,n}(\mathbb{R})$, one has that

$$\mu(A) = \int_A f,$$

where $\int_A f$ is the Lebesgue integral of f over A. The function f is referred to as the *density* of μ, as it describes where the measure is concentrated.

3.1.2. Simultaneously small linear forms. Here we are interested in Ψ-*simultaneously small linear forms*, those $X \in M_{m,n}(\mathbb{R})$ for which

$$|\mathbf{X}\mathbf{a}| := \max_{1\leq i\leq m} |x_{i,1}a_1 + x_{i,2}a_2 + \cdots + x_{i,n}a_n| \leq \Psi(\mathbf{a}) \tag{3.1.1}$$

has i.m. solutions $\mathbf{a} \in \mathbb{Z}^n \setminus \{\mathbf{0}\}$. The set of such matrices will be defined as:

$$W_0(m, n; \Psi) := \{\mathbf{X} \in M_{m,n}(\mathbb{R}) : |\mathbf{X}\mathbf{a}| \leq \Psi(\mathbf{a}) \text{ for i.m. } \mathbf{a} \in \mathbb{Z}^n \setminus \{\mathbf{0}\}\}.$$

We will be concerned with the measure of $W_0(m, n; \Psi)$.

We mention here the corresponding convergence case of the Khintchine-Groshev theorem for $W_0(m, n; \Psi)$ proved by Hussain and Levesley [**54**, Theorem 4].

THEOREM 3.1.2. *Let $m, n \in \mathbb{N}$, let $\psi : \mathbb{N} \to \mathbb{R}^+$ be an approximating function and let $\Psi : \mathbb{Z}^n \to \mathbb{R}^+$ be given by $\Psi(\mathbf{a}) := \psi(|\mathbf{a}|)$ for $\mathbf{a} \in \mathbb{Z}^n \setminus \{\mathbf{0}\}$. Then*

$$|W_0(m, n; \Psi)|_{mn} = 0 \qquad \textit{if} \qquad \sum_{q=1}^{\infty} q^{n-1}\psi(q)^m < \infty.$$

For integers $m, n \in \mathbb{N}$, an approximation function Ψ, probability measure μ absolutely continuous with respect to Lebesgue measure and $\delta \in (0,1)$, the authors of [**2**] find a formula for $\kappa = \kappa(\mu, m, \Psi)$ such that

$$\mu(\mathbf{Bad}(m, n; \Psi, \kappa)) \geq 1 - \delta.$$

We intend to find an analogue of their result over what we call the set $\mathbf{Bad}_0(m, n; \Psi, \kappa)$ of (Ψ, κ)-*badly simultaneously small linear forms* for $\kappa > 0$

$$\mathbf{Bad}_0(m, n; \Psi, \kappa) = \{\mathbf{X} \in M_{m,n}(\mathbb{R}) : |\mathbf{X}\mathbf{a}| > \kappa\Psi(\mathbf{a}) \text{ for all } \mathbf{a} \in \mathbb{Z}^n \setminus \{\mathbf{0}\}\}.$$

We are then interested in the relationship between κ and the measure of $\mathbf{Bad}_0(m, n; \Psi, \kappa)$.

3.1.3. A non-effective Khintchine-Groshev type theorem. Here we combine the effective statement in [**2**] with the convergence result in [**54**] for an effective Khintchine-Groshev type convergence theorem for simultaneously small linear forms. First, we state the non-effective version of the main result, the effective version will be stated in Section 3.3.

THEOREM 3.1.3. *Let $m, n \in \mathbb{N}$ and let μ be a probability measure on $M_{m,n}(\mathbb{R})$ that is absolutely continuous with respect to Lebesgue measure on $M_{m,n}(\mathbb{R})$. Let $\Psi : \mathbb{Z}^n \to \mathbb{R}^+$ be any function such that*

$$\Sigma_{\Psi^m} := \sum_{\mathbf{a} \in \mathbb{Z}^n \setminus \{\mathbf{0}\}} \Psi(\mathbf{a})^m < \infty. \tag{3.1.2}$$

Then for any $\delta \in (0,1)$ there is a constant $\kappa > 0$ depending only on μ, Ψ and δ such that

$$\mu(\mathbf{Bad}_0(m, n; \Psi, \kappa)) \geq 1 - \delta.$$

3.2. Supporting results

3.2.1. Slabs of hypercubes. To prove Theorem 3.1.3 we will utilise techniques from the work of Pólya [**76**] and others [**3, 68, 99**] who have found exact formulas for the intersection of hypercubes with slabs. A *slab* for $\mathbf{a} \in \mathbb{R}^n$ and $\epsilon > 0$ is defined as

$$S(\mathbf{a}, \epsilon) := \{\mathbf{x} \in \mathbb{R}^n : |\mathbf{a} \cdot \mathbf{x}| \leq \epsilon\}.$$

An *n-dimensional hypercube* $C_n(\boldsymbol{\alpha})$ with side length 1 and *corner* $\boldsymbol{\alpha} = (\alpha_i)_{i=1}^n \in \mathbb{R}^n$ is defined by

$$C_n(\boldsymbol{\alpha}) := \{\mathbf{x} \in \mathbb{R}^n : \alpha_i \leq x_i < \alpha_i + 1\}.$$

In Pólya's thesis [**76**] he discovered the following exact formula, for any $0 \leq \epsilon \leq \frac{1}{2}$, let $\boldsymbol{\alpha} = \left(-\frac{1}{2}\right)_{i=1}^{n}$ and $\mathbf{a} = (a_1, \dots, a_n) \in \mathbb{R}^n$ then

$$\mathrm{Vol}(S(\mathbf{a}, \epsilon) \cap C_n(\boldsymbol{\alpha})) = \frac{2}{\pi} \int_0^\infty \frac{\sin 2\epsilon x}{x} \prod_{i=1}^{n} \frac{\sin a_i x}{a_i x} dx.$$

Marichal and Mossinghoff [**68**] found exact formulas involving polynomials in terms of ϵ. To state their theorem we must first introduce some notation. Let V_n be the set of vertices of the hypercube of side length 2 centred at the origin, so $V_n = \{-1, 1\}^n$. For $\mathbf{s} = (s_1, \dots, s_n) \in V_n$, let $\iota_{\mathbf{s}} := \prod_{i=1}^{n} s_i$. Define $r_+^n := (\max\{r, 0\})^n$, for $r \in \mathbb{R}$. Marichal and Mossinghoff have shown

THEOREM 3.2.1. *Let* $\mathbf{a} \in \mathbb{R}^n$ *be a vector with all nonzero components. For any* $0 < \epsilon < 1/2$, *let* $\mathbf{v} = (a_1, \dots, a_n, 2\epsilon)$. *Then*

$$Vol(S(\mathbf{a}, \epsilon) \cap C_n((-1/2)_{i=1}^{n}) = \frac{1}{2^n n! \prod_{i=1}^{n} a_i} \sum_{\mathbf{s} \in V_{n+1}} \iota_s (\mathbf{v} \cdot \mathbf{s})_+^n. \qquad (3.2.1)$$

We use Theorem 3.2.1 to establish that the volume of the intersection of the unit cube with the slab $S((1)_{i=1}^{n}, \epsilon))$ through the largest diagonal has the following form.

Lemma 3.2.2. *Let* $n \in \mathbb{N}$. *There exists real numbers* p_j *for* $1 \leq j \leq n$, *with* $p_1 \neq 0$, *such that for all* $0 < \epsilon < 1/2$,

$$Vol(S((1)_{i=1}^{n}, \epsilon)) \cap C((-1/2)_{i=1}^{n}) = \begin{cases} \sum_{j=0}^{(n-1)/2} p_{2j+1} (2\epsilon)^{2j+1} & \textit{if } n \textit{ is odd}, \\ \sum_{j=0}^{n/2-1} p_{2j+1} (2\epsilon)^{2j+1} + p_n (2\epsilon)^n & \textit{if } n \textit{ is even}. \end{cases}$$

PROOF. Let $\mathbf{v} = (1, 1, \dots, 1, 2\epsilon) \in \mathbb{R}^{n+1}$. Define the set

$$K := \{k \geq 1 : k = \mathbf{v} \cdot \mathbf{s} \pm 2\epsilon\}.$$

Let $\mathbf{s}_\pm(k) \in V_{n+1}$ be those elements of V_{n+1} such that

$$\mathbf{s}_\pm(k) := k \pm 2\epsilon.$$

Then $\iota_{s_+(k)} = (-1)\iota_{s_-(k)}$, and so for each $\mathbf{s}_\pm(k)$ addend in the sum (3.2.1) will have an opposing term. Except if $k = 0$ when n is even. Note that there are multiple $\mathbf{s}$ such that $\mathbf{v} \cdot \mathbf{s} = k \pm 2\epsilon$ for any $k \in K$.

Assume n is odd then

$$\begin{aligned}\sum_{\mathbf{s}\in V_{n+1}} \iota_{\mathbf{s}}(\mathbf{v}\cdot\mathbf{s})_+^n &= \sum_{k\in K} \iota_{s_+(k)}((k+2\epsilon)^n - (k-2\epsilon)^n) \\ &= \sum_{k\in K} \iota_{s_+(k)}(k^n + \binom{n}{1}k^{n-1}(2\epsilon) + \cdots + k^0(2\epsilon)^n \\ &\qquad - (k^n + \binom{n}{1}k^{n-1}(-2\epsilon) + \cdots + k^0(-2\epsilon)^n)) \\ &= \sum_{k\in K} \iota_{s_+(k)}(2\binom{n}{1}k^{n-1}(2\epsilon) + 2\binom{n}{3}k^{n-3}(2\epsilon)^3 + \dots \\ &\qquad + 2(2\epsilon)^n).\end{aligned}$$

So for each $j \in \mathbb{N}$ the ϵ^{2j} coefficients have cancelled out.

If n is even then there is an extra term when $k = 0$, but since $0 - \epsilon < 0$ only the positive term $(2\epsilon)^n$ exists. Therefore suitable p_j exist. □

3.2.2. Proof of Theorem 3.1.3. We are now ready to prove Theorem 3.1.3.

Proof of Theorem 3.1.3. Given $\mathbf{a} \in \mathbb{Z}^n \setminus \{\mathbf{0}\}$ and $\epsilon > 0$, define

$$\mathcal{S}(\mathbf{a}, \epsilon) := \{\mathbf{X} \in M_{m,n}(\mathbb{R}) : |\mathbf{X}\mathbf{a}| \le \epsilon\}.$$

Let a *mn-dimensional corner* $\boldsymbol{\alpha}$ in $M_{m,n}(\mathbb{R})$ be written as

$$\boldsymbol{\alpha} = ((\alpha_{i,j})_{i=1}^n)_{j=1}^m \in M_{m,n}(\mathbb{R}).$$

Define an *mn-dimensional hypercube* $C(\boldsymbol{\alpha}) \subseteq M_{m,n}(\mathbb{R})$ for such a corner by

$$C(\boldsymbol{\alpha}) = \left\{ (x_{i,j}) \in M_{m,n}(\mathbb{R}) : \alpha_{i,j} \le x_{i,j} < \alpha_{i,j} + 1 \text{ for } \begin{array}{lll} 1 & \le i & \le m, \\ 1 & \le j & \le n. \end{array} \right\}.$$

Fix a corner $\boldsymbol{\alpha}_0 \in M_{m,n}(\mathbb{R})$, then let $A = A(\boldsymbol{\alpha}_0) \ni \boldsymbol{\alpha}_0$ be the set of corners such that the union of the associated mn-dimensional hypercubes form a partition of $M_{m,n}(\mathbb{R})$. That is, for any $\boldsymbol{\alpha}, \boldsymbol{\alpha}' \in A$ either $C(\boldsymbol{\alpha}) \cap C(\boldsymbol{\alpha}') = \emptyset$ or they only meet on the boundary (i.e. the measure of intersection is 0) and

$$M_{m,n}(\mathbb{R}) = \bigcup_{\boldsymbol{\alpha}\in A} C(\boldsymbol{\alpha}). \tag{3.2.2}$$

Notice that

$$M_{m,n}(\mathbb{R}) \setminus \mathbf{Bad}_0(m, n; \Psi, \kappa) = \bigcup_{\mathbf{a}\in\mathbb{Z}^n\setminus\{\mathbf{0}\}} \mathcal{S}(\mathbf{a}, \kappa\Psi(\mathbf{a})). \tag{3.2.3}$$

The set $\mathcal{S}(\mathbf{a},\epsilon)\cap C(\boldsymbol{\alpha})$ is made up of n-dimensional slabs $S(\mathbf{a}_i,\epsilon)$ through the n-dimensional hypercubes $C_n(\boldsymbol{\alpha}_i)$ that make up the mn-dimensional hypercube $C(\boldsymbol{\alpha})$. Where, $\boldsymbol{\alpha}_i$ is the ith row of $\boldsymbol{\alpha}$ and $\mathbf{a}_i$ is the ith row of $\mathbf{a}$. That is

$$C(\boldsymbol{\alpha}) = C_n(\boldsymbol{\alpha}_1) \times C_n(\boldsymbol{\alpha}_2) \times \cdots \times C_n(\boldsymbol{\alpha}_m)$$

and

$$\mathcal{S}(\mathbf{a},\epsilon) \cap C(\boldsymbol{\alpha}) = \times_{i=1}^{m} S(\mathbf{a}_i,\epsilon) \cap C_n(\boldsymbol{\alpha}_i).$$

We need to find an upper bound for each $|\mathcal{S}(\mathbf{a},\epsilon) \cap C(\boldsymbol{\alpha})|_{mn}$ for all $\boldsymbol{\alpha} \in A$. Using Theorem 3.2.1 and noticing that for $\mathbf{a},\boldsymbol{\alpha} \in \mathbb{R}^n$ the volume of $S(\mathbf{a},\epsilon) \cap C_n(\boldsymbol{\alpha})$ is upper bounded when $\mathbf{a}$ runs through the diagonal of $C_n(\boldsymbol{\alpha})$ and all n-dimensional hypercubes have the same volume. Explicitly

$$\mathrm{Vol}(S(\mathbf{a},\epsilon) \cap C_n(\boldsymbol{\alpha})) \leq \mathrm{Vol}\left(S((1)_{i=1}^{n},\epsilon) \cap C_n\left((-1/2)_{i=1}^{n}\right)\right) =: \nu(\epsilon,n),$$

implies

$$|\mathcal{S}(\mathbf{a},\epsilon) \cap C(\boldsymbol{\alpha})|_{mn} \leq \left(\nu(\epsilon,n)\right)^m. \tag{3.2.4}$$

By Lemma 3.2.2 there exists $p_j \in \mathbb{R}$ such that we can write $\nu(\epsilon,n)$ as a polynomial

$$(\nu(\epsilon,n))^m = \sum_{j=m}^{nm} p_j (2\epsilon)^j.$$

Since $\sum_{\mathbf{a}\in\mathbb{Z}^n\setminus\{\mathbf{0}\}} \Psi(\mathbf{a})^m < \infty$ there must exist M_Ψ, such that

$$M_\Psi = \sup\{\Psi(\mathbf{a}) : \mathbf{a} \in \mathbb{Z}^n \setminus \{\mathbf{0}\}\} < \infty. \tag{3.2.5}$$

Choose κ sufficiently small that $2\kappa M_\Psi \leq 1$. We can apply (3.2.4) with $\epsilon = \kappa\Psi(\mathbf{a})$.

Using (3.2.3) and (3.2.4), then for each $\boldsymbol{\alpha} \in A$,

$$\begin{aligned}
|C(\boldsymbol{\alpha}) \setminus \mathbf{Bad}_0(m,n;\Psi,\kappa)|_{mn} &\leq \sum_{\mathbf{a}\in\mathbb{Z}^n\setminus\{\mathbf{0}\}} |\mathcal{W}(\mathbf{a},\kappa\Psi(\mathbf{a})) \cap C(\boldsymbol{\alpha})|_{mn}, \\
&\leq \sum_{\mathbf{a}\in\mathbb{Z}^n\setminus\{\mathbf{0}\}} \nu(\kappa\Psi(\mathbf{a}),n)^m, \\
&= \sum_{\mathbf{a}\in\mathbb{Z}^n\setminus\{\mathbf{0}\}} \sum_{j=m}^{nm} p_j (2\kappa\Psi(\mathbf{a}))^j, \\
&= \sum_{j=m}^{nm} p_j 2^j \Sigma_{\Psi^j} \kappa^j =: v(n,m;\Psi,\kappa). \qquad (3.2.6)
\end{aligned}$$

Where $v(n,m;\Psi,\kappa)$ is a polynomial in κ with coefficients $p_j 2^j \Sigma_{\Psi^j}$ and Σ_{Ψ^j} is from 3.1.2.

Since μ is a probability measure, it follows from $\bigcup_{\boldsymbol{\alpha}\in A} C(\boldsymbol{\alpha})$ being a cover of $M_{m,n}(\mathbb{R})$, that for any $\delta \in (0,1)$ there exists a finite subset $A^* \subseteq A$

such that

$$\mu\left(\bigcup_{\boldsymbol{\alpha}\in A^*} C(\boldsymbol{\alpha})\right) > 1 - \frac{\delta}{2}. \tag{3.2.7}$$

Since μ is absolutely continuous with respect to the Lebesgue measure, for every $\boldsymbol{\alpha} \in A^*$ and any $\epsilon_1 > 0$, there exists $\epsilon_2 > 0$ such that: For any measurable subset X of $C(\boldsymbol{\alpha})$,

$$\text{if } |X|_{mn} < \epsilon_2 \quad \text{then } \mu(X) < \epsilon_1.$$

Let N be the number of elements in A^*. Letting

$$X = C(\boldsymbol{\alpha}) \setminus \mathbf{Bad}_0(m, n; \Psi, \kappa)$$

and $\epsilon_1 = \delta/(2N)$ it follows that there exists some $\epsilon_2(\boldsymbol{\alpha}, \delta, N) > 0$ such that, if

$$v(n, m; \Psi, \kappa) < \epsilon_2(\boldsymbol{\alpha}, \delta, N), \tag{3.2.8}$$

then

$$\mu(C(\boldsymbol{\alpha}) \setminus \mathbf{Bad}_0(m, n; \Psi, \kappa)) < \delta/(2N).$$

Note that there is only one root of inequality (3.2.8) to consider, since the volume is increasing for $0 < M_\Psi \kappa < 1/2$. In particular, the inequality (3.2.8) holds if κ is less than the root κ_α of the polynomial $v(n, m; \Psi, \kappa)$ from (3.2.6),

$$v(n, m; \Psi, \kappa) - \epsilon_2(\boldsymbol{\alpha}, \delta, N) = 0,$$

for all $\boldsymbol{\alpha} \in A^*$.

Since A^* is finite, we can choose κ that satisfies

$$0 < \kappa \leq \min_{\boldsymbol{\alpha}\in A^*} \kappa_{\boldsymbol{\alpha}}.$$

Clearly for such a choice of κ, the inequality (3.2.8) holds for any $\boldsymbol{\alpha} \in A^*$. Hence, by additivity of μ,

$$\begin{aligned}
\mu(M_{m,n}(\mathbb{R}) \setminus \mathbf{Bad}_0(m, n, \Psi, \kappa)) &\leq \delta/2 + \sum_{\boldsymbol{\alpha}\in A^*} \mu(C(\boldsymbol{\alpha}) \setminus \mathbf{Bad}_0(m, n, \Psi, \kappa)), \\
&\leq \delta/2 + \sum_{\boldsymbol{\alpha}\in A^*} \delta/(2N), \\
&= \delta.
\end{aligned}$$

So

$$\begin{aligned}
\mu(\mathbf{Bad}_0(m, n; \Psi, \kappa)) &= 1 - \mu(M_{m,n}(\mathbb{R}) \setminus \mathbf{Bad}_0(m, n; \Psi, \kappa)) \\
&\geq 1 - \delta.
\end{aligned}$$

□

3.3. Effective version of the Khintchine-Groshev theorem on simultaneously small linear forms

We now turn our attention to quantifying the dependence of κ on δ. To this end we will make use of the L_p-norm of a Lebesgue measurable function.

