Complex analogues of real problems

PhD thesis

Rune Esdahl-Schou — 20052073

project advisor: Simon Kristensen

October 30, 2013

Department of Mathematics
University of Aarhus

CONTENTS

Abstract

This thesis will be a mix of different problems in number theory. As such it is split into two natural parts. The first part focuses on normal numbers and construction of numbers that are normal to a given complex base. It is written in the style of a thorough and introductory paper on that subject. Certain classical theorems are stated without proof but with a reference instead, though usually a proof is given. This part of the thesis represents the pinnacle of the authors work during the first two years of his PhD study. The work presented is greatly inspired by the work of Madritsch, Thuswaldner and Tichy in [Madritsch et al., 2008] and [Madritsch, 2008] and contains a generalisation of the main theorem in [Madritsch, 2008].

The second part of the thesis focuses on Diophantine approximation, mainly on a famous conjecture by Schmidt from the 1980s. This conjecture was solved by Badziahin, Pollington and Velani, and inspired by this An gave a different proof which provides a stronger result. The conjecture is concerned with intersections of certain sets in the plane and are as such a "real problem". We will consider a slightly different setup where the real plane is replaced by the complex plane. Using geometrical interpretations we construct sets with properties similar to the sets considered in the real case. We then formulate a conjecture which can be interpreted as a complex version of Schmidt's original conjecture. Finally we construct a variant of Schmidt's game, to show a partial result leading us to believe that the complex version of Schmidt's conjecture might some day be answered in the affirmative, just as the real one has.

Resumé

Denne afhandling indeholder forskellige matematiske problemer indenfor feltet talteori. Det er af æstetiske årsager opdelt i to dele. Den første del koncentrerer sig udelukkende om normale tal og konstruktioner af tal som er normale til en given kompleks base. Afhandlingen er skrevet i samme stil som en grundlæggende artikel indenfor feltet ville være. Enkelte klassiske læresætninger bliver dog fremhævet uden bevis men med henvisninger til hvor i litteraturen et sådant kan findes, det er dog undtagelsen snarere end reglen. Første del af afhandlingen er essentielt det ypperste stykke arbejde forfatteren fik præsteret i løbet af de første to år af hans Ph.D.-forløb. Meget af det der præsenteres i denne del er inspireret af det store arbejde Madritsch, Thuswaldner og Tichy har lagt i artiklerne [Madritsch et al., 2008] og [Madritsch, 2008] og indeholder en generalisering af hovedsætningen i [Madritsch, 2008].

Den anden del af afhandlingen fokuserer på diofantisk approksimation, mere specifikt på en kendt formodning fremsat af Schmidt omkring 1980. Formodningen blev eftervist af Badziahin, Pollington og Velani, og inspireret af dette gav An et nyt bevis for selv samme formodning der viser et stærkere resultat. Formodningen omtaler fællesmængder af nogle specielle mængder i planen og er derfor et "reelt problem". Vi vil betragte en lidt anderledes opsætning, hvor den reelle plan bliver udskiftet med den komplekse plan. Ved brug af geometriske fortolkninger konstruerer vi mængder med sammenlignelige egenskaber og fremsætter en formodning, der kan fortolkes som en kompleks udgave af Schmidts oprindelige formodning. Til sidst laver vi en variant af Schmidts spil til at vise et mindre resultat der leder os til at tro at det en dag

vil lykkedes at eftervise at den komplekse udgave af Schmidts formodning er sand, ganske som det er lykkedes at gøre med den reelle.

CHAPTER 1

ACKNOWLEDGEMENTS AND INTRODUCTION

The author would like to thank numerous people, especially his fellow students for everyday inspiration, encouragement and providing other motivating factors, without which this thesis would not have seen the light of day. The author would also like to thank his project advisor Simon Kristensen for several helpful comments and suggestions and other types of advice. All images displayed can be found on wikipedia and are licensed under the Creative Commons Attribution-Share Alike 3.0 Unported license, apart from a single figure in chapter 12, which is made by the author with tikz.

Furthermore the author would like to give a special thanks to Sanju Velani and the number theory group at the university of York for a wonderful, yet short visit in the winter of 2012. The visit was an inspiration to the author and the many valuable discussions encouraged the author to consider the problem which is in the second part of this thesis.

The author will from this point on use the word "we" instead of I. This is done for the sole purpose of making the reader feel part of the process that is doing mathematics. In certain key spots the author might use "the author" to underline that this is a witty thought of the author or a remark based solely on the authors opinion. Furthermore the author will try to the utmost of his ability to avoid following the advice given in Sand-Jensen [2007] but might fail miserably.

Before moving on to the task at hand we list a brief introduction to each of the remaining chapters.

Chapters 2 and 3 are quite self explanatory. They list various definitions regarding canonical number systems and provides a basic understanding of number representations which at the end of chapter 3 allows us to state the main theorem. Chapter 4 is also self explanatory, it provides a bunch of Lemmata, most of which are classical, such as the Erdős–Turán–Koksma inequality, and most of which plays important roles later on. In chapter 5 we shift the fundamental domain by a constant, essentially requiring that we only consider numbers which start with a specific string of digits. Afterward we construct a Urysohn function for this shifted domain and finally we calculate the Fourier coefficients of the Urysohn function. In chapter 6 we consider Weyl sums and give estimates of various exponential sums in order to arrive at the key proposition. In chapter 7 we make use of all the corollaries derived in chapter 6 and use these to give the needed estimates to prove the main theorem of part one.

In the second part chapter 8 gives an introduction to Diophantine approximation by providing a few essential theorems. Chapter 9 gives an introduction to Hausdorff measure and dimension, tools often used in Diophantine approximation to describe the size of a given set. In chapter 10 we consider Schmidt games and various variants of these, while giving a short survey on why Schmidt invented them and what he used them for. Chapter 11 contains as the title indicates a survey of various conjectures in Diophantine approximation, among others Schmidt's conjecture. Chapter 12 is dedicated entirely to Schmidt's conjecture and a complex analogue of this. Finally in Chapter 13 we show a partial result leading us to believe that the complex analogue of Schmidt's conjecture might be true.

Part I

Generating normal numbers over quadratic imaginary rings

CHAPTER 2

BASIC DEFINITIONS AND CONCEPTS

Throughout this thesis we let $\mathbb{N} = \{1, 2, 3, \dots\}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ denote the natural numbers, the rational numbers, the real numbers and the complex numbers respectively. Unless otherwise mentioned, every rational p/q will be assumed to be expressed in lowest terms. As it is standard within this subject, we will use the shorthand notation $e(x) = \exp(2\pi i x)$. We shall also make use of both Vinogradov notation and Landau notation; the latter is also known as big-O notation. Letting $f, g : X \to \mathbb{R}$ be two real-valued functions defined on the same set, X, we say that $f = \mathcal{O}(g)$ if $f(x) \leq Cg(x)$ holds for all $x \in X$ and for some constant, C. We will sometimes also write this in the Vinogradov notation, i.e., $f \ll g$ if and only if $f = \mathcal{O}(g)$. In some situations the author finds that one type of notation is superior to the other and vice versa, thus we will use both of them. Later on we shall introduce the ring of integers associated to a quadratic imaginary field with discriminant d and it is customary to denote these rings with the symbols $\mathcal{O}_d$, so in order to avoid confusion the author stresses the reader to note that $\mathcal{O}$ without any index always means big-O notation, while $\mathcal{O}_d$ where there is an index always denotes a ring of integers.

In the first part of this thesis we will be constructing normal numbers over the quadratic imaginary fields for which the associated ring of integers is a Euclidean domain. In the case of the real numbers and number systems over the reals, normal numbers have been studied thoroughly.

If $b \geq 2$ is an integer, it is well-known that every real number x can be expressed in the following way:

$$x = \sum_{j=-\infty}^{k} d_j b^j.$$

Where d_j are integers from the set $D = \{0, 1, \dots, b-1\}$. We call d_j a digit and D the *set of digits* and b the base. We say that x is *simply normal to base* b if every possible digit occurs at the same frequency. i.e. every digit $d \in D$ occurs with probability $1/b$. We say that x is *normal to base* b if x is simply normal to all the bases $b, b^2, b^3, \dots$. Equivalently, this can be understood as every finite block of n consecutive digits occurs with probability $1/b^n$. Finally we say that x is *normal* if it is normal to base b for every $b \geq 2$.

One of the more interesting results is that almost every real number is normal with respect to the Lebesgue measure, a result dating back to Borel in 1909. However, given a real number it is a very difficult question to determine whether that given number is normal or not. In fact it is not even known if π, e or $\log 2$ is normal. Even worse, so far (at least to the authors knowledge) no one has been able to construct a number that is normal to two different (multiplicatively independent) bases at the same time.

There are however some methods to construct numbers that are normal to a given base. One of the first results is due to Champernowne who was able to show that

$$0.1234567891011\ldots$$

is normal to base 10. This number is therefore known as Champernowne's number. Many people have successfully extended this result. Davenport and Erdős extended the idea of Champernowne to the integer part of polynomials over the natural numbers in [Davenport and Erdős, 1952]. Schiffer [Schiffer, 1986] extended it further allowing his polynomials to have rational coefficients, Nakai and Shiokawa [Nakai and Shiokawa, 1992] allowed real coefficients and finally Madritsch, Thuswaldner and Tichy managed to prove that

$$0.\lfloor f(1)\rfloor \lfloor f(2)\rfloor \lfloor f(3)\rfloor \ldots$$

is normal if f is an entire function of bounded logarithmic order and $\lfloor x\rfloor$ denotes the expansion of the integer part of x with respect to a given base.

Madritsch later generalised this construction of normal numbers in a different direction. Namely to number systems for Gaussian integers, with the restriction that the function f is now a polynomial. Our aim in the first part of this thesis is to further generalise the result of Madritsch to number systems for other quadratic imaginary rings which are Euclidean domains.

CHAPTER 3

DEFINITIONS AND BASIC PROPERTIES OF NUMBER SYSTEMS

The Stark–Heegner Theorem states that the rings of integers in the quadratic imaginary fields of the form $\mathbb{Q}(\sqrt{d})$ have unique factorisation if and only if d is one of the so called Heegner numbers, namely $d \in \{-1, -2, -3, -7, -11, -19, -43, -67, -163\}$. Not all of these are Euclidean domains however. In fact only 5 of them are, namely when $d \in \{-1, -2, -3, -7, -11\}$. Since most of the results below are true for all 9 Heegner numbers we prove the most general case when it is possible. Another reason for this is the fact that some of the results will be used in the second part of this thesis, where the property of being a Euclidean domain is not essential as long as we have unique factorization.

The ring of integers form a lattice in the plane very similar to the one shown below. We shall later make use of this to shift the lattice into the standard lattice $\mathbb{Z}^2$.

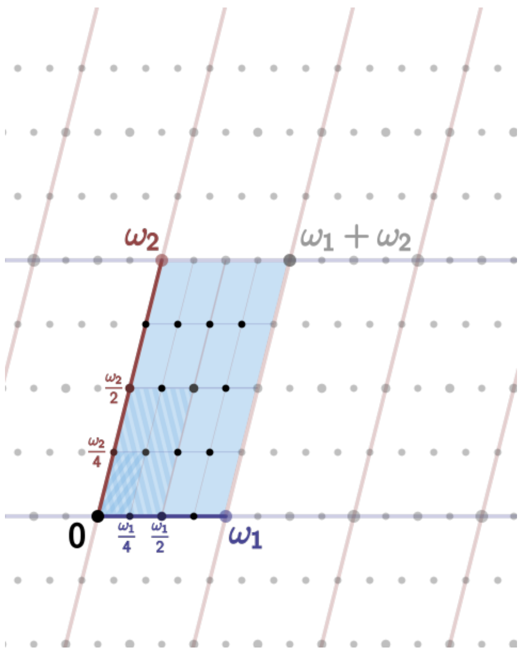

Figure 3.1: A typical lattice formed by the vectors ω_1 and ω_2, embedded in $\mathbb{R}^2$.

For each d we let $\mathcal{O}_d$ denote the ring of integers in $\mathbb{Q}(\sqrt{d})$. Note that the Heegner numbers are square-free and when d is a Heegner number, $\mathcal{O}_d$ has an integral basis consisting of $(1, c)$ where $c = \frac{1+\sqrt{d}}{2}$ if $d \equiv 1 \pmod 4$ and $c = \sqrt{d}$ if $d \equiv 2, 3 \pmod 4$. The image to keep in mind is the one above, where $\omega_1 = 1$ and $\omega_2 = c$. As an example, if $d = -7$ then every element, w of $\mathcal{O}_7$ can be written as $w = x + y\frac{1+i\sqrt{7}}{2}$, where $x, y \in \mathbb{Z}$ since $-7 \equiv 1 \pmod 4$. We shall be using not only the fact that $\mathbb{Q}(\sqrt{d})$ is a field, but also that it is a normed vector space of two real dimensions. However, for different values of d we have different spaces and each have their own norm, which we will make use of. We continue with the example and calculate the norm of w. We denote the norm of w by $N(w)$, in this case it is given by $x^2+xy+2y^2$. These facts can be found in [Lánczi, 1965]. As we shall be using the norm quite extensively we give a full list below of the various norms. Notice however that we will not intermix the different norms in the calculations below, as we will always be working with a single fixed d, as such we simply write $N(w)$ for some $w \in \mathbb{Q}(\sqrt{d})$ and not $N_d(w)$.

$$\begin{aligned}
N(x+iy) &= x^2+y^2, \text{ for } d=-1,\\
N\left(x+iy\sqrt{2}\right) &= x^2+2y^2 \text{ for } d=-2,\\
N\left(x+y\frac{1+i\sqrt{3}}{2}\right) &= x^2+xy+y^2 \text{ for } d=-3,\\
N\left(x+y\frac{1+i\sqrt{7}}{2}\right) &= x^2+xy+2y^2 \text{ for } d=-7,\\
N\left(x+y\frac{1+i\sqrt{11}}{2}\right) &= x^2+xy+3y^2 \text{ for } d=-11,\\
N\left(x+y\frac{1+i\sqrt{19}}{2}\right) &= x^2+xy+5y^2 \text{ for } d=-19,\\
N\left(x+y\frac{1+i\sqrt{43}}{2}\right) &= x^2+xy+11y^2 \text{ for } d=-43,\\
N\left(x+y\frac{1+i\sqrt{67}}{2}\right) &= x^2+xy+17y^2 \text{ for } d=-67,\\
N\left(x+y\frac{1+i\sqrt{163}}{2}\right) &= x^2+xy+41y^2 \text{ for } d=-163,
\end{aligned}$$

From this point on we consider d to be fixed unless otherwise mentioned. Now for $b \in \mathcal{O}_d$ we let $\mathcal{D}$ be a complete set of residue classes modulo b. We call the pair $(b, \mathcal{D})$ a *number system in* $\mathbb{Q}(\sqrt{d})$, sometimes just a *number system* if every $z \in \mathcal{O}_d$ has a unique representation of the form

$$z = \sum_{j=0}^{\infty} d_j(z) b^j,$$

with $d_j(z) \in \mathcal{D}$ and $d_j(z) = 0$ for $j \geq l(z)$, where $l(z)$ is an integer. We call b the *base* and we call $d_j(z)$ the *digits* of z with respect to base b and refer to $\mathcal{D}$ as the *set of digits*. Since the representation is unique $l(z)$ is well defined and we call this the *length* of z in base b. We see that $l(z) = \max\{j : d_{j-1}(z) \neq 0\}$.

As an example we consider the case $d = -1$. Here $\mathcal{O}_d$ is simply the Gaussian integers, which contains the number 3. Following Gilbert [Gilbert, 1987] we see that a possible digit set is $\{0, \pm 1, \pm i, \pm 1 \pm i\}$. We could also use the base $-1+i$ in which case the digit set can be $\{0,1\}$. The first example has a real base, but complex digits, where as the second base has a complex base, but real digits. We choose to focus our attention on the second type of number systems. So if we are in the situation where $\mathcal{D}$ is particularly nice, namely $\mathcal{D}=\{0,1,\ldots,N(b)-1\}$ where $N(b)$ denotes the norm of b, we shall refer to $(b,\mathcal{D})$ as a *canonical number system* or a CNS, as before we sometimes explicitly state the field of which the pair is a CNS, but we usually leave it out if its clear from the context.

Obviously not all $b \in \mathcal{O}_d$ can occur as the base of a CNS. This is seen by the fact that we insist on having real digits, so if the base is real we cannot possibly represent every complex number, but usually this is not the only restriction. The possible bases for the Gaussian integers were found by Kátai and Szabó in [Kátai and Szabó, 1975] who showed the following theorem.

Theorem 3.1. *The pair $(b,\mathcal{D})$ is a CNS in $\mathbb{Q}(i)$ if and only if* $\mathrm{Re}(b)<0, \mathrm{Im}(b)=\pm 1$.

Later the following theorem was shown by Kátai and Kovács in [Kátai and Kovács, 1981]

Theorem 3.2. *Let $d \geq 2, -d \not\equiv 1 \pmod 4$. $(b,\mathcal{D})$ is a CNS in $\mathbb{Q}(\sqrt{-d})$ if and only if*

$$b = A \pm i\sqrt{d}, \quad 0 \leq -2A \leq A^2+d \geq 2, \text{ where } A \text{ is an integer.}$$

Let $d \geq 2, -d \equiv 1 \pmod 4$. $(b,\mathcal{D})$ is a CNS in $\mathbb{Q}(\sqrt{-d})$ if and only if

$$b = 1/2(B \pm i\sqrt{d}), \quad -1 \leq -B \leq 1/4(B^2+d) \geq 2, \text{ where } B \text{ is an odd integer.}$$

In order to motivate the next few steps we return briefly to the real case and consider a rational p/q. It is known that no matter what positive integer $b \geq 2$ we pick, p/q when represented in base b with the digits $\{0,1,\ldots,b-1\}$ will have ultimately periodic digits and thus cannot be normal. However, we can represent every real number in base b. So we can ask whether or not a given real number is normal to a certain base. So far we have introduced the various rings of integers we will be working with and the corresponding quotient fields. But we need something to serve the purpose of the reals. We therefore extend our number system to the complex numbers.

It is well known that for every $z \in \mathbb{C}$ there is a (not neccesarily unique) representation of the shape

$$z = \sum_{j=-\infty}^{l(z)} d_j(z) b^j, \tag{3.1}$$

We give a proof of this based on the idea in [Kátai and Szabó, 1975]

Proof of (3.1). Let z be an arbitrary complex number, $z = x+iy$, where $x,y \in \mathbb{R}$. Let $b^k = U_k + iV_k$, for $k = 1,2,\ldots$. We now have

$$z = \frac{zb^k}{b^k} = \frac{(x+iy)(U_k+iV_k)}{b^k} = \frac{C_k+iD_k}{b^k} + \frac{u_k+iv_k}{b^k},$$

where $C_k + iD_k \in \mathcal{O}_d$ is picked to be the nearest lattice point (might not be unique, if not then pick any of the possible ones). Now setting

$$z_k = \frac{C_k + iD_k}{b^k}, \quad \delta_k = \frac{u_k + iv_k}{b^k}$$

we obtain two sequences $\{z_k\}_{k=1}^{\infty}, \{\delta_k\}_{k=1}^{\infty}$. We see that $|u_k|, |v_k| < |b|$. Hence $\delta_k \to 0$ and thus $z_k \to z$ as $k \to \infty$. Since $C_k + iD_k \in \mathcal{O}_d$ we have a unique representation of the form

$$C_k + iD_k = a_t^* b^t + \cdots + a_o^*, \quad t = t(k)$$

We start by proving that the sequence $t(k) - k$ has an upper bound. Dividing the above equation by b^k we obtain

$$z_k = a_t^* b^{t-k} + \cdots + a_k^* + \cdots + a_0^* b^{-k}.$$

So we get

$$a_t^* b^{t-k} + \cdots + a_k^* = z_k - \frac{a_{k-1}^*}{b} - \cdots - \frac{a_0^*}{b^k},$$

Taking absolute values, applying the triangle inequality and using that the a_i^*'s are non-negative we arrive at

$$\begin{aligned} |a_t^* b^{t-k} + \cdots + a_k^*| &\leq |z_k| + \frac{a_{k-1}^*}{|b|} + \cdots + \frac{a_0^*}{|b|^k} \\ &\leq |z| + |\delta_k| + |b|^2 \left(\frac{1}{|b|} + \frac{1}{|b|^2} + \ldots \right) \\ &\leq |z| + |\delta_k| + |b|^2 \sum_{j=0}^{\infty} \left(\frac{1}{|b|} \right)^j \\ &\leq |z| + |\delta_k| + |b|^2 \frac{|b|}{|b| - 1} \\ &\leq |z| + |\delta_k| + \frac{|b|^3}{|b| - 1}. \end{aligned}$$

From this is follows that

$$|a_t^* b^{t-k} + \cdots + a_k^*| \leq c,$$

where c is a constant depending only on z and b. The fact that $\mathcal{O}_d$ forms a lattice ensures that the circle with center 0 and radius c contains only finitely many integers from $\mathcal{O}_d$. Since each integer has a unique representation, the number $t(k) - k$ must be bounded from above. Now let K be an integer such that $t - k \leq K$. Then we can write z_k as

$$z_k = a_K^{(k)} b^K + \cdots + a_0^{(k)} + \frac{a_{-1}^{(k)}}{b} + \ldots,$$

where $a_j^{(k)} \in \{1, 2, \ldots, N(b) - 1\}$ and $j = K, K-1, \ldots, 0, -1, \ldots$. Let $m_K \in \{1, 2, \ldots, |b| - 1\}$ be an integer such that $a_K^{(k)} = m_K$ for infinitely many k. Let S_K be the subset of those integers k satisfying this. Suppose that $S_K, \ldots, S_{l+1}$ is constructed such that $S_K \supseteq \cdots \supseteq S_{l+1}$. Then there is a $m_l \in \{1, 2, \ldots |b| - 1\}$ such that for infinitely many k in S_{l+1} we have $a_l^{(k)} = m_l$. Let S_l be the set consisting of such k. then S_l has infinitely many elements. We now repeat this argument for $K, K-1, \dots, 0, -1, \dots$. Let

$$w = m_K b^K + \cdots + m_0 + \frac{m_{-1}}{b} + \dots.$$

Let $k_1 < k_2 < \dots$ be an infinite sequence, such that

$$k_v \in S_{K-v+1} \quad v = 1, 2, \dots.$$

Since we have

$$z_{k_v} = m_K b^K + \cdots + m_{K-v+1} b^{K-v+1} + a_{K-v}^{(k_v)} b^{K-v} + \dots,$$

then

$$\lim_{v \to \infty} z_{k_v} = w$$

but as $z_k \to z$ for $k \to \infty$ this means we must have $z = w$. Hence we have a representation of the desired form. □

We denote by $\lfloor z \rfloor$ the integer part of z with respect to base b, where

$$\lfloor z \rfloor = \sum_{j=0}^{l(z)} d_j(z) b^j,$$

and we denote by $\{z\}$ the fractional part of z with respect to base b, where

$$\{z\} = \sum_{j=-\infty}^{-1} d_j(z) b^j.$$

Hence, every number is given by the sum of its integer part and its fractional part. However, if the representation in base b is not unique, the integer part as well as the fractional part is not well defined. This phenomenon also occurs in the real numbers. For instance if we represent every real number in base 10 we see that $1 = 0.9999\dots$ are two different representations of the same number. The first has integer part 1 and fractional part 0 the other has integer part 0 and fractional part $0.9999\dots$. This has nothing to do with base 10, any other integer base will share similar properties. Moreover, there is nothing special about the number 1, either rational will have two representations. The good thing however is, no matter which rational we pick and what integer base we choose to represent it in, the possible representations will either terminate after a finite step or be eventually periodic. Here, terminating after a finite step simply means that we only need finitely many non-zero digits in the fractional part, and eventually periodic means that after a finite step a certain string of digits will repeat itself indefinitely.

Our situation is very much the same, however not everything behaves as nicely as in the real case. Letting $\overline{a}$ denote the string a repeated indefinitely, we can see from [Gilbert, 1981] that the base $b = (-1 + \sqrt{7}i)/2$ gives rise to the following curiosity

$$\frac{-3 - \sqrt{7}i}{8} = 0.\overline{001} = 1.\overline{010} = 11.\overline{100}.$$

Which is a number having, not only two, but in fact three different representations within the same number system. The reason for a number having several representations is due to the fact that it falls inside the intersection of different translates of the fundamental domain. To explain it in more detail we move on to discussing the fundamental domain and return to this issue later.

We define the fundamental domain $\mathcal{F}'$ as the set of all numbers which can be represented with 0 in the integer part of their b-ary expansion. i.e.,

$$\mathcal{F}' = \left\{ \gamma \in \mathbb{C} \mid \gamma = \sum_{k\geq 1} d_k b^{-k}, d_k \in \mathcal{D} \right\}. \tag{3.2}$$

Kátai and Környei in [Kátai and Környei, 1992] showed some general results about translates of the fundamental domain. First of all the proved that $\mathcal{F}'$ is a compact set, and secondly that (adapted to our special case) we have

$$\bigcup_{a\in\mathcal{O}_d} (\mathcal{F}' + a) = \mathbb{C}.$$

That is, the union of translates of the fundamental domain by every integer tile the complex plane. Furthermore they showed that the intersection of such two translations is not too large, but in fact has Lebesgue measure zero. In particular,

$$\lambda((\mathcal{F}' + a) \cap (\mathcal{F}' + c)) = 0, \quad a, c, \in \mathcal{O}_d, a \neq c,$$

where λ denotes the (2-dimensional) Lebesgue measure. Later it was shown by Akiyama and Thuswaldner, that the fundamental domain is also arcwise connected for all two-dimensional number systems, which suffices for our purpose, see [Akiyama and Thuswaldner, 2000] for more details.

Since the overlap between two translates of the fundamental domain is of Lebesgue measure zero, the fractional part of almost every number is unique, hence the integer part of almost every number is unique and we arrive at the conclusion that the representation of almost every number is unique. Now fixing a base, b, we let $d_1 \dots d_l$ be a block of digits of length l and let $\mathcal{N}(z; d_1 \dots d_l; N)$ be the number of occurences of the block $d_1 \dots d_l$ in the first N digits of z. That is

$$\mathcal{N}(z; d_1 \dots d_l; N) = \#\{1 \leq n \leq N : d_1 = d_n(z), \dots, d_l = d_{n+l-1}(z)\}.$$

We call z normal in $(b, \mathcal{D})$ if for every $l \geq 1$ we have

$$\sup_{d_1\dots d_l} \left| \frac{1}{N}\mathcal{N}(z; d_1 \dots d_l; N) - \frac{1}{|\mathcal{D}|^l} \right| = o(1).$$

Here, $o(1)$ means that the expression on the left-hand-side is bounded by 1 for all $N \geq N_0$, where N_0 is some positive integer.

It is possible that there are more than one representation of the form (3.1) of a given $z \in \mathbb{C}$. If so we call z *ambiguous*. We have the following lemma, which urges us to look elsewhere for normal numbers. Even though it is merely a lemma to us it is one of the main results found in [Madritsch, 2007] and as such the proof takes up a good amount of pages. For that reason we choose to omit it.