DEFINITION 3.3.1. Let $f : M_{m,n}(\mathbb{R}) \to \mathbb{R}^+$ be a Lebesgue measureable function, let X be a measurable subset of $M_{m,n}(\mathbb{R})$ and let $p \geq 1$. We say $f \in L_p(X)$ if the Lebesgue integral

$$\int_X |f|^p := \int_{M_{m,n}(\mathbb{R})} |f|^p \chi_X$$

exists and is finite, where χ_X is the characteristic function of X. The L_p-*norm* of f on X is then

$$||f||_{p,X} := \left(\int_X |f|^p\right)^{1/p}.$$

In the case that $p = \infty$, the L_∞-norm on X is defined as

$$||f||_{\infty,X} := \inf\{c \in \mathbb{R} : |f(x)| \leq c \text{ for almost all } x \in X\}.$$

If $||f||_{\infty,X} < \infty$, then we say $f \in L_\infty(X)$.

We then have the effective version of Theorem 3.1.3.

THEOREM 3.3.1 (Effective Theorem). *Let $m, n \in \mathbb{N}, \mu$, and Ψ be as in Theorem 3.1.3, let M_Ψ be given by* (3.2.5) *and let f denote the density of μ. Let A^* be any finite subset of A from* (3.2.2) *that satisfies* (3.2.7)*. Assume there exists $p > 1$ such that $f \in L_p(C(\boldsymbol{\alpha}))$ for any $\boldsymbol{\alpha} \in A^*$ and define the sum Σ_f by*

$$\Sigma_f := \sum_{\alpha \in A^*} ||f||_{p,C(\boldsymbol{\alpha})}.$$

Let κ^* be the root of

$$v(n, m; \Psi, x) - \left(\frac{\delta}{2\Sigma_f}\right)^{\frac{p}{p-1}} = 0, \tag{3.3.1}$$

such that $0 < x < 1/2$ where v is from (3.2.6)*. Then for any $\delta \in (0, 1)$*

$$\mu(\mathbf{Bad}_0(m, n; \Psi, \kappa)) \geq 1 - \delta$$

with

$$\kappa := \frac{1}{2} \min\left\{\frac{1}{M_\Psi}, \kappa^*\right\}, \tag{3.3.2}$$

where $p/(p-1)$ is taken to be 1 when $p = \infty$.

REMARK 3.3.1. Note that the polynomial (3.3.1) only has one real positve root. To see this, notice it is of even degree and for an $\epsilon > 0$ it is of the form

$$p(x) - \epsilon = 0.$$

To prove Theorem 3.3.1 we need two Lemmas. The first one includes two well known facts about the L_p-norm.

Lemma 3.3.2. (1) *For any $p \geq 1$ and measurable subsets $X \subseteq Y$,*

$$||f||_{p,X} \leq ||f||_{p,Y}.$$

(2) *Hölder's inequality: For any $1 \leq p, q \leq \infty$ satisfying $\frac{1}{p} + \frac{1}{q} = 1$*

$$\left| \int_X fg \right| \leq ||f||_{p,X} ||g||_{q,X}.$$

Lemma 3.3.2 implies the following bound on measures absolutely continuous to the Lebesgue measure, the details of which can be found in [**2**].

Lemma 3.3.3. *Let $p > 1$ and let μ be a probability measure on $M_{m,n}(\mathbb{R})$ with density f. Let X be a Lebesgue measurable subset of $M_{m,n}(\mathbb{R})$. If $f \in L_p(X)$, then*

$$\mu(X) \leq ||f||_{p,X} |X|_{mn}^{1-1/p}.$$

We are now in a position to prove the main Theorem 3.3.1.

Proof of Theorem 3.3.1. Let A^* be the same as in the proof of Theorem 3.1.3. Assume that there exists $p > 1$ such that for every $\boldsymbol{\alpha} \in A^*$, the density f of μ has finite L_p norm on $C(\boldsymbol{\alpha})$. Let κ be as specified by (3.3.2). Then, by Lemma 3.3.3,

$$\mu(C(\boldsymbol{\alpha}) \setminus \mathbf{Bad}_0(m, n; \Psi, \kappa)) \leq ||f||_{p,C(\boldsymbol{\alpha})} |C(\boldsymbol{\alpha}) \setminus \mathbf{Bad}_0(m, n, \Psi, \kappa)|_{mn}^{1-1/p}.$$

Using (3.2.6) we obtain:

$$\mu(C(\boldsymbol{\alpha}) \setminus \mathbf{Bad}_0(m, n; \Psi, \kappa)) \leq ||f||_{p,C(\boldsymbol{\alpha})} \left(v(n, m; \Psi, \kappa)\right)^{1-1/p}.$$

It follows that

$$\begin{aligned} &\mu(M_{m,n}(\mathbb{R}) \backslash \mathbf{Bad}_0(m, n, \Psi, \kappa)) \\ &\qquad \leq \delta/2 + \sum_{\boldsymbol{\alpha} \in A^*} \mu(C(\boldsymbol{\alpha}) \setminus \mathbf{Bad}_0(m, n; \Psi, \kappa)), \\ &\qquad \leq \delta/2 + (v(m, n; \Psi, \kappa))^{1-1/p} \sum_{\boldsymbol{\alpha} \in A^*} ||f||_{p,C(\boldsymbol{\alpha})}. \end{aligned} \tag{3.3.3}$$

If $\kappa = \kappa*$, by the assumption (3.3.1)

$$(v(m, n; \Psi, \kappa))^{1-1/p} \sum_{\boldsymbol{\alpha} \in A^*} ||f||_{p,C(\boldsymbol{\alpha})} \leq \frac{\delta}{2}.$$

Therefore substituting into (3.3.3) gives

$$\mu(M_{m,n}(\mathbb{R}) \setminus \mathbf{Bad}_0(m,n,\Psi,\kappa)) \leq \delta.$$

□

3.4. Examples

There are many variations of Theorem 3.3.1 when we know more about the measure μ. The following is a corollary for probability measures μ with bounded density f and mean value about the origin.

Corollary 3.4.1. *Let $m, n \in \mathbb{N}, \mu, \Psi$, and M_Ψ be as in Theorem 3.1.3. Let the density f of μ be bounded above by a constant $K > 0$. Furthermore, let T be the smallest integer such that*

$$\mu([-T,T)^{mn}) \geq 1 - \delta/2.$$

Let κ^* be the root of

$$v(n,m;\Psi,x) - \frac{\delta}{2K(2T)^{mn}} = 0,$$

for $0 < x < 1/2$ where v is from (3.2.6). Then for any $\delta \in (0,1)$

$$\mu(\mathbf{Bad}_0(m,n;\Psi,\kappa)) \geq 1 - \delta$$

with

$$\kappa := \frac{1}{2}\min\left\{\frac{1}{M_\Psi}, \kappa^*\right\}.$$

Proof. Let $p = \infty$ and let A^* be the smallest set of corners $\boldsymbol{\alpha}$ such that

$$[-T,T)^{mn} \subseteq \bigcup_{\boldsymbol{\alpha} \in A^*} C(\boldsymbol{\alpha}).$$

Then $\#A^* = (2T)^{mn}$ and thus $\Sigma_f \leq K(2T)^{mn}$. Substituting into (3.3.1), with $p = \infty$

$$v(n,m;\Psi,x) - \frac{\delta}{2K(2T)^{mn}} = 0.$$

□

We now introduce an example based around the one found in Section 2.3 of [**2**], with an n-dimensional Gaussian measure.

Example 3.4.1. Take $m = 1$, and let μ be the n-dimensional Gaussian measure with mean 0, variance 1 and pairwise mutual correlation 0. Therefore the density function $f : \mathbb{R}^n \to \mathbb{R}$ is given by

$$f(\mathbf{x}) = \frac{1}{\sqrt{2\pi}} e^{-(x_1^2 + x_2^2 \cdots + x_n^2)/2},$$

where $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$. Let $\Psi(\mathbf{a})$ for $\mathbf{a} \in \mathbb{Z}^n \setminus \{\mathbf{0}\}$ be the function given by

$$\Psi(\mathbf{a}) = \begin{cases} \frac{1}{2|\mathbf{a}|^n \log^2 |\mathbf{a}|} & \text{if} \quad |\mathbf{a}| \geq 2, \\ 1/2 & \text{if} \quad |\mathbf{a}| = 1, \\ 0 & \text{if} \quad |\mathbf{a}| \leq 0. \end{cases}$$

To see that Ψ satisfies the assumption of the Corollary 3.4.1, let $1 \leq j \leq n$ then

$$\begin{aligned} \Sigma_{\Psi^j} &= \sum_{\mathbf{a} \in \mathbb{Z}^n \setminus \{\mathbf{0}\}} \Psi(\mathbf{a})^j = \sum_{r=1}^{\infty} \sum_{|\mathbf{a}|=r} \Psi(\mathbf{a})^j, \\ &= \frac{1}{2^j} + \sum_{r=2}^{\infty} r^{n-1} \frac{1}{2^j r^{nj} \log^{2j} r} \\ &= \frac{1}{2^j} + \sum_{r=2}^{\infty} r^{n-nj-1} \frac{1}{2^j \log^{2j} r} < \infty \text{ as } \Sigma_\Psi < \infty. \end{aligned} \tag{3.4.1}$$

To calculate T, first define the n-dimensional Gaussian error function by

$$\begin{aligned} \operatorname{erf}(\mathbf{x}) &:= \operatorname{erf}(x_1) \operatorname{erf}(x_2) \ldots \operatorname{erf}(x_n), \\ \text{where } \operatorname{erf}(x) &:= \frac{1}{\sqrt{2\pi}} \int_{\infty}^{x} e^{-t^2/2} dt. \end{aligned}$$

The function erf is continuous, strictly increasing and

$$\lim_{x \to -\infty} \operatorname{erf}(x) = 0 \qquad \text{and} \qquad \lim_{x \to \infty} \operatorname{erf}(x) = 1.$$

Define $\operatorname{erf}^{-1}(y)$ to be the unique $x \in \mathbb{R}$ such that $\operatorname{erf}(x) = y$, with $\operatorname{erf}^{-1}(0) := -\infty$ and $\operatorname{erf}^{-1}(1) := \infty$. Now we can see that

$$\begin{aligned} \mu(\mathbb{R}^n \setminus [-T, T)^n) &= 1 - \mu([-T, T))^n, \\ &= 1 - (1 - 2\operatorname{erf}(-T))^n. \end{aligned}$$

Let r be the root of $1 - (1 - 2r)^n - \delta/2 = 0$, then

$$T = \lceil \operatorname{erf}^{-1}(1 - r) \rceil. \tag{3.4.2}$$

Here $\lceil x \rceil$ is the ceiling of $x \in \mathbb{R}$.

By (3.4.1) and (3.4.2), we have satisfied the assumptions of Corollary 3.4.1 therefore

$$\mu(\mathbf{Bad}(m, n; \Psi, \kappa)) \geq 1 - \delta,$$

when κ is the root of

$$v(n, 1; \Psi, \kappa) - \frac{\delta}{2(2T)^n} = 0,$$

where v is from (3.2.6)

We can then find κ for different δ and n. Using Mathematica we calculated the following table:

	δ				
n	0.1	0.01	0.001	0.0001	0.00001
1	0.0027	0.0002	0.000016	1.3×10^{-6}	1.1×10^{-7}
2	0.0003	0.000019	1.3×10^{-6}	1.3×10^{-7}	9.8×10^{-9}
3	0.000055	2.8×10^{-6}	1.6×10^{-7}	$1. \times 10^{-8}$	$1. \times 10^{-9}$
4	8.2×10^{-6}	3.4×10^{-7}	1.6×10^{-8}	8.7×10^{-10}	8.7×10^{-11}
5	1.2×10^{-6}	3.8×10^{-8}	1.5×10^{-9}	7.1×10^{-11}	7.1×10^{-12}
6	1.6×10^{-7}	4.2×10^{-9}	1.4×10^{-10}	5.6×10^{-12}	5.6×10^{-13}
7	2.2×10^{-8}	4.6×10^{-10}	1.3×10^{-11}	4.4×10^{-13}	1.7×10^{-14}
8	2.9×10^{-9}	4.9×10^{-11}	1.1×10^{-12}	3.3×10^{-14}	1.1×10^{-15}

TABLE 3.1. Percentage δ of "Bad" small linear forms for n variables under approximation constant κ.

It follows for instance that for 99% of the values of the random variables $(x_1, \ldots, x_n) = \mathbf{X} \in M_{1,n}(\mathbb{R})$ with normal distribution one has that

$$|\mathbf{Xa}| > \frac{1}{5000}\Psi(\mathbf{a}) \qquad \text{for all } \mathbf{a} \in \mathbb{Z}^n \setminus \{0\}.$$

CHAPTER 4

Central limit theorem for the fractional parts of imaginary parts of the Riemann zeta zeros

4.1. Historical Background of Riemann Zeta Zeros

4.1.1. The Riemann zeta function. In his epoch-making paper [**79**] Riemann in 1859 showed that there is a link between the distribution of primes and the study of the Riemann zeta function $\zeta(s)$. More than thirty years later the link was cemented with the proof of the prime number theorem by Hadamard and de la Vallée Poussin [**37**, **47**].

The Riemann zeta function for a complex number $s \in \mathbb{C}$, is defined by the sum

$$\zeta(s) := \sum_{n=1}^{\infty} \frac{1}{n^s}.$$

A second representation was found by Euler and was subsequently named Euler's product formula.

$$\zeta(s) = \prod_{p \text{ prime}} \left(1 - p^{-s}\right)^{-1}.$$

In Riemann's paper [**79**] he proved many results, here we mention only two: One, the zeta function can be analytically continued over the whole plane, and it is meromorhphic with one simple pole of residue 1 at $s = 1$. Secondly, the zeta function satisfies the functional equation

$$\pi^{-\frac{1}{2}s}\Gamma\left(\frac{1}{2}s\right)\zeta(s) = \pi^{-\frac{1}{2}(1-s)}\Gamma\left(\frac{1}{2}(1-s)\right)\zeta(1-s).$$

The functional equation shows that the properties of $\zeta(\sigma+it)$ can be inferred for $\sigma < 0$ from its properties for $\sigma > 1$. In particular, the only zeroes of $\zeta(s)$ for $\sigma < 0$ are those at the poles of $\Gamma(\frac{1}{2}s)$. The zeros are the negative even integers $s = -2, -4, -6, \dots$ and are called the *trivial zeros*. The rest of the plane i.e. $0 \leq \sigma \leq 1$, is called *the critical strip*.

Riemann made a number of remarkable conjectures in his paper [**79**], most notably the Riemann hypothesis.

Conjecture 4.1.1 (Riemann hypothesis)**.** *The only zeros of the zeta function $\zeta(s)$ in the critical strip are those with real part $\frac{1}{2}$.*

The line of complex numbers s defined by $\Re(s) = \frac{1}{2}$ is called the *critical line*. Hardy [**49**] in 1914 made progress towards the conjecture by showing infinitely many zeros lie on the line. Later in 1942, Selberg [**82**] showed that a positive proportion of the zeros lie on the line. One of the more recent results of Bui, Conrey and Young [**23**] in 2010, is that 41.05% lie on the critical line. A more up to date result of Feng [**39**] is that 41.28% lie on the critical line.

Label the non-trivial zeros (those within the critical strip) of the Riemann zeta function by $\rho = \beta + i\gamma$, with $\gamma \in \mathbb{R}$. Order the non-trivial zeros by their height on the critical strip, that is

$$\dots \gamma_{-3} \leq \gamma_{-2} \leq \gamma_{-1} < 0 < \gamma_1 \leq \gamma_2 \leq \gamma_3 \dots$$

Define $N(T)$ by

$$N(T) := \#\{n \in \mathbb{N} : 0 < \gamma_n \leq T\}.$$

Riemann found the following asymptotic of $N(T)$,

$$N(T) = \frac{T}{2\pi} \log \frac{T}{2\pi e} + \frac{7}{8} + O(T^{-1}) + S(T) \tag{4.1.1}$$

We define the smooth part of the counting function $\overline{N}(T) = N(T) - S(T)$ as the main term in the asymptotic formula. In 1895, von Mangoldt [**95**] showed that the error term $S(T)$ has asymptotic $S(T) = O(\log T)$.

One of the main tools in the first proof of the prime number theorem is the explicit formula. First, define the von Mangoldt function $\Lambda : \mathbb{N} \to \mathbb{R}$ as

$$\Lambda(n) = \begin{cases} \log p & \text{if } n = p^k, \text{ for } k \in \mathbb{N},\, p \text{ prime}, \\ \\ 0 & \text{otherwise.} \end{cases}$$

Define the function $\psi : \mathbb{R} \to \mathbb{R}$ by

$$\psi(x) = \sum_{n \leq x} \Lambda(n).$$

The explicit formula is then, for all real $x > 2$

$$\psi(x) - x = -\sum_{\rho} \frac{x^\rho}{\rho} - \frac{\zeta'(0)}{\zeta(0)} - \frac{1}{2}\log(1 - x^{-2}). \tag{4.1.2}$$

Where the $\sum_\rho$ is counted over all non-trivial zeros with positive and negative imaginary parts. Equation (4.1.2) was originally shown by von Mangoldt in [**95**], and was used in later proofs of the prime number theorem.

Let $\pi(x)$ be the number of primes $p < x$. Let $\mathrm{li}(x)$ be defined by the integral

$$\mathrm{li}(x) := \int_2^x \frac{1}{\log t}\,dt.$$

The prime number theorem states

THEOREM 4.1.2. *For a constant $c \in (0,1)$ as $t \to \infty$,*

$$\pi(t) = \mathrm{li}(t) + E(t) \qquad \textit{with} \qquad E(t) = O(te^{-c\sqrt{\log t}}). \tag{4.1.3}$$

The statement does not assume RH, for a proof and discussion on the current best error term assuming the RH see Davenport's "Multiplicative Number Theory" [**35**].

4.1.2. Distribution. Whereas a lot of work has gone into the real part of the Riemann zeta zeros (due to RH) there are a lot of interesting open problems related to the distribution of the imaginary parts. Littlewood [**67**], showed that the gap between two consecutive zeros γ_n and γ_{n+1} tends to zero, as $n \to \infty$. In particular, he obtained that

$$\gamma_{n+1} - \gamma_n \ll \frac{1}{\log\log\log \gamma_n}, \qquad \text{as } n \to \infty.$$

According to the asymptotic number of zeros at height T (4.1.1) the mean spacing between consecutive zeros at height T is given by $2\pi/\log T$, as $T \to \infty$. The *gap conjecture* predicts that there appear arbitrarily small and arbitrarily large deviations from the mean spacing: let

$$\lambda_s := \limsup_{n\to\infty} \frac{(\gamma_{n+1} - \gamma_n)\log \gamma_n}{2\pi}$$
$$\lambda_i := \liminf_{n\to\infty} \frac{(\gamma_{n+1} - \gamma_n)\log \gamma_n}{2\pi},$$

then, one expects $\lambda_s = \infty$ and $\lambda_i = 0$. Current methods are far from such results, however Fujii [**43**] after a remark by Selberg [**84**], proved

THEOREM 4.1.3 (Fujii 1975). *For each $r = 1, 2, 3, \ldots,$ a positive proportion of γ_n's satisfy*

$$\frac{\gamma_{n+r} - \gamma_n}{r} \frac{\log \gamma_N}{2\pi} < 1 - A,$$

and also a positive proportion of γ_n's satisfy

$$\frac{\gamma_{n+r} - \gamma_n}{r} \frac{\log \gamma_N}{2\pi} > 1 + A, ,$$

where A is a positive absolute constant less than 1 *which may depend on r.*

In his paper [**69**], Montgomery studied the pair correlation function. His paper brought into light that there were strong statements to be made about the Riemann zeta zeros without considering the Riemann hypothesis.