Lemma 3.3. *Let $(b, \mathcal{D})$ be a CNS. Then no number with an ambiguous representation is normal.*

Eventually we will be constructing normal numbers as a concatenation of digital expansions of a certain sequence of numbers and we therefore need an ordering on $\mathcal{O}_d$. As before we let $q = N(b)$, and define τ to be a bijection between $\mathcal{D}$ and $\{0, 1, \ldots, q-1\}$ with $\tau(0) = 0$. We can extend τ to $\mathcal{O}_d$ by letting $\tau(d_0 + d_1 b + d_2 b^2 + \cdots + d_n b^n) = \tau(d_0) + \tau(d_1)q + \tau(d_2)q^2 + \cdots + \tau(d_n)q^n$. Since expressing a number in $\mathcal{O}_d$ in base b is unique this extension of τ is a bijection. Hence we have that $\tau(\mathcal{O}_d) = \mathbb{N} \cup \{0\}$, so for each $a \in \mathcal{O}_d$ there exists an $n \in \mathbb{N} \cup \{0\}$ such that $\tau(a) = n$. Since we have a natural ordering on the image of τ we can pull this ordering back to $\mathcal{O}_d$, i.e., for $a, c \in \mathcal{O}_d$ we define an ordering on $\mathcal{O}_d$ by

$$a \leq c \Longleftrightarrow \tau(a) \leq \tau(c).$$

Now the formula $z_n = \tau^{-1}(n-1)$ defines an increasing sequence of elements from $\mathcal{O}_d$. For a function $f : \mathcal{O}_d \to \mathbb{C}$ we define

$$\theta(f) = \lfloor f(z_1) \rfloor b^{-l(f(z_1))} + \lfloor f(z_2) \rfloor b^{-l(f(z_1))-l(f(z_2))} + \ldots . \tag{3.3}$$

which is the concatenation of the integer parts of the function values evaluated on the sequence $\{z_n\}$. Now we can finally state the main theorem of this part of the thesis.

Theorem 3.4. *Let $f(z) = \alpha_n z^n + \cdots + \alpha_1 z + \alpha_0$ be a polynomial with coefficients in $\mathbb{C}$. Let $(b, \mathcal{D})$ be a CNS in $\mathcal{O}_d$, where $d \in \{-1, -2, -3, -7, -11\}$. Then for every $l \geq 1$ we have*

$$\sup_{d_1 \ldots d_l} \left| \frac{1}{N} \mathcal{N}(\theta(f); d_1 \ldots d_l; N) - \frac{1}{|\mathcal{D}|^l} \right| \ll (\log N)^{-1}$$

In particular the number given by (3.3) is normal. This extends the result of [Madritsch, 2008] by allowing us to work, not only with the canonical number systems related to the Gaussian integers, but also the others arising from the four other rings.

CHAPTER 4

A COLLECTION OF LEMMATA

We will need a couple of lemmata in order to prove the theorem. Most of these represent theorems in their own right but given that we will make use of them merely as lemmata we will call them that. The first lemma will help us with the asymptotics.

Lemma 4.1. *Let $\{a_n\}_{n\geq 1}$ and $\{b_n\}_{n\geq 1}$ be two sequences of real numbers with $0 < a_n \leq b_n$ for all n and suppose that $\lim_{n\to\infty} a_n/b_n = 0$ and $\lim_{n\to\infty} a_n \neq 0$. Then*

$$\lim_{n\to\infty} \frac{\sum_{i=1}^{n} a_i}{\sum_{i=1}^{n} b_i} = 0.$$

Before proving the lemma we note that the assumption $\lim_{n\to\infty} a_n \neq 0$ is left out in [Madritsch et al., 2008, Lemma 3.4] which makes the proof below fail. Without this assumption we cannot be certain that (see proof below) $C(N) \to \infty$. The Lemma is however corrected in [Madritsch, 2008]

Proof. Let $\varepsilon > 0$ be arbitrary. It follows from the assumptions that there exists $n_0 \in \mathbb{N}$ such that $a_n/b_n < \varepsilon/2$ for $n > n_0$ and thus that $a_n < \varepsilon b_n/2$ for $n > n_0$. Now let

$$A(N) = \sum_{n=1}^{N} a_n, \quad B(N) = \sum_{n=1}^{N} b_n, \quad C(N) = \sum_{n=n_0+1}^{N} b_n.$$

Using all of this we get

$$\frac{A(N)}{B(N)} = \frac{A(n_0) + \sum_{n=n_0+1}^{N} a_n}{B(n_0) + \sum_{n=n_0+1}^{N} b_n} < \frac{A(n_0) + \frac{\varepsilon}{2}C(N)}{B(n_0) + C(N)}.$$

Since $\lim_{n\to\infty} a_n \neq 0$ we also have $\lim_{n\to\infty} b_n \neq 0$ so the series $\sum_{n=1}^{\infty} b_n$ diverges. Thus for $N \to \infty$ we have $C(N) \to \infty$ since it is the tail of a divergent series. So we must have

$$\lim_{N\to\infty} \frac{A(N)}{B(N)} \leq \lim_{N\to\infty} \frac{A(n_0) + \varepsilon/2C(N)}{B(n_0) + C(N)} = \varepsilon/2.$$

so we end up at the desired conclusion. □

Since we need to deal with blocks of a fixed length we will need to know how the norm of a given number behaves compared to the length of the number's digital expansion, this is covered in the next lemma.

Lemma 4.2. *Let $(b, \mathcal{D})$ be a CNS in $\mathcal{O}_d$ and let $q = N(b)$. Then*

$$|l(z) - \log_q |z|^2| \leq c_b,$$

where c_b is a constant depending only on the base b and $\log_q$ is the logarithm in base q. Simply put this lemma shows that the length act like the logarithm of the norm and thus gives us a very good estimate of the length of a given number.

The above lemma is a special case of the main theorem in [Kovács and Pethő, 1992] and we choose not to prove it here.

We will need several more auxiliary results, but before giving the next lemma we need to define what the discrepancy is.

Definition 4.3. let $x_1, x_2, \ldots, x_N$ be a sequence of points in $\mathbb{R}^k$. Then the number

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

is called the discrepancy of the given sequence. Here λ_k denotes the k-dimensional Lebesgue measure, I is a Cartesian product of intervals in T and T^k is the k-dimensional torus, i.e. $T^k = \mathbb{R}^k/\mathbb{Z}^k$ and A is the function that counts the number of points from the sequence that are inside I. Formally we have

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

where χ_I is the characteristic function of I.

As we shall also need a definition of the so-called *star discrepancy* we give the definition here as well

Definition 4.4. Let $y = (y_1, \ldots, y_k)$ with $0 < y_i \leq 1$ for $i = 1, \ldots, k$ and denote by $[0, y)$ the Cartesian product of the k intervals $[0, y_i), i = 1, \ldots, k$. Then the number

$$D_N^*(x_n) = \sup_{y \in [0,1]^k} \left| \frac{A([0,y),N,x_n)}{N} - \lambda_k([0,y)) \right|.$$

is called the *star discrepancy* of the given sequence.

In order to estimate the discrepancy we will need the following much celebrated lemma.

Lemma 4.5 (Erdős–Turán–Koksma inequality)**.** *[Drmota and Tichy, 1997, Theorem 1.21] Let $x_1, x_2, \ldots, x_N$ be points in $\mathbb{R}^2$ and V an arbitrary positive integer. Then*

$$D_N(x_n) \leq \left(\frac{3}{2}\right)^2 \left(\frac{2}{V+1} + \sum_{0<\|v\|_\infty \leq V} \frac{1}{r(v)} \left| \frac{1}{N} \sum_{n=1}^{N} e(v \cdot x_n) \right| \right),$$

where $r(v) = (\max\{1, |v_1|\}) \cdot (\max\{1, |v_2|\})$, for $v = (v_1, v_2) \in \mathbb{Z}^2$, $v \cdot x_n$ denotes the usual Euclidean inner product in $\mathbb{R}^2$ and e is defined to be the function $e(z) = exp(2\pi i z)$.

In order to transform the exponential sum into an integral we have the following two lemmata

Lemma 4.6. *[Arkhipov et al., 2004, Lemma 5.4] Suppose that $F : \mathbb{R}^r \to \mathbb{R}$ is a differentiable function for $0 \leq x_j \leq P_j \leq P$ for $j = 1, 2, \ldots, r$. Assume that the function $\partial F/\partial x_j$ is piecewise monotone and of constant sign in each x_j, for $j = 1, 2, \ldots r$ and for any fixed values of the other variables. Assume furthermore* *that the number of intervals of monotonicity and constant sign does not exceed s. Finally assume that the following holds*

$$\left|\frac{\partial F(x_1, x_2, \ldots, x_r)}{\partial x_j}\right| \leq \delta, \quad j = 1, 2, \ldots, r$$

for some δ satisfying $0 < \delta < 1$. Then

$$\sum_{x_1=0}^{P_1} \cdots \sum_{x_r=0}^{P_r} e(F(x_1, x_2, \ldots, x_r))$$
$$= \int_0^{P_1} \ldots \int_0^{P_r} e(F(x_1, x_2, \ldots, x_r)) dx_1 dx_2 \ldots dx_r + \theta_1 r s P^{r-1} \left(3 + \frac{2\delta}{1-\delta}\right),$$

for some θ_1 with $|\theta_1| \leq 1$.

Lemma 4.7. *[Titchmarsh, 1986, Lemma 4.2] Let $F : \mathbb{R} \to \mathbb{R}$ be a differentiable function such that F' is monotonic and $F'(x) \geq m > 0$ or $F'(x) \leq -m < 0$ for all $x \in [a, b]$. Then*

$$\left|\int_a^b e(F(x))\, dx\right| \leq \frac{4}{m}.$$

We will now apply these last two lemmata in order to obtain the following

Lemma 4.8. *Let M and N be positive integers with M much smaller than N. Let $F : \mathbb{C} \to \mathbb{C}$ be such that the composition of tr with F satisfy the conditions of Lemmata 4.6 and 4.7. Where $tr(x)$ denotes the trace of the matrix corresponding to multiplication by x in $\mathbb{R}^2$ Then*

$$\left|\sum_{M \leq |z|^2 < M+N} e(\mathrm{tr}(F(z))\right| \ll \frac{\sqrt{N}}{m} + \frac{N}{(\log N)^{\sigma/2}} + s\frac{3-\delta}{1-\delta}\sqrt{N(\log N)^{\sigma}},$$

for any positive real number σ.

Proof. This is a generalization of [Madritsch, 2008, Lemma 3.6] or [Gittenberger and Thuswaldner, 2000, Lemmata 2.1 and 2.2]. We start by noting the composition $(x, y) \mapsto \mathrm{tr}(F(x+iy))$ is a function from $\mathbb{R}^2$ to $\mathbb{R}$, and hence satisfies the assumptions of Lemma 4.6. The main idea is to cover the annulus $M \leq |z|^2 < M+N$ by translates of the lattice associated with the ring of integers $\mathcal{O}_d$, get an estimate for how many translates are needed for this, and finally estimate the exponential sum on each of these. Consider the parallelograms $D_{v,d} = \{z \in \mathcal{O}_d \mid z = x + cy, -v \leq x \leq v, -v \leq y \leq v\}$. We shall need to estimate a sum of the form

$$\sum_{z \in D_{v,d}} e(tr(F(z))) = \sum_{x=-v}^{v} \sum_{y=-v}^{v} e(tr(F(x + cy)))$$

We can split up the sum on the right hand side, so each variable runs over a range allowing us to use Lemma 4.6 on each part and then joining the integrals afterwards. However this will give us a slight error-term which can be absorbed into the error-term of Lemma 4.6 by noting that the integration over a line where $x = 0$ or $y = 0$

is bounded by some constant times $2v$. So by choosing s big enough we can make the following hold.

$$\sum_{z\in D_{v,d}} e(tr(F(z))) = \sum_{x=-v}^{v}\sum_{y=-v}^{v} e(tr(F(x+cy))) \tag{4.1}$$
$$= \int_{-v}^{v}\int_{-v}^{v} e(tr(F(x+cy)))dxdy + 2\theta_1 sv\left(3+\frac{2\delta}{1-\delta}\right) \tag{4.2}$$
$$= \int_{-v}^{v}\int_{-v}^{v} e(tr(F(x+cy)))dxdy + 2\theta_1 sv\left(\frac{3-3\delta+2\delta}{1-\delta}\right) \tag{4.3}$$
$$= \int_{-v}^{v}\int_{-v}^{v} e(tr(F(x+cy)))dxdy + 2\theta_1 sv\left(\frac{3-\delta}{1-\delta}\right). \tag{4.4}$$

Taking absolute values allows us to apply Lemma 4.7

$$\left|\sum_{z\in D_{v,d}} e(tr(F(z)))\right| \le \int_{-v}^{v}\left|\int_{-v}^{v} e(tr(F(x+cy)))dx\right|dy + 2\theta_1 sv\left(\frac{3-\delta}{1-\delta}\right) \tag{4.5}$$
$$\le 2v \max_{-v\le y\le v}\left|\int_{-v}^{v} e(tr(F(x+cy)))dx\right| + 2\theta_1 sv\left(\frac{3-\delta}{1-\delta}\right) \tag{4.6}$$
$$\le \frac{8v}{m} + 2\theta_1 sv\left(\frac{3-\delta}{1-\delta}\right). \tag{4.7}$$

Now we will begin covering the annulus with parallelograms. We start by estimating how many parallelograms are needed to cover the annulus by giving a lower and an upper bound. Consider a disk of radius R centered at the origin. As noted earlier $\mathcal{O}_d$ has an integral basis formed by $(1, \frac{1+\sqrt{d}}{2})$ or $(1, \sqrt{d})$, depending on the value of d. Denote by c either $\frac{1+\sqrt{d}}{2}$ or $\sqrt{d}$. Now notice that the lattice in $\mathcal{O}_d$ divides the plane into parallelograms with side lengths 1 and $|c|$. And that $|c| \ge 1$ for any value of d. Denote by D the diameter of such a parallelogram and notice that it is also the length of the diagonal extending from the lower left corner to the upper right corner. It is obvious that we can fit in $\lfloor R/D\rfloor^2$ parallelograms in the first quadrant of the disc, and similarly in the 3rd quadrant. Consider next the antidiagonal of the parallelogram, extending from the lower right corner to the upper left corner. Let this have length A, now in the space remaining in the 1st and 3rd quadrant and all that of the 2nd quadrant we can fill in $\lfloor R/A\rfloor^2$ parallelograms, similarly in the 4th quadrant. This in total gives us a lower bound on the amount of disjoint parallelograms we can place fully inside a disc, since we are only interested in asymptotics we notice that it is comparable to C_1R^2, where C_1 is some constant depending only on the diameter of the parallelogram.

As for the upper bound we simply give estimates using the areas of parallelograms and the disc, hence the maximum amount of disjoint parallelograms needed to cover the disc is $\lfloor R^2/C_2\rfloor$, where C_2 is the a constant depending only on the area of the parallelogram. Now consider the annulus $\{z \in \mathbb{C} : M \le |z|^2 \le M+N\}$, we can estimate the amount of disjoint parallelograms that fits inside this by using the estimates we found for the disc. We can get an upper bound by using a packing of disc of radius M, extending this to a cover of a disc of radius $M+N$ and then removing the packing. This gives us $\frac{(M+N)^2}{C_2} - C_1M^2$.

As for the lower bound we do the opposite, that is we take a packing for the disc of radius $M+N$ and remove every parallelogram needed to cover the cocentric disc of radius M. We then get $C_1(M+N)^2 - \frac{M^2}{C_2}$ for the lower bound. Of these parallelograms some will be contained entirely in the annulus and some will intersect the boundary. The ones intersecting the boundary will provide an error term which we want to minimize, so we will not be using parallelograms of this size, instead we will shrink them by a factor of $\sqrt{N/\log N^\sigma}$. Now define the sets I and B to be the sets of parallelograms completely inside the annulus respective the set of parallelograms intersecting the boundary. Let C_I and C_B denote their contributions to the sum, respectively. Since we are looking for asymptotic estimates we can get an upper bound on the size of I by placing each parallelogram inside a square of side length $\sqrt{N/\log N^\sigma}$, now up to some constant, there are as many parallelograms as there are squares in I, and there are $\mathcal{O}((\log N^\sigma)$ of squares in I, inserting this into the estimate (4.7) we get

$$C_I \ll \frac{\sqrt{N}}{m} + s\frac{3-\delta}{1-\delta}\sqrt{N(\log N)^\sigma}.$$

As for the boundary we can do the same as before, that is place squares around each parallelogram and notice that the boundary is covered by two annuli of width $\mathcal{O}(\sqrt{M/(\log M)^\sigma})$ and $\mathcal{O}(\sqrt{(M+N)/(\log(M+N))^\sigma})$. From this and the fact that M is much smaller than N we get

$$C_B \ll \frac{N}{(\log N^{\sigma/2})}.$$

Adding these two we get the desired result. □

Lemma 4.9. *Let f be a k-th degree polynomial with coefficients in any algebraic field which contains the rational field and q be the least common multiple of its coefficients. If $\Lambda(q)$ is a complete set of residues modulo q, then, for any $\varepsilon > 0$,*

$$\left|\sum_{\lambda\in\Lambda(q)} e(\mathrm{tr}(f(\lambda)))\right| \ll (N(q))^{1-1/k+\varepsilon},$$

where the implied constant depends only on f and ε.

A proof of this can be found in [Hua, 1951, Theorem 1]

CHAPTER 5

THE FUNDAMENTAL DOMAIN AND ITS PROPERTIES

In this chapter we follow the paper [Gittenberger and Thuswaldner, 2000] closely, as with the preceding chapters most of the results below can also be found in [Madritsch, 2008]. The first part of the following result was shown by Kátai and Kóvacs in [Kátai and Kovács, 1981] and the second part was later shown by Kóvacs in [Kóvacs, 1981].

Theorem 5.1. *1. For any imaginary quadratic number field K and a CNS $(b, \mathcal{D})$. The pair $(1, b)$ form an integral basis for K.*

2. For any algebraic number field K of degree n and a CNS $(b, \mathcal{D})$. The n-tuple $(1, b, b^2, \dots, b^{n-1})$ form an integral basis for K.

Even though we shall only need the first part explicitly, the second part is equally important. In fact it is due to this result that the connection between the number systems we work with and matrix number systems was established. Without matrix number systems the construction below would not be possible. See [Madritsch, 2007] for more details on this.

Now let $(b, \mathcal{D})$ be a CNS in $\mathcal{O}_d$. Then by the above result every $\gamma \in \mathbb{C}$ has a unique representation of the form $\gamma = \alpha + \beta b$ where $\alpha, \beta \in \mathbb{R}$. Thus we can define the map

$$\varphi : \mathbb{C} \to \mathbb{R}^2, \quad \varphi(\alpha + \beta b) = (\alpha, \beta)$$

Since the representation is unique the map is well defined. Now we can use φ to transfer the fundamental domain to $\mathbb{R}^2$.

$$\mathcal{F} = \varphi(\mathcal{F}') = \left\{ \gamma \in \mathbb{R}^2 \mid \gamma = \sum_{k \geq 1} B^{-k} d_k, d_k \in \varphi(\mathcal{D}) \right\},$$

where B is the matrix corresponding to multiplication by b in $\mathbb{R}^2$, represented in the base $(1, b)$. For every b we consider the defining polynomial of b

$$x^2 + b_1 x + b_0.$$

The matrix B is then given by

$$B = \begin{pmatrix} 0 & -b_0 \\ 1 & -b_1 \end{pmatrix}$$

We notice that $b_0 = N(b) = q$ and that $b_1 = 2\,\mathrm{Re}(b)$. So the determinant of B is $b_0 \neq 0$ and thus B is invertible, which in turn allows us to define B^{-k}. More information about this can be found in [Scheicher and Thuswaldner, 2002].

Before we move on we shall need to introduce the Hausdorff metric.

For any metric space (X,d) we let $C(X)$ denote the class of all non-empty and compact subsets of X. For each positive δ and each $A \in C(X)$ we denote by A_δ the δ-neighbourhood of A, i.e. $A_\delta = \{x \in X : d(x,a) \le \delta \text{ for some } a \in A\}$. By abuse of notation we use d for the following function from $C(X) \times C(X)$ into $[0,\infty)$:

$$d(A,B) = \inf\{\delta : A \subset B_\delta \text{ and } B \subset A_\delta\}.$$

We call this function the Hausdorff metric on the space $C(X)$. It can be shown that this is a metric and in fact it turns $C(X)$ into a complete metric space, (more information can be found in [Falconer, 2003]). We now return to our study of the fundamental domain.

For every $a \in \mathcal{O}_d$ we define the domain corresponding to the elements of $\mathcal{F}$ whose digit representation after the point starts with the digits of the expansion of a. Thus we define

$$\mathcal{F}_a = B^{-l(a)}(\mathcal{F} + \varphi(a)). \tag{5.1}$$

It is well known that, see for instance [Müller et al., 2001] $\mathcal{F}$ can be approximated by the sets

$$Q_0 = \{z \in \mathbb{R}^2 \mid \|z\|_\infty \le \frac{1}{2}\}, \text{ and}$$
$$Q_k = \bigcup_{a \in \mathcal{D}} B^{-1}(Q_{k-1} + \varphi(a)).$$

The approximation satisfies the following

$$d(\partial Q_k, \partial \mathcal{F}) \ll |b|^{-k},$$

where $d(\cdot,\cdot)$ denotes the Hausdorff metric, and ∂Q_k denotes the boundary of the set Q_k. We need to be careful from now on, though the fundamental domain is just a parallelogram for some bases, it can be much more obscure for other bases. In fact it can have a fractal structure, if we pick $1-i$ to be our base, we notice that $N(1-i) = 2$. Hence $(1-i, \{0,1\})$ is a CNS in the complex plane. But the fundamental domain, when shifted to $\mathbb{R}^2$ is the famous Davis-Knuth dragon or twin dragon which can be seen below.

As shown in [Akiyama and Thuswaldner, 2000], Q_k is connected for all k. It can also be shown that Q_k can be covered by $|\mathcal{D}|^k$ parallelograms. Furthermore there exists a μ with $1 < \mu < |b|^2$ such that $\mathcal{O}(\mu^k)$ of these parallelograms intersect the boundary of Q_k, [Müller et al., 2001].

In order to keep track of $\mathcal{F}_a$ we will need a Urysohn function, which is a continuous function that is identically 1 on some closed set and identically 0 on some other disjoint closed set. In order to construct it we need tubes covering the boundary of $\mathcal{F}_a$, with certain properties. By a tube we simply mean a cylinder, i.e. a product of intervals. We use the notion tube, since it is the standard for higher dimensions. Specificly we have the following lemma which was shown for $a \in \mathcal{D}$ by Gittenberger and Thuswaldner. We will follow their ideas in the proof.

Lemma 5.2. *For all $a \in \mathcal{O}_d$ and all $k \in \mathbb{N}$ there exists a lattice-parallel tube $P_{k,a}$ with the following properties:*

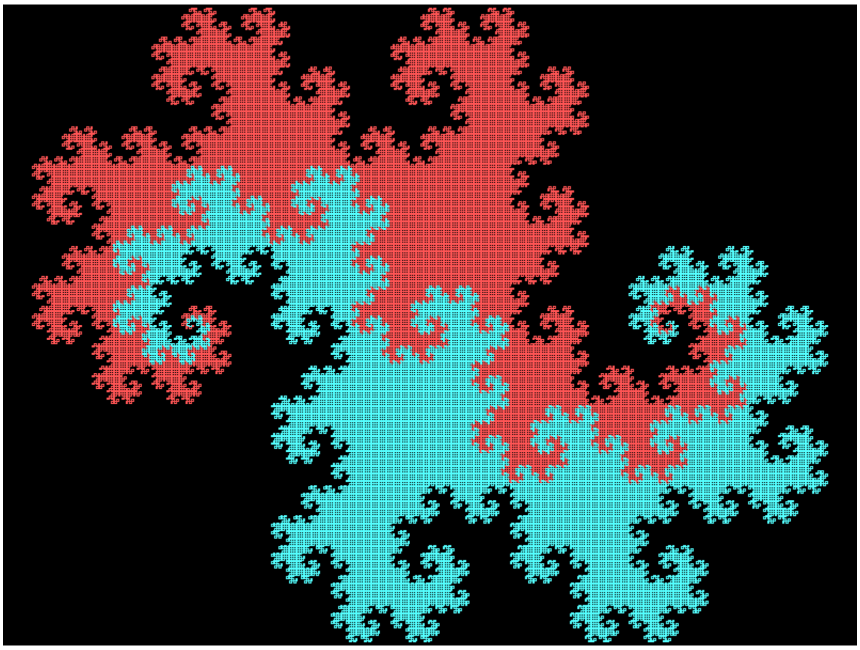

Figure 5.1: Davis-Knuth dragon

- *$\partial\mathcal{F}_a \subset P_{k,a}$ for all $k \in \mathbb{N}$.*
- *$\lambda_2(P_{k,a}) = \mathcal{O}(\mu^k/|b|^{2k})$.*
- *$P_{k,a}$ consists of $\mathcal{O}(\mu^k)$ lattice-parallel parallelograms, each of which has Lebesgue measure $\mathcal{O}(|b|^{-2k})$.*

Here the word lattice refers to the lattice formed by the ring of integers, $\mathcal{O}_d$, and a lattice-parallel tube is thus a parallelogram with sides parallel to the vectors forming the lattice.

Proof. This is a generalisation of [Gittenberger and Thuswaldner, 2000, Lemma 3.1]. We will construct a tube with the required properties. Let $Q_{k,a} = B^{-1}(Q_k + \varphi(a))$, then $Q_{k,a}$ will approximate $\mathcal{F}_a$ just as Q_k approximates $\mathcal{F}$. Denote by $II'_{k,a} = \partial Q_{k,a}$ the boundary of $Q_{k,a}$. This is seen to be a polygon, since $Q_{k,a}$ is a union of parallelograms. Let $R_{k,a}$ be the set consisting of the $|\mathcal{D}|^k$ parallelograms that forms $Q_{k,a}$. Then $\mathcal{O}(\mu^k)$ of these intersects the boundary $II'_{k,a}$, hence the number of edges in $II'_{k,a}$ is bounded by $\mathcal{O}(\mu^k)$. Since every element in $R_{k,a}$ has a diameter of length $c|b|^{-k}$, for some constant c the length of $II'_{k,a}$ is $\mathcal{O}(\mu^k|b|^{-k})$. We now construct a new polygon $II''_{k,a}$ in the following way:

Let $E_{II'_{k,a}}$ denote the edges of $II'_{k,a}$, and define $II_{k,a}$ by

$$II_{k,a} = \bigcup_{\substack{\overline{(\alpha_1,\alpha_2)(\beta_1,\beta_2)} \in E_{II'_{k,a}} \\ \alpha_2 \le \beta_2}} \left(\overline{(\alpha_1,\alpha_2)(\beta_1,\alpha_2)} \cup \overline{(\beta_1,\alpha_2)(\beta_1,\beta_2)}\right),$$

where $\overline{(\alpha_1,\alpha_2)(\beta_1,\alpha_2)}$ denotes the edge from the point (α_1,α_2) to the point (β_1,α_2).

It is easily seen that $II_{k,a}$ has lattice-parallel sides. We note that the length, number of edges and maximal distrance from $\mathcal{F}_a$ is comparable for $II_{k,a}$ and $I_{k,a}$. Since $d(II_{k,a}, \partial\mathcal{F}_a) < c'|b|^{-k}$, for some constant c' we can define the tube $P_{k,a}$ by

$$P_{k,a} = \{z \in \mathbb{R}^2 \mid \|z - II_{k,a}\|_\infty \leq 2c'|b|^{-k}\},$$

where $\|z - II_{k,a}\|_\infty = \inf_{z' \in II_{k,a}} |z - z'|_\infty$. We see that this tube has the required properties. □

Now to each pair (k, a) we fix the corresponding polygon $II_{k,a}$ and the corresponding tube $P_{k,a}$ constructed in the lemma above. Let $I_{k,a}$ denote the set of inner points in $II_{k,a}$ and define f_a by

$$f_a(x, y) = \frac{1}{\Delta^2} \int_{-\Delta/2}^{\Delta/2} \int_{-\Delta/2}^{\Delta/2} \Psi_a(x + x_1, y + y_1)\, dx_1\, dy_1,$$

where

$$\Delta = 2c'|b|^{-k}, \tag{5.2}$$

and

$$\Psi_a(x, y) = \begin{cases} 1 & \text{if } (x, y) \in I_{k,a} \\ 1/2 & \text{if } (x, y) \in II_{k,a} \\ 0 & \text{otherwise.} \end{cases}$$

Thus f_a is the Urysohn function which equals 1 for $z \in I_{k,a} \setminus P_{k,a}$, and 0 for $z \in \mathbb{R}^2 \setminus (I_{k,a} \cup P_{k,a})$ and linear interpolation in between. Now we move on to give estimates for the Fourier coefficients of this function, as we shall need these in chapter 7.

Lemma 5.3. *Let $f_a(x, y) = \sum_{n_1, n_2} c_{n_1, n_2} e((n_1, n_2) \cdot_d (x, y))$ be the Fourier expansion of f_a, where $(n_1, n_2) \cdot_d (x, y)$ denotes the inner product that comes from the quadratic form corresponding to the given lattice. Later on when d is clear from the context we shall supress this and simply use the notation $\cdot$ to denote the inner product. Then we have the following estimates for the Fourier coefficients*

$$\begin{aligned} c_{n_1,n_2} &= \mathcal{O}\left(\frac{\mu^k}{\Delta^2 n_1^2 n_2^2}\right) \quad (n_1, n_2 \neq 0), \\ c_{n_1,0} &= \mathcal{O}\left(\frac{\mu^k}{\Delta n_1^2}\right) \quad (n_1 \neq 0), \\ c_{0,n_2} &= \mathcal{O}\left(\frac{\mu^k}{\Delta n_2^2}\right) \quad (n_2 \neq 0), \\ c_{0,0} &= \frac{1}{|b|^2}. \end{aligned}$$

The proof is an exercise in straight forward calculations, however it takes up some space and is therefore given in the appendix.

CHAPTER 6

WEYL SUMS

In the following chapter we let f denote a polynomial with coefficients in $\mathbb{C}$. As the letter d is already used for other purposes we denote the degree of the polynomial f by d'. Thus

$$f(z) = \alpha_{d'} z^{d'} + \alpha_{d'-1} z^{d'-1} + \cdots + \alpha_1 z + \alpha_0. \quad \text{with } \alpha_i \in \mathbb{C} \text{ and } \alpha_{d'} \neq 0.$$

We will now need a generalisation of [Nakai and Shiokawa, 1992, Lemma 2], which in chapter 7 will be needed to give us some estimates on exponential sums.

Proposition 6.1. *Let $G > 0$ and $N \geq 2$. Let s be an integer with $1 \leq s < d'$, let H_i, K_i be any positive constants, where $i = s+1, \ldots, d'$ and let H_s^*, K_s^* be constants such that*

$$H_s^* \geq 2^{3(s+2)} + 2^{2+3}\left(G + \max_{s<i\leq d'} H_i\right) + s \sum_{i=s+1}^{d'} K_i,$$

$$K_s^* \geq 2^{3(s+2)} + 2^{2+3}\left(G + \max_{s<i\leq d'} H_i\right) + 2s \sum_{i=s+1}^{d'} K_i.$$

Suppose that there are integers a_i and q_i from $\mathcal{O}_d$ for $s < i \leq d'$ such that

$$1 \leq |q_i|^2 \leq (\log N)^{K_i} \quad \text{and} \quad \left|\alpha_i - \frac{a_i}{q_i}\right| \leq \frac{(\log N)^{H_i}}{q_i N^{i/2}},$$

but there exists no integers a_s and q_s from $\mathcal{O}_d$ that are relatively prime such that

$$1 \leq |q_s|^2 \leq (\log N)^{K_s^*} \quad \text{and} \quad \left|\alpha_s - \frac{a_s}{q_s}\right| \leq \frac{(\log N)^{H_s^*}}{q_s N^{s/2}}.$$

Then we have

$$\left|\sum_{|z|^2 \leq N} e(\operatorname{tr}(f(z)))\right| \ll N(\log N)^{-G}.$$

In order to prove the above proposition we shall need two lemmata. During the proof of the first Lemma we shall need the following classic theorem due to Minkowski, namely Minkowski's linear forms theorem, which along with a proof can be found in [Cassels, 1959].

Theorem 6.2 (Minkowski's linear forms theorem)**.** *Let Λ be an n-dimensional lattice of determinant $d(\Lambda)$ and let $a_{ij} (1 \leq i, j \leq n)$ be real numbers. Suppose that $c_j > 0$ for $1 \leq j \leq n$, are numbers such that*

$$c_1 \ldots c_n \geq |\det(a_{ij})|\, d(\Lambda).$$

Then there is a point $u = (u_1, u_2, \ldots, u_n) \in \Lambda$ *other than* 0 *satisfying*

$$\left|\sum_{j=1}^{n} a_{1j}u_j\right| \leq c_1$$

and

$$\left|\sum_{j=1}^{n} a_{ij}u_j\right| \leq c_i \quad (2 \leq i \leq n).$$

The first lemma deals with approximation of complex numbers by ratios of integers from $\mathcal{O}_d$.

Lemma 6.3. *Given any* $z = x + iy \in \mathbb{C}$ *and* $N \in \mathbb{N}$, *there exists integers* p *and* q *in* $\mathcal{O}_d$ *with* $0 < |q| \leq N$ *such that*

$$\left|z - \frac{p}{q}\right| < \frac{2\sqrt{-d+1}}{|q|N}. \tag{6.1}$$

The result can be found in [Esdahl-Schou and Kristensen, 2010] as well as in [Dodson and Kristensen, 2004] where a proof is given in the case of $\mathcal{O}_d$ being the Gaussian integers.

Proof. We split the proof into two parts depending on the value of d. For $d \in \{-1, -2\}$ the lattice is axis-parallel and elements in $\mathcal{O}_d$ can be written as $p = p_1 + i\sqrt{-d}p_2$. Let $z = x + iy, p = p_1 + i\sqrt{-d}p_2, q = q_1 + i\sqrt{-d}q_2$. We want to show that the inequality

$$\left|z - \frac{p}{q}\right| < \frac{c'}{|q|N} \tag{6.2}$$

holds, where c' is some constant to be chosen later, satisfying $c' < 2\sqrt{-d+1}$. This holds if and only if

$$|qz - p| < \frac{c'}{N},$$

which holds if and only if

$$\left|(q_1 + i\sqrt{-d}q_2)(x + iy) - (p_1 + i\sqrt{-d}p_2)\right| < \frac{c'}{N}.$$

Calculating the left-hand-side we see this is true if and only if

$$\left|q_1x - \sqrt{-d}q_2y - p_1 + i(q_1y + \sqrt{-d}q_2x - \sqrt{-d}p_2)\right| < \frac{c'}{N}.$$

Which is the case if

$$\max\{|q_1x - \sqrt{-d}yq_2 - p_1|, |\sqrt{-d}xq_2 + yq_1 - \sqrt{-d}p_2|\} < \frac{c'}{\sqrt{2}N}.$$

We will write this in a slightly different manner (the purpose becomes clear in a little while)

$$\max\{|-p_1 + 0 \cdot p_2 + xq_1 - \sqrt{d}yq_2|, |0 \cdot p_1 - \sqrt{d}p_2 + yq_1 + \sqrt{d}xq_2|\} < \frac{c'}{\sqrt{2}N}. \tag{6.3}$$

Since we are trying to obtain $|q| < N$. The best result possible is obtained by looking at $|\sqrt{q_1^2 + dq_2^2}| < N$ and minimising q_1 and q_2 at the same time. This corresponds to finding the minimum absolute value of an ellipsis and is not very practical to work with. Instead we will use the maximum norm and thereby minimising q_1 and q_2 seperately. As a result the constant we obtain in this way might not be the best possible but it will suffice. The key is using Theorem 6.2. From (6.3) it is clear that we will need four inequalities. The linear form will be expressed in the following way:

$$\begin{pmatrix} -1 & 0 & x & -\sqrt{-d}y \\ 0 & -\sqrt{-d} & y & \sqrt{-d}x \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ q_1 \\ q_2 \end{pmatrix} = \begin{pmatrix} -p_1 + 0 \cdot p_2 + xq_1 - \sqrt{d}yq_2 \\ 0 \cdot p_1 - \sqrt{d}p_2 + yq_1 + \sqrt{d}xq_2 \\ q_1 \\ q_2 \end{pmatrix} \tag{6.4}$$

Before we apply Minkowski's Theorem on linear forms we summarise the relevant inequalities. We have

$$|q| < N$$

if and only if

$$|q_1^2 + dq_2^2| < N^2. \tag{6.5}$$

Assuming $|q_1| < aN$ and $|q_2| < bN$ we see that

$$|q_1^2 + dq_2^2| \leq |q_1|^2 + |d||q_2|^2 \leq (a^2 + |d|b^2)N^2.$$

So (6.5) holds if we can pick a and b such that $a^2 + |d|b^2 = 1$. We start by letting $b = \frac{a}{\sqrt{-d}}$ to make the expression symmetric. Thus we now need to pick a such that

$$1 = a^2 + |d|b^2 = a^2 + a^2 = 2a^2.$$

Picking $a = \frac{1}{\sqrt{2}}$ we see this is true. This means we can pick $c_3 = aN = \frac{N}{\sqrt{2}}$ and $c_4 = bN = \frac{N}{\sqrt{-2d}}$. The remaining two constants, c_1 and c_2 need to fulfill that the product $c_1c_2c_3c_4$ is greater than the absolute value of the determinant of the matrix in (6.4). We denote this matrix A. Since A is an upper triangular matrix the absolute value of its determinant is clearly $|\det(A)| = \sqrt{-d}$. From (6.3) we see it is possible to pick $c_1 = c_2$ and we want

$$c_1c_2c_3c_4 = c_1^2 \frac{N^2}{2\sqrt{-d}} \geq \sqrt{-d}.$$

So we can pick

$$c_1 = \frac{\sqrt{-2d}}{N}.$$

Taking another glance at (6.3) we can pick $c' = \sqrt{2}Nc_1 = 2\sqrt{-d}$. Now we have our constants and applying Minkowski's Theorem on linear forms gives us a non-zero solution in integers p_1, p_2, q_1, q_2. Hence (6.1) has a solution with $|q| = |q_1 + i\sqrt{d}q_2| \leq N$.

We now let $d \in \{-3, -7, -11, -19, -43, -67, -163\}$. For $p \in \mathcal{O}_d$ we have the representation $p = p_1 + (1 + i\sqrt{-d})/2p_2$, where $p_1, p_2 \in \mathbb{Z}$. So the lattice is twisted, but we can use the same procedure as above.

Assume that
$$\left|z - \frac{p}{q}\right| < \frac{c''}{|q|N},$$
where c'' is some constant to be chosen later satisfying $c'' < 2\sqrt{-d+1}$. This holds if and only if
$$|qz - p| < \frac{c''}{N}.$$
Calculating the left-hand-side, we see this holds if and only if
$$\left|\left(q_1 + \frac{1+i\sqrt{-d}}{2}q_2\right)(x+iy) - \left(p_1 + \frac{1+i\sqrt{-d}}{2}p_2\right)\right| < \frac{c''}{N}$$
which is if and only if
$$\left|q_1x + \frac{q_2}{2}x - \frac{\sqrt{-d}}{2}q_2y - p_1 - \frac{p_2}{2} + i\left(\frac{\sqrt{-d}}{2}q_2x + q_1y + \frac{q_2}{2}y - \frac{\sqrt{-d}}{2}p_2\right)\right| < \frac{c''}{N},$$
which is the case if
$$\max\{\left|q_1x + \frac{q_2}{2}x - \frac{\sqrt{-d}}{2}q_2y - p_1 - \frac{p_2}{2}\right|, \left|\frac{\sqrt{-d}}{2}q_2x + q_1y + \frac{q_2}{2}y - \frac{\sqrt{-d}}{2}p_2\right|\} < \frac{c''}{\sqrt{2}N}.$$
Once more we write up the matrix
$$\begin{pmatrix} -1 & -1/2 & x & x/2 - \sqrt{-d}y/2 \\ 0 & -\sqrt{-d}/2 & y & \sqrt{-d}x/2 + y/2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ q_1 \\ q_2 \end{pmatrix} = \begin{pmatrix} -p_1 - p_2/2 + q_1x + q_2x/2 - \sqrt{-d}yq_2/2 \\ -\sqrt{-d}p_2/2 + q_1y + \sqrt{-d}q_2x/2 + q_2y/2 \\ q_1 \\ q_2 \end{pmatrix} \tag{6.6}$$
We once again consider $|q| < N$ and notice that this holds if and only if
$$\left|\left(q_1 + \frac{q_2}{2}\right)^2 + \frac{-d}{4}q_2^2\right| < N^2.$$
Assuming once more that $|q_1| < aN$ and $|q_2|bN$ we see that
$$|q| \leq |q_1|^2 + \frac{1}{4}|q_2|^2 + |q_1||q_2| + \frac{-d}{4}|q_2|^2 < \left(a^2 + \frac{b^2}{4} + ab + \frac{d}{4}b^2\right)N^2.$$
We now pick a and b such that the right-hand-side is equal to N^2. We start by picking $b = \frac{2}{\sqrt{-d+1}}a$ and arrive at that
$$1 = a^2 + ab + \frac{-d+1}{4}b^2 = a^2 + \frac{2}{\sqrt{-d+1}}a^2 + a^2 = \left(2 + \frac{2}{\sqrt{-d+1}}\right)a^2.$$
So we can pick
$$a = \frac{1}{\sqrt{2 + \frac{2}{\sqrt{-d+1}}}}.$$
As before this allows us to pick $c_3 = aN$ and $c_4 = bN$ so
$$c_3 = \frac{N}{\sqrt{2 + \frac{2}{\sqrt{-d+1}}}}.$$

$$c_4 = \frac{N}{\sqrt{2 + \frac{2}{\sqrt{-d+1}}}} \frac{2}{\sqrt{-d+1}}.$$

We note that the matrix has determinant $\sqrt{-d}/2$. So we can pick $c_1 = c_2$ and as long as

$$c_1c_2c_3c_4 \geq \frac{\sqrt{-d}}{2} \tag{6.7}$$

We get a non-zero integer solution p_1, p_2, q_1, q_2. We get from picking $c_1 = c_2$ we get from (6.7)

$$c_1^2 \geq \frac{\sqrt{-d}\sqrt{-d+1}}{4N^2}\left(2 + \frac{2}{\sqrt{-d+1}}\right),$$

noting that $\sqrt{-d+1} > \sqrt{-d}$ and that $\sqrt{-d} > 2$ we can pick

$$c_1 = c_2 = \frac{\sqrt{-d+1}}{N}.$$

Thus we can choose $c'' = \sqrt{2}Nc_1$ so

$$c'' = \sqrt{2}\sqrt{-d+1}.$$

To finish the proof we need only let notice that $c = 2\sqrt{-d+1}$ is bigger than both $c' = 2\sqrt{-d}$ and $c'' = \sqrt{2}\sqrt{-d+1}$. □

The second lemma we will need considers the case $s = d'$ and can be found in [Gittenberger and Thuswaldner, 2000, Proposition 2.1]. Notice that the proof given there is not valid in our case, but the ideas can be adopted to provide a proof that works in the generality needed, before we can do that we shall need [Gittenberger and Thuswaldner, 2000, Lemmata 2.3,2.4 an 2.5].

Lemma 6.4. *Let $h, q \in \mathcal{O}_d$ with $|q| > 2$, and let h and q be relatively prime. Let*

$$S = \sum_{|z|^2 < N} e\left(\mathrm{tr}\left(\frac{h}{q}z\right)\right).$$

Then

$$|S| \ll \sqrt{N}|q|.$$

Proof. We see that for fixed d there exists a residue system R modulo q satisfying

$$R \subset \{z \in \mathcal{O}_d \mid |z| \leq c|q|\},$$

where c is some positive fixed constant depending only on the ring $\mathcal{O}_d$. We now tesselate the open disc $K_N = \{z \mid |z|^2 < N\}$ with translates of R. Let T be this tesselation and define

$$E_N = \{R \in T \mid R \subset K_N\},$$
$$F_N = \{R \in T \mid R \not\subset K_N\}.$$

We have by [Hua, 1951, Theorem 3] that

$$\sum_{z \in R} e\left(\mathrm{tr}\left(\frac{h}{q}z\right)\right) = 0 \quad \text{for } R \in E_N.$$

Thus

$$S = \sum_{R \in F_N} \sum_{z \in R \cap K_N} e\left(\operatorname{tr}\left(\frac{h}{q} z\right)\right).$$

But this sum has at most $\mathcal{O}(\sqrt{N}|q|)$ summands, so the proof is complete. □

We adopt Hua's notation and let $\sum_x^{c'}$ denote the sum over all integers, x in a set of the form $\cap_{j=1}^{J}(a_j + \{y \mid |y|^2 \leq c_j N\})$ with $a_j \in \mathcal{O}_d$ and $0 < c_j < c'$. The exact values of a_j, c_j and c' are not important but can be found in [Hua, 1965, Lemma 3.3 and 3.4].

Lemma 6.5. *Let $f(x) = \sum_{j=0}^{k} a_j x^j$ be a polynomial of degree k and let*

$$S = \sum_{|z|^2 < N} e(\operatorname{tr}(f(z))).$$

Then we have the estimate

$$|S|^{2^{k-1}} \leq c N^{2^{k-1}-1} \Big|\sum_{y_1}^{c'} \cdots \sum_{y_k}^{c'} e(\operatorname{tr}(y_1 \ldots y_{k-1}(k! a_k y_k + \beta)))\Big|$$

with certain constants c and β.

Let $d_k(z)$ be the number of representations of z as a product of k non-zero integers from $\mathcal{O}_d$. Then we have

$$\sum_{|z|^2 < N} \leq N(\log N)^{k-1}$$

as found in [Narkiewicz, 1990, page 514]. Using [Hua, 1965, Lemma 6.1] we get the following lemma

Lemma 6.6. *For $\sigma_2 \geq 2^{3k} - 1$ the estimate*

$$\sum_{|z|^2 < N}{}' \, d_k(z) = \mathcal{O}(N(\log N)^{-\sigma_2})$$

holds. Here the prime ($'$) indicates that the sum is taken over all z in the range of summation satisfying

$$(\log N)^{\sigma_2} \leq c d_k(z).$$

Lemma 6.7. *Let h and q be relatively prime and suppose that*

$$f(x) = \frac{h}{q} x^{d'} + \alpha_{d'-1} x^{d'-1} + \cdots + \alpha_1 x + \alpha_0,$$

where $(\log N)^H \leq |q|^2 \leq N^{d'} (\log N)^{-H}$. Then

$$\left| \sum_{|z|^2 \leq N} e(\operatorname{tr}(f(z))) \right| \ll N(\log N)^{-G}, \quad \text{with} \quad H \geq 2^{d'+2} G + 2^{3(d'+2)}.$$

Proof. For ease of notation we set $k = d'$ and now consider different cases, depending on the value of k. For $k = 1$ we can apply Lemma 6.4 and obtain

$$S = \left| \sum_{|z|^2 < N} e\left(\operatorname{tr}\left(\frac{h}{q} + \alpha_{d'-1}\right)\right) \right| \le N(\log N)^{-H/2} \ll N(\log N)^{-G}.$$

Suppose now that $k > 1$. We start by noting that the calculations done in (4.7) remains valid if the summation is shifted to be of the shape $\sum_x^{c'}$.

Following the idea in Hua's proof of [Hua, 1965, Lemma 3.6] we apply Lemma 6.5 to estimate $|S|^{2^{k-1}}$

$$\begin{aligned} |S|^{2^{k-1}} &\le cN^{2^{k-1}-k} \left| \sum_{y_1}^{c'} \cdots \sum_{y_k}^{c'} e(\operatorname{tr}(y_1 \dots y_{k-1}(k! a_k y_k + \beta))) \right| \\ &\le cN^{2^{k-1}-k} \sum_{y_1}^{c'} \cdots \sum_{y_{k-1}}^{c'} \left| \sum_{y_k}^{c'} e(\operatorname{tr}(y_1 \dots y_{k-1}(k! a_k y_k + \beta))) \right| \\ &\ll N^{2^{k-1}-k} \sum_{y_1}^{c'} \cdots \sum_{y_{k-1}}^{c'} \left| k! y_1 \dots y_{k-1} \sum_{y_k}^{c'} e(\operatorname{tr}(a_k y_k + \beta)) \right| \end{aligned}$$

Using (4.7) we can estimate the inner sum

$$\ll N^{2^{k-1}-k} \sum_{y_1}^{c'} \cdots \sum_{y_{k-1}}^{c'} \left| k! y_1 \dots y_{k-1} \left(\sqrt{N} + \frac{N}{(\log N)^{\sigma/2}} + \sqrt{N(\log N)^{\sigma}} \right) \right|$$

By setting

$$\zeta = k! y_1 \dots y_{k-1} \tag{6.8}$$

we have $|\zeta|^2 \le M = c'^k k! N^{k-1}$. For a fixed $\zeta \neq 0$ the number of solutions to equation (6.8) is less than or equal to $d_{k-1}(\zeta)$. For $\zeta = 0$ the number of solutions is $\mathcal{O}(N^{k-2})$. We now get

$$|S|^{2^{k-1}} \ll N^{2^{k-1}-k} \left(\sum_{|\zeta|^2 \le M}' d_{k-1}(\zeta) + N^{k-2} \right) \left(\sqrt{N} + \frac{N}{(\log N)^{\sigma/2}} + \sqrt{N(\log N)^{\sigma}} \right)$$

Using Lemma 6.6 we see that

$$\begin{aligned} &\ll N^{2^{k-1}-k} \left(N(\log N)^{-\sigma_2} + N^{k-2} \right) \left(\sqrt{N} + \frac{N}{(\log N)^{\sigma/2}} + \sqrt{N(\log N)^{\sigma}} \right) \\ &\ll N^{2^{k-1}-1} \end{aligned}$$

Taking the 2^{k-1}th root concludes the proof. □

We will apply the last lemma recursively and therefore we need a tool to rewrite it, which is the following lemma.

Lemma 6.8. *[Korobov, 1992, Lemma 26] Let f_1 and f_2 be real-valued functions defined on a finite set M. Then*

$$\sum_{x \in M} e(f_1(x) + f_2(x)) = \sum_{x \in M} e(f_1(x)) + 2\pi\theta \sum_{x \in M} |f_2(x)|,$$

for some θ with $|\theta| \leq 1$.

Corollary 6.9. *Let $g(x) = \alpha_{d'}x^{d'} + \alpha_{d'-1}x^{d'-1} + \cdots + \alpha_1 x + \alpha_0 \in \mathbb{C}[X]$. If there exist $h, q \in \mathcal{O}_d[X]$ relatively prime and*

$$\left|\alpha_{d'} - \frac{h}{q}\right| \leq \frac{(\log N)^H}{|q| \leq N^{d'/2}},$$

with $(\log N)^{H+G} \leq |q| \leq N^{d'/2}(\log N)^{-H}$ and $H \geq 2^{d'+1}G + 2^{3(d'+2)-1}$, then

$$\left|\sum_{|z|^2 \leq N} e(\operatorname{tr}(g(z)))\right| \ll N(\log N)^{-G}$$

We first show that the corollary follows from the lemmata, and then we prove the proposition.

of Corollary 6.9. Using first Lemma 6.8 and then the estimate in Lemma 6.7 together with the assumption of the corollary we get

$$\begin{aligned}
\left|\sum_{|z|^2 \leq N} e(\operatorname{tr}(g(z)))\right| &= \left|\sum_{|z|^2 \leq N} e\left(\operatorname{tr}\left(\alpha_{d'}z^{d'} + \alpha_{d'-1}z^{d'-1} + \cdots + \alpha_1 z + \alpha_0\right)\right)\right| \\
&\leq \quad \left|\sum_{|z|^2 \leq N} e\left(\operatorname{tr}\left(\left(\alpha_{d'-1}z^{d'-1} + \cdots + \alpha_1 z + \alpha_0\right)\right)\right)\right| \\
&\quad + \left|2\pi \sum_{|z|^2 \leq N} \left|\operatorname{tr}\left(\left(\alpha_{d'} - \frac{h}{q}\right) z^{d'}\right)\right|\right| \\
&\ll N(\log N)^{-G} + N^{1/2}.
\end{aligned}$$

□

Proof of Proposition 6.1. The proof goes much along the same lines as that of [Madritsch, 2008, Proposition 5.1], which is inspired by [Nakai and Shiokawa, 1992, Lemma 2]. We consider the different possible values for s. The case $s = d'$ is dealt with in Corollary 6.9, so assume that $s < d'$. Denote by k the least common multiple of $q_{s+1}, \ldots, q_{d'}$. Obviously $k \in \mathcal{O}_d$ since $\mathcal{O}_d$ is a unique factorisation domain. Denote by Q the integer such that

$$|k|^2 Q \leq N < |k|^2(Q+1).$$

From our assumptions we now get that

$$1 \leq |k|^2 \leq (\log N)^K, \quad \text{where} \quad K = \sum_{i=s+1}^{d'} K_i,$$

and

$$N(\log N)^{-K} \ll Q \ll N/|k|^2.$$

Now since $\mathcal{O}_d$ is a Euclidean domain, for every $s \in \mathcal{O}_d$ there exist unique $q, r \in \mathcal{O}_d$ with $|r|^2 < |k|^2$ such that $s = qk + r$. Thus there exists a complete residue system R modulo k with

$$R \subset \{z \in \mathcal{O}_d \mid |z| \leq |k|\}.$$

We will use this system to tessellate the open disc D, defined by $D = \{z : |z|^2 < N\}$, with translates of R. Let

$$\begin{aligned} T &= \{t \in \mathcal{O}_d \mid (R + tk) \cap D \neq \emptyset\}, \\ I &= \{t \in T \mid R + tk \subset D\}. \end{aligned}$$

Since there are $\mathcal{O}(\sqrt{N})$ points on the circumference of D and $\mathcal{O}(|k|)$ points inside R, there will be $\mathcal{O}(\sqrt{N}|k|)$ on a neighbourhood of the circumference. So we get

$$\sum_{|z|^2 \leq N} e(\mathrm{tr}(f(z))) = \sum_{t \in I} \sum_{r \in R} e(\mathrm{tr}(f(tk + r))) + \mathcal{O}(\sqrt{N}|k|).$$

Notice that if the assumption of $\mathcal{O}_d$ being Euclidean is weakened to that of only having unique factorisation, it might still be possible to find a complete set of residues such that $z = tk + r$, but where r is not unique. In this situation we will however pick up an error term since we would be summing up over all possible residues. There is, as far as the author can tell no simple way to control the size of this error term, hence we limit ourselves to the case where $\mathcal{O}_d$ is Euclidean.