Conjecture 4.1.4 (Montgomery's pair correlation conjecture 1973)**.** *For any fixed* $0 < \alpha < \beta$*,*

$$\lim_{T\to\infty} \frac{\#\{\gamma,\gamma' \in (0,T] : \alpha \leq \frac{(\gamma-\gamma')\log T}{2\pi} \leq \beta\}}{N(T)} = \int_\alpha^\beta \left(1 - \left(\frac{sin \pi u}{\pi u}\right)^2\right) du.$$

Montgoemery's conjecture implies that a positive proportion of zeros are simple. Dyson famously pointed out to Montgomery that eigenvalues of random Hermitian matrices have exactly the same pair correlation function. The observation was the basis for the GUE hypothesis that states that the re-scaled zeros have the distribution of the Gaussian unitary ensemble. Odlyzko [**72**] showed for a large number of zeros, that experimentally and heuristically this is true.

4.1.3. Uniform Distribution. The definition of uniformly distributed mod 1 is as follows:

Definition 4.1.1. Given a sequence of real numbers $(x_n)_{n\in\mathbb{N}}$, define the counting function for an interval $[a,b) \subseteq [0,1)$

$$A([a,b),N) = \#\{\{x_n\} \in [a,b) : \text{ for all } n \leq N\}.$$

Then $(x_n)_{n\in\mathbb{N}}$ is *uniformly distributed mod* 1 *(u.d. mod 1.)* if and only if:

$$\lim_{N\to\infty} \frac{A([a,b);N)}{N} = b - a \qquad \forall\, [a,b) \subseteq [0,1).$$

Weyl developed the theory of uniform distribution in his paper [**98**] where he proved the following equivalent statement to uniformly distributed mod 1.

Theorem 4.1.5 (Weyl's Criterion 1916). *A sequence of real numbers* $(x_n)_{n\in\mathbb{N}}$ *is uniformly distributed mod* 1 *if and only if:*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} e^{2\pi i h x_n} = 0, \qquad \textit{for all integers } h \neq 0.$$

Rademacher [**77**] was the first to prove the following theorem, assuming the Riemann Hypothesis (RH). Subsequently, Hlawka showed that they are u.d. mod 1 without the need for RH [**50**].

Theorem 4.1.6 (Rademacher 1956 assuming RH, Hlawka 1975 without RH). *The imaginary parts of the non-trivial zeros of the Riemann zeta function are uniformly distributed mod* 1 *(u.d. mod* 1*).*

The main result required to prove Theorem 4.1.6 is Landau's formula [**63**].

THEOREM 4.1.7 (Landau's formula 1912). *Let $\rho = \beta + i\gamma$ be the non-trivial zeros of the Riemann zeta function, then for $x > 1$ as $T \to \infty$*

$$\sum_{0<\gamma<T} x^{\rho} = -\frac{\Lambda(x)T}{2\pi} + O(\log T).$$

PROOF OF THEOREM 4.1.6 UNDER RH. Assuming the Riemann hypothesis, we apply Landau's formula to exponential sums. Let $x(h) = e^{2\pi h}$

$$\begin{aligned}\frac{1}{N(T)} \sum_{0\leq\gamma\leq T} e^{2\pi i h\gamma} =& \frac{1}{N(T)} \sum_{0\leq\gamma\leq T} \frac{x(h)^{\rho}}{x(h)^{1/2}}, \\ \ll& \frac{1}{T\log T}\left(\frac{\Lambda(x(h))T}{x(h)^{1/2}} + \frac{\log T}{x(h)^{1/2}}\right), \\ \ll& \frac{\log x(h)}{x(h)^{1/2}\log T} \to 0,\end{aligned}$$

as $N \to \infty$. By Weyl's criterion Theorem 4.1.5 the sequence is then u.d. mod 1. □

4.1.3.1. *Discrepancy.* We now describe a way of quantifying the difference between u.d. mod 1 sequences. The discrepancy tells you how "fast" a u.d. mod 1 sequence tends towards being uniform in the unit interval.

DEFINITION 4.1.2. The *discrepancy* D_N, of a sequence of real numbers $(x_n)_{n\in\mathbb{N}}$ is defined by

$$D_N((x_n)_{n\in\mathbb{N}}) := \sup_{0<a\leq b\leq 1}\left|\frac{A([a,b);N)}{N} - (b-a)\right|.$$

An alternative related measure, is the star discrepancy D_N^* defined by

$$D_N^*((x_n)_{n\in\mathbb{N}}) := \sup_{0<a\leq 1}\left|\frac{A([0,a);N)}{N} - a\right|.$$

There are some well known bounds on the Discrepancy which can be found in Kuipers and Niederreiter [**61**].

(1) For any sequence of N numbers, we have

$$\frac{1}{N} \leq D_N((x_n)_{n\in\mathbb{N}}) \leq 1. \tag{4.1.4}$$

(2) There exists a constant $C > 0$ such that for all $N > 1$:

$$C\log N < ND_N^*((x_n)_{n\in\mathbb{N}}).$$

(3) Le Veque's Inequality: For any finite sequence $x_1, x_2, \ldots, x_N$,

$$D_N((x_n)_{n\in\mathbb{N}}) \leq \left(\frac{6}{\pi^2}\sum_{h=1}^{\infty}\frac{1}{h^2}\left|\frac{1}{N}\sum_{n=1}^{N} e^{2\pi i h x_n}\right|^2\right)^{1/3}.$$

(4) Erdős-Turán Inequality: For any finite sequence $x_1, x_2, \ldots, x_N$ and real positive integer m, we have

$$D_N((x_n)_{n\in\mathbb{N}}) \leq \frac{6}{m+1} + \frac{4}{\pi}\sum_{h=1}^{m}\left(\frac{1}{h} - \frac{1}{m+1}\right)\left|\frac{1}{N}\sum_{n=1}^{N} e^{2\pi i h x_n}\right|.$$

(5) The inequality is often referred to by the weaker statement:

$$D_N((x_n)_{n\in\mathbb{N}}) = O\left(\frac{1}{m} + \sum_{h=1}^{m}\frac{1}{h}\left|\frac{1}{N}\sum_{n=1}^{N} e^{2\pi i h x_n}\right|\right).$$

Let γ_n be the imaginary parts of the Riemann zeta zeros with positive imaginary part in order of height. For $\alpha \in \mathbb{R}$, define the discrepancy of the sequence $(\{\alpha\gamma_n\})_{n\in\mathbb{N}}$ by

$$D_\alpha(T) := \sup_{[a,b)\subseteq[0,1)}\left|\frac{1}{N(T)}\#\{\{\alpha\gamma\} \in [a,b) : 0 < \gamma \leq T\} - (b-a)\right|.$$

Unconditionally, Fujii [**44**] proved that $D_\alpha(T) \ll \frac{\log\log T}{\log T}$ for every α. Assuming RH, Hlawka [**50**] showed that $D_\alpha(T) \ll \frac{1}{\log T}$.

A stronger version of Landau's formula was found by Ford and Zaharescu [**41**] based on work of Gonek [**46**].

Lemma 4.1.8 (Ford and Zaharescu 2005)**.** *Let $x, T > 1$, then uniformly in x, T*

$$\sum_{0<\gamma\leq T} x^\rho = -\frac{\Lambda(n_x)}{2\pi}\frac{e^{iT\log(x/n_x)} - 1}{i\log(x/n_x)} + O\left(x\log^2(2xT) + \frac{\log 2T}{\log x}\right),$$

where n_x is the nearest prime power to x and if $x = n_x$ and the first term is $-T\frac{\Lambda(n_x)}{2\pi}$.

In further work with Soundararajan [**40**], they used the formula to find a lower bound of the $D_\alpha(T)$.

Define the function g_α on the torus $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ to the complex numbers $\mathbb{C}$. When α is not a rational multiple of $\frac{\log p}{2\pi}$ for some prime p, let $g_\alpha(t) = 0$. Otherwise, if $\alpha = \frac{a\log p}{q 2\pi}$ for some rational a/q where $(a,q) = 1$ define g_α by:

$$g_\alpha(t) := \frac{-\log p}{\pi}\Re\sum_{k=1}^{\infty}\frac{e^{-2\pi i q h x}}{p^{ak/2}}.$$

Here we state Ford, Soundararajan and Zaharescu's full conjecture in [**40**] on the formula for the discrepeancy of the zeros of which they showed was a lower bound for the discrepancy.

Conjecture 4.1.9 (Ford, Soundararajan and Zaharescu 2009)**.** *For $\alpha \in \mathbb{R}$, the discrepancy of the sequence $(\{\alpha\gamma_n\})_{n\in\mathbb{N}}$ is*

$$D_\alpha(T) = \frac{T}{N(T)} \sup_{\mathbb{I}} \left| \int_{\mathbb{I}} g_\alpha(x)dx \right| + o\left(\frac{1}{\log T}\right).$$

The authors showed in [**40**] that Conjecture 4.1.9 is true if the following conjecture (which is an analogue to one of Gonek's [**46**]) is true.

Conjecture 4.1.10. *Let $A > 1$ be a fixed real number. Uniformly for all $\frac{T^2}{(\log T)^5} \le x \le T^A$ we have*

$$\sum_{0\le\gamma\le T} x^{i\gamma} = o(T). \tag{4.1.5}$$

Conjecture 4.1.10 can then again be implied by a stronger version of the pair correlation Conjecture 4.1.4. First define the following pair correlation function, which is different to the one Montgomery studied in [**69**].

$$\mathcal{F}(x,T) = \sum_{0<\gamma,\gamma'\le T} \frac{4x^{i(\gamma-\gamma')}}{4+(\gamma-\gamma')^2}. \tag{4.1.6}$$

Conjecture 4.1.11. *Fix a real number $A > 1$. Uniformly for all $\frac{T^2}{(\log T)^6} \le x \le T^A$ we have*

$$\mathcal{F}(x,T) = N(T) + o\left(\frac{T}{\log T}\right).$$

Conjecture 4.1.11 implies the pair correlation function contains more information than the sum (4.1.5).

Steuding's survey paper [**88**], gives a comprehensive overview of results regarding the distribution of fractional parts of imaginary parts of zeta zeros. He also discusses his results on the distribution of the fractional parts of $\Im(s)$ such that $\zeta(s) = a$ for a fixed $a \in \mathbb{C}$.

4.1.4. Central limit theorem. Selberg [**83**] has been the main inspiration for a sequence of works studying the remainder term in the counting function $S(T)$ as t varies between T and $2T$. He proved that $S(T)$ has Gaussian moments, essentially showing that when $T \to \infty$,

$$\frac{1}{T}\int_T^{2T} \left| \frac{S(t)}{\sqrt{(\log\log T)/2\pi^2}} \right|^{2k} dt \to \frac{(2k)!}{k!2^k}.$$

Fujii [**42**] built on Selberg's work to look at the remainder term $S(T,\chi)$ of the counting function of Dirichlet L-functions.

Later Hughes and Rudnick [**51**], looked at the linear statistics of the normalised zeros. In particular, for a real-valued even function f, and real numbers τ and $T > 1$, set

$$N_f(\tau) := \sum_{j=\pm 1,\pm 2,,\ldots} f\left(\frac{\log T}{2\pi}(\gamma_j - \tau)\right).$$

If f is the characteristic function of an interval $[-1, 1]$ and if all the γ_j are real, then $N_f(\tau)$ counts the number of zeros in the interval

$$[\tau - 2\pi/\log T, \tau + 2\pi/\log T].$$

However, we will take f so that its Fourier transform,

$$\hat{f}(u) := \int_{-\infty}^{\infty} f(x)e^{-2\pi ixu}dx,$$

is smooth and of compact support, and will not assume RH. Hughes and Rudnick proved the following statement.

THEOREM 4.1.12. *Let $H = T^a$ with $0 < a \leq 1$, and let $f \in C_\infty(\mathbb{R})$ be such that the support of $f \subseteq (-2a/m, 2a/m)$. Then the first m moments of N_f converge as $T \to \infty$ to those of a Gaussian random variable with*

$$\mathbb{E} = \int_{-\infty}^{\infty} f(x)dx \quad \textit{and} \quad \sigma^2 = \int_{-\infty}^{\infty} \min(|u|, 1)\hat{f}(u)^2 du.$$

4.2. Central limit theorem for the fractional parts of zeta zeros

We have already seen that the imaginary parts of the zeta zeros are u. d. mod 1, with bounds on the discrepancy. Now we ask the question of how like identically independent distributed (i.i.d) points are they? In particular we ask do they obey a central limit theorem?

For $0 < x < 1$, choose an integer t randomly in $[T, T + H]$. The number of zeros between $t + 1$ and $t + M$ with fractional part less than or equal to x is

$$\sum_{n=1}^{M} N(t + n + x) - N(t + n)$$
$$= \sum_{n=1}^{M} \overline{N}(t + n + x) - \overline{N}(t + n) + \sum_{n=1}^{M} S(t + n + x) - S(t + n).$$

We are interested in the sum

$$S_M(x, t) := \sum_{n=1}^{M} S(t + n + x) - S(t + n), \tag{4.2.1}$$

which has 0 mean, and whether or not it obeys a central limit theorem. In other words, our result is that the zeros do not act like i.i.d points, but rather when "layering" up the blocks of zeros from intervals $[t+n, t+n+x)$ for $1 \leq n \leq M$ and $T \leq t \leq 2T$, the blocks act like i.i.d points. In the sense that, the average of $S_M(x,t)$ converges in distribution to the Gaussian distribution. We will average over all $t \in [T, 2T)$, since the continuous average and the average over the integers is expected to be close.

Let $|\cdot|_1$ be the 1-dimensional Lebesgue measure. Our main result is then

THEOREM 4.2.1. *Let* $0 < x < 1$. *For* $T \geq 1$ *and an interval* $\mathcal{A} \subseteq \mathbb{R}$, *if* $1 \leq M \leq \sigma^{1-\epsilon}$ *for* $\epsilon > 0$ *then*

$$\frac{1}{T}\left|\left\{t \in [T,2T] : \frac{S_M(x,t)}{\sigma M} \in \mathcal{A}\right\}\right|_1 \to \frac{1}{\sqrt{2\pi}}\int_{\mathcal{A}} e^{-u^2/2}du \tag{4.2.2}$$

weakly as $T \to \infty$, *where*

$$\sigma^2 = \frac{1}{\pi^2}\log(x\log T) + O(1)$$

is the variance of $S(t+x) - S(t)$.

Note that, we expect the limit $1 \leq M \leq \sigma^{1-\epsilon}$ to not be optimal. In fact we expect that $M = T^a$ for some $a \ll 1$.

The local statistics of critically scaled zeros of the Riemann zeta function around height T are believed [**69**, **72**] to behave like eigenvalues of a random unitary matrix, when scaled by their mean density. A similar result to Theorem 4.2.1 in random matrix theory, is Rains' theorem [**78**]. Rains' result says that if U is a Haar-distributed $N \times N$ unitary matrix then the eigenvalues of U^m are distributed as those of m independent Haar-distributed unitary matrices of size $\lfloor N/m \rfloor$, so long as $m \leq N$. For $m \geq N$ the eigenvalues of U^m are distributed as N independent random variables uniformly distributed on the unit circle (that is, N independent 1×1 Haar-distributed unitary matrices' eigenvalues). This replicates our effect of "layering" up zeros by "layering" eigenvalues of a matrix instead.

4.2.1. Model of $S(T)$ and supporting statements. The structure of the argument is as follows: For all t around height T we replace the function $S(t)$ by a truncated version

$$\overline{S}(t,X) = \frac{1}{\pi}\Im\sum_{p\leq X}\frac{1}{p^{1/2+it}}, \tag{4.2.3}$$

with a parameter X dependent on the height T. To calculate the variance and moments of the truncated version of $S_M(x,t)$, we use a smooth version of Fujii's Lemma [**42**], which finds asymptotics for integrals of the form

$$\int_T^{T+H}\left|\Im\sum_{p<X}\frac{a(p)}{p^{1/2+it}}\right|^{2k} dt.$$

Using another approximation of $S(t)$ by Selberg, we then show the variance between the moments of the truncated version $\overline{S}_M(x,t,X)$ and $S_M(x,t)$ is small. Finally, we use an approximation argument from [**52**] to pass from smooth measures to the one dimensional Lebesgue measure used in Theorem 4.2.1.

The symbols $\ll$ and $\gg$ will be used to indicate an inequality with an unspecified positive multiplicative constant. The symbol $\asymp$ means $\ll$ and $\gg$ at the same time. We will use the notation $f(x) = O(g(x))$ to mean $f \ll g$, also $f(x) = o(g(x))$ if for all $\epsilon > 0$, $|f(x)| \le \epsilon g(x)$ as x tends to a limit.

We are going to use two models of $S(t)$ from Selberg in [**83**, Theorem 2], the truncated one $\overline{S}(t,X)$ has the nice property.

Lemma 4.2.2. *Let $k \in \mathbb{N}$ and $0 < a \le 1$. For $T^a \le H \le T^2$ and $T^{a/k} \le X \le H^{1/k}$*

$$\frac{1}{H}\int_T^{T+H} \left|S(t) - \overline{S}(t,X)\right|^{2k} dt = O(1),$$

as $T \to \infty$.

We also use an asymptotic approximation of $S(t)$. Define the function $\Lambda_X(n) : \mathbb{N} \to \mathbb{R}^+$ by

$$\Lambda_X(n) := \begin{cases} \Lambda(n) & 1 \le n \le \sqrt{X}, \\ \Lambda(n)\frac{\log \frac{X}{n}}{\log \sqrt{X}} & \sqrt{X} \le n \le X. \end{cases} \tag{4.2.4}$$

where $\Lambda(n)$ is the Von-Mangoldt function

$$\Lambda(n) := \begin{cases} \log p & \text{if } n = p^k \text{ for some } k \in \mathbb{N}, \\ 0 & \text{otherwise.} \end{cases}$$

Selberg's approximation $S(t)$ from Theorem 1 of [**83**], states

Theorem 4.2.3. *For $t > 2$, $2 \le X \le t$, $\sigma_1 = \frac{1}{2} + \frac{2}{\log X}$ we have that,*

$$S(t) = -\frac{1}{\pi}\sum_{n<X}\frac{\Lambda_X(n)}{n^{\sigma_1}}\frac{\sin(t\log n)}{\log n} + O\left(\frac{1}{\log\sqrt{X}}\left|\sum_{n<X}\frac{\Lambda_X(n)}{n^{\sigma_1+it}}\right|\right) + O\left(\frac{\log t}{\log\sqrt{X}}\right). \tag{4.2.5}$$

We will make much use of the Fejér kernel $K_M : \mathbb{R} \to \mathbb{R}$ for $M \in \mathbb{N}$, defined as follows

$$K_M(t) := \sum_{k=-M}^{M}\frac{(M-|k|)}{M}e^{-ikt}. \tag{4.2.6}$$

It has several nice properties such as

$$K_M(t) = \frac{1}{M}\frac{1-\cos(Mt)}{1-\cos t} = \frac{1}{M}\frac{\sin^2\left(\frac{Mt}{2}\right)}{\sin^2\left(\frac{t}{2}\right)}, \tag{4.2.7}$$

also for all $t \in \mathbb{R}$

$$K_M(t) \leq K_M(0) = M. \tag{4.2.8}$$

The Fejér kernel is periodic with period 2π and has the following integral along its period

$$\int_0^{2\pi} K_M(t)dt = 2\pi. \tag{4.2.9}$$

4.2.2. Definition of the smooth average and related Lemmas. We will use a smooth weight function $\frac{1}{H}w\left(\frac{t-T}{H}\right)$ for averaging. We choose the weight function $w \geq 0$ such that $\int_{-\infty}^{\infty} w(x)dx = 1$, and the Fourier transform $\hat{w}(k)$ is compactly supported in $(-\frac{1}{2\pi}, \frac{1}{2\pi})$, where the Fourier transform of a complex function $f : \mathbb{C} \to \mathbb{C}$ will be defined by:

$$\hat{f}(k) = \int_{-\infty}^{\infty} f(x)e^{-2\pi ikx}dx. \tag{4.2.10}$$

Note that the conditions on w imply that it is rapidly decaying in the sense that for any $A > 2$

$$w(t) \ll \frac{1}{(1+|t|)^A},$$

for all t.

The smooth average of a complex function $W : \mathbb{C} \to \mathbb{C}$ is defined by

$$\langle W\rangle_{T,H} = \frac{1}{H}\int_{-\infty}^{\infty} w\left(\frac{t-T}{H}\right)W(t)dt. \tag{4.2.11}$$

Let f be a real function and $\mathcal{A}$ a measurable subset $\mathcal{A} \subseteq \mathbb{R}$. We define the associated probability measure with an attached weight function w by

$$\mathbb{P}_{w,T,H}\left\{f \in \mathcal{A}\right\} = \frac{1}{H}\int_{-\infty}^{\infty} \mathbb{1}_{\mathcal{A}}(f(t))w\left(\frac{t-T}{H}\right)dt. \tag{4.2.12}$$

Where $f \in \mathcal{A}$ if there exists $t \in \mathbb{R}$ such that $f(t) \in \mathcal{A}$.

We will then need the following claim:

Lemma 4.2.4. *Let $k \in \mathbb{N}$ and $0 < a \leq 1$. Let $T \geq 1$, $T^a \leq H \leq T^2$ and $T^{a/k} \leq X \leq H^{1/k}$. Let w be a weight function with a Fourier transform $\hat{w}(k)$ compactly supported in $(-\frac{1}{2\pi}, \frac{1}{2\pi})$. For $\mathbf{p} = p_1p_2\ldots p_m$ and $\mathbf{q} = q_1q_2\ldots q_n$ with p_i and q_j primes less than X for $1 \leq i \leq m$ and $1 \leq j \leq n$ with $m+n = 2k$, we have*

$$\left\langle\left(\frac{\mathbf{p}}{\mathbf{q}}\right)^{it}\right\rangle_{T,H} = \begin{cases} 1 & \text{if } \mathbf{p} = \mathbf{q}, \\ 0 & \text{otherwise.}\end{cases}$$

PROOF. If $\mathbf{p} = \mathbf{q}$ it is obvious. So consider $\mathbf{p} \neq \mathbf{q}$. First note that

$$\left\langle \left(\frac{\mathbf{p}}{\mathbf{q}}\right)^{it} \right\rangle_{T,H} = \frac{1}{H}\int_{-\infty}^{\infty} w\left(\frac{t-T}{H}\right) e^{it\log\frac{\mathbf{p}}{\mathbf{q}}}dt, \tag{4.2.13}$$

$$= \left(\frac{\mathbf{p}}{\mathbf{q}}\right)^{iT} \hat{w}\left(-\frac{H}{2\pi}\log\frac{\mathbf{p}}{\mathbf{q}}\right). \tag{4.2.14}$$

Since $\hat{w}(t)$ is supported in $(-1/2\pi, 1/2\pi)$, we are interested in how small $\left|\frac{H}{2\pi}\log\frac{\mathbf{p}}{\mathbf{q}}\right|$ is.

The smallest $\left|\log\frac{\mathbf{p}}{\mathbf{q}}\right|$ can be is when $\frac{\mathbf{p}}{\mathbf{q}}$ is close to 1. This happens when $k = n$ and p_i, q_i are as big as possible, but $\mathbf{p}$ and $\mathbf{q}$ are not equal. Therefore we may bound it by

$$\frac{\mathbf{p}}{\mathbf{q}} \leq \frac{X^k - 1}{X^k},$$

and since for $0 < y < 1$ we have $|\log(1-y)| > y$ then

$$\left|\frac{H}{2\pi}\log\frac{\mathbf{p}}{\mathbf{q}}\right| > \frac{H}{2\pi}\frac{1}{X^k} > \frac{1}{2\pi}$$

since by assumption $X^k \leq H$. Due to the support condition on $\hat{w}$ this means $\hat{w}(-\frac{H}{2\pi}\log\frac{\mathbf{p}}{\mathbf{q}})$ is equal to 0. ☐

Now we present a variation of Lemma 3 in [**42**]. Define $F_\alpha : \mathbb{R}^+ \to \mathbb{R}^+$ for $\alpha \in \mathbb{R}^+$, by

$$F_\alpha(X) = \sum_{p<X} \frac{|a(p)|^{2\alpha}}{p^\alpha}, \tag{4.2.15}$$

where the sum is taken over primes less than X.

Lemma 4.2.5. *Assume that*

$$F_1(X) \to \infty \quad \textit{as} \quad X \to \infty$$

and $F_2(X) \ll F_1(X)^{2-\delta}$ *for* $0 < \delta < 2$. *Let* $0 < a \leq 1$ *and* $k \in \mathbb{N}$. *For* $T \geq 1$, $T^a \leq H \leq T^2$ *and* $T^{a/k} \leq X \leq H^{1/k}$,

$$\left\langle \left| \Im \sum_{p<X} \frac{a(p)}{p^{1/2+it}} \right|^{2k} \right\rangle_{T,H} = \frac{(2k)!}{2^{2k}k!}F_1(X)^k + O\left(F_1(X)^{\max(0,k-\delta)}\right), \tag{4.2.16}$$

as $T \to \infty$.

If we let m_k be the Gaussian moments, i.e $m_k = \frac{(2k)!}{2^k k!}$, then we can rewrite equation (4.2.16) as

$$\left\langle \left| \Im \sum_{p<X} \frac{a(p)}{p^{1/2+it}} \right|^{2k} \right\rangle_{T,H} = m_k\left(\frac{1}{2}F_1(X)\right)^k + O\left(F_1(X)^{\max(0,k-\delta)}\right).$$

Proof. Let

$$r = \sum_{p<X} \frac{a(p)}{p^{1/2+it}}, \qquad \text{then} \quad \Im \sum_{p<X} \frac{a(p)}{p^{1/2+it}} = \frac{r-\bar{r}}{2i}.$$

Substituting into the left hand side of (4.2.16) via (4.2.11)

$$\begin{aligned}
&\left\langle \left| \Im \sum_{p<X} \frac{a(p)}{p^{1/2+it}} \right|^{2k} \right\rangle_{T,H} \qquad (4.2.17)\\
&= \frac{1}{H}\int_{-\infty}^{\infty} w\left(\frac{t-T}{H}\right)\left(\frac{r-\bar{r}}{2i}\right)^{2k} dt,\\
&= \frac{1}{H}\int_{-\infty}^{\infty} w\left(\frac{t-T}{H}\right)\frac{1}{2^{2k}i^{2k}}\sum_{b=0}^{2k}(-1)^b\binom{2k}{b} r^b\bar{r}^{2k-b}dt,\\
&= \frac{(-1)^k}{2^{2k}}\sum_{b=0}^{2k}\binom{2k}{b}(-1)^b\frac{1}{H}\int_{-\infty}^{\infty} w\left(\frac{t-T}{H}\right) r^b\bar{r}^{2k-b}dt. \qquad (4.2.18)
\end{aligned}$$

Expanding out the integrand gives

$$r^b\bar{r}^{2k-b} = \sum_{p_1,\dots,p_{2k}<X} \frac{a(p_1)\dots a(p_b)\overline{a(p_{b+1})}\dots\overline{a(p_{2k})}}{\sqrt{p_1\dots p_{2k}}}\left(\frac{p_{b+1}\dots p_{2k}}{p_1\dots p_b}\right)^{it}.$$

Substituting back into (4.2.18) we need to consider the integral

$$\frac{1}{H}\int_{-\infty}^{\infty} w\left(\frac{t-T}{H}\right)\left(\frac{p_{b+1}\dots p_{2k}}{p_1\dots p_b}\right)^{it} dt. \qquad (4.2.19)$$

Let $\mathbf{p} = p_{b+1}\dots p_{2k}$ and $\mathbf{q} = p_1\dots p_b$. It then follows from Lemma 4.2.4 that, (4.2.19) is 0 unless $b=k$ and $\mathbf{p}=\mathbf{q}$. Therefore,

$$\left\langle \left| \Im \sum_{p<X} \frac{a(p)}{p^{1/2+it}} \right|^{2k} \right\rangle_{T,H} = \frac{1}{2^{2k}}\binom{2k}{k}\sum_{p_i<X}^{*}\frac{|a(p_1)|^2\dots|a(p_k)|^2}{p_1p_2\dots p_k}, \qquad (4.2.20)$$

where $\sum^*$ is a sum over all primes $p_1, p_2, \dots, p_{2k} < X$ such that

$$p_1p_2\dots p_k = p_{k+1}\dots p_{2k}. \qquad (4.2.21)$$

Since the p_i are all prime, note that in the case when all the $p_1, \dots, p_k$ are distinct, there are $k!$ ways of arranging the product (that is, there are $k!$ ways of pairing primes on the LHS of (4.2.21) with primes on the RHS).

Putting terms where at least two of the $p_1, \dots, p_k$ are equal into an error term we see that

$$\sum_{p_i<X}^{*}\frac{|a(p_1)|^2\dots|a(p_k)|^2}{p_1p_2\dots p_k} = k!\sum_{\substack{p_1,\dots,p_k\\ p_i<X}}\frac{|a(p_1)|^2\dots|a(p_k)|^2}{p_1p_2\dots p_k}$$

$$+ O_k\left(\sum_{\substack{p_1,p_3,\dots,p_k\\ p_i<X}} \frac{|a(p_1)|^4|a(p_3)|^2\dots|a(p_k)|^2}{p_1^2p_3\dots p_k}\right)$$
$$= k!F_1(X)^k + O_k\left(F_2(X)F_1(X)^{\max(0,k-2)}\right)$$

Note that the error term does not exist for $k = 1$.

Since $F_2(X) \ll F_1(X)^{2-\delta}$ by assumption, we have

$$\left\langle\left|\Im\sum_{p<X}\frac{a(p)}{p^{1/2+it}}\right|^{2k}\right\rangle_{T,H} = \frac{1}{2^{2k}}\binom{2k}{k}k!F_1(X)^k + O_k\left(F_1(X)^{\max(0,k-\delta)}\right).$$

□

Here we look at a similar result, proved in a non-smooth form by Selberg [**83**, Lemma 3].

Lemma 4.2.6. *Let* $0 < a \leq 1$ *and* $k \in \mathbb{N}$*. Let* $T \geq 1$*,* $T^a \leq H \leq T^2$*,* $T^{a/k} \leq X \leq H^{1/k}$*,* $A = A(T)$ *and* $|a(p)| < A\frac{\log p}{\log X}$ *for* $p < X$*; then*

$$\left\langle\left|\sum_{p<X}\frac{a(p)}{p^{1/2+it}}\right|^{2k}\right\rangle_{T,H} = O_k(A^{2k}) \tag{4.2.22}$$

and if $|a'(p)| < A$ *for* $p < \sqrt{X}$*; then*

$$\left\langle\left|\sum_{p<\sqrt{X}}\frac{a'(p)}{p^{1+2it}}\right|^{2k}\right\rangle_{T,H} = O(A^{2k}). \tag{4.2.23}$$

Proof. Let

$$\left(\sum_{p<X}\frac{a(p)}{p^s}\right)^k = \sum_{n<X^k}\frac{b_n}{n^s}.$$

Using Lemma 4.2.4 again, we can easily see that

$$\left\langle\left|\sum_{p<X}\frac{a(p)}{p^{1/2+it}}\right|^{2k}\right\rangle_{T,H} = \sum_{m<X^k}\sum_{n<X^k}\frac{b_m\bar{b}_n}{\sqrt{mn}}\left\langle\left(\frac{m}{n}\right)^{it}\right\rangle_{T,H},$$
$$= \sum_{n<X^k}\frac{|b_n|^2}{n},$$
$$= \left(\sum_{p<X}\frac{|a(p)|^2}{p}\right)^k.$$

By using the bound in the assumption of the statement,

$$\left\langle \left| \sum_{p<X} \frac{a(p)}{p^{1/2+it}} \right|^{2k} \right\rangle_{T,H} \leq \left(A^2 \sum_{p<X} \frac{\log p}{p \log X} \right)^k, \\ \ll A^{2k},$$

where we used the well known result $\sum_{p<X} \frac{\log p}{p} = O(\log X)$.

To prove Equation (4.2.23), we follow a similar method, although to apply Lemma 4.2.4 we need the primes to be restricted to be less than $\sqrt{X}$.

$$\begin{aligned} \left\langle \left| \sum_{p<\sqrt{X}} \frac{a'(p)}{p^{1+2it}} \right|^{2k} \right\rangle_{T,H} &= \sum_{m<X^{k/2}} \sum_{n<X^{k/2}} \frac{b'_m \bar{b}'_n}{mn} \left\langle \left(\frac{m}{n}\right)^{2it} \right\rangle_{T,H}, \\ &= \sum_{n<X^{k/2}} \frac{|b'_n|^2}{n^2}, \\ &= \left(\sum_{p<\sqrt{X}} \frac{|a'(p)|^2}{p^2} \right)^k. \end{aligned}$$

By using the bound in the assumption of the statement,

$$\left\langle \left| \sum_{p<\sqrt{X}} \frac{a'(p)}{p^{1+2it}} \right|^{2k} \right\rangle_{T,H} \leq \left(A^2 \sum_{p<\sqrt{X}} \frac{1}{p^2} \right)^k, \\ \ll A^{2k}.$$

□

4.2.3. Truncated average of smooth moments of $S_M(x,T)$. We define the truncated version of $S_M(x,t)$ for $0<x<1$ by

$$\overline{S}_M(x,t,X) = \sum_{n=1}^{M} \overline{S}(t+n+x,X) - \overline{S}(t+n,X).$$

We now calculate a truncated smooth version of the variance of $S(t+x) - S(t)$, which we define by

$$\overline{\sigma}^2 := \left\langle |\overline{S}(t+x,X) - \overline{S}(t,X)|^2 \right\rangle_{T,H}. \tag{4.2.24}$$

Lemma 4.2.7. *Let $x>0$ and $0<a\leq 1$. For $T\geq 1$, $T^a \leq H \leq T^2$ and $T^a \leq X \leq H$,*

$$\overline{\sigma}^2 = \frac{1}{\pi^2} \int_{x\log 2}^{x \log X} \frac{1-\cos u}{u} du + O(1),$$

as $T\to\infty$.

As a corollary we have the following asymptotic expansion of $\overline{\sigma}^2$,

$$\overline{\sigma}^2 = \frac{1}{\pi^2}\log(x\log X) + O(1) = \frac{1}{\pi^2}\log\log T + O(1). \tag{4.2.25}$$

Proof. We intend on applying Lemma 4.2.5. Let $a(p) := \frac{1}{\pi}(p^{-ix} - 1)$, we get that

$$\overline{\sigma}^2 = \left\langle \left| \Im \sum_{p<X} \frac{a(p)}{p^{1/2+it}} \right|^2 \right\rangle_{T,H}.$$

Now taking the square of $a(p)$ gives,

$$|a(p)|^2 = \frac{1}{\pi^2}|1-p^{ix}|^2 = \frac{1}{\pi^2}(2 - 2\cos(x\log p)). \tag{4.2.26}$$

Therefore $|a(p)|^2$ is bounded above and hence $F_2(X)$ is bounded. To apply Lemma 4.2.5, we now need to study $F_1(X)$ as $X \to \infty$.

Substituting (4.2.26) into (4.2.15)

$$\begin{aligned} F_1(X) &= \frac{1}{\pi^2}\sum_{p<X} 2\frac{1-\cos(x\log p)}{p}, \\ &= \frac{2}{\pi^2}\int_2^X \frac{1-\cos(x\log p)}{p} d(\pi(p)), \end{aligned}$$

where $\pi(p)$ is the number of primes less than p. By applying the prime number theorem (4.1.3) and making a change of variable $y = x\log p$,

$$\begin{aligned} F_1(X) =& \frac{2}{\pi^2}\int_2^X \frac{1-\cos(x\log p)}{p}\frac{1}{\log p}dp + \frac{2}{\pi^2}\int_2^X \frac{1-\cos(x\log p)}{p}dE(p), \\ =& \frac{2}{\pi^2}\int_{x\log 2}^{x\log X} \frac{1-\cos(y)}{y}dy + \frac{2}{\pi^2}\frac{1-\cos(x\log p)}{p}E(p)\Big|_2^X \\ &- \frac{2}{\pi^2}\int_2^X \frac{x\sin(x\log p) - 1 + \cos(x\log p)}{p^2}E(p)dp, \\ =& \frac{2}{\pi^2}\int_{x\log 2}^{x\log X} \frac{1-\cos(y)}{y}dy + I_{E1} + I_{E2}. \end{aligned} \tag{4.2.27}$$

The first error integral has asymptotics

$$\begin{aligned} I_{E1} &= \frac{2}{\pi^2}\frac{1-\cos(x\log p)}{p}E(p)\Big|_2^X, \\ &\ll \frac{2}{\pi^2}(1 + |\cos(x\log X)|)e^{-c\sqrt{\log X}} + O(1), \\ &= O(1). \end{aligned}$$

The second error integral is

$$\begin{aligned}
I_{E2} &= \frac{2}{\pi^2}\int_2^X \frac{x\sin(x\log p) - 1 + \cos(x\log p)}{p^2}E(p)dp,\\
&\ll \int_2^X \frac{e^{-c\sqrt{\log p}}}{p}dp,\\
&= O(1).
\end{aligned}$$

By substituting I_{E1} and I_{E2} back into (4.2.27)

$$F_1(X) = \frac{2}{\pi^2}\int_{x\log 2}^{x\log X} \frac{1-\cos(u)}{u}du + O(1).$$

It is now obvious that $F_1(X) \to \infty$ as $X \to \infty$, hence, by Lemma 4.2.5

$$\overline{\sigma}^2 = \frac{2!}{2^2}F_1(X) + O\left(1\right) = \frac{1}{2}F_1(X) + O(1). \tag{4.2.28}$$

□

The following Lemma is a pure analysis result that we use for our central limit theorem.