Now we will want to perform Abel summation, in order to do this we need an ordering on I. For $x, y \in I$, we define

$$x \prec y \Leftrightarrow \begin{cases} |x| < |y| & \text{or} \\ |x| = |y| & \text{and } \arg(x) < \arg(y). \end{cases}$$

From the polar representation of every complex number it is clear that this ordering is well defined. Let $M = |I|$, and let σ be a bijection from the set $\{1, 2, \ldots, M\}$ to I such that $\sigma(1) = 0$, and $\sigma(|I|) = \max I$ where the maximum is taken with respect to the newly defined ordering on I. Now define

$$\sigma(x) \prec \sigma(y) \Leftrightarrow x \leq y.$$

Then we have

$$\sum_{|z|^2 \leq N} e(\mathrm{tr}(f(z))) = \sum_{n=1}^{M} \sum_{r \in R} e(\mathrm{tr}(f(\sigma(n)k + r))) + \mathcal{O}(\sqrt{N}|k|). \tag{6.9}$$

Before we perform Abel summation we define the following quantities for ease of notation. Let

$$\begin{aligned} \psi_r(x) &= \sum_{i=s+1}^{d'} \gamma_i (xk + r)^i, \quad \gamma_i = \alpha_i - \frac{a_i}{q_i}, \\ \varphi_r(x) &= \sum_{i=1}^{s} \alpha_i (xk + r)^i, \quad T_r(l) = \sum_{n=1}^{l} e(\mathrm{tr}(\varphi_r(\sigma(n)))). \end{aligned}$$

Using the linearity of the trace we get

$$
\begin{aligned}
&\sum_{n=1}^{M}\sum_{r\in R} e(\operatorname{tr}(f(\sigma(n)k+r)))\\
&=\sum_{r\in R}\sum_{n=1}^{M} e\left(\operatorname{tr}\left(\sum_{i=1}^{d'}\alpha_i(\sigma(n)k+r)^i\right)\right)\\
&=\sum_{r\in R}\sum_{n=1}^{M} e\left(\operatorname{tr}\left(\sum_{i=1}^{s}\alpha_i(\sigma(n)k+r)^i+\sum_{i=s+1}^{d'}\alpha_i(\sigma(n)k+r)^i\right)\right)\\
&=\sum_{r\in R}\sum_{n=1}^{M} e\left(\operatorname{tr}\left(\varphi_r(\sigma(n))+\sum_{i=s+1}^{d'}\left(\gamma_i+\frac{a_i}{q_i}\right)(\sigma(n)k+r)^i\right)\right)\\
&=\sum_{r\in R}\sum_{n=1}^{M} e\left(\operatorname{tr}\left(\varphi_r(\sigma(n))\right)\right) e\left(\operatorname{tr}\left(\sum_{i=s+1}^{d'}\left(\gamma_i+\frac{a_i}{q_i}\right)(\sigma(n)k+r)^i\right)\right)\\
&=\sum_{r\in R}\sum_{n=1}^{M} e\left(\operatorname{tr}\left(\varphi_r(\sigma(n))\right)\right) e\left(\operatorname{tr}\left(\sum_{i=s+1}^{d'}\gamma_i(\sigma(n)k+r)^i\right)\right) e\left(\operatorname{tr}\left(\sum_{i=s+1}^{d'}\frac{a_i}{q_i}(\sigma(n)k+r)^i\right)\right)\\
&=\sum_{r\in R}\sum_{n=1}^{M} e\left(\operatorname{tr}\left(\varphi_r(\sigma(n))\right)\right) e\left(\operatorname{tr}\left(\psi_r(\sigma(n))\right)\right) e\left(\operatorname{tr}\left(\sum_{i=s+1}^{d'}\frac{a_i}{q_i}(\sigma(n)k+r)^i\right)\right).
\end{aligned}
$$

Then we note that $\sigma(n)k+r$ runs over the entire set of possible residues r, so we can relabel these as r in the last sum

$$
\begin{aligned}
&=\sum_{r\in R} e\left(\operatorname{tr}\left(\sum_{i=s+1}^{d'}\frac{a_i}{q_i}r^i\right)\right)\sum_{n=1}^{M} e\left(\operatorname{tr}\left(\varphi_r(\sigma(n))\right)\right) e\left(\operatorname{tr}\left(\psi_r(\sigma(n))\right)\right)\\
&=\sum_{r\in R} e\left(\operatorname{tr}\left(\sum_{i=s+1}^{d'}\frac{a_i}{q_i}r^i\right)\right)\sum_{n=1}^{M} e\left(\operatorname{tr}\left(\psi_r(\sigma(n))\right)\right)\left(T_r(n)-T_r(n-1)\right)\\
&=\sum_{r\in R} e\left(\operatorname{tr}\left(\sum_{i=s+1}^{d'}\frac{a_i}{q_i}r^i\right)\right)\\
&\quad\cdot\left(e(\operatorname{tr}(\psi_r(\sigma(M))))T_r(M)+\sum_{n=1}^{M}\Big(e(\operatorname{tr}(\psi_r(\sigma(n))))-e(\operatorname{tr}(\psi_r(\sigma(n+1))))\Big)T_r(n)\right).
\end{aligned}
$$

Taking absolute values in all of the above we get the following estimate

$$
\ll\left|\sum_{r\in R}\left(|T_r(M)|+\sum_{n=1}^{M}|e(\operatorname{tr}(\psi_r(\sigma(n))))-e(\operatorname{tr}(\psi_r(\sigma(n+1))))||T_r(n)|\right)\right|.
$$

Since the trace is a linear functional we get

$$
\frac{d}{dx}\operatorname{tr}(f(x))=\operatorname{tr}\left(\frac{df}{dx}\right).
$$

We notice that $\operatorname{tr}(a)\ll|a|$ for $a\in\mathbb{C}$ and $|\sigma(n)-\sigma(n+1)|\ll N^{1/2}$ for $1<n\leq M$, so by applying the mean value theorem we get

$$
|e(\operatorname{tr}(\psi_r(\sigma(n))))-e(\operatorname{tr}(\psi_r(\sigma(n+1))))|\ll|k|\sum_{i=s+1}^{d'}|\gamma_i|N^{i/2-1}\ll|k|\frac{(\log N)^H}{N},
$$

where $H = \max\{H_i \mid i = 1, 2, \ldots, s\}$. Combining this with the above calculation we get

$$\sum_{n=1}^{M} \sum_{r \in R} e(\mathrm{tr}(f(\sigma(n)k + r))) \ll \sum_{r \in R} \left(|T_r(M)| + |k| \frac{(\log N)^H}{N} \sum_{n=1}^{M} |T_r(n)| \right). \quad (6.10)$$

In order to complete the proof we need only show that

$$|T_r(n)| \ll \frac{N}{|k|(\log N)^{G+H}}. \quad (6.11)$$

Since $T_r(n)$ is an exponential sum with n terms, $|T_r(n)|$ is at most n, so we can assume that

$$n \gg \frac{N}{|k|(\log N)^{G+H}}. \quad (6.12)$$

We split the estimation of $T_r(n)$ into two cases, according to whether or not there exists a and q relatively prime such that

$$(\log N)^{H'} \leq |q|^2 \leq N^s (\log N)^{-H'} \quad (6.13)$$

and

$$\left| k^s \alpha_s - \frac{a}{q} \right| \leq \frac{1}{|q|^2},$$

with $H' = 2^{3(s+2)} + 2^{s+3}(G + H) + sK$.

Assuming such a and q exist we first notice that $n \leq M = |I| \leq |D| \leq C_2 N^2$. In other words, n is bounded from above by the number of integer points in the disc D, which in turn is clearly bounded by the area of the disc up to some constant. This gives us the upper bound

$$n \leq C_2 N^2,$$

for some explicit constant C_2. Taking logarithms now yield

$$\log n \leq \log C_2 + 2 \log N.$$

We let $h' = 2^{3(s+2)} + 2^{s+2}(G + H)$. By taking h' powers we obtain

$$(\log n)^{h'} \leq (\log C_2 + 2 \log N)^{h'}. \quad (6.14)$$

Now from the definition of H' and (6.13) it follows that we have

$$(\log n)^{h'} \leq (\log N)^{H'} \leq |q|^2.$$

At the same time we get from (6.12)

$$C_1 \frac{N}{(\log N)^{G+H}} \leq n,$$

for some explicit constant C_1, where we absorbed $|k|$ into it. Taking s powers we get

$$\frac{C_1^s N^s}{(\log N)^{s(G+H)}} \leq n^s,$$

and using (6.14) we get

$$\frac{C_1^s N^s}{(\log N)^{s(G+H)} (\log C_2 + \log N)^{h'}} \le \frac{n^s}{(\log n)^{h'}}.$$

Now from the definition of H' and h' it follows that

$$\frac{C_1^s N^s}{(\log N)^{s(G+H)} (\log C_2 + \log N)^{h'}} \ge \frac{N^s}{(\log N)^{H'}}.$$

Combining these last two inequalities and using (6.13) yields that

$$|q|^2 \le \frac{N^s}{(\log N)^{H'}} \le \frac{n^s}{(\log n)^{h'}}$$

Lemma 6.7 now implies that

$$\begin{aligned} |T_r(n)| &\ll |n(\log n)^{-(G+H)}| \\ &\ll \frac{N}{|k|(\log N)^{G+H}}. \end{aligned}$$

And so (6.11) holds subject to assumption (6.13). Assume now that there are no a, q such that (6.13) holds. Then by Lemma 6.3 we find a, q such that $0 \le |q|^2 \le N^s(\log N)^{-H'}$ we have that

$$|k^s\alpha_s - \frac{a}{q}| < \frac{2d^{3/2}}{|q|\sqrt{(N^s(\log N)^{-H'}}} = \frac{2d^{3/2}(\log N)^{H'/2}}{|q|N^{s/2}}. \tag{6.15}$$

Since (6.13) does not hold we must have $1 \le |q|^2 \le (\log N)^{H'}$ and therefore we have

$$|k^s q|^2 = |k|^{2s}|q|^2 \le (\log N)^{sK}(\log N)^{H'} \le (\log N)^{K_s^*}.$$

Dividing (6.15) by $|k|^s$ we get

$$|\alpha_s - \frac{a}{k^s q}| \le \frac{2d^{3/2}(\log N)^{H'/2}}{|k^s q|N^{s/2}} \le \frac{(\log N)^{H_s^*}}{|k^s q|N^{s/2}},$$

where H_s^* is chosen so the last inequality holds. But this contradicts the assumption on α_s in the proposition, thus (6.11) holds.

Now using (6.9) and (6.10) we obtain

$$\begin{aligned} \sum_{|z|^2 \le N} e(\operatorname{tr}(f(z))) &\ll \sum_{r\in R} \left(|T_r(M)| + |k|\frac{(\log N)^H}{N} \sum_{n=1}^{M} |T_r(n)| \right) + \sqrt{N}|k| \\ &\ll \sum_{r\in R} \left(\frac{N}{|k|(\log N)^{G+H}} + \frac{1}{(\log N)^G} M \right) + \sqrt{N}|k| \\ &\ll \frac{N}{(\log N)^G}, \end{aligned}$$

which proves the proposition. □

CHAPTER 7

PROOF OF THE MAIN THEOREM

We dedicate, as indicated by the name, this entire chapter to proving the main theorem. We start out with defining some essential parameters and prove useful connections between them. Using these parameters we rewrite the problem into one of estimating exponential sums. Finally we use Proposition (6.1) or Lemma (6.7) to calculate these exponential sums, depending on their type. In order to do so we will need to consider sums according to the b-adic length of their arguments. Essentially we divide these into three cases; those of short length, those of medium length and those of long length, since a different technique is required for each case.

From this point on we fix N and the block $d_1 \dots d_l$. We also define a by

$$a = \sum_{i=1}^{l} d_i b^{l-i} \tag{7.1}$$

We restate the main theorem of this part of the thesis here to remind us what we need to show.

Theorem 7.1. *Let $f(z) = \alpha_n z^n + \cdots + \alpha_1 z + \alpha_0$ be a polynomial with coefficients in $\mathbb{C}$. Let $(b, \mathcal{D})$ be a CNS in $\mathcal{O}_d$, where $d \in \{-1, -2, -3, -7, -11\}$. Then for every $l \geq 1$ we have*

$$\sup_{d_1 \dots d_l} \left| \frac{1}{N} \mathcal{N}(\theta(f); d_1 \dots d_l; N) - \frac{1}{|\mathcal{D}|^l} \right| \ll (\log N)^{-1}$$

7.1 Essential parameters

Let m be the unique positive integer such that

$$\sum_{n \leq m-1} \ell(f(z_n)) < N \leq \sum_{n \leq m} \ell(f(z_n)), \tag{7.2}$$

where $z_n = \tau^{-1}(n-1)$ for $n \geq 1$. We let M and J denote the maximum norm and the maximum length of the $(b, \mathcal{D})$-ary expansion of $\lfloor f(z_n) \rfloor$ for $n \geq 1$, respectively. In other words,

$$M = \max_{n \leq m} |z_n|^2, \quad J = \max_{n \leq m} \ell(f(z_n)).$$

We now use Lemma 4.2 to obtain a relationship between m and M.

$$\begin{aligned} |\log_{|b|^2} \max_{n \leq m} |z_n|^2 - \ell(\max_{n \leq m} z_n)| &= |\log_{|b|^2} M - \ell(z_m)| \\ &= |\log_{|b|^2} M - \lfloor \log_{|b|^2} m \rfloor| \leq c. \end{aligned}$$

Hence we have

$$M \asymp m,$$

where $\asymp$ means both $\ll$ and $\gg$ holds simultanously.

Now we consider the connection between M and J and we start by noting that $|f(z)| \asymp |z|^{d'}$. For some absolute constant c_1, Lemma 4.2 implies that

$$\begin{aligned}|\log_{|b|^2} \max_{n\le m}|f(z_n)|^2 - J| &= |\log_{|b|^2} \max_{n\le m}|z_n|^{2d'} - J| + c_1\\ &= |\log_{|b|^2} M^{d'} - J| + c_1,\end{aligned}$$

and so

$$M \asymp |b|^{2J/d'} \le c.$$

From (7.2) we get the following relation between M and N:

$$N = mJ + \mathcal{O}(m) = c_0 M \log_{|b|^2} M + \mathcal{O}(M),$$

where c_0 is a positive constant depending only on d and b.

Now we want to split the sum on the right hand side of (7.2) into parts where $f(z_n)$ has the same b-ary length. Let $I_l, I_{l+1}, \ldots, I_J \subset \{1,2,\ldots,m\}$ be a partition of the set $\{1,2,\ldots,m\}$ such that

$$n \in I_j \Leftrightarrow \ell(f(z_n)) = j.$$

In order to estimate the size of these subsets we define some more parameters. Define M_j, where $j = l, l+1, \ldots, J$ to be the least integer such that any $z \in \mathbb{C}$ of norm greater than or equal to M_j has length at least j, i.e.,

$$M_j = \max_{\ell(z)<j}|z|^2 = \max_{n<|b|^{2(j-1)}}|z_n|^2.$$

Using the same arguments as we did when we found a relation between M and b we now get $M_j \asymp |b|^{2j/d'}$. Finally define X_j by

$$X_j = M - M_j. \tag{7.3}$$

7.2 Rewriting the problem

Using the parameters we defined above we can rewrite our problem. Let $\mathcal{N}(f(z_n))$ be the number of occurences of the block $d_1 \ldots d_l$ in the b-ary expansion of the integer part of $f(z_n)$. Then we have

$$|\mathcal{N}(\theta_q(f); d_1 \ldots d_l, N) - \sum_{n\le m} \mathcal{N}(f(z_n))| \le 2lm.$$

So, it suffices to show

$$\sum_{n\le m} \mathcal{N}(f(z_n)) = \frac{N}{|\mathcal{D}|^l} + \mathcal{O}\left(\frac{N}{\log N}\right). \tag{7.4}$$

To count the occurences of $d_1 \ldots d_l$ in $\lfloor f(z_n) \rfloor$ properly we use the indicator function of $\mathcal{F}_a$, where a is the number defined in (7.1) and $\mathcal{F}_a$ is then defined as in (5.1).

For a fixed $n \in \{1,2,\ldots,m\}$ we write out $f(z_n)$ in its b-ary expansion, i.e.,

$$f(z_n) = a_r b^r + \cdots + a_0 + a_{-1} b^{-1} + \ldots,$$

where $a_i \in \mathcal{D}$ for $i = r, r-1, \dots$, we see that

$$\mathcal{I}(\{b^{-j-1}f(z_n)\}) = 1 \Leftrightarrow d_1 \dots d_l = a_{j-1} \dots a_{j-l},$$

where $\{x\}$ denotes the fractional part of x. Now since every I_j, where $l \leq j \leq J$, consist of exactly those n such that $f(z_n)$ has b-ary expansion of length at least j, we get

$$\sum_{n \leq m} \mathcal{N}(f(z)) = \sum_{l \leq j \leq J} \sum_{n \in I_j} \mathcal{N}(f(z)) = \sum_{l \leq j \leq J} \sum_{n \in I_j} \mathcal{I}\left(\left\{\frac{f(z_n)}{b^{j+1}}\right\}\right).$$

For every j there may be elements $z \in \mathcal{O}_d$ with $|z|^2 < M_j$ but $\ell(z) \geq j$. We want to ignore these in the calculations to come, so we need to show that they provide an error-term which is small enough to be disregarded. From Lemma 4.2

$$\sum_{n \in I_j} 1 = \sum_{|z_n|^2 < M_j} 1 + \sum_{M_j \leq |z_n|^2 < M} 1.$$

Lemma 4.1 now gives us that that the first sum can be considered as an error-term and so we are left with

$$\sum_{M_j \leq |z_n|^2 < M} 1.$$

Thus we can assume that no z has $\ell(z) \geq j$ and $|z|^2 < M_j$. In order to estimate $\mathcal{I}(z)$ we make use of various sets and functions introduced in Lemma 5.2 and we note that $\mathcal{F}_a$ can be covered by a set $I_{k,a}$ and an axis-parallel tube $P_{k,a}$. Thus the question at hand is reduced to estimating how often the sequence $\{b^{-j-1}f(z_n)\}_{n \in I_j}$ hits each of these sets. The first one $I_{k,a}$ is characterized by the Urysohn function f_a also introduced in Lemma 5.2. As for the axis-parallel tube $P_{k,a}$ we define

$$\mathcal{E}_j = \#\left\{n \in I_j \,\middle|\, \varphi\left(\frac{f(z_n)}{b^{j+1}}\right) \in P_{k,a}\right\},$$

where $\#\{A\}$ denotes the amount of elements in the set A. For every $j \in \{l, l+1, \dots, J\}$ we get

$$\sum_{n \in I_j} \mathcal{I}\left(\frac{f(z_n)}{b^{j+1}}\right) = \sum_{n \in I_j} f_a\left(\varphi\left(\frac{f(z_n)}{b^{j+1}}\right)\right) + \mathcal{O}(\mathcal{E}_j). \tag{7.5}$$

We consider the terms on the right right hand side seperately. We start with f_a, from Lemma 5.3 we get

$$f_a\left(\varphi\left(\frac{f(z_n)}{b^{j+1}}\right)\right) = |b|^{-2\ell(a)} + \sum_{\mathbf{0} \neq \boldsymbol{v} \in \mathbb{Z}^2} c(v_1, v_2) e\left(\boldsymbol{v} \cdot_d \varphi\left(\frac{f(z_n)}{b^{j+1}}\right)\right), \tag{7.6}$$

where $\boldsymbol{v} = (v_1, v_2)$ and $c(\cdot, \cdot)$ is the Fourier coefficients of f_a. We split the sum into those $\boldsymbol{v}$ with $|\boldsymbol{v}|_\infty \leq \Delta^{-1}$ and the rest. For $|\boldsymbol{v}|_\infty > \Delta^{-1}$ we can apply Lemma 5.3 and estimate the exponential function trivally to get

$$\begin{aligned}&\sum_{n \in I_j} f_a\left(\varphi\left(\frac{f(z_n)}{b^{j+1}}\right)\right) \\ &\ll \frac{X_j}{|b|^{2l}} + X_j \mu^k \Delta^2 + \mu^k \sum_{\mathbf{0} < |\boldsymbol{v}|_\infty \leq \Delta^{-1}} \frac{1}{r(\boldsymbol{v})} \sum_{n \in I_j} e\left(\boldsymbol{v} \cdot_d \varphi\left(\frac{f(z_n)}{b^{j+1}}\right)\right).\end{aligned} \tag{7.7}$$

Here, X_j is as defined in (7.3). To estimate $\mathcal{E}_j$ we use Lemma 5.2 to split $P_{k,a}$ into a family $\boldsymbol{R}_j$ of μ^k rectangles. Since the discrepancy is defined on a rectangle we can now use the Erdős–Turán–Koksma-inequality to get

$$\begin{aligned}\mathcal{E}_j &\ll \sum_{R\in\boldsymbol{R}_j} X_j\lambda_2(R) + X_j D_{X_j}(\{x_n\}) \qquad (7.8)\\ &\ll X_j \sum_{R\in\boldsymbol{R}_j}\left(\lambda_2(R) + \frac{2}{H+1} + \sum_{0<|\boldsymbol{v}|_\infty\le H}\frac{1}{r(\boldsymbol{v}}\left|\frac{1}{X_j}\sum_{n\in I_j} e\left(\boldsymbol{v}\cdot_d\varphi\left(\frac{f(z_n)}{b^{j+1}}\right)\right)\right|\right).\end{aligned}$$

Even if some of the rectangles in $\boldsymbol{R}_j$ overlap, it follows from Lemma 5.2 property (3) that

$$\sum_{R\in\boldsymbol{R}_j}\lambda_2(R) \ll \left(\frac{\mu}{|b|^2}\right)^k.$$

So we can simplify (7.8) to

$$\begin{aligned}\mathcal{E}_j \ll X_j&\left(\left(\frac{\mu}{|b|^2}\right)^k + \frac{\mu^k}{H+1}\right. \qquad (7.9)\\ &\left.+\frac{\mu^k}{X_j}\sum_{0<|\boldsymbol{v}|_\infty\le H}\frac{1}{r(\boldsymbol{v})}\sum_{n\in I_j} e\left(\boldsymbol{v}\cdot_d\varphi\left(\frac{f(z_n)}{b^{j+1}}\right)\right)\right).\end{aligned}$$

As we can see, the exponential sums in (7.7) and (7.9) are of the same type, so we define

$$S(\boldsymbol{v},j) = \sum_{n\in I_j} e\left(\boldsymbol{v}\cdot_d\varphi\left(\frac{f(z_n)}{b^{j+1}}\right)\right). \qquad (7.10)$$

We now insert (7.7), (7.9) and (7.10) into (7.5) in order to obtain

$$\begin{aligned}&\sum_{n\in I_j}\mathcal{I}\left(\frac{f(z_n)}{b^{j+1}}\right)\\ &\ll \frac{X_j}{|b|^{2l}} + X_j\mu^k\Delta^2 + \mu^k\sum_{\boldsymbol{0}<|\boldsymbol{v}|_\infty\le\Delta^{-1}}\frac{1}{r(\boldsymbol{v})}S(\boldsymbol{v},j)\\ &+X_j\left(\left(\frac{\mu}{|b|^2}\right)^k + \frac{\mu^k}{H+1} + \frac{\mu^k}{X_j}\sum_{0<|\boldsymbol{v}|_\infty\le H}\frac{1}{r(\boldsymbol{v})}S(\boldsymbol{v},j)\right).\end{aligned}$$

By rearranging we get

$$\begin{aligned}&\left|\sum_{n\in I_j}\mathcal{I}\left(\frac{f(z_n)}{b^{j+1}}\right) - \frac{X_j}{|b|^{2l}}\right| \qquad (7.11)\\ &\ll X_j\left(\mu^k\Delta^2 + \frac{\mu^k}{H+1} + \left(\frac{\mu}{|b|^2}\right)^k\right) + 2\sum_{0<|\boldsymbol{v}|_\infty\le H}\frac{\mu^k}{r(\boldsymbol{v})}S(\boldsymbol{v},j).\end{aligned}$$

The next thing we want to do is to transfer the exponential sum from $\mathbb{Z}^2$ to $\mathcal{O}_d$ and we use the same idea as Gittenberger and Thuswaldner. Hence we let V denote the Vandermonde matrix
$$V = \begin{pmatrix} 1 & 1 \\ b & \bar{b} \end{pmatrix}.$$
Now let
$$\omega(z) = (\mathrm{tr}(z), \mathrm{tr}(bz))^T = W\varphi(z),$$
where T denotes the transposed vector and $W = VV^T$. Using this notation we get
$$\boldsymbol{v} \cdot \varphi\left(\frac{f(z)}{b^{j+1}}\right) = \boldsymbol{v}W^{-1}\omega\frac{f(z)}{b^{j+1}} = \mathrm{tr}\left((\tilde{v}_1 + b\tilde{v}_2)\frac{f(z)}{b^{j+1}}\right),$$
where $(\tilde{v_1}, \tilde{v_2}) = \boldsymbol{v}W^{-1}$. Using that $|I_j| \asymp X_j$ and the definition of X_j we can now transfer (7.10) to

$$\begin{aligned} S(\boldsymbol{v}, j) &= \sum_{n\in I_j} e\left(\mathrm{tr}\left((\tilde{v}_1 + b\tilde{v}_2)\frac{f(z_n)}{b^{j+1}}\right)\right) \qquad (7.12) \\ &\ll \sum_{M_j \le |z_n|^2 < M_j + X_j} e\left(\mathrm{tr}\left((\tilde{v}_1 + b\tilde{v}_2)\frac{f(z_n)}{b^{j+1}}\right)\right). \end{aligned}$$

We assume that k and H, which we will choose later on depending on j, are such that $\Delta^{-1}, H \ll \log N$. This is possible since Δ depends on k as seen in (5.2). Now we will want to consider the sums $S(\boldsymbol{v}, j)$ according to the size of j. We therefore split the size of j into three intervals

$$l \le j \le l + C_l \log\log N, \qquad (7.13)$$
$$l + C_l \log\log N < j \le J - C_u \log\log N, \qquad (7.14)$$
$$J - C_u \log\log N < j \le J, \qquad (7.15)$$

where C_l and C_u are sufficiently large constants. We refer to there as short, medium and long b-ary expansions, respectively.