Lemma 4.2.8. *Let $f : \mathbb{R} \to \mathbb{R}$ be a monotonically decreasing function such that*

$$\int_c^T |f(t)|dt \ll 1 \qquad \text{as } T \to \infty. \tag{4.2.29}$$

Then

$$\int_c^T f(t)K_M'(t)dt \ll M \qquad \text{as } T \to \infty.$$

Proof. Using the sine representation of $K_M(t)$ at (4.2.7)

$$\begin{aligned}
K_M'(t) =& \frac{1}{M}\left(\frac{\sin^2(Mt/2)}{\sin^2(t/2)}\right)',\\
=& \frac{1}{M}\left(M\frac{\sin(Mt/2)\cos(Mt/2)}{\sin^2(t/2)} - \frac{\cos(t/2)\sin^2(Mt/2)}{\sin^3(t/2)}\right).
\end{aligned}$$

Using the known inequalities $|\sin(Mt/2)| \leq CM|t|$ and $|\sin(t/2)| \geq c|t|$ for $|t| < \pi$ and constants $c, C \in \mathbb{R}^+$.

$$\begin{aligned}
\left|M\frac{\sin(Mt/2)\cos(Mt/2)}{\sin^2(t/2)}\right| &\leq M\frac{1}{|\sin^2(t/2)|},\\
&\leq M\frac{1}{c^2|t|^2},
\end{aligned}$$

and

$$\begin{aligned}
\left|\frac{\cos(t/2)\sin^2(Mt/2)}{\sin^3(t/2)}\right| &\leq CM|t|\left|\frac{\sin(Mt/2)}{\sin^3(t/2)}\right|,\\
&\leq \frac{MC'|t|}{|t|^3} \leq \frac{C'M}{|t|^2}.
\end{aligned}$$

This implies for some constant $A \in \mathbb{R}^+$ and $|t| \le \pi$,

$$|K'_M(t)| \le \frac{A}{t^2}. \tag{4.2.30}$$

However we can obtain a better bound when t is close to 0. Consider $\epsilon > 0$ and t in the range $[-\epsilon/M, \epsilon/M]$

$$\begin{aligned}
&\int_{-\epsilon/M}^{\epsilon/M} |K'_M(t)| dt \\
&= \int_{-\epsilon/M}^{\epsilon/M} \frac{1}{M} \left| \frac{M \sin(Mt/2)}{\sin^3(t/2)} (\cos(Mt/2)\sin(t/2) - \cos(t/2)\sin(Mt/2) \right| dt \\
&= \frac{2}{M^2} \int_{-\epsilon}^{\epsilon} \left| \frac{M\sin(x)(\cos(x)\sin(x/M) - \cos(x/M)\sin(x)}{\sin^3(x/M)} \right| dx
\end{aligned}$$

The Taylor expansion of the inside term is

$$\frac{M\sin(x)(\cos(x)\sin(x/M) - \cos(x/M)\sin(x)}{\sin^3(x/M)} = \frac{1}{3}(M - M^3)x + O(x^3).$$

Therefore

$$\int_{-\epsilon/M}^{\epsilon/M} |K'_M(t)| dt = \frac{2}{3M}(M^2 - 1) \int_{-\epsilon}^{\epsilon} |x| dx + O\left(\int_{-\epsilon}^{\epsilon} |x|^3 dx \right), \tag{4.2.31}$$

$$= \frac{4}{3M}(M^2 - 1)\epsilon^2 + O(\epsilon^4). \tag{4.2.32}$$

Since $K'_M(t)$ is 2π-periodic this is true around every integer multiple of 2π. We can then rewrite the integral as

$$\int_c^T f(t) K'_M(t) dt = \sum_{k=1}^{T/(2\pi)} \left(\int_{2k\pi - \epsilon/M}^{2k\pi + \epsilon/M} + \int_{2k\pi - 1/2}^{2k\pi - \epsilon/M} + \int_{2k\pi + \epsilon/M}^{2k\pi + 1/2} f(t) K'_M(t) dt \right).$$

Using (4.2.30), (4.2.32), the assumption (4.2.29) and that f is monotonically decreasing, then implies

$$\begin{aligned}
\int_c^T f(t) K'_M(t) dt &= \sum_{k=1}^{T/(2\pi)} f(2k\pi + 1/2) O(M\epsilon^2 + M/\epsilon), \\
&= O(M\epsilon^2 + M/\epsilon) \int_c^T f(t) dt.
\end{aligned}$$

Which by the assumptions on f

$$\int_c^T f(t) K'_M(t) dt \ll M,$$

and completes the proof. □

We are now ready to calculate the smooth truncated moments.

Lemma 4.2.9. *Let $k \in \mathbb{N}$, $0 < x < 1$, $0 < \delta < 2$ and $0 < a \le 1$. For $T \ge 1$, $T^a \le H \le T^2$, $T^{a/k} \le X \le H^{1/k}$ and $1 \le M < \overline{\sigma}^{2-\delta}$,*

$$\left\langle \left(\frac{\overline{S}_M(x,t,X)}{M} \right)^{2k} \right\rangle_{T,H} = \frac{(2k)!}{2^{2k}k!}(2\overline{\sigma}^2 + O(M))^k + O\left((2\overline{\sigma}^2 + O(M))^{\max(0,k-\delta)} \right),$$

as $T \to \infty$.

Note we assume that $x \neq 1$. If $x = 1$ then $\overline{S}_M(x,t,X)$ would be a telescoping sum, i.e.

$$\overline{S}_M(1,t,X) = \sum_{n=1}^{M} \overline{S}(t+n+1,X) - \overline{S}(t+n,X) = \overline{S}(x+M,X).$$

PROOF. Again we intend on applying Lemma 4.2.5, let

$$\begin{aligned} a(p) :=& \frac{1}{\pi M} \sum_{n=1}^{M} (p^{-i(n+x)} - p^{-in}) \\ =& (p^{-ix} - 1) \frac{1}{\pi M} \sum_{n=1}^{M} p^{-in}. \end{aligned} \tag{4.2.33}$$

Therefore,

$$\overline{S}_M(x,t,X) = \Im \sum_{p<X} \frac{a(p)}{p^{1/2+it}}.$$

Now taking the modulus and square of $a(p)$ gives,

$$\begin{aligned} |a(p)|^2 &= \frac{2}{\pi^2}(1 - \cos(x \log p)) \frac{1}{M} \sum_{1 \le j,n \le M} p^{i(j-n)}, \\ &= \frac{2}{\pi^2}(1 - \cos(x \log p)) \frac{1}{M} \sum_{j=-(M-1)}^{M-1} (M - |j|) p^{-ij}, \\ &= \frac{2}{\pi^2}(1 - \cos(x \log p)) K_M(\log p), \end{aligned}$$

where the kernel K_M is defined in (4.2.6). We now consider $F_1(X)$ as $X \to \infty$.

Using the technique from the proof of Lemma 4.2.7 and applying the prime number theorem (4.2.27)

$$\begin{aligned} F_1(X) =& \frac{2}{\pi^2} \sum_{p<X} \frac{(1 - \cos(x \log p))}{p} K_M(\log p), \\ =& \frac{2}{\pi^2} \int_2^X \frac{(1 - \cos(x \log p))}{p} K_M(\log p) d\pi(p), \\ =& \frac{2}{\pi^2} \int_2^X \frac{(1 - \cos(x \log p))}{p} K_M(\log p) \frac{1}{\log p} dp \\ &+ \frac{2}{\pi^2} \int_2^X \frac{(1 - \cos(x \log p))}{p} K_M(\log p) dE(p), \end{aligned}$$

$$=I_1 + I_2. \tag{4.2.34}$$

Analysing the main integral I_1 first

$$\begin{aligned} I_1 &= \frac{2}{\pi^2}\int_{\log 2}^{\log X} \frac{1-\cos xu}{u} K_M(u)du, \\ &= \frac{2}{\pi^2}\int_{\log 2}^{\log X} \frac{1-\cos xu}{u} du + \frac{2}{\pi^2}\sum_{j=1}^{M-1}\frac{M-j}{M}\int_{\log 2}^{\log X}\frac{1-\cos xu}{u}\cos(ju)du, \\ &= 2\overline{\sigma}^2 + \frac{2}{\pi^2}\sum_{j=1}^{M-1}\left(1-\frac{j}{M}\right) I_{(j,x)}. \end{aligned} \tag{4.2.35}$$

Expanding out each $I_{(j,x)}$

$$\begin{aligned} I_{(j,x)} &= \int_{\log 2}^{\log X} \frac{1-\cos xu}{u}\cos(ju)du, \\ &= \int_{\log 2}^{\log X} \frac{1}{u}\left(\cos(ju) - \frac{1}{2}(\cos(x+j)u) + \cos((j-x)u)\right) du. \end{aligned}$$

The cosine integral has the following definition

$$\mathrm{Ci}(t) := -\int_t^{\infty} \frac{\cos u}{u} du. \tag{4.2.36}$$

Since $x \neq k$, we can make a change of variables and extend the sum to upper bound $I_{(k,x)}$ by

$$I_{(j,x)} \ll -\mathrm{Ci}(j\log 2) + \frac{1}{2}\left(\mathrm{Ci}((j-x)\log 2) + \mathrm{Ci}((j+x)\log 2)\right).$$

Using the auxiliary functions representation of the cosine integral $\mathrm{Ci}(t)$, there exists the following asymptotic expansion from [**1**],

$$\begin{aligned} \mathrm{Ci}(t) =& \frac{\sin t}{t}\left(1 - \frac{2!}{t^2} + \frac{4!}{t^4} - \dots\right) - \frac{\cos t}{t}\left(\frac{1}{t} - \frac{3!}{t^3} + \frac{5!}{t^5} + \dots\right), \\ \ll& \frac{1}{t}, \end{aligned}$$

as $t \to \infty$. Therefore, applying to $I_{(j,x)}$ gives

$$\begin{aligned} I_{(j,x)} \ll& \frac{1}{j} + \frac{1}{j-x} + \frac{1}{j+x}, \\ \ll& \frac{1}{j}. \end{aligned}$$

Substituting the above asymptotic of $I_{(j,x)}$ back into the sum (4.2.35)

$$\begin{aligned} \sum_{j=1}^{M-1}\left(1-\frac{j}{M}\right) I_{(j,x)} &\ll \sum_{j=1}^{M-1}\left(1-\frac{j}{M}\right)\frac{1}{j}, \\ &\ll \log M, \end{aligned} \tag{4.2.37}$$

as $M \to \infty$.

Now we will deal with the error integral I_2. Using the error function from (4.1.3) and the upper bound for the Fejér kernel (4.2.8)

$$\begin{aligned} I_2 =& \frac{2}{\pi^2}\int_2^X \frac{(1-\cos(x\log p))}{p} K_M(\log p) dE(p), \\ &\ll \int_2^X \frac{1}{p} K_M(\log p) dE(p) \quad (4.2.38) \\ &= \int_{\log 2}^{\log X} \frac{1}{e^y} K_M(y) dE(e^y). \quad (4.2.39) \end{aligned}$$

Starting with simple integration by parts

$$\begin{aligned} I_2 &= e^{-t}K_M(t)E(e^t)\Big|_{\log 2}^{\log X} - \int_{\log 2}^{\log X} (e^{-t}K_M(t))' E(e^t) dt, \\ &= e^{-t}K_M(t)E(e^t)\Big|_{\log 2}^{\log X} - \int_{\log 2}^{\log X} -e^{-t}K_M(t)E(e^t) dt \\ &\quad - \int_{\log 2}^{\log X} e^{-t}K_M'(t)E(e^t) dt, \\ &= I_{2.1} - I_{2.2} - I_{2.3}. \quad (4.2.40) \end{aligned}$$

We can now analyse each integral separately.

$$\begin{aligned} I_{2.1} &= e^{-t}K_M(t)E(e^t)\Big|_{\log 2}^{\log X} \\ &\ll M(e^{-\log X + \log X - c\sqrt{X}} + O(1)) \\ &\ll M. \quad (4.2.41) \end{aligned}$$

as $X \to \infty$.

The second integral

$$\begin{aligned} I_{2.2} &= \int_{\log 2}^{\log X} -e^{-t}e^t e^{-c\sqrt{t}} K_M(t) dt \\ &\ll M \int_{\log 2}^{\log X} e^{-c\sqrt{t}} dt \\ &\ll M. \quad (4.2.42) \end{aligned}$$

Let $f(y) = e^{-c\sqrt{t}}$, then

$$\int_{\log 2}^{\log X} |f(y)| dy \ll 1.$$

Therefore we can apply Lemma 4.2.8 to get

$$I_{2.3} \ll M,$$

and subsequently due to (4.2.41) and (4.2.42)

$$I_2 \ll M.$$

Putting together with (4.2.35) and (4.2.37), gives

$$F_1(X) = 2\overline{\sigma}^2 + O(M).$$

Since $\overline{\sigma}^2 \to \infty$ as $X \to \infty$, $F_1(X) \to \infty$ as $X \to \infty$.

Now, by the bound on the Fejér kernel (4.2.8),

$$|a(p)|^2 \ll M,$$

and the assumptions in the Lemma,

$$F_2(X) \ll M^2 \ll (\overline{\sigma}^2)^{\delta-2}.$$

Therefore $F_2(X)$ is suitably bounded and, we can therefore apply Lemma 4.2.5 to get our result. □

4.2.4. Detruncation of smooth moments. We detruncate by using the asymptotic expansion of $S(t)$ from (4.2.5).

Lemma 4.2.10. *Let $0 < x < 1$ and $0 < a < 1$. For $T^a \leq H \leq T^2$, $T^{a/k} < X < H^{1/k}$ and $1 \leq M = M(T)$,*

$$\left\langle \left| \frac{S_M(x,t) - \overline{S}(x,t,X)}{M} \right|^{2k} \right\rangle_{T,H} = O(M^{2k}),$$

as $T \to \infty$.

As a Corollary we can say that the variance is the same asymptotically to the truncated version

$$\sigma^2 = \overline{\sigma}^2 + O(1). \tag{4.2.43}$$

Proof. First note that by (4.2.5) and (4.2.1) we have

$$\begin{aligned}\frac{S_M(x,t)}{M} &= \frac{1}{\pi}\Im \sum_{n<X} \frac{\Lambda_X(n)}{n^{\sigma_1}} \frac{\sum_{k=1}^M \left(n^{-i(t+k+x)} - n^{-i(t+k)}\right)}{M \log n} \\ &\quad + O\left(\frac{1}{\log\sqrt{X}} \left| \sum_{n<X} \frac{\Lambda_X(n)}{n^{\sigma_1}} \frac{1}{M} \sum_{k=1}^M \left(n^{-i(t+k+x)} - n^{-i(t+k)}\right) \right| \right) \\ &\quad + O\left(\frac{1}{M} \frac{\sum_{k=1}^M (\log(t+k+x) - \log(t+k))}{\log\sqrt{X}}\right).\end{aligned}$$

To make things a little easier let

$$G_M(n,x) := \frac{1}{M}(n^{-ix} - 1) \sum_{k=1}^M n^{-ik}. \tag{4.2.44}$$

We can then reduce to

$$\frac{S_M(x,t)}{M} = \frac{1}{\pi}\Im \sum_{n<X} \frac{\Lambda_X(n)}{n^{\sigma_1+it}} \frac{1}{\log n} G_M(n,x) + O\left(\frac{1}{\log\sqrt{X}}\left|\sum_{n<X}\frac{\Lambda_X(n)}{n^{\sigma_1+it}}G_M(n,x)\right|\right) + O\left(\frac{1}{\log\sqrt{X}}\log\prod_{k=1}^{M}\left(1-\frac{x}{t+k}\right)\right). \quad (4.2.45)$$

By combining (4.2.45) and (4.2.3)

$$\left|\frac{S_M(x,t)-\overline{S}_M(x,t,X)}{M}\right| = O\left(\left|\sum_{p<X}\frac{\Lambda(p)-\Lambda_X(p)p^{1/2-\sigma_1}}{\sqrt{p}\log p}\frac{G_M(p,x)}{p^{it}}\right|\right) + O\left(\frac{1}{\log\sqrt{X}}\left|\sum_{p<X}\frac{\Lambda_X(p)}{p^{\sigma_1}}\frac{G_M(p,x)}{p^{it}}\right|\right) + O\left(\left|\sum_{p^2<X}\frac{\Lambda_X(p^2)}{p^{2\sigma_1}\log p}\left(\frac{G_M(p,x)}{p^{it}}\right)^2\right|\right) + O\left(\frac{1}{\log\sqrt{X}}\left|\sum_{p^2<X}\frac{\Lambda(p^2)}{p^{2\sigma_1}}\left(\frac{G_M(p,x)}{p^{it}}\right)^2\right|\right) + O(1).$$

Therefore

$$\left\langle\left|\frac{S_M(x,t)-\overline{S}_M(x,t,X)}{M}\right|^{2k}\right\rangle_{T,H} = O\left(\left\langle\left|\sum_{p<X}\frac{\Lambda(p)-\Lambda_X(p)p^{1/2-\sigma_1}}{\sqrt{p}\log p}\frac{G_M(p,x)}{p^{it}}\right|^{2k}\right\rangle_{T,H}\right) + O\left(\left\langle\frac{1}{\log\sqrt{X}}\left|\sum_{p<X}\frac{\Lambda_X(p)}{p^{\sigma_1}}\frac{G_M(p,x)}{p^{it}}\right|^{2k}\right\rangle_{T,H}\right) + O\left(\left\langle\left|\sum_{p^2<X}\frac{\Lambda_X(p^2)}{p^{2\sigma_1}\log p}\left(\frac{G_M(p,x)}{p^{it}}\right)^2\right|^{2k}\right\rangle_{T,H}\right) + O\left(\left\langle\frac{1}{\log\sqrt{X}}\left|\sum_{p^2<X}\frac{\Lambda_X(p^2)}{p^{2\sigma_1}}\left(\frac{G_M(p,x)}{p^{it}}\right)^2\right|^{2k}\right\rangle_{T,H}\right) + O(1).$$

(4.2.46)

Define the terms in the sums $a_1(p), a_2(p), a_3(p)$, and $a_4(p)$ by

$$\left\langle \left| \frac{S_M(x,t) - \overline{S}_M(x,t,X)}{M} \right|^{2k} \right\rangle_{T,H} =$$
$$= O\left(\left\langle \left| \sum_{p<X} \frac{a_1(p)}{p^{1/2+it}} \right|^{2k} \right\rangle_{T,H} \right) + O\left(\left\langle \left| \sum_{p<X} \frac{a_2(p)}{p^{1/2+it}} \right|^{2k} \right\rangle_{T,H} \right)$$
$$+ O\left(\left\langle \left| \sum_{p^2<X} \frac{a_3(p)}{p^{1+2it}} \right|^{2k} \right\rangle_{T,H} \right) + O\left(\left\langle \left| \sum_{p^2<X} \frac{a_4(p)}{p^{1+2it}} \right|^{2k} \right\rangle_{T,H} \right). \quad (4.2.47)$$

By the bound (4.2.8), $|G_M(p,x)|$ is bounded,

$$\begin{aligned} |G_M(p,x)|^2 &= 2(1-\cos\log p)K_M(\log p), \\ &< 4M, \end{aligned} \quad (4.2.48)$$

where $K_M(t)$ is the Fejér kernel from (4.2.6). Therefore

$$\begin{aligned} a_1(p) <& \quad \sqrt{M}\frac{(\Lambda(p) - \Lambda_X(p))}{p^{\frac{2}{\log X}}\log p} &<& \quad \sqrt{M}\frac{\log p}{\log\sqrt{X}}, \\ a_2(p) <& \quad \frac{\sqrt{M}}{\log\sqrt{X}}\frac{\Lambda_X(p)}{p^{\frac{2}{\log X}}} &<& \quad \sqrt{M}\frac{\log p}{\log\sqrt{X}}. \end{aligned}$$

And the second two $a(p)$

$$\begin{aligned} a_3(p) <& \quad M\frac{\Lambda_X(p^2)}{p^{\frac{4}{\log X}}\log p} &<& \quad M, \\ a_4(p) <& \quad M\frac{\Lambda_X(p^2)}{\log\sqrt{X}p^{\frac{4}{\log X}}} &<& \quad M. \end{aligned}$$

We can see that the functions a_1, a_2, a_3 and a_4 satisfy the conditions of Lemma 4.2.6. It follows that

$$\left\langle \left| \frac{S_M(x,t) - \overline{S}(x,t,X)}{M} \right|^{2k} \right\rangle_{T,H} = O(M^{2k}).$$

□

THEOREM 4.2.11. *Let $0 < x < 1$ and $0 < a < 1$. For $T^a \leq H \leq T^2$, $T^{a/k} < X < H^{1/k}$ and $1 \leq M \leq \sigma^{1-\epsilon}$ for some $\epsilon > 0$,*

$$\left\langle \left| \frac{S_M(x,t)}{M} \right|^{2k} \right\rangle_{T,H} = \frac{(2k)!}{2^{2k}k!}(2\sigma^2 + O(M))^k + O\left(M(2\sigma^2 + O(M))^{k-\frac{1}{2}} \right),$$

as $T \to \infty$.

PROOF. First note that

$$\begin{aligned}
\left|\frac{S_M(x,t)}{M}\right|^{2k} &= \left|\frac{\overline{S}_M(x,t,X) + S_M(x,t) - \overline{S}_M(x,t,X)}{M}\right|^{2k}, \\
&= \left|\frac{\overline{S}_M(x,t,X)}{M}\right|^{2k} \\
&\quad - 2k\left(\frac{S_M(x,t) - \overline{S}_M(x,t,X)}{M}\right)\left(\frac{\overline{S}_M(x,t,X)}{M}\right)^{2k-1} \\
&\quad + O\left(\left|\frac{S_M(x,t) - \overline{S}(x,t,X)}{M}\right|^{2k}\right) \\
&\quad + O\left(\left(\frac{S_M(x,t) - \overline{S}_M(x,t,X)}{M}\right)^2\left(\frac{\overline{S}_M(x,t,X)}{M}\right)^{2k-2}\right), \\
&= B_1 - B_2 + B_3 + B_4.
\end{aligned}$$

By using the well known Hölder's inequality, Lemma 4.2.10 and Theorem 4.2.9

$$\begin{aligned}
&\langle B_2\rangle_{T,H} \\
&\ll 2k\left\langle\left|\frac{S_M(x,t) - \overline{S}_M(x,t,X)}{M}\right|^{2k}\right\rangle_{T,H}^{\frac{1}{2k}}\left\langle\left|\frac{\overline{S}_M(x,t,X)}{M}\right|^{2k}\right\rangle_{T,H}^{1-\frac{1}{2k}}, \\
&\ll M(2\sigma^2 + O(M))^{k-\frac{1}{2}}.
\end{aligned} \tag{4.2.49}$$

By using Hölder's inequality, Lemma 4.2.10 and Theorem 4.2.9

$$\begin{aligned}
&\langle B_4\rangle_{T,H} \\
&\ll \left\langle\left|\frac{S_M(x,t) - \overline{S}_M(x,t,X)}{M}\right|^{2k}\right\rangle_{T,H}^{\frac{1}{k}}\left\langle\left|\frac{\overline{S}_M(x,t,X)}{M}\right|^{2k}\right\rangle_{T,H}^{1-\frac{1}{k}}, \\
&\ll M^2(2\sigma^2 + O(M))^{k-1}.
\end{aligned}$$

By using Lemma 4.2.10

$$\langle B_3\rangle_{T,H} \ll M^{2k}.$$

Combining the asymptotic bounds of B_1, B_2, B_3 and B_4 gives you the required result. □

4.2.5. Central limit theorem of $S_M(x,T)$.

Lemma 4.2.12. *Let $0 \le a \le 1$. For $T^a \le H \le T^2$, $1 \le M \le \sigma^{1-\epsilon}$ for $\epsilon > 0$ and any interval $\mathcal{A}$,*

$$\mathbb{P}_{\omega,T,H}\left\{\frac{S_M(x,t)}{\sigma M} \in \mathcal{A}\right\} \to \frac{1}{\sqrt{2\pi}}\int_{\mathcal{A}} e^{-x^2/2}dx,$$

as $T \to \infty$, where

$$\sigma^2 = \frac{1}{\pi^2}\log\log T + O(1).$$

Proof. We need to calculate the mean, variance and moments of $S_M(x,t)/(\sigma M)$. Firstly the mean, from Lemma 4.2.4

$$\left\langle \frac{1}{n^{it}} \right\rangle_{T,H} = 0,$$

for all $n \neq 1$. We can then apply to (4.2.45)

$$\left\langle \frac{S_M(x,t)}{\sigma M} \right\rangle_{T,H} = 0 + O\left(\frac{1}{\log\sqrt{X}}\sum_{k=1}^{M}\left\langle \log\left(1-\frac{x}{t+k}\right)\right\rangle_{T,H}\right).$$

Using the Taylor expansion of $\log(1-x)$ again, we get

$$\left\langle \log\left(1-\frac{x}{t+k}\right)\right\rangle_{T,H} \ll \frac{x}{k+T}.$$

Therefore

$$\left\langle \frac{S_M(x,t)}{\sigma M} \right\rangle_{T,H} \to 0 \tag{4.2.50}$$

as $T \to \infty$.

Lemma 4.2.10 tells us the variance. Theorem 4.2.9 tells us the even moments of $S_M(x,t)/(\sigma M)$ are those of the standard normal distribution.

The odd moments are

$$\begin{aligned}\left\langle \left|\frac{S_M(x,t)}{\sigma M}\right|^{2k+1} \right\rangle_{T,H} &\ll \left\langle \left|\frac{S_M(x,t)}{\sigma M}\right|^{2k} \right\rangle_{T,H} \left\langle \frac{S_M(x,t)}{\sigma M} \right\rangle_{T,H} \\ &\ll \quad 1 \cdot \frac{x}{k+T} \to 0.\end{aligned}$$

as $T \to \infty$ by Theorem 4.2.9 and (4.2.50).

We see this is sufficient to prove that the distribution of $S_M(x,t)/(\sigma M)$ weakly converges as $T \to \infty$ to a Gaussian with mean zero and variance 1. □

Finally we use an argument from [**52**] to prove our main Theorem.

Proof of Theorem 4.2.1. Fix $\epsilon > 0$, and approximate the indicator function $\mathbb{1}_{[0,1]}$ above and below by smooth functions $\chi_\pm \ge 0$ so that $\chi_- \le \mathbb{1}_{[0,1]} \le \chi_+$, where both $\chi_\pm$ and their Fourier transforms are smooth and of rapid decay, and so their total masses are within ϵ of unity: $|\int \chi_\pm(x)dx - 1| <$

ϵ. Now set $\omega_\pm := \chi_\pm / \int \chi_\pm$. Then $\omega_\pm$ are "admissable" and for all t,

$$(1-\epsilon)\omega_-(t) \le \mathbb{1}_{[0,1]}(t) \le (1+\epsilon)\omega_+(t) \tag{4.2.51}$$

for all t.

Now

$$\frac{1}{H}\left|\left\{t \in [T, T+H] : \frac{S_M(x,t)}{\sigma M} \in \mathcal{A}\right\}\right|_1 = \frac{1}{H}\int_{-\infty}^{\infty} \mathbb{1}_{\mathcal{A}}\left(\frac{S_M(x,t)}{\sigma M}\right) \mathbb{1}_{[0,1]}\left(\frac{t-T}{H}\right) dt$$

since (4.2.51) holds we find

$$(1-\epsilon)\mathbb{P}_{\omega_-,T,H}\left\{\frac{S_M(x,t)}{\sigma M} \in \mathcal{A}\right\} \le \frac{1}{H}\left|\left\{t \in [T, T+H] : \frac{S_M(x,t)}{\sigma M} \in \mathcal{A}\right\}\right|_1 \le (1+\epsilon)\mathbb{P}_{\omega_+,T,H}\left\{\frac{S_M(x,t)}{\sigma M} \in \mathcal{A}\right\}.$$

By Lemma 4.2.12 the two extreme sides of the inequality have a limit as $T \to \infty$, of

$$(1 \pm \epsilon)\frac{1}{\sqrt{2\pi}}\int_{\mathcal{A}} e^{-x^2/2} dx$$

and so

$$(1-\epsilon)\frac{1}{\sqrt{2\pi}}\int_{\mathcal{A}} e^{-x^2/2} dx \le \liminf_{T\to\infty} \frac{1}{H}\left|\left\{t \in [T, T+H] : \frac{S_M(x,t)}{\sigma M} \in \mathcal{A}\right\}\right|_1$$

with a similar statement for lim sup, since ϵ is arbituary this shows that the limit exists and is as required. □

CHAPTER 5

Distribution of sequences inspired by Champernowne's number

5.1. Background of Uniform Distribution and Normal numbers

5.1.1. Uniform Distribution mod 1. Let $\lfloor x \rfloor$ denote the greatest integer less than or equal to a real number x. Let $\{x\}$ denote the fractional part of x, that is $\{x\} = x - \lfloor x \rfloor$. In 1914-16 Weyl produced some of the founding papers [**97**, **98**] of the theory of uniform distribution.

DEFINITION 5.1.1. Given a sequence of real numbers $(x_n)_{n\in\mathbb{N}}$, define the associated counting function for an interval $[a, b) \subseteq [0, 1)$ by

$$A([a, b), N; (x_n)_{n\in\mathbb{N}}) = \#\{x_n : \{x_n\} \in [a, b) \text{ and } n \leq N\}.$$

The sequence $(x_n)_{n\in\mathbb{N}}$ is *uniformly distributed mod* 1 (u.d. mod 1.) if and only if for all intervals $[a, b) \subseteq [0, 1)$

$$\lim_{N\to\infty} \frac{A([a, b), N; (x_n)_{n\in\mathbb{N}})}{N} = b - a.$$

Note that some authors use the term *equidistributed* instead of uniformly distributed.

In 1916 [**98**], Weyl introduced the following equivalent statement to u.d. mod 1, which is a fundamental result in the area.

THEOREM 5.1.1 (Weyl's Criterion 1916). *A sequence of real numbers* $(x_n)_{n\in\mathbb{N}}$ *is u.d. mod* 1 *if and only if*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} e^{2\pi i h x_n} = 0, \qquad \text{for all integers } h \neq 0.$$

Here, we display a simple example of the use of Weyl's criterion.

EXAMPLE 5.1.1. Weyl's criterion can be used to show that for any irrational α the sequence $(\{n\alpha\})_{n\in\mathbb{N}}$ is u.d. mod 1.

For each $N \in \mathbb{N}$ and integer $h \neq 0$, we have

$$\left| \sum_{n=1}^{N} e^{2\pi i h n\alpha} \right| \leq \left| \frac{e^{2\pi i h N\alpha} - 1}{e^{2\pi i h\alpha} - 1} \right| \leq \left| \frac{2}{e^{2\pi i h\alpha} - 1} \right|. \tag{5.1.1}$$

Note that the denominator of the right hand side of (5.1.1) is a non zero constant for any $h \neq 0$ as long as α is a irrational. Thus,

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

So, by Weyl's criterion the sequence $(\{n\alpha\})_{n\in\mathbb{N}}$ is u.d. mod 1.

In his paper [**98**], Weyl proved a much stronger statement for all polynomials of $n\alpha$.

Theorem 5.1.2 (Weyl 1916). *Let $f : \mathbb{R} \to \mathbb{R}$ be a polynomial with expansion*

$$f(x) = c_d x^d + \cdots + c_0.$$

If at least one the coefficients c_i is irrational then the sequence $(f(n))_{n\in\mathbb{N}}$ is uniformly distributed mod 1.

It is an obvious consequence from the definition, that any u.d. mod 1 sequence of real numbers $(x_n)_{n\in\mathbb{N}}$ is dense in the unit interval $\mathbb{I}$. As an illustrating example see the following sequence which consists of all of the positive rationals in the interval $(0, 1]$.

Example 5.1.2. Let $(x_n)_{n\in\mathbb{N}}$ be the sequence defined by

$$(x_n)_{n\in\mathbb{N}} = \frac{1}{2}, 1, \frac{1}{3}, \frac{2}{3}, 1, \frac{1}{4}, \frac{2}{4}, \frac{3}{4}, \ldots$$

To see that the sequence is u.d. mod 1, first note that for any $\epsilon > 0$ there exists $N \in \mathbb{N}$ and $Q \in \mathbb{N}$ such that

$$\sum_{n=1}^{N} e^{2\pi i h x_n} = \sum_{q=1}^{Q} \sum_{p=1}^{q} e^{2\pi i h \frac{p}{q}} + O(Q) = \sum_{q=1}^{Q} 0 + O(Q),$$

and $Q/N < \epsilon$. Therefore,

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

for all $h \neq 0$. By Weyl's criterion it follows that the sequence is u.d. mod 1.

Note that, although all u.d. mod 1 sequences are dense, the opposite implication does not hold, as we will see in a later section.

5.1.2. Normal Numbers. The concept of normal numbers was introduced by Borel in 1909 in his substantial contribution to the field of probability theory [**18**]. In essence, a real number α is said to be normal to base $b \geq 2$ if the frequencies of strings of digits in the b-ary expansion are as would be expected if the digits were completely random.

More explicitly: For an integer $b \geq 2$, the b-ary expansion of a real number x is

$$x = \lfloor x \rfloor + \sum_{k=1}^{\infty} \frac{a_k}{b^k} = \lfloor x \rfloor + 0.a_1 a_2 \ldots$$

where the digits $a_1, a_2, \ldots,$ are integers from $\{0, 1, \ldots, b-1\}$ and infinitely many of the a_k are not equal to $b-1$.

DEFINITION 5.1.2. Let $b \geq 2$ be an integer. The *frequency* of a digit d in the b-ary expansion of a real number x is equal to the limit of the sequence

$$\left(\frac{\#\{j : 1 \leq j \leq N, a_j = d\}}{N} \right)_{N \in \mathbb{N}},$$

if this sequence converges and does not exist otherwise. A real number x is called *simply normal* to base b if every digit $0, 1, \ldots, b-1$ occurs in its b-ary expansion with the same frequency $\frac{1}{b}$. That is, if

$$\lim_{N \to \infty} \frac{\#\{j : 1 \leq j \leq N, a_j = d\}}{N} = \frac{1}{b}, \quad \text{for } d = 0, 1, \ldots, b-1.$$

It is called *normal* to base b if for all $k \in \mathbb{N}$, all sequences of k digits appear with the frequency $\frac{1}{b^k}$. That is, if

$$\lim_{N \to \infty} \frac{\#\{j : 1 \leq j \leq N, a_j a_{j+1} \ldots a_{j+k} = d_1 d_2 \ldots d_k\}}{N} = \frac{1}{b^k},$$

for $d_1 d_2 \ldots d_k \in \{0, 1, \ldots, b-1\}^k$.

Borel used the word "normal" to describe these numbers due the following result from his paper [**18**].

THEOREM 5.1.3 (Borel 1909). *Almost all real numbers are normal to all integer bases.*

5.1.2.1. *Constructed normal numbers.* Despite that almost all numbers are normal, no explicit examples were found until 1933 when Champernowne presented his famous paper [**26**].

THEOREM 5.1.4 (Champernowne 1933). *The real number*

$$x_c := 0.12345678910111121314\ldots,$$

whose decimal expansion is the increasing sequence of all positive integers, is normal to base ten.

The number x_c is often called Champernowne's number. As well as proving the normality of x_c Champernowne, also highlights a few other very natural constructions of decimals which turn out to be normal. For example the number

$$0.46891012141516182021\ldots \tag{5.1.2}$$

formed by concatenating all of the composite numbers in ascending order. For $k \in \mathbb{N}$, the numbers

$$0 \cdot \lfloor k \rfloor \lfloor 2k \rfloor \lfloor 3k \rfloor \lfloor 4k \rfloor \lfloor 5k \rfloor \ldots, \tag{5.1.3}$$

are normal to the base ten. We remark that Champernowne does not provide an explicit proof of the normality of (5.1.3) (or indeed of the normality of (5.1.2)) in [**26**]. He also conjectured that the number

$$0.3571113171923\ldots \tag{5.1.4}$$

composed of the digits of the increasing sequence of prime numbers is normal to the base ten.

It was not until much later that the numbers (5.1.2), (5.1.3) and (5.1.4) were shown to be normal by Copeland and Erdős [**31**]. In fact, they proved the following much more general result.

THEOREM 5.1.5 (Copeland and Erdős 1945). *If $a_1, a_2, \ldots$ is an increasing sequence of integers such that for every $\theta > 1$*

$$\#\{a_i : a_i \leq N\} > N^{\theta}$$

provided N is sufficiently large, then the infinite decimal

$$0.a_1 a_2 a_3 \ldots$$

is normal with respect to the base b in which these integers a_i are expressed.

Copeland and Erdős' result is based on the more general concept of normality introduced by Besicovitch [**17**] in 1935.

DEFINITION 5.1.3. A number x expressed in base b is said to be (ϵ, k)-normal if any combination of k digits appears consecutively among the digits of x with a relative frequency between $b^{-k} - \epsilon$ and $b^{-k} + \epsilon$.

Besicovitch [**17**], used this definition to show that the number

$$0.149162536496481\ldots$$

consisting of the concatenation of all the integer squares is normal. Erdős also worked with Davenport [**34**] to build upon Besicovistch's work to a general polynomial.

THEOREM 5.1.6 (Davenport and Erdős 1951). *For a polynomial $f : \mathbb{N} \to \mathbb{N}$ that produces natural numbers, the number*

$$0.f(1)f(2)f(3)\ldots$$

is normal to base ten.

In Wall's thesis [**96**], he showed that there is an direct link between normal numbers and u.d. mod 1 sequences.

THEOREM 5.1.7 (Wall 1949). *Let $b \geq 2$ be an integer. A real number x is a normal number to base b if and only if the sequence $(b^n x)_{n\in\mathbb{N}}$ is uniformly distributed mod 1.*

Wall's result can then be used to contstruct the more general definition of normality found in Bugeaud's book [**22**].

DEFINITION 5.1.4. Let β be a real number with $|\beta| > 1$. The real number x is said to be *normal* to base β if the sequence $(\beta^n x)_{n\in\mathbb{N}}$ is uniformly distributed mod 1.

This definition leads to the following extension of Theorem 5.1.3, which can be found in Bugeaud's book [**22**].

THEOREM 5.1.8. *Let β be a real number with $|\beta| > 1$. Then almost all real numbers x are normal to base β.*

5.2. Distribution mod 1 over a subinterval

Motivated by Walls' Theorem 5.1.7, we ask the following questions:

Is the sequence

$$x_n = 0.(n)(n+1)(n+2)(n+3)\ldots \qquad \text{for } n \in \mathbb{N} \tag{5.2.1}$$

u.d. mod 1? Note that this sequence differs from $(\{10^n x_c\})_{n\in\mathbb{N}}$, for example the 20th term of the sequence $(\{10^n x_c\})_{n\in\mathbb{N}}$ would be $0.516171819\ldots$ whereas the 20th term of $(x_n)_{n\in\mathbb{N}}$ is

$$x_{20} = 0.202122232425\ldots.$$

Since the sequence only contains numbers in the interval $[0.1, 1)$, the sequence is not u.d. mod 1. Therefore we ask, is this sequence u.d. mod 1 over a suitable subinterval? This idea will be made precise in the following subsection.

More generally, given an increasing sequence of integers $(a_n)_{n\in\mathbb{N}}$, we attempt to answer the question of what is the *distribution* of sequences of the following form?

$$x_n = 0.a_n a_{n+1} a_{n+2}\ldots \qquad \text{for } n \in \mathbb{N}$$

For instance, for a real number k, we consider how is the following sequence distributed?

$$x_n = 0.(kn)(k(n+1))(k(n+2))\ldots \qquad \text{for } n \in \mathbb{N} \tag{5.2.2}$$

Inspired by Theorem 5.1.6, we consider, for a given polynomial f with real coefficients. How is the following sequence distributed?

$$x_n = 0.\lfloor f(n)\rfloor\lfloor f(n+1)\rfloor\lfloor f(n+2)\rfloor\dots \qquad \text{for } n \in \mathbb{N} \tag{5.2.3}$$

5.2.1. Notation and Definitions. Before stating our first result, we introduce some necessary terminology and notation which will be used throughout.

Returning to the sequence $(x_n)_{n\in\mathbb{N}}$ from (5.2.1), we note that although

$$\{x_n : n \in \mathbb{N}\} \cap [0, 0.1) = \emptyset$$

it is dense in the interval $[0.1, 1)$. So, we consider u.d. mod 1 over such a subinterval of $[0, 1)$ using the following definition.