7.3 Estimation of medium length b-ary expansions

We start with the easiest of the three cases and therefore consider those j satisfying (7.14). First we assume that there are integers $a, q \in \mathcal{O}_d$ satisfying
$$\left|\frac{\tilde{v}_1 + b\tilde{v}_2}{b^j}\alpha_{d'} - \frac{a}{q}\right| \le \frac{1}{|q|^2}, \qquad (7.16)$$
and
$$(\log X_j)^H \le |q|^2 \le X_j^{d'}(\log X_j)^{-H},$$
with $G = 3$ and $H = 2^{d'+2}G + 2^{3(d'+2)}$. Then we apply Lemma 6.7 to get
$$S(\boldsymbol{v}, j) \ll X_j(\log X_j)^{-G}.$$

Now if j satisfies (7.14) then (7.16) holds. Since by Lemma 6.3 there are $a, q \in \mathcal{O}_d$ such that

$$(a,q) = 1, \quad 1 \leq |q|^2 \leq X_j^{d'} (\log X_j)^{-H},$$

and

$$\left| \frac{\tilde{v}_1 + b\tilde{v}_2}{b^j} \alpha_{d'} - \frac{a}{q} \right| \leq \frac{(\log X_j)^H}{|q| X_j^{d'/2}} \leq \frac{1}{|q|^2}.$$

We now split this case into two subcases depending on the size of $|q|^2$. Assume first that $2 \leq |q|^2 \leq (\log X_j)^H$. Then we have

$$\left| \frac{\tilde{v}_1 + b\tilde{v}_2}{b^j} \alpha_{d'} \right| > \frac{1}{|q|} - \frac{1}{|q|^2} \geq \frac{1}{2|2|} \gg (\log X_j)^{-H},$$

and thus by the assumption $\Delta^{-1} \ll \log N$ we must have

$$|b|^j \ll |(\tilde{v}_1 + b\tilde{v}_2)\alpha_{d'}| (\log X_j)^H \ll (\log N)(\log X_j)^H,$$

Taking logarithms and dividing by $\log|b|$ we get a contradiction with (7.14) for sufficiently large C_l. For $z \in \mathbb{C}$ we write

$$\|z\| = \min_{n \in \mathbb{Z}} ||z|^2 - n|.$$

Now, for $q \in \mathcal{O}_d$, if $|q|^2 = 1$ then we must have $q = 1$ and $\|(\tilde{v}_1 + b\tilde{v}_2)\alpha_{d'}\| < X_j^{d'} (\log X_j)^{-2H}$. If $|(\tilde{v}_1 + b\tilde{v}_2) b^{-j} \alpha_{d'}|^2 > \sqrt{2}/2$ then we get

$$|b|^{2j} \ll |(\tilde{v}_1 + b\tilde{v}_2)\alpha_{d'}| \ll \log N,$$

this contradicts (7.14) for C_l large enough. On the other hand, if $|(\tilde{v}_1 + b\tilde{v}_2) b^{-j} \alpha_{d'}|^2 < \sqrt{2}/2$ we get

$$|(\tilde{v}_1 + b\tilde{v}_2) b^{-j} \alpha_{d'}|^2 = \left\|(\tilde{v}_1 + b\tilde{v}_2) b^{-j} \alpha_{d'}\right\| < X_j^{d'} (\log X_j)^{-2H},$$

which implies that

$$|b|^{2j} \gg |(\tilde{v}_1 + b\tilde{v}_2)\alpha_{d'}|^2 X_j^{d'} (\log X_j)^{-2H}.$$

This contradicts (7.14) for C_u sufficiently large. Thus for j such that (7.14) holds we get

$$S(\boldsymbol{v}, j) \ll X_j (\log X_j)^{-G}. \tag{7.17}$$

Inserting this into (7.11) gives us

$$\left| \sum_{n \in I_j} \mathcal{I}\left(\frac{f(z_n)}{b^{j+1}} \right) - \frac{X_j}{|b|^{2l}} \right| \ll \tag{7.18}$$

$$X_j \left(\mu^k \Delta^2 + \frac{\mu^k}{V+1} + \left(\frac{\mu}{|b|^2} \right)^k + \frac{\mu^k}{(\log X_j)^3} \left\{ \sum_{0 < |\boldsymbol{v}|_\infty \leq \Delta^{-1}} + \sum_{0 \leq |\boldsymbol{v}|_\infty \leq V} \right\} \frac{1}{r(\boldsymbol{v})} \right).$$

Now we can choose k and H so that $\Delta^{-1}, H \ll \log N$ as we claimed earlier. For if j satisfies (7.14), using the definition of Δ given in (5.2) we set

$$k = C_k \log\log X_j, \quad H = \mu^k \log X_j, \quad \text{and } \Delta^{-1} = \frac{(\log X_j)^{C_k \log|b|}}{2c_\Delta}. \tag{7.19}$$

Where C_k is an arbitrary constant. Furthermore we define $C_\mu > 1$ to be such that $C_\mu \mu = |b|^2$. Then we have for those j satisfying (7.14) that

$$\left|\sum_{n\in I_j}\mathcal{I}\left(\frac{f(z_n)}{b^{j+1}}\right)-\frac{X_j}{|b|^{2l}}\right| \ll X_j((\log X_j)^{-1}+(\log X_j)^{-2}(\log\log X_j)^2) \ll \frac{X_j}{j}. \tag{7.20}$$

We now move on to prove the same estimate for larger and smaller j.

7.4 Estimation related to long b-ary expansions

In this section we focus entirely on those j satisfying (7.15). We will start with the same assumptions on Δ^{-1} and H, i.e., $\Delta^{-1}, H \ll \log N$. For those j where (7.16) holds, we get by Lemma 6.7

$$S(\boldsymbol{v}, j) \le X_j(\log X_j)^{-G}.$$

If (7.16) does not hold, then for every j satisfying (7.15) with $|b|^{j/d'} \ll X_j \ll |b|^{J/d'}$, we have

$$0 \ll |\tilde{v}_1 + b\tilde{v}_2||b|^{-j/2d'} \ll |f'(z)| \ll |\tilde{v}_1 + b\tilde{v}_2||b|^{J-j-j/2d'} \ll |\tilde{v}_1 + b\tilde{v}_2||b|^{-j/2d'}(\log N)^{\tilde{C}_2}. \tag{7.21}$$

We will use these inequalities to apply Lemma 4.8 with

$$F = \operatorname{tr}\left((\tilde{v}_1 + b\tilde{v}_2)\frac{f(z_n)}{b^{j+1}}\right),$$

$m = |\tilde{v}_1 + b\tilde{v}_2||b|^{-j/d'}$, and $\delta = |\tilde{v}_1 + b\tilde{v}_2||b|^{-j/d'}(\log N)^{\tilde{C}_2}$. So for j as in (7.15) and for $\sigma = 2G$ we get

$$S(\boldsymbol{v}, j) \ll \frac{\sqrt{X_j}}{|\tilde{v}_1 + b\tilde{v}_2||b|^{-j/d'}} + \frac{X_j}{(\log X_j}^{\sigma/2} + s\frac{3-\delta}{1-\delta}\sqrt{X_j(\log X_j)^\sigma} \ll \frac{\sqrt{X_j}|b|^{j/d'}}{\tilde{v}_1 + b\tilde{v}_2|} + \frac{X_j}{(\log X_j)^G}. \tag{7.22}$$

We now insert this into (7.11) which yields

$$\left|\sum_{n\in I_j}\mathcal{I}\left(\frac{f(z_n)}{b^{j+1}}\right)-\frac{X_j}{|b|^{2l}}\right| \ll X_j\left(\mu^k\Delta^2 + \frac{2\mu^k}{H+1} + \left(\frac{\mu}{|b|^2}\right)^k + \frac{\mu^k}{X_j}\left\{\sum_{0<\|\boldsymbol{v}\|_\infty\le\Delta^{-1}} + \sum_{0<\|\boldsymbol{v}\|_\infty\le H}\right\}\frac{1}{r(\boldsymbol{v})}\left(\frac{\sqrt{X_j}|b|^{j/d'}}{|\tilde{v}_1 + b\tilde{v}_2} + \frac{X_j}{(\log X_j)^3}\right)\right). \tag{7.23}$$

Now we define k, H and Δ^{-1} as follows

$$k = \max\left(1, \frac{1/2\log X_j + \log 4C_\Delta^2 - j/d'\log|b|}{\log C_\mu}\right), \quad H = \mu^k \log X_j, \quad \Delta^{-1} = \frac{|b|^k}{2c_\Delta}.$$

Which gives us

$$\mu^k\Delta^2 = \frac{|b|^{j/d'}}{\sqrt{X_j}}, \quad \mu^k \le |b|^{2k} \ll \left(\frac{X_j}{|b|^{2j/d'}}\right)^{\frac{\log|b|}{\log C_\mu}}, \quad \left(\frac{\mu}{|b|^2}\right)^k = \frac{1}{C_\mu^k} \ll \frac{|b|^{j/d'}}{\sqrt{X_j}}.$$

We also get the following estimate

$$|\tilde{v}_1 + b\tilde{v}_2| = |(1,b)(v_1,v_2)^t W^{-1}| \gg |(v_1,v_2)^t| \gg \sqrt{v_1 v_2}.$$

By inserting the above into (7.23) gives us

$$\left|\sum_{n\in I_j} \mathcal{I}\left(\frac{f(z_n)}{b^{j+1}}\right) - \frac{X_j}{|b|^{2l}}\right| \tag{7.24}$$
$$\ll \sqrt{X_j}|b|^{j/d'} + \frac{X_j}{j} + \left(\frac{X_j}{|b|^{2j/d'}}\right)^{\frac{\log|b|}{\log C_\mu}} (\sqrt{X_j}|b|^{j/d'} + X_j(\log X_j)^{-3}),$$

for j satisfying (7.15).

7.5 Iterative estimates for the short b-ary expansions

We now consider the last case, namely those j satisfying (7.13). This is the hardest case and will take some time. Notice that in this case our assumptions on H and Δ^{-1} we have

$$|\tilde{v}_1 + b\tilde{v}_2| \asymp |b|^j.$$

We use the ideas of Nakai and Shiokawa, namely applying Proposition 6.1 iteratively. If there is no such s as assumed in the proposition, we instead apply Lemmata 4.8 and 4.9.

By we assumption in (7.13) we immediately get

$$|b|^j \le (\log N)^{C_l \log|b| + o(1)}. \tag{7.25}$$

We now define g to be the polynomial

$$g(z) = \frac{\tilde{v}_1 + b\tilde{v}_2}{b^j} f(z),$$

so it has the coefficients

$$\beta_i = \frac{\tilde{v}_1 + b\tilde{v}_2}{b^j}\alpha_i, \quad i = 0,1,\ldots,d'. \tag{7.26}$$

Now we start applying Proposition 6.1. We first assume that $1 \le s < d'$. Then we set

$$H_{d'} = H_{d'}^* + C_1\log|b| + 1, \quad H_{d'}^* = 2^{3(d'+2)} + 2^{d'+3}G,$$

and inductively define H_r^*, H_r, and h_r, where $1 \le r < d'$, by the following

$$H_r^* = 2^{3(r+2)} + 2^{r+3}(G + H_{r+1}) + 2r\sum_{i=r+1}^{d'} H_r,$$
$$H_r = H_r^* + 2(C_1\log|b| + 1),$$

and

$$h_r = H_r^* + C_1 \log|b| + 1.$$

Now let j satisfy (7.13 and assume that there are coprime pairs $(a_{d'}, q_{d'}), \ldots, (a_{s+1}, q_{s+1})$ of integers, i.e., pairs of elements from $\mathcal{O}_d$ such that

$$1 \leq |q|^2 \leq (\log X_j)^{2h_r} \quad \text{and} \quad \left|\alpha_r - \frac{a_r}{q_r}\right| \leq \frac{(\log X_j)^{h_r}}{|q_r| X_j^{r/2}} \quad (s < r \leq d'),$$

but there is no pair (a_s, q_s) such that

$$1 \leq |q_s|^2 \leq (\log X_j)^{2h_s} \quad \text{and} \quad \left|\alpha_s - \frac{a_s}{q_s}\right| \leq \frac{(\log X_j)^{h_s}}{|q_s| X_j^{s/2}}.$$

We denote the set of all j satisfying these conditions by $\mathbb{J}_s$. It follows from (7.25) and (7.26) that for every $j \in \mathbb{J}_s$ we have

$$1 \leq |b^j q_r| \leq (\log X_j)^{2H_r} \quad \text{and} \quad \left|\beta_r - \frac{(\tilde{v}_1 + b\tilde{v}_2)a_r}{b^j q_r}\right| \leq \frac{(\log X_j)^{H_r}}{|b^j q_r| X_j^{r/2}},$$

for $s < r \leq d'$, but there is no pair (A_s, Q_s) of coprime integers such that

$$1 \leq |Q_s| \leq (\log X_j)^{2H_s^*} \quad \text{and} \quad \left|\beta_r - \frac{A_s}{Q_s}\right| \leq \frac{(\log X_j)^{H_s^*}}{|Q_s| X_j^{s/2}},$$

since if there were such A_s and Q_s, we would have

$$1 \leq |(\tilde{v}_1 + b\tilde{v}_2)Q_s|^2 \leq (\log X_j)^{2H_s^* + t} \leq (\log X_j)^{2h_s},$$

together with (7.25) we end up with

$$\left|\alpha_s - \frac{b^j A_s}{\tilde{v}_1 + b\tilde{v}_2 Q_s}\right| \leq \frac{(\log X_j)^{H_s^* + C_1 \log|b| + 1}}{|(\tilde{v}_1 + b\tilde{v}_2)Q_s| X_j^{s/2}} \leq \frac{(\log X_j)^{h_s}}{|\tilde{v}_1 + b\tilde{v}_2)Q_s| X_j^{s/2}},$$

which contradicts the assumption that $j \in \mathbb{J}_s$. So we can apply Proposition 6.1 with H_i, H_i^* and $K_i = 2H_i$, $K_i^* = 2H_i^*$ which gives us

$$S(\boldsymbol{v}, j) \ll X_j (\log X_j)^{-G},$$

for all $j \in \mathbb{J}_1 \bigcup \cdots \bigcup \mathbb{J}_d$. We denote by $\mathbb{J}_0$ the set of all positive integers j satisfying (7.13) with $j \notin \mathbb{J}_1 \bigcup \cdots \bigcup \mathbb{J}_d$. It remains to estimate $S(\boldsymbol{v}, j)$ for such j. In order to do so we shall apply Lemmata 4.8 and 4.9. For $j \in \mathbb{J}_0$ there exist coprime pairs (a_r, q_r) of integers such that

$$1 \leq |q_r|^2 \leq (\log X_j)^{2h_r} \quad \left|\alpha_r - \frac{a_r}{q_r}\right| \leq \frac{(\log X_j)^{h_r}}{|q_r| X_j^{r/2}} \quad \text{for } 1 \leq r \leq d'.$$

We now set $\Omega_r = \alpha_r - a_r/q_r$ for $r = 1, \ldots, d'$, and $a = \gcd(a_1, \ldots, a_{d'})$ and $q = \operatorname{lcm}(q_1, \ldots, q_{d'})$. Furthermore, for $r = 1, \ldots, d'$ we define c_r by the equation

$$\frac{a_r}{q_r} = \frac{a}{q} c_r$$

This allows us to rewrite the exponential sum in the following way

$$\begin{aligned} S(\boldsymbol{v}, j) &= \sum_{n\in I_j} e\left(\operatorname{tr}\left((\tilde{v}_1 + b\tilde{v}_2)\frac{f(z_n)}{b^{j+1}}\right)\right) \\ &= \sum_{\lambda\in r(b^{j+1}q)} e\left(\operatorname{tr}\left(\frac{\hat{v}a}{b^{j+1}q}\sum_{k=1}^{d'} c_k\lambda^k\right)\right) \\ &\times \sum_{\substack{\mu \\ \exists n\in I_j : \mu q+\lambda = z_n}} e\left(\operatorname{tr}\left(\frac{\hat{v}}{b^{j+1}}\sum_{k=1}^{d'} \Omega_k(\mu q+\lambda)^k\right)\right), \end{aligned}$$

where $r(b^{j+1})$ denotes a complete system of residues modulo $b^{j+1}q$ and $\hat{v} = \tilde{v}_1 + b\tilde{v}_2$.

We start by considering the second sum. Let $R_0 = R_0(j,q) = \{(b^{j+1}q)\cdot(\alpha+\beta_i) : 0 \le \alpha, \beta \le 1\}$ and let T_0 be the set of translation vectors such that R_0 tiles $\mathbb{Z}^2$, i.e., $T_0 = \{(b^{j+1}q)z : z \in \mathbb{Z}[i]\}$. Furthermore we define

$$T = \{t \in T_0 \mid (R_0 + t) \cap \{z_n \mid n \in I_j\} \neq \emptyset\}. \tag{7.27}$$

We then have $|T| \ll X_j|b^{j+1}q|^{-2}$. Furthermore, we let $\mathcal{T}$ denote the area covered by the translates of R_0 by elements of T, i.e.,

$$\mathcal{T} = \bigcup_{t\in T}(R_0 + t).$$

For a fixed $\lambda \in R_0 \cap \mathbb{Z}[i]$ we then get

$$\begin{aligned} &\sum_{\substack{\mu \\ \exists n\in I_j : \mu q+\lambda = z_n}} e\left(\operatorname{tr}\left(\frac{\hat{v}}{b^{j+1}}\sum_{k=1}^{d'} \Omega_k(\mu q+\lambda)^k\right)\right) \\ &\le \sum_{\mu\in T} e\left(\operatorname{tr}\left(\frac{\hat{v}}{b^{j+1}}\sum_{k=1}^{d'} \Omega_k(\mu q+\lambda)^k\right)\right). \end{aligned}$$

We now define F_λ by

$$F_\lambda(x,y) = e\left(\operatorname{tr}\left(\frac{\hat{v}}{b^{j+1}}\sum_{k=1}^{d'} \Omega_k((x+iy)q+\lambda)^k\right)\right).$$

Then we have

$$\begin{aligned} \frac{\partial F_\lambda(x,y)}{\partial x} \asymp \frac{\partial F_\lambda(x,y)}{\partial y} &\ll \frac{\hat{v}}{|b|^j}\sum_{k=1}^{d'} k|q|\frac{(\log X_j)^{H_k}}{q_k X_j^{k/2}} X_j^{(k-1)/2} \\ &\ll \frac{\hat{v}}{|b|^j} X_j^{-1/2}|q|(\log X_j)^{H_1^*}. \end{aligned}$$

We will use the same idea as in the proof of Lemma 4.8, so we first consider a single square. Let $D_\nu = \{z = x+iy \in \mathcal{O}_d : -\nu \le x, y \le \nu\}$. We now apply Lemma 4.7 which yields

$$\sum_{x+iy\in D_\nu} F_\lambda(x,y) = \sum_{x=-\nu}^{v}\sum_{y=-\nu}^{\nu} F_\lambda(x,y) = \int_{-\nu}^{\nu}\int_{-\nu}^{\nu} F_\lambda(x,y)\,dx\,dy + \mathcal{O}(\nu).$$

Now we again want to split $\mathcal{T}$ into squares. We notice that we have assumed that $|I_j| = X_j$ so we can consider I_j as an annulus, namely $\{z \in \mathbb{C} : M_j \leq |z|^2 < M\}$. So we choose a $\sigma > 0$ and tessellate $\mathcal{T}$ by squares of sidelength $\sqrt{|T|/(\log|T|)^\sigma}$. Then we glue all the the squares in the interior of $\mathcal{T}$ together and estimate their contribution on the boundary to the error term. This gives us

$$\sum_{x+iy\in T} F_\lambda(x,y) = \iint_{\mathcal{T}} F\lambda(x,y)\,dx\,dy + \mathcal{O}\left(\frac{|T|}{(\log|T|)^{\sigma/2}}\right).$$

We now put all of this together and we get

$$\begin{aligned}
S(\boldsymbol{v},j) &= \sum_{n\in I_j} e\left(\operatorname{tr}\left((\tilde{v}_1 + b\tilde{v}_2)\frac{f(z_n)}{b^{j+1}}\right)\right)\\
&= \sum_{\lambda\in r(b^j q)} e\left(\operatorname{tr}\left(\frac{\nu a}{b^j q}\sum_{k=1}^{d'} c_k\lambda^k\right)\right)\left\{\iint_{\mathcal{T}} F_\lambda(x,y)\,dx\,dy + \mathcal{O}\left(\frac{|T|}{(\log|T|)^{\sigma/2}}\right)\right\}\\
&= \sum_{\lambda\in r(b^j q)} e\left(\operatorname{tr}\left(\frac{\nu a}{b^j q}\sum_{k=1}^{d'} c_k\lambda^k\right)\right)\frac{1}{|b^{j+1}q|^2}\iint_{M_j\leq|z|^2<M} G(z)\,dz + \mathcal{O}\left(\frac{X_j}{(\log X_j)^\sigma}\right),
\end{aligned}$$

where

$$G(z) = e\left(\operatorname{tr}\left(\frac{\hat{v}}{b^{j+1}}\sum_{k=1}^{d'}\Omega_k z^k\right)\right).$$

Finally, we define quotients of integers R_i/Q, where $R_i, Q \in \mathcal{O}_d$, for $i = 1, \ldots, d'$ by

$$\frac{R_i}{Q} = \frac{\hat{v}}{b^j}\frac{ac_i}{q}.$$

We note that

$$N(\hat{v}q) = N(b^{j+1}R_i q_i/a_i) \asymp N(b^{j+1}R_i\alpha_i^{-1}) \asymp N(b^{j+1}R_i) \gg N(b^{j+1}),$$

So by estimating the integral trivially and applying Lemma 4.9 we get

$$\begin{aligned}
S(\boldsymbol{v},j) &= \sum_{n\in I_j} e\left(\operatorname{tr}\left((\tilde{v}_1 + b\tilde{v}_2)\frac{f(z_n)}{b^{j+1}}\right)\right) && (7.28)\\
&\ll \frac{|b^j q|^2}{N(Q)}(N(Q))^{1-1/d'+\epsilon}\frac{X_j}{|b^j q|^2} + \frac{X_j}{(\log X_j)^\sigma} && (7.29)\\
&\ll X_j((N(\hat{v}^{-1}b^{j+1}))^{-1/d'+\epsilon} + (\log X_j)^{/\sigma}). && (7.30)
\end{aligned}$$

By plugging this into (7.11) we obtain

$$\begin{aligned}
&\left|\sum_{n\in I_j}\mathcal{I}\left(\frac{f(z_n)}{b^{j+1}}\right) - \frac{X_j}{|b|^{2l}}\right| && (7.31)\\
&\ll X_j\left(\mu^k\Delta^2 + \frac{2\mu^k}{H+1} + \left(\frac{\mu}{|b|^2}\right)^k\right.\\
&+ \mu^k\left\{\sum_{0<\|\boldsymbol{v}\|_\infty\leq\Delta^{-1}} + \sum_{0<\|\boldsymbol{v}\|_\infty\leq H}\right\}\frac{1}{r(\boldsymbol{v})}\left((N(\hat{v}^{-1}b^{j+1}))^{-1/d'+\epsilon}\right.\\
&\left.\left.+(\log X_j)^{-\sigma}\right)\right).
\end{aligned}$$

Setting σ, k and H to the same values as in (7.19) and using (5.2) we get

$$\sigma = G \quad , k = C_k \log\log X_j, \quad H = \mu^k \log X_j, \quad \text{and } \Delta^{-1} = \frac{(\log X_j)^{C_k \log|b|}}{2c_\Delta},$$

where C_k is an arbitrary constant. We see that

$$|\tilde{v}_1 + b\tilde{v}_2| = |(1,b)(v_1,v_2)^t W^{-1}| \ll |(v_1,v_2)| \ll r(\boldsymbol{v}).$$

We now split up into two cases depending on the size of d'.

- $d' = 1$: We note that $\Delta^{-1}, H \ll \log N$ and so we get

$$\sum_{0<\|\boldsymbol{v}\|_\infty \log N} \frac{1}{r(\boldsymbol{v})}(N(\hat{v}^{-1}b^{j+1}))^{-1+\epsilon} \ll \sum_{0<\|\boldsymbol{v}\|_\infty \leq \log N} \frac{|\tilde{v}_1 + b\tilde{v}_2|}{|b|^{(2-\epsilon)(j+1)/d'}} \ll \frac{(\log N)^4}{|b|^{2j/d'}}.$$

- $d' \geq 2$: In this case we have

$$r(\boldsymbol{v})^{-1} \ll |\tilde{v}_1 + b\tilde{v}_2|^{-1} \ll |\tilde{v}_1 + b\tilde{v}_2|^{-2/d'}.$$

Using this together with the fact that $\Delta^{-1}, H \ll \log N$ yields us

$$\sum_{0<\|\boldsymbol{v}\|_\infty \leq \log N} \frac{1}{r(\boldsymbol{v})}(N(\hat{v}^{-1}b^{j+1}))^{-1/d'+\epsilon} \ll \sum_{0<\|\boldsymbol{v}\|_\infty \leq \log N} \frac{1}{|b|^{(2-\epsilon)(j+1)/d'}} \ll \frac{(\log N)^2}{|b|^{2j/d'}}.$$

So in any case we have

$$\sum_{0<\|\boldsymbol{v}\|_\infty \leq \log N} \frac{1}{r(\boldsymbol{v})}(N(\hat{v}^{-1}b^{j+1}))^{-1/d'+\epsilon} \ll \frac{(\log N)^4}{|b|^{2j/d'}}.$$

Putting this into (7.31) gives us

$$\begin{aligned} &\left|\sum_{n\in I_j} \mathcal{I}\left(\frac{f(z_n)}{b^{j+1}}\right) - \frac{X_j}{|b|^{2l}}\right| \\ &\ll X_j\left((\log X_j)^{-1} + \frac{(\log X_j)^4 \log X_j}{|b|^{2j/d'}}\right) \ll \frac{X_j}{j} + X_j\frac{(\log N)^5}{|b|^{2j/d'}}. \end{aligned} \tag{7.32}$$

7.6 Summing up

Now all that is left is to sum everything up using the estimates we obtained in the last section. So putting (7.20), (7.24) and (7.32) together while considering the corresponding intervals we arrive at

$$\sum_{l\leq j\leq J} \left|\sum_{n\in I_j} \mathcal{I}\left(\frac{f(z_n)}{b^{j+1}}\right) - \frac{X_j}{|b|^{2l}}\right| \ll S_1 + S_2 + S_3, \tag{7.33}$$

where

$$S_1 = \sum_{l \le j \le J} \frac{X_j}{j},$$

$$S_2 = \sum_{l \le j \le l + C_l \log\log N} X_j \frac{(\log N)^5}{|b|^{2j/d'}},$$

$$S_3 = \sum_{J - C_u \log\log N \le j \le J} \sqrt{X_j}|b|^{j/d'}$$

$$+ \left(\frac{X_j}{|b|^{2j/d'}}\right)^{\frac{\log|b|}{\log C_\mu}} (\sqrt{X_j}|b|^{j/d'} + X_j(\log X_j)^{-3}).$$

Firstly, it is easy to see that

$$S_1 \ll M.$$

Secondly, after a quick estimation we also see

$$S_2 \ll \sum_{l \le j \le l + C_l \log\log N} M \frac{(\log N)^5}{|b|^{2j/d'}} \ll M \frac{(\log N)^5}{|b|^{2/d'(C_l \log\log N)}} \ll M,$$

assuming $C_l \ge 5$. For the last sum we have the following estimate:

$$S_3 \ll \sum_{J - C_u \log\log N \le j \le J} \sqrt{M}|b|^{j/d'} + \left(\frac{M}{|b|^{2j/d'}}\right)^{\frac{\log|b|}{\log C_\mu}} (\sqrt{M}|b|^{j/d'} + M)$$

$$\ll \sqrt{M}|b|^{J/d'} + \left(\frac{M}{|b|^{2J/d'}}\right)^{\frac{\log|b|}{\log C_\mu}} (\sqrt{M}|b|^{J/d'} + M) \ll M.$$

By inserting this into (7.33) we get

$$\sum_{l \le j \le J} \left| \sum_{n \in I_j} \mathcal{I}\left(\frac{f(z_n)}{b^{j+1}}\right) - \frac{X_j}{|b|^{2l}} \right| \ll M \ll \frac{N}{\log N},$$

which proves the main theorem.

Remark 7.2. As the reader should now be aware every single relevant calculation in the above should hold if we lessen the restriction of being Euclidean to that of having unique factorisation, apart from a single place, i.e., in the proof of Proposition 6.1. In fact, as was pointed out by M. Risager during the authors qualification exam, if 6.1 could be shown to hold without any assumption of being Euclidean, it would be possible to generalise the main result to any imaginary quadratic field.

Part II

Diophantine approximation and Schmidt's conjecture

CHAPTER 8

INTRODUCTION TO DIOPHANTINE APPROXIMATION

Simply put, Diophantine approximation is the study of how well it is possible to approximate real numbers with rational numbers. We all know that the reals can be constructed as a complete metric space having the rationals as a dense subset, where the metric on the reals extends that of the rationals. Hence every real number can be approximated arbitrarily well by a rational, in the sense that given any real number x and any positive ε, we can find $p \in \mathbb{Z}$ and $q \in \mathbb{N}$ such that $|x - p/q| < \varepsilon$. However this approximation does not tell us a lot about the complexity of the rational used. The standard thing to do is to assign a weight to each rational and then see how well it is possible to approximate a given real number with rationals having a small weight. Usually one chooses the weight of p/q to be q. If we now ask the same question of how well we can approximate a given real, it is easy to see that given any $x \in \mathbb{R}$ and any positive integer q there is an integer p which is closest to qx, hence we end up with $|qx - p| < 1$. But this is a very crude estimate and in fact we can say a lot more in general using Dirichlet's theorem. For the sake of completeness we state Dirichlet's theorem here. Note that p/q need not be in reduced form in Dirichlet's theorem nor in the two corollaries.

Theorem 8.1. *For every real number x and any positive integer $N \geq 1$, there exists a rational p/q with denominator satisfying $1 \leq q \leq N$, such that*

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

The following corollary is sometimes more interesting to work with, even though the bound is no longer uniform.

Corollary 8.2. *For every real number x and any positive integer $q \geq 1$, there exists an integer p such that*

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

Or put in another way

Corollary 8.3. *For every irrational number x there are infinitely many rationals such that*

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

Some natural questions to ask at this point is, can we do anything about the exponent 2 and can we do anything about the constant 1 to gain more information? The exponent 2 is the *approximation exponent* or *irrationality exponent* and we shall not go into too much detail on it. Except to say that almost all real numbers have approximation exponent 2, so we will focus on changing the constant 1 in the numerator on the right hand side, and see what information we can get. We start by recalling the following theorem due to Hurwitz.

Theorem 8.4 (Hurwitz)**.** *For every irrational number x there are infinitely many rationals such that*

$$\left|x - \frac{p}{q}\right| < \frac{1}{\sqrt{5}q^2}.$$

It turns out that there are actually real numbers x where

$$\left|x - \frac{p}{q}\right| \geq \frac{c}{q^2}$$

holds infinitely often for some positive constant c depending only on x. The set of numbers for which there is a positive constant c, such that the above inequality holds is called the *badly approximable numbers* and we denote them by **Bad**. We shall often make use of $\|\cdot\|$ to indicate the distance to the nearest integer, using this terminology we can multiply the above inequality with q to obtain

$$\|qx\| \geq \frac{c}{q}.$$

The set **Bad** is of Lebesgue measure zero, yet it is a rather large set in some other sense, as it has Hausdorff dimension 1, more information on what this means is given in chapter 9. This fact implies that the set is uncountable. The fact that the set **Bad** is of Lebesgue measure zero is due to Borel [Borel, 1909]. Later Jarník used the theory of continued fractions to show that **Bad** has Hausdorff dimension 1 [Jarník, 1929]. We shall not replicate their proofs here, but instead consider how to generalise the notion of being badly approximable to the n-dimensional Euclidean space and give higher dimensional results.

One possible way of generalising badly approximable numbers to higher dimensions is to consider what the correct analogue of Dirichlet's theorem would look like in several dimensions.

In order to do so we shall need a version Minkowski's convex body theorem, which we now state and prove.

Theorem 8.5 (Minkowski 1896)**.** *Let S be a convex set in $\mathbb{R}^n$ that is symmetric around* 0*, bounded and with volume $\lambda(S)$. Assume that $\lambda(S) > 2^n$. Then S contains a non-zero integer point.*

Proof. The following neat proof is due to Mordell and can be found in [Schmidt, 1980]. For every positive integer m let S_m denote the set of points in S with rational coordinates and common denominator m. As $m \to \infty$ we see that the amount of points in S_m will be asymptotically equal to $\lambda(S)m^n$, which for certain large m is bigger than $(2m)^n$ since $\lambda(S) > 2^n$. As a consequence we must have two distinct points $(a_1/m, a_2/m, \ldots, a_n/m)$ and $(b_1/m, b_2/m, \ldots, b_n/m)$ both lying in S_m and satisfying

$$a_i \equiv b_i \mod 2m \quad i = 1, 2, \ldots, n. \tag{8.1}$$

By symmetry and convexity we see that

$$g = \frac{1}{2}(a_1/m, a_2/m, \ldots, a_n/m) - \frac{1}{2}(b_1/m, b_2/m, \ldots, b_n/m)$$

is in S. It follows from (8.1) that g has integer coefficients, and it is obviously non-zero, being a non-zero scaling of the difference of two distinct points. □

We shall now proof the higher dimensional version of Dirichlet's theorem.

Theorem 8.6. *Given n real numbers $\alpha_1, \alpha_2, \ldots, \alpha_n$ and a natural number N there exists integers $p_1, p_2, \ldots, p_n, q$ with $1 \leq q \leq N$ such that*

$$\left|\alpha_i - \frac{p_i}{q}\right| \leq \frac{1}{qN^{1/n}} \quad \text{for } i = 1, 2, \ldots, n.$$

Proof. Let α be a fixed real number and consider the set

$$S = \left\{(x, y_1, y_2, \ldots, y_n) \in \mathbb{R}^{n+1} : -N + 1/2 \leq x \leq N + 1/2, |\alpha x - y_i| \leq \frac{1}{N^{1/n}}\right\}$$

We can calculate the volume of this set easily by noticing that it is a product of intervals. We see that x can range through an interval of length $2N$ and each y_i can range through an interval of length $2/N^{1/n}$ centered around αx. Hence

$$\lambda(S) = 2N \prod_{i=1}^{n} \frac{2}{N^{1/n}} = 2^{n+1},$$

where $|S|$ denotes the volume of S.

Applying Minkowski's convex body theorem now gives us the existence of a non-trivial point with integer coefficients inside S and we are done. □

We note that since $q \leq N$ we must have

Corollary 8.7. *For every real n-tuple $(x_1, x_2, \ldots, x_n)$ any $N \geq 1$ and any positive integer $N \geq q \geq 1$, there exists an n-tuple of integer $(p_1, p_2, \ldots, p_n)$ such that*

$$\left|x_i - \frac{p_i}{q}\right| \leq \frac{1}{q^{1+1/n}} \text{ for } 1 \leq i \leq n.$$

This allows us to extend our notion of badly approximable to several dimensions. We let $\mathbf{Bad}_n$ denote the set of *badly approximable vectors in* $\mathbb{R}^n$. i.e those $x \in \mathbb{R}^n$ satisfying

$$\left|x - \frac{p}{q}\right| \geq \frac{c}{q^{1+1/n}},$$

for some $c = c(x) > 0$ and for all $p \in \mathbb{Z}^n, q \in \mathbb{N}$. Or more suitable for our purposes, those $x \in \mathbb{R}^n$ satisfying

$$\|qx\| \geq \frac{c}{q^{1/n}},$$

for some $c = c(x) > 0$ and for all $q \in \mathbb{N}$.

We will now show that $\mathbf{Bad}_n$ has Lebesgue measure zero using a theorem of Khintchine dating back to 1926, known in the litterature simply as Khintchine's Theorem [Khintchine, 1926].

Theorem 8.8 (Khintchine's Theorem)**.** *Let $\Psi_1, \ldots, \Psi_n : \mathbb{N} \to (0, 1]$ and suppose that $\Psi(q) = \prod_{i=1}^{n} \Psi_i(q)$ is non-increasing. If the sum $\sum_{q=1}^{\infty} \Psi(q)$ is convergent, then for almost all n-tuples $(\alpha_1, \ldots, \alpha_n)$ there are only finitely many q with*

$$\|q\alpha_i\| < \Psi_i(q) \quad i = 1, 2, \ldots, n. \tag{8.2}$$

But if the sum $\sum_{q=1}^{\infty} \Psi(q)$ is divergent, then for almost all n-tuples $(\alpha_1, \ldots, \alpha_n)$ there are infinitely many q satisfying the condition (8.2).

We shall not prove Khintchine's Theorem here, but a proof can be found in [Bugeaud, 2004]. Note however that the notation differs slightly, and we have chosen to use the notation from [Schmidt, 1980].

With Khintchine's Theorem it is now easy to show the following

Corollary 8.9. *The set* $\mathbf{Bad}_n$ *has Lebesgue measure zero.*

Proof. For each $c > 0$ let $\mathbf{Bad}_n(c)$ denote the set of $x \in \mathbb{R}^n$ for which there are infinitely many $p \in \mathbb{Z}^n$ and $q \in \mathbb{N}$.

$$\left| x - \frac{p}{q} \right| \geq \frac{c}{q^{1+1/n}}.$$

It is now clear that $\mathbf{Bad}_n = \bigcup_{c>0} \mathbf{Bad}_n(c)$. Now consider the sequence $\{c_j\}_{j=1}^{\infty}$ where $c_j = 1/j$ we see that $\mathbf{Bad}_n = \cup_{j=1}^{\infty} \mathbf{Bad}_n(c_j)$. Now we let j be any positive integer and let $\Psi_j(q) = c_j q^{-1/n}$. Then

$$\sum_{q=1}^{\infty} \prod_{j=1}^{n} c_j q^{-1/n} = \sum_{q=1}^{\infty} c_j^n q^{-1}$$

is divergent. So by Khintchine's Theorem, for almost all n-tuples $(\alpha_1, \ldots, \alpha_n)$ there are infinitely many q with

$$\|q\alpha_j\| < \Psi_j(q) \quad i = 1, 2, \ldots, n.$$

That is , almost all n-tuples are not in $\mathbf{Bad}_n(c_j)$, and so $\mathbf{Bad}_n(c_j)$ is a Lebesgue null set. Taking the union over all j completes the proof. □

As a final comment from the author before moving on to the next chapter, the reader should be aware that the title "Zur Metrischen Theorie der diophantischen Approximationen." was very popular in the later part of the 1920s and has been used by both Khintchine and Jarník.

CHAPTER 9

HAUSDORFF MEASURE AND DIMENSION

Hausdorff dimension is a useful tool in Diophantine approximation, as it intuitively gives us some information about the size of a given set. Usually the sets studied in Diophantine approximation are of Lebesgue measure zero, meaning Lebesgue measure will not give us any new information or any way of distinguishing between two given sets. We proceed to give a short and self-contained account of the Hausdorff measure and dimension. More details and further references can be found in [Falconer, 2003]

9.1 Hausdorff measure

Let us consider a set $E \subset \mathbb{R}^n$. And for any set $V \in \mathbb{R}^n$, let $|V|$ denote the diameter of the set V, i.e. $|V| = \sup_{x,y \in V}\{|x-y|\}$, where $|x-y|$ is the standard Euclidean norm on $\mathbb{R}^n$ of the vector $x - y$. Given any positive δ, and an index set I which is finite or countable, we say that the collection $\{V_i \subset \mathbb{R}^n \mid i \in I\}$ is a *δ-cover of E* if $|V_i| < \delta$ and $E \subset \cup_{i \in I} V_i$. Let

$$\mathcal{H}^s_\delta(E) = \inf\left\{\sum_{i \in I} |V_i|^s\right\},$$

where the infimum is taken over all possible δ-covers of the set E. This set function is an outer measure on $\mathbb{R}^n$ and as $\delta \to 0$ we notice that there are less possible δ-covers, hence $\mathcal{H}^s_\delta(E)$ increases to a possibly infinite limit given by

$$\mathcal{H}^s(E) = \lim_{\delta \to 0} \mathcal{H}^s_\delta(E) = \sup_{\delta > 0} \mathcal{H}^s_\delta(E) \in [0, \infty].$$

As we notice from its construction of covers $\mathcal{H}^s(E) \leq \mathcal{H}^s(F)$ for any $E \subset F$. The set function $\mathcal{H}^s$ is indeed subadditive and a regular outer measure. Restricting this to the σ-algebra of $\mathcal{H}^s$ measurable sets we obtain the s-dimensional Hausdorff measure. We should note that $\mathcal{H}^s$ measurable sets contain among others, the Lebesgue-measurable sets. If s is a positive integer, say n, then the n-dimensional Hausdorff measure and the n-dimensional Lebesgue measure are comparable. This is quite obvious from the construction of covers and it is an elementary fact that we will come back to in a bit. Another important fact about the Hausdorff-measure is that for any s, $\mathcal{H}^s$ is unchanged under isometries and behaves very well by scaling; in fact, for any $\lambda \geq 0$ we have that $\mathcal{H}^s(\lambda E) = \lambda^s \mathcal{H}^s(E)$.

9.2 Hausdorff dimension

For a fixed set E we can consider $\mathcal{H}^s(E)$ as a function of s and see what happens as s increases. For $s = 0$, $\mathcal{H}^s$ is simply the counting measure. We see that if E is a finite set $\mathcal{H}^s(E)$ is going to be zero for $s > 0$. We also see that if E is an infinite set, then $\mathcal{H}^s(E)$ is either 0 or ∞, except for possibly one value of s. In essence this means that there is at most one value of s where $\mathcal{H}^s$ is finite, no matter if E is finite or not. We will make a short justification of the last fact. It follows directly from the definition using covers that there is a δ-cover of E such that

$$\sum_{i\in I} |V_i|^s \leq \mathcal{H}^s_\delta(E) + 1 \leq \mathcal{H}^s(E) + 1 \leq \infty.$$

Now suppose that $\mathcal{H}^{s_0}(E) > 0$ and $s = s_0 + \varepsilon$, where $\varepsilon > 0$. Then for each V_i we have $|V_i|^s = |V_i|^{s_0+\varepsilon} \leq \delta^\varepsilon |V_i|^{s_0}$. So we see that

$$\sum_{i\in I} |V_i|^s = \sum_{i\in I} |V_i|^{s_0+\varepsilon} \leq \delta^\varepsilon \sum_{i\in I} |V_i|^{s_0}.$$

Thus we arrive at

$$\mathcal{H}^s_\delta(E) = \mathcal{H}^{s_0+\varepsilon}_\delta(E) = \sum_{i\in I} |V_i|^{s_0+\varepsilon} \leq \delta^\varepsilon \sum_{i\in I} |V_i|^{s_0} \leq \delta^\varepsilon(\mathcal{H}^{s_0}(E) + 1).$$

The conclusion of the above calculations is

$$0 \leq \mathcal{H}^s(E) = \mathcal{H}^{s_0+\varepsilon}(E) = \lim_{\delta\to 0} \mathcal{H}^{s_0+\varepsilon}_\delta(E) \leq \lim_{\delta\to 0} \delta^\varepsilon(\mathcal{H}^{s_0}(E) + 1) = 0.$$

So for all values of s bigger than s_0 we have $\mathcal{H}^s(E) = 0$. Now suppose once more that $\mathcal{H}^{s_0}(E) > 0$. If for any $\varepsilon > 0$ we have $\mathcal{H}^{s_0-\varepsilon}(E) > 0$, then by the above we see that $\mathcal{H}^{s_0}(E) = 0$, a clear contradiction. Hence $\mathcal{H}^{s_0-\varepsilon}(E) = \infty$. So for any infinite set $E \subset \mathbb{R}^n$ there is a unique non-negative number s_0 such that

$$\mathcal{H}^s(E) = \begin{cases} \infty, & 0 \leq s < s_0, \\ 0, & s_0 < s < \infty. \end{cases}$$

This critical exponent s_0 is called the *Hausdorff dimension* of the set E and we denote it by $\dim_{\mathrm{H}}(E)$. The Hausdorff dimension has certain nice properties and we list a few of them, which are valid for sets in $\mathbb{R}^n$.

(i) If $E \subset F$ then $\dim_{\mathrm{H}}(E) \leq \dim_{\mathrm{H}}(F)$.

(ii) $\dim_{\mathrm{H}}(E) \leq n$.

(iii) If E has positive Lebesgue measure, then $\dim_{\mathrm{H}}(E) = n$.

(iv) If $\dim_{\mathrm{H}}(E) < n$ then E has Lebesgue measure 0.

(v) $\dim_{\mathrm{H}} \cup_{j=1}^{\infty} E_j = \sup\{\dim_{\mathrm{H}}(E_j) \mid j \in \mathbb{N}\}$.

The properties $(ii) - (iv)$ are the noteworthy ones for us in this thesis. They tell us that when we consider a given set in $\mathbb{R}^n$ of Lebesgue measure 0, the biggest a set can get is one having Hausdorff dimension n, that of the ambient space. When this happens to be the case for a set E we say that E has *full dimension*.

There are different ways of determining the Hausdorff dimension of a given set and there are methods developed for particular types of sets. We will not delve too much into the different methods but focus solely on one which serves our purpose, namely Schmidt games, which will be introduced in the next chapter.

CHAPTER 10

SCHMIDT GAMES

Schmidt's game was introduced in 1965 in a paper by Wolfgang Schmidt (though not published till 1966) [Schmidt, 1966]. It is an important tool when studying certain types of sets of numbers, as it can give information about their Hausdorff dimension. The game is played between two players usually refered to in the literature as Black and White, or Alice and Bob. We shall choose the latter names for our players and describe how the game is played in much the same way as McMullen does, since we will be using his terminologi later on. Alice and Bob take turns in choosing a nested sequence of closed Euclidean balls in $\mathbb{R}^n$

$$B_1 \supset A_1 \supset B_2 \supset A_2 \supset B_3 \ldots,$$

whose diameter satisfy, for fixed $0 < \alpha, \beta < 1$,

$$|A_i| = \alpha\,|B_i| \text{ and } |B_{i+1}| = \beta\,|A_i|\,. \tag{10.1}$$

Notice that since the radii of the balls tend to 0 and all the balls are closed in $\mathbb{R}^n$ and hence compact, the intersection will be a unique point which we denote by x. We say $E \subset \mathbb{R}^n$ is an (α, β)-*winning set* if Alice has a strategy that ensures that $x \in E$. We say E is α-*winning* if it is (α, β)-winning for all $0 < \beta < 1$. Finally we say that E is a *winning set* if it is α-winning for some $\alpha > 0$. Winning sets have many useful properties. We list the ones that are most important to us here.

1. Any winning set in $\mathbb{R}^n$ has Hausdorff dimension n.

2. A countable intersection of α-winning sets is α-winning.

3. The image of a winning set under a bi-Lipschitz homeomorphism of $\mathbb{R}^n$ is again winning.

In Schmidt's original paper from 1965, Schmidt describes the games and uses them to prove several interesting results. For instance property 1 and 2 above. He also showed that the set of badly approximable numbers is winning and states that similar results hold in higher dimensions.

In his paper, [McMullen, 2010], McMullen suggests two variants of Schmidt's game. The first is to change (10.1) to

$$|A_i| \geq \alpha\,|B_i| \text{ and } |B_{i+1}| \geq \beta\,|A_i|\,.$$

If the game is played like this, and Alice has a winning strategy for the set E. We say that E is an (α, β)-*strong winning set*. The rest of the terminology is made in the same manner by putting *strong* in front. The second variant is the absolute version. Here the sequence of balls chosen need no longer only be nested, but should satisfy the following

$$B_1 \supset (B_1 \setminus A_1) \supset B_2 \supset (B_2 \setminus A_2) \supset B_3 \ldots \tag{10.2}$$

On top of this we must have for fixed $0 < \beta < 1/3$

$$|B_{i+1}| \geq \beta |B_i| \text{ and } |A_i| \leq \beta |B_i| . \quad (10.3)$$

Notice that the radii of the balls here do not have to tend to 0, so there is not necessarily a unique point in the intersection. Therefore we say that Alice has a winning strategy for the set E if she can play such that E contains any point from the intersection of all the B_i. If this is the case for all $0 < \beta < 1/3$, we say that E is an *absolutely winning set*. The important thing to note in this variant of the game is that Alice has very little control over Bob's choices. Essentially she can only prevent him from picking his favorite ball in each turn.

We will not always be playing games on all of $\mathbb{R}^n$ but sometimes restrict ourselves to certain subsets. We therefore require some more terminology.

Definition 10.1. Given a closed set $K \subset \mathbb{R}^n$ and a set $S \subset K$ we say that S is winning on K if Alice has a winning strategy for the set S with the further restriction that every ball should now be placed with center in K. i.e., Bob picks B_1 with center $b_1 \in K$, then Alice picks A_1 with center $a_1 \in K \cap B_1$ and the game is played in exactly the same manner as before, apart from this restriction. The fact that K is closed and the centres $b_i \in K$ for all i ensures that $\cap B_i \subset K$.

In other words, we completely ignore every point in $\mathbb{R}^n \setminus K$ when playing the game.

10.1 Hyperplane absolute winning sets

Based on the ideas of McMullen another variant of the game was invented by Broderick, Fishman, Kleinbock, Reich and Weiss in [Broderick et al., 2012]. They observed that many sets of interest to people in the Diophantine approximation branch tended to be sets that avoided certain hyperplanes. As such they are almost guarenteed not to be absolutely winning, since Alice can only block a single ball each turn and not an entire hyperplane. Thus Bob could come up with a strategy ensuring that his next ball would always be placed on this very hyperplane. The variant they propose is to change the game, so that Alice instead of blocking a ball, blocks a small neighbourhood of a hyperplane. To make this more precise, we let $k \in \{0, 1, \ldots, n-1\}$, where n is the dimension of the ambient space ($\mathbb{R}^n$). We let $0 < \beta < 1/3$, and define the *k-dimensional β-absolute game* in the following way. Bob initially picks $x_1 \in \mathbb{R}^n$ and $r_1 > 0$ defining the closed ball $B(x_1, r_1)$. At each stage of the game, after Bob has chosen x_i, r_i, Alice chooses an affine subspace of dimension k and removes its $\varepsilon(i)$-neighbourhood, denoted by $A_i = \mathcal{L}^{\varepsilon(i)}$ from $B_i = B(x_i, r_i)$. Then Bob chooses x_{i+1} and r_{i+1} such that $r_{i+1} \geq \beta r_i$ and

$$B_{i+1} \subset (B_i \setminus A_i) .$$

A set E is said to be *k-dimensionally β-absolute winning* if Alice has a strategy guarenteeing that $\cap_i B_i$ intersects E. In the special case that $k = n-1$ we say that E is a *Hyperplane Absolute Winning set*, or a HAW set. In order to show just how strong the HAW property is we need to recall a few definitions from measure theory.

10.2 Measures on $\mathbb{R}^n$

We begin with the notion of being absolutely decaying, a concept introduced by Kleinbock, Lindenstrauss and Weiss in [Kleinbock et al., 2004], though in a slightly different manner than given here.

Definition 10.2. If μ is a locally finite Borel measure on $\mathbb{R}^n$ and $C, \gamma > 0$ we say that μ is (C,γ)-*absolutely decaying* if there exists $\rho_0 > 0$ such that, for all $0 < \rho < \rho_0$, all $x \in \operatorname{supp}\mu$, all affine hyperplanes $\mathcal{L} \subset \mathbb{R}^n$ and all $\varepsilon > 0$, one has

$$\mu\left(B(x,\rho) \cap \mathcal{L}^{(\varepsilon)}\right) < C(\varepsilon/\rho)^{\gamma}\mu\left(B(x,\rho)\right).$$

Here $\mathcal{L}^{(\varepsilon)}$ denotes the ε-neighbourhood of the hyperplane $\mathcal{L}$. We say that μ is absolutely decaying if it is (C,γ)-absolutely decaying for some positive C and some positive γ. Intuitively this means that the measure μ is not concentrated on any hyperplane.

Definition 10.3. A measure μ is called D-Federer (or D-doubling) if there exists $\rho_0 > 0$ such that

$$\mu\left(B(x,2\rho)\right) < D\mu\left(B(x,\rho)\right) \forall x \in \operatorname{supp}\mu, \forall 0 < \rho < \rho_0.$$

We say that μ is *Federer* if it is D-Federer for some $D > 0$.

Measures that are both absolutely decaying and Federer are said to be *absolutely friendly*, a notion coined in [Pollington and Velani, 2005]. An important thing to note is that the Federer condition always holds when the measure is Ahlfors regular.

Definition 10.4. μ is said to be δ-*Ahlfors regular*, if there exists positive δ, c_1, c_2, ρ_0 such that

$$c_1\rho^{\delta} \le \mu\left(B(x,\rho)\right) \le c_2\rho^{\delta} \quad \forall x \in \operatorname{supp}\mu, \forall 0 < \rho < \rho_0.$$

We say that μ is *Ahlfors regular* if it is δ-Ahlfors regular for some positive δ. It can easily be shown that if μ is a δ-Ahlfors regular measure, then the Hausdorff dimension of $\operatorname{supp}\mu$ is equal to δ. There are many examples of measures that are both absolutely decaying and Ahlfors regular, for instance limit measures of irreducible families of contracting self-similar transformations of $\mathbb{R}^n$ satisfying the open set condition. The open set condition, as well as much of the base study on fractals and measures supported on fractals, are due to Hutchinson and the work in his famous paper [Hutchinson, 1981]. The reader may also consult [Falconer, 2003] for more background information.

To show how strong the HAW property is we list a few definitions and some theorems given in [Broderick et al., 2012].

Definition 10.5. Given $E \subset \mathbb{R}^n$ we say that E is *strongly (resp. strongly C^1, strongly affinely incompressible) on* K if the following holds

$$\dim_{\mathrm{H}}\left(\bigcap_{i=1}^{\infty} f_i^{-1}(E) \cap K\right) = \dim_{\mathrm{H}}(U \cap K),$$

for any open set $U \subset \mathbb{R}^n$ with $U \cap K \neq \emptyset$ and any sequence $\{f_i\}$ of bi-Lipschitz maps (resp. C^1 diffeomorphisms, affine nonsingular maps) of U onto (possibly different) open subsets of $\mathbb{R}^n$.