DEFINITION 5.2.1. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence of real numbers. If for any pair of real numbers $\alpha \le a < b \le \beta$

$$\lim_{N\to\infty} \frac{A([a,b), N; (x_n)_{n\in\mathbb{N}})}{N} = \frac{b-a}{\beta-\alpha}. \tag{5.2.4}$$

then we define the sequence to be *uniformly distributed mod* 1 *over* $[\alpha, \beta) \subseteq [0, 1)$ (u.d. mod 1 over $[\alpha, \beta)$),

It follows directly from this definition that a sequence $(x_n)_{n\in\mathbb{N}}$ of real numbers, is u.d. mod 1 over $[\alpha, \beta)$ if and only if

$$\lim_{N\to\infty} \frac{A([\alpha,c), N; (x_n)_{n\in\mathbb{N}})}{N} = \frac{c-\alpha}{\beta-\alpha}, \text{ for each } \alpha \le c \le \beta.$$

We will also be interested in the weaker notion of being almost uniformly distributed mod 1. Thus, inspired by [**61**, Definition 7.2] we make the following definition.

DEFINITION 5.2.2. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence of real numbers. If there exists an increasing sequence of natural number $N_1, N_2, \dots$ such that

$$\lim_{i\to\infty} \frac{A([\alpha,x), N_i; (x_n)_{n\in\mathbb{N}})}{N_i} = \frac{x-\alpha}{\beta-\alpha} \qquad \text{for } \alpha \le x \le \beta, \tag{5.2.5}$$

then define the sequence $(x_n)_{n\in\mathbb{N}}$ to be *almost uniformly distributed mod* 1 *over* $[\alpha, \beta)$ (we shall abbreviate this as a.u.d. mod 1 over $[\alpha, \beta)$).

REMARK 5.2.1. It is easily seen from definition 5.2.1, that if $(x_n)_{n\in\mathbb{N}}$ is u.d. mod 1 over $[\alpha, \beta)$ then it is a.u.d. mod 1 over $[\alpha, \beta)$.

More generally we will be interested in the distribution mod 1 of a sequence. To this end we introduce the analogues of upper and lower distribution functions mod 1 given by [**61**, Definition 7.1].

DEFINITION 5.2.3. The *upper and lower distribution functions* (u.d.f and l.d.f) over $[\alpha, \beta)$, $\Phi : [\alpha, \beta) \mapsto [\alpha, \beta)$ and $\phi : [\alpha, \beta) \mapsto [\alpha, \beta)$ respectively of a sequence of real numbers $(x_n)_{n\in\mathbb{N}}$ are defined by

$$\phi(x) := \liminf_{N\to\infty} \frac{A([\alpha, x), N; (x_n)_{n\in\mathbb{N}})}{N} \qquad \text{for } \alpha \leq x \leq \beta,$$

$$\Phi(x) := \limsup_{N\to\infty} \frac{A([\alpha, x), N; (x_n)_{n\in\mathbb{N}})}{N} \qquad \text{for } \alpha \leq x \leq \beta.$$

Note that the functions ϕ and Φ are non-decreasing, with $\phi(\alpha) = \Phi(\alpha) = 0$ and $\phi(\beta) = \Phi(\beta) = 1$, while $0 \leq \phi(x) \leq \Phi(x) \leq 1$ for $\alpha \leq x \leq \beta$.

REMARK 5.2.2. A sequence $(x_n)_{n\in\mathbb{N}}$ is u.d. mod 1 over $[\alpha, \beta)$ if and only if $\phi(x) = \Phi(x) = (x - \alpha)/(\beta - \alpha)$.

5.2.2. The distribution of sequences inspired by Champernowne's number. Our first result is inspired by the normality of Champernowne's number and Wall's Theorem 5.1.7.

THEOREM 5.2.1. *The sequence* $(x_n)_{n\in\mathbb{N}}$ *of real numbers from* (5.2.1)*, is not u.d. mod* 1 *over* $[0.1, 1)$.

One can prove this quite easily by observing that, for any natural number J, upon reaching the term x_{10^J} the next 10^J terms in the sequence will begin with a 1 immediately after the decimal point. That is; for each $J \in \mathbb{N}$ at least half of the terms up to the term $x_{2\times 10^J}$ begin with a first decimal digit 1. More precisely, for $J \in \mathbb{N}$;

$$\frac{A([0.1, 0.2), 2 \times 10^J; (x_n)_{n\in\mathbb{N}})}{2 \times 10^J} = \frac{\#\{x_n \in [0.1, 0.2) : n \leq 2 \times 10^J\}}{2 \times 10^J} \geq \frac{1}{2}.$$

Comparing this with Definition 5.2.1 the result of Theorem 5.2.1 follows. The point is that there are too many terms of the sequence defined at (5.2.1) in the interval $[0.1, 0.2)$ infinitely often.

Despite the fact that this sequence is not u.d. mod 1 over $[0.1, 1)$, it turns out that it is still a.u.d. mod 1 over $[0.1, 1)$. This and Theorem 5.2.1 are easy corollaries of our next result.

THEOREM 5.2.2. *Consider a fixed real number k. The sequence* (5.2.2) *when expressed in base $b > 2$ is a.u.d. mod* 1 *over* $[b^{-1}, 1)$*, but is not u.d. mod* 1 *over* $[b^{-1}, 1)$.

In fact Theorems 5.2.1 and 5.2.2 are the corollaries of the following more general result.

THEOREM 5.2.3. *Let $f:\mathbb{N}\to\mathbb{R}^+$ be a polynomial of degree $d\geq 1$ with real coefficients, then the sequence* (5.2.3) *expressed in base $b>2$ is not u.d. mod* 1 *over* $[b^{-1},1)$ *and is a.u.d. mod* 1 *over* $[b^{-1},1)$ *if and only if* $d=1$.

The methods used to prove such a theorem additionally provides the u.d.f and l.d.f mod 1 over $[b^{-1},1)$.

THEOREM 5.2.4. *Let $f:\mathbb{N}\to\mathbb{R}^+$ be a polynomial of degree $d\geq 1$ with real coefficients, the l.d.f. and u.d.f. mod* 1 *over* $[b^{-1},1)$ *of the sequence* (5.2.3) *expressed in base $b>2$, are respectively*

$$\phi(t)=\frac{t^{1/d}-b^{-1/d}}{1-b^{-1/d}},\qquad \Phi(t)=\frac{t^{1/d}-b^{-1/d}}{t^{1/d}(1-b^{-1/d})},\qquad \textit{for } b^{-1}\leq t\leq 1.$$

5.2.3. Supporting Statements. In this section we gather some auxiliary statements which we will require for the proof of Theorem 5.2.4. First of all we require the following observation on the asymptotic size of the inverse function of polynomials.

Lemma 5.2.5. *Let $f:\mathbb{R}\to\mathbb{R}^+$ be a polynomial with real coefficients given by*

$$f(n)=c_dn^d+c_{d-1}n^{d-1}+\cdots+c_1n+c_0. \tag{5.2.6}$$

Let $g(m)$ be the inverse of f (i.e. $f(g(m))=m$). Then $g(m)=m^{1/d}c_d^{-1/d}+O(1)$ as $m\to\infty$.

PROOF. First, substitute $n=m^{1/d}c_d^{-1/d}+\varepsilon$ into (5.2.6) to get

$$f(m^{1/d}c_d^{-1/d}+\varepsilon)=c_d(m^{1/d}c_d^{-1/d}+\varepsilon)^d+c_{d-1}(m^{1/d}c_d^{-1/d}+\varepsilon)^{d-1}+\cdots+c_0.$$

Then, a combination of the Binomial theorem and Taylor expansions establishes that

$$\varepsilon=-\frac{c_{d-1}}{dc_d^{1-2/d}}+O(m^{-2/d}).$$

□

The following Lemma is the count of numbers x_n in the interval $[b^{-1},t)$ from the sequence (5.2.3).

Lemma 5.2.6. *Let $f:\mathbb{R}\to\mathbb{R}^+$ be a polynomial*

$$f(n)=c_dn^d+c_{d-1}n^{d-1}+\cdots+c_1n+c_0,$$

let $(x_n)_{n\in\mathbb{N}}$ *be the sequence* (5.2.3) *expressed in base* $b > 2$. *For all* $\alpha, t \in [b^{-1}, 1)$,

$$A([b^{-1},t), \lfloor f^{-1}(\alpha b^J)\rfloor; (x_n)_{n\in\mathbb{N}})$$
$$= \begin{cases} \dfrac{1}{c_d^{1/d}}\left((\alpha^{1/d} - b^{-1/d})b^{J/d} + (t^{1/d} - b^{-1/d})\displaystyle\sum_{i=1}^{J-1} b^{i/d} + O(J)\right) & \text{if } \alpha < t, \\ \dfrac{1}{c_d^{1/d}}\left((t^{1/d} - b^{-1/d})\displaystyle\sum_{i=1}^{J} b^{i/d} + O(J)\right) & \text{if } \alpha \geq t, \end{cases}$$

as $J \to \infty$.

The principal idea of the argument is that the leading digits of the α are the leading digits of the terms of the sequence $x_{\lfloor f^{-1}(\alpha b^J)\rfloor}$.

PROOF. First note that $x_n \in [b^{-1}, t)$ whenever $b^K \leq f(n) < tb^{K+1}$ for some K sufficiently large. By Lemma 5.2.5 we have $f^{-1}(n) = c_d^{-1/d} n^{1/d} + O(1)$. Also note that, for sufficiently large n, $f(n)$ is increasing.

We will consider the cases $\alpha < t$ and $\alpha \geq t$ separately. In the case that $\alpha < t$,

$$\begin{aligned} &A([b^{-1},t), \lfloor f^{-1}(\alpha b^J)\rfloor; (x_n)_{n\in\mathbb{N}}) \\ &\quad =(f^{-1}(\alpha b^J) - f^{-1}(b^{J-1}) + O(1)) + (f^{-1}(tb^{J-1}) - f^{-1}(b^{J-2}) + O(1)) + \dots \\ &\qquad\quad + (f^{-1}(tb) - f^{-1}(b^0) + O(1)) \\ &\quad =(f^{-1}(\alpha b^J) - f^{-1}(b^{J-1})) + \sum_{i=1}^{J-1}(f^{-1}(tb^i) - f^{-1}(b^{i-1})) + O(J). \end{aligned}$$

Where the $O(1)$ terms and subsequently the $O(J)$ terms come from the error in approximation of the floor function.

Using Lemma 5.2.5.

$$\begin{aligned} &A([b^{-1},t), \lfloor f^{-1}(\alpha b^J)\rfloor; (x_n)_{n\in\mathbb{N}}) \\ &\quad =\frac{\alpha^{1/d} b^{J/d}}{c_d^{1/d}} - \frac{b^{(J-1)/d}}{c_d^{1/d}} \\ &\qquad\quad + \sum_{i=1}^{J-1}\left(\left(\frac{tb^i}{c_d}\right)^{1/d} + O(1) - \left(\frac{b^{i-1}}{c_d}\right)^{1/d} - O(1)\right) + O(J), \\ &\quad =\frac{1}{c_d^{1/d}}\left((\alpha^{1/d} - b^{-1/d})b^{J/d} + (t^{1/d} - b^{-1/d})\sum_{i=1}^{J-1} b^{i/d}\right) + O(J), \end{aligned}$$

as required.

Consider the case of $\alpha \geq t$. When $\alpha > t$, $x_n \in [t, 1)$ for all n such that $tb^J \leq f(n) < b^J$ and $x \notin [b^{-1}, t)$ for all n such that $tb^J \leq f(n) < \alpha b^J$. Therefore, if we let $\alpha = t$ and substitute into the case of $\alpha < t$, the result follows. □

We can then use Lemma 5.2.6 to prove the following Lemma on the frequency of elements $x_n \in [b^{-1}, t)$ of subsequences $(x_{\lfloor f^{-1}(\alpha b^J)\rfloor})_{J \in \mathbb{N}}$.

Lemma 5.2.7. *Let $f : \mathbb{N} \to \mathbb{R}^+$ be a polynomial of degree $d \geq 1$ with real coefficients, let $(x_n)_{n \in \mathbb{N}}$ be the sequence* (5.2.3) *expressed in base $b > 2$. For all $\alpha, t \in [b^{-1}, 1)$,*

$$\lim_{J\to\infty} \frac{A([b^{-1}, t), \lfloor f^{-1}(\alpha b^J)\rfloor; (x_n)_{n\in\mathbb{N}})}{\lfloor f^{-1}(\alpha b^J)\rfloor} = \begin{cases} \frac{\alpha^{1/d} + b^{-1/d}(t^{1/d} - \alpha^{1/d} - 1)}{\alpha^{1/d}(1 - b^{-1/d})} & \text{if } \alpha < t, \\ \frac{t^{1/d} - b^{-1/d}}{\alpha^{1/d}(1 - b^{-1/d})} & \text{if } \alpha \geq t. \end{cases}$$

Proof. Let c_d be the coefficient of n^d in the polynomial $f(n)$. Using Lemma 5.2.6 for $J \in \mathbb{N}$ and $\alpha < t$,

$$\lim_{J\to\infty} \frac{A([b^{-1}, t), \lfloor f^{-1}(\alpha b^J)\rfloor, (x_n)_{n\in\mathbb{N}})}{\lfloor f^{-1}(\alpha b^J)\rfloor} = \lim_{J\to\infty} \frac{b^{J/d}\left(\frac{\alpha^{1/d} - b^{-1/d}}{c_d^{1/d}}\right) + \left(\frac{t^{1/d} - b^{-1/d}}{c_d^{1/d}}\right)\sum_{i=1}^{J-1}(b^{i/d}) + O(J)}{\left(\frac{\alpha b^J}{c_d}\right)^{1/d} + O(1)}. \quad (5.2.7)$$

By dividing through by the $b^{J/d}$ and taking the limit of terms that tend to 0.

$$\lim_{J\to\infty} \frac{A([b^{-1}, t), \lfloor f^{-1}(\alpha b^J)\rfloor; (x_n)_{n\in\mathbb{N}})}{\lfloor f^{-1}(\alpha b^J)\rfloor} = \lim_{J\to\infty}\left(\frac{\alpha^{1/d} - b^{-1/d}}{\alpha^{1/d}} + \frac{(t^{1/d} - b^{-1/d})(\sum_{i=1}^{J-1} b^{-i/d}) + O(J/b^{J/d})}{\alpha^{1/d}}\right).$$

Using the formula for an infinite geometric series, and again taking the limit of terms that tend to 0.

$$\begin{aligned}
&\lim_{J\to\infty}\frac{A([b^{-1},t),\lfloor f^{-1}(\alpha b^J)\rfloor,(x_n)_{n\in\mathbb{N}})}{\lfloor f^{-1}(\alpha b^J)\rfloor}\\
=&\frac{\alpha^{1/d}-b^{-1/d}}{\alpha^{1/d}}+\frac{t^{1/d}-b^{-1/d}}{\alpha^{1/d}}\left(\frac{1}{1-b^{-1/d}}-1\right),\\
=&\frac{1}{\alpha^{1/d}}\left(\alpha^{1/d}-b^{-1/d}+(t^{1/d}-b^{-1/d})\frac{b^{-1/d}}{(1-b^{-1/d})}\right),\\
=&\frac{1}{\alpha^{1/d}(1-b^{-1/d})}\left((\alpha^{1/d}-b^{-1/d})(1-b^{-1/d})+(t^{1/d}-b^{-1/d})b^{-1/d}\right).
\end{aligned}$$

After some more rearranging we have the result.

By a similar argument, using Lemma 5.2.6, for $\alpha \geq t$ it follows that:

$$\begin{aligned}
&\lim_{J\to\infty}\frac{A([b^{-1},t),\lfloor f^{-1}(\alpha b^J)\rfloor,(x_n)_{n\in\mathbb{N}})}{\lfloor f^{-1}(\alpha b^J)\rfloor}\\
&=\lim_{J\to\infty}\frac{\left(t^{1/d}-b^{-1/d}\right)\sum_{i=0}^{J-1}(b^{-i/d})+O(J/b^{J/d})}{\alpha^{1/d}},\\
&=\frac{t^{1/d}-b^{-1/d}}{\alpha^{1/d}(1-b^{-1/d})}. \qquad (5.2.8)
\end{aligned}$$

□

5.3. Proof of Main Theorems

Using the previous Lemma 5.2.7, we can calculate the u.d.f and l.d.f mod 1 over $[b^{-1},1)$ for sequences (5.2.3) to prove Theorem 5.2.4.

PROOF OF THEOREM 5.2.4. Let $X_J(\alpha,t)$ be defined by

$$X_J(\alpha,t):=\frac{A([b^{-1},t),\lfloor f^{-1}(\alpha b^J)\rfloor,(x_n)_{n\in\mathbb{N}})}{\lfloor f^{-1}(\alpha b^J)\rfloor}, \qquad (5.3.1)$$

We will, find

$$\inf_{\alpha\in[b^{-1},1)}\lim_{J\to\infty}X_J(\alpha,t) \quad \text{and} \quad \sup_{\alpha\in[b^{-1},1)}\lim_{J\to\infty}X_J(\alpha,t),$$

then we show that they are the relevant u.d.f and l.d.f.

Suppose $\alpha < t$. Let $\alpha^{1/d} = t^{1/d} - \epsilon$ for $0 \le \epsilon < t^{1/d} - b^{-1/d}$. Substitute into Lemma 5.2.7

$$\lim_{J\to\infty} X_J(\alpha, t) = \frac{(t^{1/d} - \epsilon + b^{-1/d}(\epsilon - 1))}{(t^{1/d} - \epsilon)(1 - b^{-1/d})}, \tag{5.3.2}$$
$$= \frac{(t^{1/d} - \epsilon(1 - b^{-1/d}) - b^{-1/d})}{(t^{1/d} - \epsilon)(1 - b^{-1/d})},$$
$$= \frac{1}{1 - b^{-1/d}} - \frac{b^{-1/d}}{1 - b^{-1/d}} \frac{1 - \epsilon}{t^{1/d} - \epsilon}.$$

Let $\psi(\epsilon) = \frac{1-\epsilon}{t^{1/d}-\epsilon}$, since

$$\frac{d\psi}{d\epsilon} = \frac{1 - t^{1/d}}{(t^{1/d} - \epsilon)^2} > 0$$

The function $\psi(\epsilon)$ is increasing over $\epsilon \in [0, t^{1/d} - b^{-1/d}]$, so has its minimum when $\epsilon = 0$, and maximum when $\epsilon = t^{1/d} - b^{-1/d}$. Equivalently, the maximum of the limit (5.3.2) is reached at $\alpha = t$ and minimum when $\alpha = b^{-1}$.

Now consider $\alpha \ge t$. Let $\alpha^{1/d} = t^{1/d} + \epsilon$ again for $0 < \epsilon$,

$$\lim_{J\to\infty} X_J(\alpha, t) = \frac{t^{1/d} - b^{-1/d}}{(t^{1/d} + \epsilon)(1 - b^{-1/d})}.$$

This has its maximum when $\epsilon = 0$, and minimum when $\epsilon = b^{-1/d} - t^{1/d}$. Equivalently, maximum is reached at $\alpha = t$ and minimum when $\alpha = 1$.