We note that if E is strongly C^1 incompressible, then E behaves as well as any open set, with respect to the Hausdorff dimension, when intersected with any sufficiently nice fractal. Here sufficiently nice means those supporting an Ahlfors regular measure which is absolutely decaying, two concepts we shall define in a short while. The main theorem of [Broderick et al., 2012] is the following, stating in particular that $\mathbf{Bad}_n$ is strongly C^1 incompressible.

Theorem 10.6. *[Broderick et al., 2012, Theorem 1.1] Let μ be a measure on $\mathbb{R}^n$ which is absolutely decaying and Ahlfors regular. Then $\mathbf{Bad}_n$ is strongly C^1 incompressible on $K = \operatorname{supp}\mu$.*

For our purpose the following corollary is of interest.

Corollary 10.7. *[Broderick et al., 2012, Corollary 5.4] Let μ be absolutely decaying and Ahlfors regular, and let $E \subset \mathbb{R}^n$ be k-dimensionally absolute winning. Then E is strongly C^1 incompressible on $\operatorname{supp}\mu$.*

As the corollary is interesting to us we shall now replicate the proof of it as given in [Broderick et al., 2012] and explain the various concepts needed. The following lemma provides a condition on a set K to estimate the Hausdorff dimension from below of $S \cap K$ when S is any set which is winning on K. The lemma is a modified version of a corollary found in Schmidt's paper [Schmidt, 1966, Corollary 1]. First a bit of notation. For $K \subset \mathbb{R}^n, x \in K, \rho > 0$ and $0 < \beta < 1$, let $N_K(\beta, x, \rho)$ denote the maximum number of disjoint balls of radius $\beta\rho$ centered on K and contained in $B(x, \rho)$.

Lemma 10.8. *[Broderick et al., 2012, Lemma 4.1] Suppose there exists positive M, δ, ρ_0 and β_0 such that*

$$N_K(\beta, x, \rho) \geq M\beta^{-\delta},$$

for all $x \in K, \rho < \rho_0$ and $\beta < \beta_0$. Then $\dim_{\mathrm{H}}(S \cap K \cap U) \geq \delta$ for all open sets U with $U \cap K \neq \emptyset$ and any set S, which is winning on K.

Note that the condition in the lemma can be satisfied with $\delta = \dim_{\mathrm{H}} K$ whenever K supports an Ahlfors regular measure.

We shall use this lemma to prove another lemma which gives us an estimation from below on the Hausdorff dimension of the support of an absolutely decaying measure.

Lemma 10.9. *[Broderick et al., 2012, Lemma 5.2] Let μ be (C, γ)-absolutely decaying, let $K = \operatorname{supp}\mu$, and let $S \subset \mathbb{R}^n$ be winning on K. Then $\dim_{\mathrm{H}}(S \cap K \cap U) \geq \gamma$ for any open set U with $K \cap U \neq \emptyset$.*

Proof. For any $\beta < 1$, let k be chosen such that for some sufficiently small but fixed ρ we have $\inf_{x \in K} N_K(\beta, x, \rho) \geq k$. For any $x \in K$ we can pick k disjoint balls of radius $\beta\rho$ inside the ball $B(x, \rho)$. For each of these balls pick a hyperplane neighbourhood of width 2β contained in the ball around x of radius $(1 - \beta)\rho$. The total measure of these hyperplane neighbourhoods is at most $kC(2\beta)^\gamma/(1 - \beta)^\gamma$ of the measure of the ball $B(x, (1 - \beta)\rho)$. So if $k < (1 - \beta)^\gamma/C(2\beta)^\gamma$, there is a point of K outside the union of these hyperplane neighbourhoods and in $B(x, (1 - \beta)\rho)$. This point can be the center of a new $\beta\rho$-ball in $B(x, \rho)$ and we will now have $k + 1$ disjoint balls, since each hyperplane neighbourhood contains a ball. Hence we must have $N_K(\beta, x, \rho) \geq (1-\beta)^\gamma / C(2\beta)^\gamma$, so for sufficiently small β we get $N_K(\beta, x, \rho) \geq M\beta^{-\gamma}$ for some constant M. Applying Lemma 10.8 yields us the estimate. □

Similar arguments to the one above allowed Fishman in [Fishman, 2009] to prove the following

Lemma 10.10. *[Fishman, 2009, Theorem 5.1] Let μ be δ-Ahlfors regular, let $K = \operatorname{supp}\mu$, and let $S \subset \mathbb{R}^n$ be winning on K. Then $\dim_{\mathrm{H}}(S \cap K \cap U) = \delta = \dim_{\mathrm{H}}(K \cap U)$ for every open set $U \subset \mathbb{R}^n$ with $U \cap K \neq \emptyset$.*

Note that a more specialized version of the above theorem was proven independently in [Esdahl-Schou and Kristensen, 2010, Theorem 1].

We shall need another theorem before we can conclude that Corollary 10.7 holds. We choose not to prove the theorem though, as it is beyond the scope of this thesis.

Theorem 10.11. *[Broderick et al., 2012, Theorem 2.4] Let $S \subset \mathbb{R}^n$ be k-dimensionally absolute winning, $U \subset \mathbb{R}^n$ open, and $f : U \to \mathbb{R}^n$ a C^1 nonsingular map. Then $f^{-1}(S) \cup U^c$ is k-dimensionally absolute winning. Consequently, any k-dimensionally absolute winning set is strongly C^1 incompressible.*

Combining Theorem 10.11 with 10.10 we can prove Corollary 10.7.

We continue with a short historical survey.

CHAPTER 11

A SURVEY OF FAMOUS CONJECTURES IN DIOPHANTINE APPROXIMATION

Littlewood's conjecture is (at least to the knowledge of the author) still an open problem. It was suggested by Littlewood in the 1930s and can be explained quite simply. If $\|x\|$ denotes the distance between x and the nearest integer. Then Littlewood's conjecture states that for any two real numbers x and y we have

$$\liminf_{q\to\infty} q\,\|qx\|\,\|qy\| = 0.$$

We shall not delve too deeply into why Littlewood's conjecture is interesting, other than the fact that after roughly 80 years of several mathematicians trying to attack it, it has remained unsolved. In fact it has led to other conjectures both solved and unsolved. It is worth noting that Borel in 1909 (before the conjecture was made) showed that the set of points $(x, y) \in \mathbb{R}^2$ that violates Littlewood's conjecture, which we denote by *the exceptional set*, is of Lebesgue measure zero. Much later Einsiedler, Katok and Lindenstrauss showed that the exceptional set has Hausdorff dimension zero, see [Einsiedler et al., 2006]. The work done to achieve this result is part of what earned Lindenstrauss his Field's medal in 2010.

Note that if the exceptional set had positive dimension, it would be non-empty and the conjecture would have been answered in the negative straight away. But since it has dimension zero, we are in some sense, none the wiser. It can still be uncountable, though everyone hopes that it is empty.

Somewhere in between this, namely in the 1980s Wolfgang Schmidt proposed a famous conjecture, which in essence shows why Littlewood's conjecture is such a hard nut to crack. Schmidt showed that if his conjecture is false, then there is a counterexample to Littlewood's conjecture. Schmidt's conjecture was solved in 2010 by Badziahin, Pollington and Velani, see [Badziahin et al., 2011]. Later another proof was given by An in 2012, [An, 2013], which improves the result of Badziahin, Pollington and Velani.

So much for the history lesson. The rest of this thesis is concerned with explaining what Schmidt's conjecture actually states and what exactly was shown by Badziahin, Pollington and Velani, what An managed to improve upon, and finally to consider an analogue of Schmidt's conjecture in the complex case.

11.1 Schmidt's conjecture

For any $0 \leq i, j \leq 1$ with $i + j = 1$, let $\mathbf{Bad}(i, j)$ denote the set of points $(x, y) \in \mathbb{R}^2$ for which $\max\{\|qx\|^{1/i}, \|qy\|^{1/j}\} > c/q$ for all $q \in \mathbb{N}$, where $c = c(x, y)$ is a positive constant.

Schmidt noted that if there were just two pairs (i, j) and (i', j') with $i \neq i'$ such that $\mathbf{Bad}(i, j) \cap \mathbf{Bad}(i', j') = \emptyset$, then he would have a counter-example to Littlewood's conjecture. After observing this Schmidt's original question was the following: Is the intersection of $\mathbf{Bad}(1/3, 2/3)$ and $\mathbf{Bad}(2/3, 1/3)$ non-empty? This is a very specific problem and there is nothing special about 1/3 and 2/3. So it seemed natural to quickly consider the following question, which became known commonly as Schmidt's conjecture: Is $\mathbf{Bad}(i, j) \cap \mathbf{Bad}(i', j') = \emptyset$?

Badziahin, Pollington and Velani managed to show that any finite intersection of sets of the type $\mathbf{Bad}(i, j)$ is non-empty and in fact have full dimension. They also managed to show that certain infinite intersections had full dimension as long as they satisfied a technical condition. An was able to remove the technical condition and in fact show that any countable infinite intersection would, not only have full dimension, but be a winning set.

CHAPTER 12

A COMPLEX ANALOGUE OF SCHMIDT'S CONJECTURE

In this chapter we will study an analogue of Schmidt's conjecture, not in $\mathbb{R}^2$ but instead in quadratic imaginary fields, where the associated ring of integers is a unique factorization domain. As already mentioned, there are 9 such fields and they are described by the following 9 numbers known as the Heegner-Stark numbers $d \in \{-1, -2, -3, -7, -11, -19, -43, -67, -163\}$. For each d we let $\mathcal{O}_d$ denote the ring of integers associated with the field $\mathbb{Q}(\sqrt{d})$. We now consider the inequality

$$\left|z - \frac{p}{q}\right| \geq \frac{C(z)}{|q|^2}, \text{ for all } p, q \in \mathcal{O}_d, q \neq 0. \tag{12.1}$$

We shall refer to p/q as rationals, and call the set of $z \in \mathbb{C}$ for which this inequality holds for *the set of badly approximable complex numbers with respect to* $\mathcal{O}_d$ and we denote it by $\mathbf{Bad}_d$. Hopefully this will not cause any confusion with the sets $\mathbf{Bad}_n$, even though it is merely a dummy index that puts the notations apart. The reason why we require unique factorization of the associated ring of integers is twofold. To ensure that there is no ambiguity in the choice of q in the inequality (12.1) and also because we shall need unique factorization in the proof of the main theorem, which is yet to be stated.

We now consider the set $\mathbf{Bad}(i, j)$ in the real case again and notice that if $(x, y) \in \mathbf{Bad}(i, j)$, then $\max\{\|qx\|^{1/i}, \|qy\|^{1/j}\} > c/q$. Assume that $\|qx\|^{1/i} > c/q$. Then we have the following

$$\|qx\|^{1/i} > \frac{c}{q}$$

which implies

$$\|qx\| > \frac{c^i}{q^i}$$

which gives us

$$|qx - p| > \frac{c^i}{q^i} \quad \text{for some } p \in \mathcal{O}_d$$

so by dividing by q we get

$$\left|x - \frac{p}{q}\right| > \frac{c^i}{q^{1+i}}.$$

If we consider this geometrically, then $(x, y) \in \mathbf{Bad}(i, j)$ is trying to avoid every rectangle with sides of size $\asymp q^{-(1+i)}$ and $q^{-(1+j)}$, which has volume $q^{-(1+i)-(1+j)} = q^{-3}$.

However, if we tried using the same approach for the complex numbers, as An used in the real case, we would quickly run into trouble. The complex numbers behave differently than the real numbers when it comes to approximation. It turns out that the right volume for a rectangle is not q^{-3} but q^{-4}. The reason behind this is that we want the proper analogues of the theorems of Dirichlet, Khintchine and Jarník. We have already seen the complex version of Dirichlet's Theorem, namely Lemma 6.3. The complex version of Khintchine's theorem is due to LeVeque, [LeVeque, 1952] who in 1952 combined the original ideas of Khintchine, using continued fractions and some ideas from hyperbolic geometry was able to show the following. See [Dodson and Kristensen, 2004] for further references.

Theorem 12.1. *Suppose $|q|^2\Psi(|q|)$ is decreasing. Let $\|qz\|$ denote the distance from qz to the nearest element of $\mathcal{O}_d$. Then the Lebesgue measure of the set $W^\star(\Psi)$ is null or full according as the sum*

$$\sum_{|q|=1}^{\infty} |q|^3\, \Psi(|q|)^2$$

converges or diverges. Where

$$W^\star(\Psi) = \{z \in \mathbb{C} \mid \|qz\| < |q|\, \Psi(|q|) \text{ for infinitely many } q \in \mathbb{Z}(i)\}.$$

The theorem generalises to the cases where the Gaussian integers is replaced by one of the remaining ring of integers $\mathcal{O}_d$, a result due to Sullivan. See [Sullivan, 1982] for more details.

We can now easily conclude the following

Theorem 12.2. *The set $\mathbf{Bad}_d$ has Lebesgue measure* 0.

Proof. As in the real case we let $c_j = 1/j$ and define $\mathbf{Bad}_d(c)$ in the same manner as in the the proof of Corollary 8.9. For each fixed j we consider the function $\Psi(q) = \frac{c_j}{|q|^2}$ We then see that the sum

$$\sum_{|q|=1}^{\infty} |q|^3\, \frac{c_j^2}{|q|^4} = \sum_{|q|=1}^{\infty} \frac{c_j^2}{|q|}$$

is divergent and thus the set $\{z \in \mathbb{C} \mid \|qz\| < |q|\, \Psi(|q|) \text{ for infinitely many } q \in \mathbb{Z}(i)\}$ has full Lebesgue measure. Hence $\mathbf{Bad}_d(c_j)$ is a null set, and taking the union over all j we see that $\mathbf{Bad}_d$ is a null set. □

The complex analogue of Jarníks theorem in the case of the Gaussian integers could, as noted in [Dodson and Kristensen, 2004, page 329], be seen as easy consequences of the work of [Bishop and Jones, 1997] or [Fernandez and Melián, 1995]. However a self-contained proof can be found in [Dodson and Kristensen, 2004]. The slightly more general case where the Gaussian integers are replaced by $\mathcal{O}_d$ can be found in [Esdahl-Schou and Kristensen, 2010].

If we return to the concept of a complex analogue of Schmidt's conjecture and the idea of avoiding rectangles we can now for each $0 < i, j < 1$ with $i + j = 1$ consider the $z \in \mathbb{C}$ for which

$$\max\left\{ \left|\mathrm{Re}\left(z - \frac{p}{q}\right)\right|^{1/(1+2i)}, \left|\mathrm{Im}\left(z - \frac{p}{q}\right)\right|^{1/(1+2j)} \right\} > \frac{c}{|q|} \text{ for all } p, q \in \mathcal{O}_d, q \neq 0. \tag{12.2}$$

We denote the set of $z \in \mathbb{C}$ satisfying the inequality (12.2) by $\mathbf{Bad}_d(i,j)$. We note that in the case where $i = j = 1/2$ has already been shown to be winning in [Esdahl-Schou and Kristensen, 2010]. We also note that unlike in the real case, where $i = 1, j = 0$ or vice versa $\mathbf{Bad}(i,j)$ corresponds to $\mathbf{Bad} \times \{0\}$, respectively $\{0\} \times \mathbf{Bad}$. We do not have that in the complex case.

It follows from [Kristensen et al., 2006, Theorem 2] that $\mathbf{Bad}_d(i,j)$ is of full dimension, but what can we say about the intersections? Using the same approach as in the reals one might show that $\mathbf{Bad}_d(i,j)$ is winning for every pair of i and j satisfying $0 \le i,j \le 1$ and $i+j=1$. Using the fact that any countable intersection of winning sets is again winning and hence has full dimension would answer the following conjecture in the affirmative.

Conjecture 12.3. *The set $\mathbf{Bad}_d(i,j)$ is winning for all i,j with $0 \le i,j \le 1$ and $i+j=1$.*

One might even show that $\mathbf{Bad}_d(i,j)$ is not only winning but perhaps hyperplane absolutely winning, in which case the following stronger conjecture would be true.

Conjecture 12.4. *The set $\mathbf{Bad}_d(i,j)$ is hyperplane absolute winning for all i,j with $0 \le i,j \le 1$ and $i+j=1$.*

We shall now assume that $i \ge j$. That is we consider rectangles which are tall and thin. One of the standard tools to show that something is winning is the following lemma known as the simplex lemma or the Davenport trick.

Lemma 12.5. *For every $\beta \in (0,1)$ and for every $k \in \mathbb{N}$ let*

$$U_{k,i} = \left\{ \frac{p}{q} : p,q \in \mathcal{O}_d \text{ and } \beta^{-\frac{(k-1)}{1+2i}} \le |q| < \beta^{-\frac{k}{1+2i}} \right\}$$

Then for all $x \in \mathbb{C}$ and for k big enough, there is only one point in $U_{k,i} \cap B(x, \beta^{k-1} r)$ as long as r satisfies $0 < r < \frac{1}{2}\beta$.

Proof.

$$\left| \frac{p}{q} - \frac{p'}{q'} \right| \ge \frac{1}{|qq'|} > R^{-2k}.$$

But the diameter of $B(x, \beta^{k-1} r)$ is $2\beta^{k-1} r < \beta^{k-1}\beta = \beta^k \le \beta^{2k/(1+2i)} = R^{-2k}$, where we used that for $i \ge j$ we have $1+2i \ge 2$. So for k big enough so the result follows. □

The lemma is the key ingredient in what is known as the Davenport method. Simply put the Davenport method is a way of constructing a Cantor type set within the set we are interested in. Davenport originally used this method to show that certain sets contained continuum many points. The method works as follows:

Consider a square of size $R^{-n} \times R^{-n}$ and call this B_n and split this up into R^2 smaller squares of size $R^{-(n+1)} \times R^{-(n+1)}$. Consider the rational points with denominator at least $R^{-(n+1)}$. If there is only one such point inside B_n it falls within one of the smaller squares. Place a square neighbourhood of size $R^{-(n+1)} \times R^{-(n+1)}$ around the point and remove the squares that intersect this neighbourhood. Assume that not all of the R^2 squares have been removed, then any of the remaining ones will be used in the Cantor set construction. Within each of the squares that we did not remove we can continue this process. If there are two points inside B_n we place neighbourhoods around each of the points and continue in the same manner. If there is three or more points we could risk that the points were spread out all over B_n and that their neighbourhoods would cover everything. This is where the lemma comes in, if there is three or more points that are not colinear, pick any three of them. Their convex hull forms a triangle and estimating the volume of this triangle will show that it is simply too big to fit inside B_n, so the points have to be colinear. Now we place a neighbourhood around each point on the line passing through these points and remove any of the squares intersecting this neighbourhood. Even though Davenport used this method only to show continuum many points with certain properties (i.e. simultanously badly approximable points in $\mathbb{R}^2$), the method can give us much more. If enough squares survive this elimination process at each step, we end up with a Cantor type set of which we can easily estimate the Hausdorff-dimension, which gives us a lower bound on the Hausdorff-dimension of the set we are interested in. Often this method is all that is required to show full dimension.

This lets us believe that the above conjectures are true since there is only one problematic point at each step and hence it should be easy to win the Schmidt game, even in the absolute version. There is a slight problem though. If we consider a square of size $R^{-2k} \times R^{-2k}$ we need not only avoid a rectangle centered around the sole point inside the square, but also the part of the rectangles coming from neighbouring squares, which intersect the square we are considering.

Figure 12.1 describes the situation quite well. The rectangle made up of fully drawn lines comes from a point inside the square we are considering, where as the rectangle made up of dashes lines comes from a point in the square above the one we consider. Potentially a situation like this can happen, but it can get alot worse. A rough upper estimate of how many rectangles we might end up having to deal with is given by simply dividing the height of the rectangle with the height of the square. So we might risk having as many as $R^{-(1+2j)n}/R^{-2n}$ points each of which generates a rectangle which will intersect the square we are considering. If $i = j = 1/2$ there is no problem, but as soon as $i > j$ we can potentially have as many as $R^{\varepsilon n}$ points, where ε is some positive constant. That is simple too many points to consider all at once. It seems that with the tools avaliable to us we have little to no chance of proving any of the conjectures. We can however show a partial result. We present this in the next chapter.

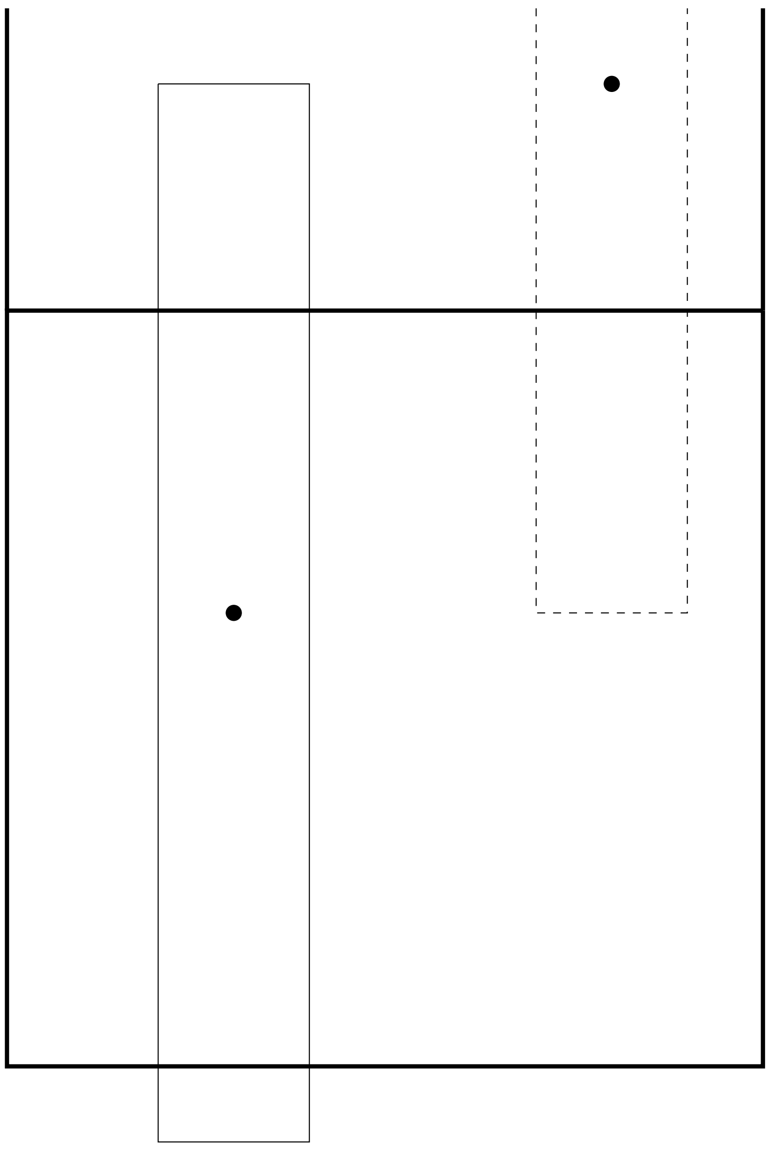

Figure 12.1: An unfortunate, yet very likely situation

CHAPTER 13

A PARTIAL RESULT

Before things become too technical let us note that the problems we encountered was the fact that neighbouring squares could contain points generating rectangles that intersect the square we were considering. We shall get around this problem by changing our setup to that of [Kristensen et al., 2006]. For each fixed i and j instead of considering squares of size $R^{-2n} \times R^{-2n}$ we consider rectangles of size $R^{-(1+2i)2n} \times R^{-(1+2j)2n}$. We will show a version of the Davenport trick, forcing problematic points to lie on a line and hence avoiding a neighbourhood around each point translates to avoiding a neighbourhood of the line. This enables us to show that $\mathbf{Bad}_d(i,j)$ has full dimension and is stable under intersection with suitably nice fractals. We proceed by giving the necessary terminology needed for showing this, essentially following the ideas in [Kristensen et al., 2006, Example 2.3]. Note however that we use w as a dummy index instead of i as we shall need that letter later on.

Let (X,d) be the product space of the t metric spaces (X_w, d_w). Let (Ω, d) be a compact subspace of X which contains the support of a non-atomic finite measure m. Let $\mathcal{R} = \{R_\alpha \in X : \alpha \in J\}$ be a family of subsets R_α indexed by an infinite countable set J. We call R_α a resonant set and notice that each resonant set can be split into its t components $R_{\alpha,w} \subset (X_w, d_w)$. We let $\beta : J \to \mathbb{R}^+$, $\alpha \mapsto \beta_\alpha$ be a positive function on J and we assume that the number of $\alpha \in J$ with β_α bounded above is finite, this last assumption is known as the Northcott property, named after Douglas Northcott who showed that this assumption holds for algebraic numbers. We let $\rho_w : \mathbb{R}^+ \to \mathbb{R}^+, r \mapsto \rho_w(r)$ be a real positive function such that $\rho_w(r) \to 0$ as $r \to \infty$ and that ρ_w is decreasing for r large enough. We also assume that $\rho_1(r) \geq \rho_2(r) \geq \cdots \geq \rho_t(r)$ for large r. Given R_α. Let

$$F_\alpha(\rho_1, \ldots, \rho_t) = \{x \in X : d_w(x_w, R_{\alpha,w}) \leq \rho_w(\beta_\alpha) \text{ for all } 1 \leq w \leq t\},$$

denote the rectangular $(\rho_1, \ldots, \rho_t)$-neighbourhood of R_α and now consider the set

$$\begin{aligned}&\mathbf{Bad}^*(\mathcal{R}, \beta, \rho_1, \ldots, \rho_t)\\&= \{x \in \Omega : \exists c(x) > 0 \text{ such that } x \notin c(x)F_\alpha(\rho_1, \ldots, \rho_t) \text{ for all } \alpha \in J\}.\end{aligned}$$

For $l_1, \ldots, l_t \in \mathbb{R}^+$ and $c \in \Omega$ we let

$$F(c; l_1, \ldots, l_t) = \{x \in X : d_w(x_w, c_w) \leq l_w \text{ for all } 1 \leq w \leq t\},$$

denote the closed rectangle centred at c with sidelengths determined by $l_1, \ldots, l_t$. For any $k > 1$ and $n \in \mathbb{N}$ we let F_n denote a generic rectangle $F(c; \rho_1(k^n), \ldots, \rho_t(k^n)) \cap \Omega$, in Ω centered at a point $c \in \Omega$. We let $B(c,r)$ denote the closed ball with centre c and radius r. We shall need no less than five conditions on the measure m and the functions ρ_w in order for this framework to work out the way we want it to.

A There exists a strictly positive constant δ such that for any $c \in \Omega$ we have

$$\liminf_{r \to 0} \frac{\log m(B(c,r))}{\log r} = \delta.$$

This is a weaker notion than that of being δ-Ahlfors regular and for our purpose Ahlfors regularity will suffice.

B For $k > 1$ sufficiently large, any integer $n \geq 1$ and any $w \in \{1, \ldots, t\}$, we have the following bounds

$$\lambda_w^l(k) \leq \frac{\rho_w(k^n)}{\rho_w(k^{n+1})} \leq \lambda_w^u(k),$$

where λ_w^l and λ_w^u are lower and upper constants tending to infinity as k does.

C There exists constants $0 < a \leq 1 \leq b$ and $l_0 > 0$ such that

$$a \leq \frac{m(F(c; l_1, \ldots, l_t))}{m(F(c'; l_1, \ldots, l_t))} \leq b$$

for any $c, c' \in \Omega$ and any $l_1, \ldots, l_t \leq l_0$. This ensures that the measure of two rectangles of the same "size" but with different centres is not too different.