Therefore

$$\sup_{\alpha\in[b^{-1},1)} \lim_{J\to\infty} X_J(\alpha, t) = \lim_{J\to\infty} X_J(t, t) \tag{5.3.3}$$
$$= \frac{t^{1/d} - b^{-1/d}}{t^{1/d}(1 - b^{-1/d})},$$

and

$$\inf_{\alpha\in[b^{-1},1)} \lim_{J\to\infty} X_J(\alpha, t) = \lim_{J\to\infty} X_J(b^{-1}, t), \tag{5.3.4}$$
$$= \frac{t^{1/d} - b^{-1/d}}{1 - b^{-1/d}}.$$

To show that the quantities (5.3.4) and (5.3.3) are the upper and lower distribution functions, first let

$$X_N(t) := \frac{A([b^{-1}, t), N, (x_n)_{n\in\mathbb{N}})}{N}. \tag{5.3.5}$$

It then follows that $\phi(t) = \liminf_{N\to\infty} X_N(t)$ and $X_{\lfloor f^{-1}(\alpha b^J)\rfloor}(t) = X(\alpha, t)$. Therefore

$$\begin{aligned}\phi(t) = \liminf_{N\to\infty} X_N(t) &= \lim_{N\to\infty}\left\{\inf_{m\geq N} X_m(t)\right\},\\ &= \lim_{N\to\infty}\left\{\inf_{\alpha\in[b^{-1},1)}\inf_{J\geq J(N)} X_J(\alpha,t)\right\},\end{aligned}$$

where $J(N)$ is the biggest integer such that $N \geq b^{J-1}$. We have already found the infimum over α, so

$$\begin{aligned}\phi(t) &= \lim_{N\to\infty}\left\{\inf_{J\geq J(N)} X_J(b^{-1},t)\right\},\\ &= \lim_{J\to\infty} X_J(b^{-1},t).\end{aligned}$$

A similar argument can be used for the limsup of (5.3.5). □

Proof of Theorem 5.2.3. By Lemma 5.2.7 if we let $\alpha = b^{-1}$ and $d = 1$ then

$$\lim_{J\to\infty}\frac{A([b^{-1},t),\lfloor f^{-1}(b^J)\rfloor,(x_n)_{n\in\mathbb{N}})}{\lfloor f^{-1}(b^J)\rfloor} = \frac{t-b^{-1}}{1-b^{-1}}.$$

Therefore the subsequence with index $N_J = \lfloor f^{-1}(b^J)\rfloor$ satisfies 5.2.5 and the sequence is a.u.d. mod 1.

Note that $\phi(x) \neq \Phi(x)$ for any d, therefore we can apply Remark 5.2.2 to show the seqeunces are not u.d. mod 1 over $[b^{-1}, 1)$.

Now for the opposite part of the implication, notice that for all subsequences N_i the u.d.f and l.d.f bound the limit

$$\phi(t) \leq \lim_{i\to\infty}\frac{A([\alpha,t),N_i;(x_n)_{n\in\mathbb{N}})}{N_i} \leq \Phi(t).$$

Therefore, by Theorem 5.2.4

$$\frac{t^{1/d}-b^{-1/d}}{1-b^{-1/d}} \leq \lim_{i\to\infty}\frac{A([\alpha,t),N_i;(x_n)_{n\in\mathbb{N}})}{N_i} \leq \frac{t^{1/d}-b^{-1/d}}{t^{1/d}(1-b^{-1/d})}.$$

We claim for all $d \geq 1$ the l.d.f. is greater than or equal to the LHS of (5.2.5) for all $t \in [b^{-1}, 1)$, with equality being reached when $d = 1$. That is:

$$\phi(t) = \frac{t^{1/d}-b^{-1/d}}{1-b^{-1/d}} \geq \frac{t-b^{-1}}{1-b^{-1}}.$$

To prove the claim, first consider for a fixed $t \in [b^{-1}, 1)$ and $b > 2$ the function

$$\psi(x) = \frac{t^x - b^{-x}}{1-b^{-x}}.$$

The claim is equivalent to $\psi(x)$ is increasing as $x \to 0$. That is again equivalent to $\psi(x)$ is decreasing as $x \to \infty$. Since $t^x - b^{-x}$ is decreasing to 0 as $x \to \infty$ and $1 - b^{-x}$ is increasing to 1, $\psi(x)$ is decreasing as $x \to \infty$ and the claim is true.

Therefore, no subsequence exists to satisfy (5.2.5) and the result is proven. $\square$

Bibliography

[1] M. Abramowitz, I. A Stegun, et al., *Handbook of mathematical functions*, Applied mathematics series **55** (1966), no. 62, 39.

[2] F. Adiceam, V. Beresnevich, J. Levesley, S. L. Velani, and E. Zorin, *Diophantine approximation and applications in interference alignment*, Advances in Mathematics **302** (2016), 231 –279.

[3] K. Ball, *Cube slicing in* $\mathbb{R}^n$, Proceedings of the American Mathematical Society (1986), 465–473.

[4] A. F. Beardon and B. Maskit, *Limit points of Kleinian groups and finite sided fundamental polyhedra*, Acta Mathematica **132** (1974), no. 1, 1–12.

[5] V. Beresnevich, *Application of the concept of regular systems of points in metric number theory*, Vestsī Nats. Akad. Navuk Belarusī Ser. Fīz.-Mat. Navuk **1** (2000), 35–39, 140.

[6] V. Beresnevich, *A Groshev type theorem for convergence on manifolds*, Acta Mathematica Hungarica **94** (2002), no. 1-2, 99–130.

[7] ———, *Rational points near manifolds and metric Diophantine approximation*, Annals of Mathematics **175** (2012), no. 1, 187–235.

[8] V. Beresnevich, D. Dickinson, and S. L. Velani, *Measure theoretic laws for lim sup sets*, Vol. 179, American Mathematical Soc., 2006.

[9] V. Beresnevich, A. Ghosh, D. Simmons, and S. L. Velani, *Diophantine approximation in Kleinian groups: singular, extremal and bad limit points* (2016), to appear.

[10] V. Beresnevich, A. Haynes, and S. L. Velani, *Sums of reciprocals of fractional parts and multiplicative Diophantine approximation*, arXiv preprint arXiv:1511.06862 (2015).

[11] V. Beresnevich, F. Ramírez, and S. L. Velani, *Metric Diophantine approximation: aspects of recent work*, Dynamics and analytic number theory, 2016, pp. 1–96. arXiv:1601.01948.

[12] V. Beresnevich, R. C Vaughan, S. Velani, and E. Zorin, *Diophantine approximation on manifolds and the distribution of rational points: contributions to the convergence theory*, International Mathematics Research Notices **2017** (2016), no. 10, 2885–2908.

[13] V. Beresnevich, R. C. Vaughan, S. L. Velani, and E. Zorin, *Diophantine approximation on manifolds and the distribution of rational points: contributions to the divergence theory* (2016).

[14] V. Beresnevich and S. L. Velani, *Classical metric Diophantine approximation revisited: the Khintchine–Groshev theorem*, International Mathematics Research Notices **2010** (2010), no. 1, 69–86.

[15] V. Beresnevich and E. Zorin, *Explicit bounds for rational points near planar curves and metric Diophantine approximation*, Advances in Mathematics **225** (2010), no. 6, 3064–3087.

[16] V. Beresnevish and S. L. Velani, *Ubiquity and a general logarithm law for geodesics*, Dynamical systems and Diophantine Approximation **20** (2009), 21–36.

[17] A. S. Besicovitch, *The asymptotic distribution of the numerals in the decimal representation of the squares of the natural numbers*, Mathematische Zeitschrift **39** (1935), no. 1, 146–156.

[18] É. Borel, *Les probabilités dénombrables et leurs applications arithmétiques*, Rendiconti del Circolo Matematico di Palermo (1884-1940) **27** (1909), no. 1, 247–271.

[19] R. P. Brent, A. J. van der Poorten, and H. J. J. te Riele, *A comparative study of algorithms for computing continued fractions of algebraic numbers*, Algorithmic number theory (Talence, 1996), 1996, pp. 35–47.

[20] S. Brewer, *Source code for computing and analysing the contnued fraction of algebric numbers.* (2017), available at `http://github.com/jacomago/continuedFractions`.

[21] Y. Bugeaud, *Approximation by algebraic numbers*, Vol. 160, Cambridge University Press, 2004.

[22] ———, *Distribution modulo one and Diophantine approximation*, Vol. 193, Cambridge University Press, 2012.

[23] H. Bui, J. Conrey, and M. P Young, *More than 41% of the zeros of the zeta function are on the critical line*, Acta Arithmetica (2010).

[24] F. P. Cantelli, *Sulla determinazione empirica delle leggi di probabilita*, Giorn. Ist. Ital. Attuari **4** (1933), no. 421-424.

[25] J. W. S. Cassels, *Some metrical theorems in Diophantine approximation*, Mathematical Proceedings of the Cambridge Philosophical Society, 1951, pp. 18–21.

[26] D. G. Champernowne, *The Construction of Decimals Normal in the Scale of Ten*, J. London Math. Soc. **S1-8** (1933), no. 4, 254.

[27] Y. Cheung, *Hausdorff dimension of the set of singular pairs*, Annals of mathematics (2011), 127–167.

[28] Y. Cheung and N. Chevallier, *Hausdorff dimension of singular vectors*, preprint (2014).

[29] K. L. Chung and P. Erdos, *On the application of the Borel-Cantelli lemma*, Transactions of the American Mathematical Society **72** (1952), no. 1, 179–186.

[30] R. F. Churchhouse and S. T. E. Muir, *Continued fractions, algebraic numbers and modular invariants*, J. Inst. Math. Appl. **5** (1969), 318–328.

[31] A. H. Copeland and P. Erdős, *Note on normal numbers*, Bull. Amer. Math. Soc. **52** (1946), 857–860.

[32] T. Das, L. Fishman, D. Simmons, and M. Urbański, *Extremality and dynamically defined measures, part I: Diophantine properties of quasi-decaying measures*, arXiv preprint arXiv:1504:04778 (2015).

[33] ———, *Extremality and dynamically defined measures, part II: Measures from conformal dynamical systems*, arXiv preprint arXiv:1508.05592 (2015).

[34] H. Davenport and P. Erdős, *Note on normal decimals*, Canadian J. Math. **4** (1952), 58–63.

[35] H. Davenport, *Multiplicative number theory*, Third, Graduate Texts in Mathematics, vol. 74, Springer-Verlag, New York, 2000. Revised and with a preface by Hugh L. Montgomery.

[36] H. Davenport and W. Schmidt, *Dirichlet's theorem on Diophantine approximation. II*, Acta Arithmetica **4** (1970), no. 16, 413–424.

[37] C.-J. De la Vallée Poussin, *Recherches analytiques sur la théorie des nombres (première partie)*, Ann. Soc. Sci. Bruxelles [I] **20** (1896), 183–256.

[38] H. G. Diamond and J. D. Vaaler, *Estimates for partial sums of continued fraction partial quotients*, Pacific J. Math. **122** (1986), no. 1, 73–82.

[39] S. Feng, *Zeros of the riemann zeta function on the critical line*, Journal of Number Theory **132** (2012), no. 4, 511–542.

[40] K. Ford, K. Soundararajan, and A. Zaharescu, *On the distribution of imaginary parts of zeros of the Riemann zeta function. II*, Math. Ann. **343** (2009), no. 3, 487–505.

[41] K. Ford and A. Zaharescu, *On the distribution of imaginary parts of zeros of the Riemann zeta function*, J. Reine Angew. Math. **579** (2005), 145–158.

[42] A. Fujii, *On the zeros of Dirichlet L-functions. I*, Transactions of the American Mathematical Society **196** (1974), 225–235.

[43] ———, *On the distribution of the zeros of the Riemann zeta function in short intervals*, Bull. Amer. Math. Soc. **81** (1975), 139–142.

[44] ———, *On the zeros of Dirichlet L-functions. III*, Trans. Amer. Math. Soc. **219** (1976), 347–349.

[45] P. Gallagher, *Metric simultaneous Diophantine approximation (II)*, Mathematika **12** (1965), no. 02, 123–127.

[46] S. M. Gonek, *An explicit formula of Landau and its applications to the theory of the zeta-function*, A tribute to Emil Grosswald: number theory and related analysis, 1993, pp. 395–413.

[47] J. Hadamard, *Sur la distribution des zéros de la fonction $\zeta(s)$ et ses conséquences arithmétiques*, Bulletin de la Société Mathématique de France **24** (1896), 199–220 (fre).

[48] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers, 6th edn, revised by D.R. Heath-Brown and J.H. Silverman*, Oxford University Press, Oxford, 2008.

[49] G. H. Hardy, *Sur les zéros de la fonction $\zeta(s)$ de Riemann*, CR Acad. Sci. Paris **158** (1914), 1012–1014.

[50] E. Hlawka, *Über die Gleichverteilung gewisser Folgen, welche mit den Nullstellen der Zetafunktion zusammenhängen*, Österreich. Akad. Wiss. Math.-Naturwiss. Kl. S.-B. II **184** (1975), no. 8-10, 459–471.

[51] C. Hughes and Z. Rudnick, *Linear statistics for zeros of Riemann's zeta function*, Comptes Rendus Mathematique **335** (2002), no. 8, 667–670.

[52] ———, *On the distribution of lattice points in thin annuli*, International Mathematics Research Notices **2004** (2004), no. 13, 637–658.

[53] A. Hurwitz, *Über die angenäherte darstellung der irrationalzahlen durch rationale brüche*, Mathematische Annalen **39** (1891), no. 2, 279–284.

[54] M. Hussain and J. Levesley, *The metrical theory of simultaneously small linear forms*, Functiones et Approximatio Commentarii Mathematici **48** (2013), no. 2, 167–181.

[55] A. Khintchine, *Einige Sätze über Kettenbrüche, mit Anwendungen auf die theorie der Diophantischen Approximationen*, Mathematische Annalen **92** (1924), no. 1, 115–125.

[56] ———, *Über eine Klasse linearer Diophantischer Approximationen*, Rendiconti del Circolo Matematico di Palermo (1884-1940) **50** (1926), no. 2, 170–195.

[57] ———, *Continued fractions*, Russian, Dover Publications Inc., Mineola, NY, 1997. With a preface by B. V. Gnedenko, Reprint of the 1964 translation.

[58] D. Y. Kleinbock, E. Lindenstrauss, and B. Weiss, *On fractal measures and Diophantine approximation*, Selecta Mathematica **10** (2005), no. 4, 479–523.

[59] D. Y. Kleinbock and G. A. Margulis, *Flows on homogeneous spaces and Diophantine approximation on manifolds*, Annals of mathematics (1998), 339–360.

[60] S. Kristensen, R. Thorn, and S. L. Velani, *Diophantine approximation and badly approximable sets*, Advances in Mathematics **203** (2006), no. 1, 132–169.

[61] L. Kuipers and H. Niederreiter, *Uniform distribution of sequences*, Wiley-Interscience [John Wiley & Sons], New York, 1974. Pure and Applied Mathematics.

[62] R. Kuzmin, *Ob odnoi zadache gaussa*, Doklady akad. nauk, ser a, 1928, pp. 375–380.

[63] E. Landau, *Über die Nullstellen der Zetafunktion*, Mathematische Annalen **71** (December 1912), no. 4, 548–564 (de).

[64] S. Lang and H. Trotter, *Continued fractions for some algebraic numbers*, Introduction to Diophantine Approximations (1972), 93.

[65] D. H. Lehmer, *Euclid's algorithm for large numbers*, American Mathematical Monthly (1938), 227–233.

[66] J. Liouville, *Nouvelle démonstration d'un théoreme sur les irrationnelles algébriques inséré dans le compte rendu de la derniere séance*, CR Acad. Sci. Paris **18** (1844), 910–911.

[67] J. E. Littlewood, *On the zeros of the Riemann zeta-function*, Mathematical Proceedings of the Cambridge Philosophical Society **22** (1924), no. 3, 295–318.

[68] J.-L. Marichal and M. J. Mossinghoff, *Slices, slabs, and sections of the unit hypercube*, Online J. Anal. Comb. **3** (2008), Art. 1, 11.

[69] H. L. Montgomery, *The pair correlation of zeros of the zeta function*, Analytic number theory (Proc. Sympos. Pure Math., Vol. XXIV, St. Louis Univ., St. Louis, Mo., 1972), 1973, pp. 181–193.

[70] A. S. Motahari, S. Oveis-Gharan, M.-A. Maddah-Ali, and A. K. Khandani, *Real interference alignment: Exploiting the potential of single antenna systems*, IEEE Transactions on Information Theory **60** (2014), no. 8, 4799–4810.

[71] B. Nazer and O. Ordentlich, *Diophantine approximation for network information theory: A survey of old and new results*, Communication, control, and computing (allerton), 2016 54th annual allerton conference on, 2016, pp. 990–996.

[72] A. Odlyzko, *On the distribution of spacings between zeros of the zeta function*, Mathematics of Computation **48** (1987), no. 177, 273–308.

[73] S. J. Patterson, *Diophantine approximation in Fuchsian groups*, Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences **282** (1976), no. 1309, 527–563.

[74] ———, *The limit set of a Fuchsian group*, Acta Math. **136** (1976), 241–273.

[75] A. Pollington and S. L. Velani, *Metric Diophantine approximation and 'absolutely friendly' measures*, Selecta Mathematica **11** (2005), no. 2, 297–307.

[76] G. Pólya, *On a few questions in probability theory and some definite integrals related to them*, Ph.D. Thesis, 1912.

[77] H. Rademacher, *Fourier analysis in number theory, symposium on harmonic analysis and related integral transforms: Final technical report*, H. Rademacher, Collected Works, 1956, pp. 434–458.

[78] E. M. Rains, *Images of eigenvalue distributions under power maps*, Probability theory and related fields **125** (2003), no. 4, 522–538.

[79] B. Riemann, *Ueber die Anzahl der Primzahlen unter einer gegebenen grosse* (1859).

[80] K. F. Roth, *Rational approximations to algebraic numbers*, Mathematika **2** (1955), no. 1, 1–20.

[81] W. M. Schmidt, *A metrical theorem in Diophantine approximation*, Canad. J. Math **12** (1960), 619–631.

[82] A. Selberg, *On the zeros of Riemann's zeta-function*, Dybwad in Komm., 1943.

[83] ———, *On the remainder in the formula for $N(T)$, the number of zeros of $\zeta(s)$ in the strip* $0 < t < T$, Collected Works **1** (1944), 179–203.

[84] ———, *The zeta-function and the Riemann hypothesis*, Skandinaviske Matematikerkongres **10** (1946), 187–200.

[85] P. Shiu, *Computation of continued fractions without input values*, Mathematics of computation **64** (1995), no. 211, 1307–1317.

[86] D. Simmons, *Some manifolds of Khinchin type for convergence*, arXiv preprint arXiv:1602.01727 (2016).

[87] H. M. Stark, *A complete determination of the complex quadratic fields of class-number one*, The Michigan Mathematical Journal **14** (1967), no. 1, 1–27.

[88] J. Steuding, *One hundred years uniform distribution modulo one and recent applications to Riemann's zeta-function*, Topics in mathematical analysis and applications, 2014, pp. 659–698.

[89] B. Stratmann, *Diophantine approximation in Kleinian groups*, Mathematical proceedings of the cambridge philosophical society, 1994, pp. 57–78.

[90] B. Stratmann and S. L. Velani, *The Patterson measure for geometrically finite groups with parabolic elements, new and old*, Proceedings of the London Mathematical Society **3** (1995), no. 1, 197–220.

[91] D. Sullivan, *Disjoint spheres, approximation by imaginary quadratic numbers, and the logarithm law for geodesics*, Acta Mathematica **149** (1982), no. 1, 215–237.

[92] ———, *Entropy, Hausdorff measures old and new, and limit sets of geometrically finite Kleinian groups*, Acta Mathematica **153** (1984), no. 1, 259–277.

[93] R. C. Vaughan and S. L. Velani, *Diophantine approximation on planar curves: the convergence theory*, Inventiones mathematicae **166** (2006), no. 1, 103–124.

[94] S. L. Velani, *An application of metric Diophantine approximation in hyperbolic space to quadratic forms*, Publicacions Matemàtiques (1994), 175–185.

[95] H. von Mangoldt, *Zu Riemanns Abhandlung*, Journal für die reine und angewandte Mathematik **114** (1895), 255–305.

[96] D. D. Wall, *Normal numbers*, University of California, Berkeley, 1949.

[97] H. Weyl, *Über ein problem aus dem gebiete der Diophantischen approximationen*, Nachr. Ges. Wiss. Göttingen, Math.-phys. K **1** (1914), 234–244.

[98] ———, *Über die Gleichverteilung von Zahlen mod. Eins*, Math. Ann. **77** (1916), no. 3, 313–352.

[99] C. Zong, *What is known about unit cubes*, Bulletin of the American Mathematical Society **42** (2005), no. 2, 181–211.