D There exist strictly positive constants D and l_0 such that

$$\frac{m(2F(c; l_1, \ldots, l_t))}{m(F(c; l_1, \ldots, l_t))} \leq D$$

for any $c \in \Omega$ and any $l_1, \ldots, l_t \leq l_0$. This is simply the D-Federer condition for rectangles.

E For $k > 1$ sufficiently large and any integer $n \geq 1$

$$\frac{m(F_n)}{m(F_{n+1})} \geq \lambda(k),$$

where $\lambda(k) \to \infty$ as $k \to \infty$.

We now have the following theorem

Theorem 13.1. *[Kristensen et al., 2006, Theorem 3] For $1 \leq w \leq t$, let (X_w, d_w) be a metric space and (Ω_w, d_w, m_w) be a compact measure subspace of X_w where the measure m_w is δ_w-Ahlfors regular. Let (X, d) be the product space of the spaces (X_w, d_w) and let (Ω, d, m) be the product measure space of the measure spaces (Ω_w, d_w, m_w). Let the functions ρ_w satisfy condition B. For $k \geq k_0 > 1$, suppose there exists some $\theta \in \mathbb{R}^+$ so that for $n \geq 1$ and any rectangle F_n there exists a disjoint collection $\mathcal{C}(\theta F_n)$ of rectangles $2\theta F_{n+1}$ contained within θF_n satisfying the following two conditions*

$$\#\mathcal{C}(\theta F_n) \geq \kappa_1 \prod_{w=1}^{t} \left(\frac{\rho_w(k^n)}{\rho_w(k^{n+1})} \right)^{\delta_w}$$

and

$$\#\left\{2\theta F_{n+1} \subset \mathcal{C}(\theta F_n) : \min_{\alpha \in J(n+1)} d_w(c_w, R_{\alpha,w}) \leq 2\theta\rho_w(k^{n+1}) \text{ for any } 1 \leq w \leq t\right\}$$
$$\leq \kappa_2 \prod_{w=1}^{t} \left(\frac{\rho_w(k^n)}{\rho_w(k^{n+1})}\right)^{\delta_w},$$

where $0 < \kappa_2 < \kappa_1$ *are absolute constants independent of* k *and* n*, and where* $J(n) = \{\alpha \in J : k^{n-1} \leq \beta_\alpha < k^n\}$*. Furthermore suppose that* $\dim_{\mathrm{H}}(\cup_{\alpha \in J} R_\alpha) < \sum_{w=1}^{t} \delta^w$*. Then*

$$\dim_{\mathrm{H}} \mathbf{Bad}^*(\mathcal{R}, \beta, \rho_1, \ldots, \rho_t) = \sum_{w=1}^{t} \delta^w.$$

We shall not go into the proof of this theorem as it is quite technical. We shall instead apply it on the sets $\mathbf{Bad}_d(i,j)$.

Proposition 13.2. *The sets* $\mathbf{Bad}_d(i,j)$ *have dimension* 2 *for all* $0 < i, j < 1$ *with* $i + j = 1$*.*

Proof. Let $I = [0,1]$ and assume that $i \leq j$. Let $\rho_1(r) = r^{-(1+2i)}, \rho_2(r) = r^{-(1+2j)}$ and let

$$X = \Omega = I^2, J = \{(p,q) \in \mathcal{O}_d \times \mathcal{O}_d \setminus \{0\} : |p| \leq |q|\},$$
$$\alpha = (p,q) \in J, \beta_\alpha = |q|, R_\alpha = (\mathrm{Re}(p/q), \mathrm{Im}(p/q)).$$

$d_1 = d_2$ is the standard Euclidean metric on I and $m_1 = m_2$ is the one-dimensional Lebesgue measure on I. Hence d is the product metric on I^2 and m is the two-dimensional Lebesgue measure on I^2. Clearly the Lebesgue measure is Ahlfors regular with $\delta_1 = \delta_2 = 1$, and the functions ρ_1, ρ_2 satisfy condition B. We shall now establish the existence of the collection $\mathcal{C}(\theta F_n)$, where F_n is any rectangle of size $2k^{-n(1+2i)} \times 2k^{-n(1+2j)}$. We note that $m(\theta F_n) = 4\theta^2 k^{-4n}$. We shall now give establish a version of the Davenport trick to help us on the way. Assume that there are at least three rational points $a = p_1/q_1, b = p_2/q_2, c = p_3/q_3$ with $k^n \leq |q_1|, |q_2|, |q_3| < k^{n+1}$ lying inside θF_n. Suppose they do not lie on a line and thus form a triangle T. The area of the triangle can be calculated using the so-called shoelace formula, which is based solely on the geometry of the complex numbers and therefore not dependant on the value of d.

$$m(T) = \frac{i}{4}\begin{vmatrix} a & \bar{a} & 1 \\ b & \bar{b} & 1 \\ c & \bar{c} & 1 \end{vmatrix} \geq \frac{1}{4}k^{-4(n+1)}$$

setting θ to any value $< 1/4k^{-2}$ we see that $m(T) > m(\theta F_n)$ but that is not possible as $T \subset \theta F_n$. Hence the triangle cannot exist. So if there are two or more rational points inside θF_n with denominators between k^n and k^{n+1} they must lie on a line, $\mathcal{L}$. We now partition the rectangle θF_n into rectangles $2\theta F_{n+1}$ of size $4k^{-(n+1)(1+2i)} \times 4k^{-(n+1)(1+2j)}$ and denote by $\mathcal{C}(\theta F_n)$ the collection of rectangles $2\theta F_{n+1}$ we obtain. We now have that

$$\#\mathcal{C}(\theta F_n) \geq \frac{2\theta k^{-n(1+2i)}}{4\theta k^{-(n+1)(1+2i)}} \frac{2\theta k^{-n(1+2j)}}{4\theta k^{-(n+1)(1+2j)}} \geq \frac{k^4}{16}.$$

The Davenport trick also shows us that

$$\#\left\{2\theta F_{n+1} \subset \mathcal{C}(\theta F_n) : \min_{\alpha \in J(n+1)} d_w(c_w, R_{\alpha,w}) \leq 2\theta\rho_w(k^{n+1}) \text{ for any } 1 \leq w \leq 2\right\} \tag{13.1}$$

$$\leq \#\{2\theta F_{n+1} \subset \mathcal{C}(\theta F_n) : 2\theta F_{n+1} \cap \mathcal{L} \neq \emptyset\}, \tag{13.2}$$

where $\mathcal{L}$ is any line passing through θF_n. As we have assumed that $i \leq j$ the quantity on the right hand side can be shown to be less than or equal to $(2\theta k^{-n(1+2j)})/(4\theta k^{-(n+1)(1+2j)})$ which is seen to be less than $k^4/32$. This ensures that the collection of rectangles satisfy the conditions required. Which in turn allows us to use Theorem 13.1 to conclude that $\dim_{\mathrm{H}} \mathbf{Bad}_d(i,j) \cap I^2 = 2$. Since $I^2 \subset \mathbb{R}^2$ we conclude that $\dim_{\mathrm{H}} \mathbf{Bad}_d(i,j) = 2$ and thus have full dimension. □

It is possible to tweak the framework of [Kristensen et al., 2006] even more and show that $\mathbf{Bad}_d(i,j)$ behaves well under intersection with suitably nice fractals, i.e. those being the support of a Ahlfors regular measure which is absolutely decaying. We shall not do that though as it would require another two theorems to justify the result. Instead we will show that $\mathbf{Bad}_d(i,j)$ is winning in yet another variant of Schmidt's game.

13.1 A new game

Just as we have defined the k-dimensionally β-absolute game we can for each fixed pair of i and j define the k-dimensionally (β, i, j)-absolute game. We explain how it works in the case $k = n-1$ which we call the hyperplane (β, i, j)-absolute game.

Bob initially picks $x_1 \in \mathbb{R}^2$ and $r_1 > 0$ defining the closed rectangle with centre x_1 and sidelengths r_1^{1+2i}, r_1^{1+2j}, denoted by B_1. At each stage of the game, after Bob has chosen x_k, r_k Alice chooses a hyperplane and removes an $\varepsilon(k)$-neighbourhood of it, denoted by A_k from B_k. Where $0\varepsilon(k) \leq \beta r_k$. Then Bob chooses x_{k+1} and r_{k+1} such that $r_{k+1} \geq \beta r_k$ and

$$B_{k+1} \subset (B_k \setminus A_k)\,.$$

A set E is said to be *hyperplane* (β, i, j)*-absolute winning* if Alice has a strategy that ensures that $\cap_k B_k$ intersects E.

Essentially the only difference between the hyperplane absolute game and this is the fact that we play on rectangles rather than balls. Indeed the case where $i = j = 1/2$ reduces to the original hyperplane absolute game, only that the original radius is changed from r_1 to r_1^2.

Theorem 13.3. *The set* $\mathbf{Bad}_d(i,j)$ *is hyperplane* (β, i, j)*-absolute winning*

Before we show this we have extracted the Davenport trick from Proposition 13.2 and modified it to fit into the (β, i, j)-game. We will from now on assume that $i \geq j$.

Lemma 13.4. *For every* $\beta \in (0,1)$ *and for every* $k \in \mathbb{N}$ *let*

$$U_{k,i} = \left\{\frac{p}{q} : p, q \in \mathcal{O}_d \text{ and } \beta^{-\frac{(k-1)}{1+2i}} \leq |q| < \beta^{-\frac{k}{1+2i}}\right\}$$

Then there exists a line $\mathcal{L}$ *containing* $U_{k,i} \cap B_k$.

Proof of Theorem 13.3. Let $\beta < 1/3$, let i, j be fixed and assume once more that $i \geq j$. When the hyperplane β-absolute game begins, Alice makes dummy moves untill the radius, r, is small enough to be less than $\frac{1}{2}\beta^{-1}$. We then set $c = \beta^2 r$. Now let B_{lk} be the subsequence of moves where the radius r_{lk} first satisfies the inequalities: $\beta^{k-1}r \geq r_{lk} > \beta^k r$. Alice continues to make dummy moves outside this subsequence, and on the turns in this subsequence Alice can choose the line as described by the above lemma, denoted by L_k and pick

$$A_{lk+1} = \mathcal{L}_k^{\beta^{k+1}r}.$$

Here the right hand side denotes the $\beta^{k+1}r$-neighbourhood of the line $\mathcal{L}_k$. By choice of c we have

$$\beta^{k+1}r = c\beta^{k-1}.$$

Now since $|q| \geq \beta^{-\frac{(k-1)}{1+2i}}$ we immediately get that the right hand side of the above has to be bigger than $c\,|q|^{1+2i}$. Thus Alice can black out the entire rectangular neighbourhood in her next turn. □

At this time it is not entirely clear what a game of this type would bring to the table. Clearly we cannot hope to show any theorems on the intersections between sets winning the (β, i, j)-game and the (β', i', j')-game, at least not in general. This is also true for the modified Schmidt games invented by Kleinbock and Weiss in [Kleinbock and Weiss, 2010], and for this very reason they could not prove Schmidt's conjecture with that setup. We can get a nice intersection theorem for fractals though, as we have already mentioned earlier in this chapter.

As a concluding remark the author would like to stress that much of the structure of the sets $\mathbf{Bad}_d(i, j)$ is still unknown, as long as $(i, j) \neq (1/2, 1/2)$. We can show full dimension of $\mathbf{Bad}_d(i, j)$ and get the best possible intersection properties when it comes to suitably nice fractals. Yet we still cannot show that the set is winning in Schmidt's original game, only in this modified (i, j) version.

APPENDIX

Proof of Lemma 5.3. We start by the additional assumption that $II_{k,a}$ consists of only one parallelogram with lattice-parallel sides and denote the lower left corner by (α_1, β_1) and the upper right corner by (α_2, β_2). First we note that $(n_1, n_2) \cdot_d (x, y) = (n_1, n_2)A(x, y)$, where A as before is the matrix that transforms our lattice to the standard lattice $\mathbb{Z}^2$. Now we begin with the case $n_1 = n_2 = 0$ and note that Ψ_a agrees with the indicator function of $I_{k,a}$ almost everywhere, hence using Fubini's Theorem:

$$
\begin{aligned}
c_{0,0} &= \int_{\mathbb{R}^2} f_a(x, y) e(0)\, d(x, y) \\
&= \int_{\mathbb{R}^2} \frac{1}{\Delta^2} \int_{-\Delta/2}^{\Delta/2} \int_{-\Delta/2}^{\Delta/2} \Psi_a(x + x_1, y + y_1)\, dx_1\, dy_1\, d(x, y) \\
&= \frac{1}{\Delta^2} \int_{-\Delta/2}^{\Delta/2} \int_{-\Delta/2}^{\Delta/2} \int_{\mathbb{R}^2} \Psi_a(x + x_1, y + y_1)\, d(x, y)\, dx_1\, dy_1 \\
&= \frac{1}{\Delta^2} \int_{-\Delta/2}^{\Delta/2} \int_{-\Delta/2}^{\Delta/2} \int_{\mathbb{R}^2} 1_{I_{k,a}}(x + x_1, y + y_1)\, d(x, y)\, dx_1\, dy_1 \\
&= \frac{1}{\Delta^2} \int_{-\Delta/2}^{\Delta/2} \int_{-\Delta/2}^{\Delta/2} \int_{\mathbb{R}^2} 1_S(x + x_1, y + y_1) \det A\, d(x, y)\, dx_1\, dy_1 \\
&= \det A \frac{1}{\Delta^2} \int_{-\Delta/2}^{\Delta/2} \int_{-\Delta/2}^{\Delta/2} \int_{\mathbb{R}} \int_{\mathbb{R}} 1_{(\alpha_1 - x_1, \alpha_2 - x_1)}(x) 1_{(\beta_1 - y_1, \beta_2 - y_1)}(y)\, dx\, dy\, dx_1\, dy_1 \\
&= \det A \frac{1}{\Delta^2} \int_{-\Delta/2}^{\Delta/2} \int_{-\Delta/2}^{\Delta/2} (\alpha_2 - \alpha_1)(\beta_2 - \beta_1)\, dx_1\, dy_1 \\
&= \det A (\alpha_2 - \alpha_1)(\beta_2 - \beta_1)
\end{aligned}
$$

where S is the square with lower left and upper right corners given by (α_1, β_1), respectively (α_2, β_2).

In the case that $n_1 \neq 0$ we have

$$
\begin{aligned}
c_{n_1,0} &= \int_{\mathbb{R}^2} f_a(x,y) e(-(n_1 0) A(xy))\, d(x,y) \\
&= \int_{\mathbb{R}^2} \frac{1}{\Delta^2} \int_{-\Delta/2}^{\Delta/2} \int_{-\Delta/2}^{\Delta/2} \Psi_a(x+x_1, y+y_1) e(-(n_1 0) A(xy))\, dx_1\, dy_1\, d(x,y) \\
&= \frac{1}{\Delta^2} \int_{-\Delta/2}^{\Delta/2} \int_{-\Delta/2}^{\Delta/2} \int_{\mathbb{R}^2} \Psi_a(x+x_1, y+y_1) e(-(n_1 0) A(xy))\, d(x,y)\, dx_1\, dy_1 \\
&= \frac{1}{\Delta^2} \int_{-\Delta/2}^{\Delta/2} \int_{-\Delta/2}^{\Delta/2} \int_{\mathbb{R}^2} 1_{I_{k,a}}(x+x_1, y+y_1) e(-(n_1 0) A(xy))\, d(x,y)\, dx_1\, dy_1 \\
&= \frac{1}{\Delta^2} \int_{-\Delta/2}^{\Delta/2} \int_{-\Delta/2}^{\Delta/2} \int_{\mathbb{R}^2} 1_S(x+x_1, y+y_1) e(-n_1 x) \det A\, d(x,y)\, dx_1\, dy_1 \\
&= \det A \frac{1}{\Delta^2} \int_{-\Delta/2}^{\Delta/2} \int_{-\Delta/2}^{\Delta/2} \int_{\mathbb{R}} \int_{\mathbb{R}} 1_{(\alpha_1 - x_1, \alpha_2 - x_1)}(x) 1_{(\beta_1 - y_1, \beta_2 - y_1)}(y) e(-n_1 x)\, dx\, dy\, dx_1\, dy_1 \\
&= \det A \frac{\beta_2 - \beta_1}{\Delta} \int_{-\Delta/2}^{\Delta/2} \int_{\mathbb{R}} 1_{(\alpha_1 - x_1, \alpha_2 - x_1)}(x) e(-n_1 x)\, dx\, dx_1 \\
&= \det A \frac{\beta_2 - \beta_1}{\Delta} \int_{-\Delta/2}^{\Delta/2} \left[\frac{-1}{2\pi i n_1} e(-n_1 x) \right]_{x = \alpha_1 - x_1}^{\alpha_2 - x_1} dx_1 \\
&= \det A \frac{\beta_2 - \beta_1}{\Delta} \int_{-\Delta/2}^{\Delta/2} \frac{-1}{2\pi i n_1} \left(e(n_1(\alpha_2 - x_1)) - e(n_1(\alpha_1 - x_1)) \right) dx_1 \\
&= \det A \frac{\beta_2 - \beta_1}{\Delta} \left(\left[\frac{-1}{2\pi i n_1} \frac{-1}{2\pi i n_1} e(n_1 \alpha_2) e(-n_1 x_1) \right]_{-\Delta/2}^{\Delta/2} - \left[\frac{-1}{2\pi i n_1} \frac{-1}{2\pi i n_1} e(n_1 \alpha_1) e(-n_1 x_1) \right]_{-\Delta/2}^{\Delta/2} \right) \\
&= \det A \frac{\beta_2 - \beta_1}{\Delta} \frac{-1}{4\pi^2 n_2^2} \left(e(n_1 \alpha_2) - e(n_1 \alpha_1) \right) \left(e(-n_1 \Delta/2) - e(n_1 \Delta/2) \right).
\end{aligned}
$$

Analogously we find for $n_2 \neq 0$ that

$$
c_{0,n_2} = \det A \frac{\alpha_2 - \alpha_1}{\Delta} \frac{-1}{4\pi^2 n_1^2} \left(e(n_2 \beta_2) - e(n_2 \beta_1) \right) \left(e(-n_2 \Delta/2) - e(n_2 \Delta/2) \right).
$$

Now for $n_1, n_2 \neq 0$ we have

$$
\begin{aligned}
c_{n_1,n_2} &= \int_{\mathbb{R}^2} f_a(x,y)e(-(n_1n_2)A(xy))\,d(x,y)\\
&= \int_{\mathbb{R}^2} \frac{1}{\Delta^2}\int_{-\Delta/2}^{\Delta/2}\int_{-\Delta/2}^{\Delta/2} \Psi_a(x+x_1,y+y_1)e(-(n_1n_2)A(xy))\,dx_1\,dy_1\,d(x,y)\\
&= \frac{1}{\Delta^2}\int_{-\Delta/2}^{\Delta/2}\int_{-\Delta/2}^{\Delta/2}\int_{\mathbb{R}^2} 1_{I_{k,a}}(x+x_1,y+y_1)e(-(n_1n_2)A(xy))\,d(x,y)\,dx_1\,dy_1\\
&= \frac{1}{\Delta^2}\int_{-\Delta/2}^{\Delta/2}\int_{-\Delta/2}^{\Delta/2}\int_{\mathbb{R}^2} 1_S(x+x_1,y+y_1)e(-(n_1x+n_2y))\det A\,d(x,y)\,dx_1\,dy_1\\
&= \frac{\det A}{\Delta^2}\int_{-\Delta/2}^{\Delta/2}\int_{-\Delta/2}^{\Delta/2}\int_{\mathbb{R}}\int_{\mathbb{R}} 1_{(\alpha_1-x_1,\alpha_2-x_1)}(x)1_{(\beta_1-y_1,\beta_2-y_1)}(y)e(-(n_1x+n_2y))\,dx\,dy\,dx_1\,dy_1\\
&= \frac{\det A}{\Delta^2}\int_{-\Delta/2}^{\Delta/2}\int_{-\Delta/2}^{\Delta/2}\int_{\mathbb{R}}\int_{\mathbb{R}} 1_{(\alpha_1-x_1,\alpha_2-x_1)}(x)1_{(\beta_1-y_1,\beta_2-y_1)}(y)e(-(n_1x+n_2y))\,dx\,dy\,dx_1\,dy_1\\
&= \frac{\det A}{\Delta^2}\int_{-\Delta/2}^{\Delta/2}\int_{-\Delta/2}^{\Delta/2}\int_{\mathbb{R}} 1_{(\beta_1-y_1,\beta_2-y_1)}(y)e(-n_2y)\int_{\mathbb{R}} 1_{(\alpha_1-x_1,\alpha_2-x_1)}(x)e(-n_1x)\,dx\,dy\,dx_1\,dy_1\\
&= \frac{\det A}{\Delta^2}\int_{-\Delta/2}^{\Delta/2}\int_{-\Delta/2}^{\Delta/2}\int_{\mathbb{R}} 1_{(\beta_1-y_1,\beta_2-y_1)}(y)e(-n_2y)\left[\frac{-1}{2\pi i n_1}e(-n_1x)\right]_{x=\alpha_1-x_1}^{\alpha_2-x_1}dy\,dx_1\,dy_1\\
&= \frac{\det A}{\Delta^2}\int_{-\Delta/2}^{\Delta/2}\int_{-\Delta/2}^{\Delta/2}\left[\frac{-1}{2\pi i n_2}e(-n_2y)\right]_{y=\beta_1-y_1}^{\beta_2-y_1}\left[\frac{-1}{2\pi i n_1}e(-n_1x)\right]_{x=\alpha_1-x_1}^{\alpha_2-x_1}dx_1\,dy_1\\
&= \frac{\det A}{\Delta^2}\int_{-\Delta/2}^{\Delta/2}\int_{-\Delta/2}^{\Delta/2}\left(\frac{-1}{2\pi i n_2}\left(e(-n_2\beta_2+n_2y_1)-e(-n_2\beta_1+n_2y_1)\right)\right)\\
&\qquad\cdot\left(\frac{-1}{2\pi i n_1}\left(e(-n_1\alpha_2+n_1x_1)-e(-n_1\alpha_1+n_1x_1)\right)\right)dx_1\,dy_1\\
&= \frac{-\det A}{4\Delta^2\pi^2}\int_{-\Delta/2}^{\Delta/2}\frac{1}{n_1}\left(e(-n_1\alpha_1+n_1x_1)-e(-n_1\alpha_1+n_1x_1)\right)\,dx_1\\
&\qquad\cdot\int_{-\Delta/2}^{\Delta/2}\frac{1}{n_2}\left(e(-n_2\beta_1+n_2y_1)-e(-n_2\beta_1+n_2y_1)\right)\,dy_1\\
&= \frac{-\det A}{4\Delta^2\pi^2}\left(\left[\frac{1}{n_1}\frac{1}{2\pi i n_1}e(-n_1\alpha_2+n_1x_1)\right]_{x_1=-\Delta/2}^{\Delta/2}-\left[\frac{1}{n_1}\frac{1}{2\pi i n_1}e(-n_1\alpha_1+n_1x_1)\right]_{x_1=-\Delta/2}^{\Delta/2}\right)\\
&\qquad\cdot\left(\left[\frac{1}{n_2}\frac{1}{2\pi i n_2}e(-n_2\beta_2+n_2y_1)\right]_{y_1=-\Delta/2}^{\Delta/2}-\left[\frac{1}{n_2}\frac{1}{2\pi i n_2}e(-n_2\beta_1+n_2y_1)\right]_{y_1=-\Delta/2}^{\Delta/2}\right)\\
&= \frac{-\det A}{4\Delta^2\pi^2}\left(\frac{1}{2\pi i n_1^2}\Big(e(-n_1\alpha_2+n_1\Delta/2)-e(-n_1\alpha_2-n_1\Delta/2)\right.\\
&\qquad\left.-e(-n_1\alpha_1+n_1\Delta/2)+e(-n_1\alpha_1-n_1\Delta/2)\Big)\right)\\
&\qquad\cdot\left(\frac{1}{2\pi i n_2^2}\Big(e(-n_2\beta_2+n_2\Delta/2)-e(-n_2\beta_2-n_2\Delta/2)\right.\\
&\qquad\left.-e(-n_2\beta_1+n_2\Delta/2)+e(-n_2\beta_1-n_2\Delta/2)\Big)\right).\\
&= \frac{\det A}{16\Delta^2\pi^4 n_1^2n_2^2}\Big(e(-n_1\alpha_2)-e(-n_1\alpha_1))(e(n_1\Delta/2)-e(-n_1\Delta/2))\Big)\\
&\qquad\cdot\Big(e(-n_2\beta_2)-e(-n_2\beta_1))(e(n_2\Delta/2)-e(-n_2\Delta/2))\Big).
\end{aligned}
$$

So we see that

$$\begin{aligned} c_{n_1,n_2} &= \mathcal{O}\left(\frac{1}{\Delta^2 n_1^2 n_2^2}\right), && (n_1, n_2 \neq 0), \\ c_{n_1,0} &= \mathcal{O}\left(\frac{1}{\Delta^2 n_1^2}\right), && (n_1 \neq 0), \\ c_{0,n_2} &= \mathcal{O}\left(\frac{1}{\Delta^2 n_2^2}\right), && (n_2 \neq 0), \\ c_{0,0} &= \det A(\alpha_2 - \alpha_1)(\beta_2 - \beta_1). \end{aligned} \tag{13.3}$$

From equation (13.3) it is obvious that the contribution to the Fourier coefficients from each rectangle depends only on the vertices of the rectangle. Since the cases $n_1 \neq 0, n_2 \neq 0$ and $n_1, n_2 \neq 0$ are quite similar we only consider the last of these. For $n_1, n_2 \neq 0$ observer that the Fourier coefficients are of the form

$$c_{n_1,n_2} = C(n_1, n_2) \sum_{\alpha_i,\beta_i} \operatorname{sgn}((\alpha_i, \beta_i)) e(n_1\alpha_i + n_2\beta_i),$$

where the sum is taken over all vertices (α_i, β_i) of the parallelogrammical subdomains, and where sgn is the sign of the given vertex, which is taken to be positive if the vertex is the upper right or lower left vertex of such a parallelogram, and negative if it is the upper left or lower right vertex. Now with this convention it is easily seen that the contributions from all vertices that are not vertices in $II_{k,a}$ cancel. Hence in the general setting we can estimate the Fourier coefficients with the amount of vertices in $II_{k,a}$ times the contribution from each parallelogram. There are $\mathcal{O}(\mu^k)$ vertices in $II_{k,a}$ so we get

$$c_{n_1,n_2} = \mathcal{O}(\frac{\mu^k}{\Delta^2 n_1^2 n_2^2}) \quad (n_1, n_2 \neq 0).$$

As already meantioned we can treat the other cases in a similar fashion and arrive at the desired result. □

BIBLIOGRAPHY

S. Akiyama and J. M. Thuswaldner. Topological properties of two-dimensional number systems. *Journal de Théorie des Nombres de Bordeaux, vol 12. p. 69-79*, 2000.

J. An. Badziahin-pollington-velani's theorem and schmidt's game. *Bull. London. Math. Soc. vol 45*, 2013.

G. I. Arkhipov, V. N. Chubarikov, and A. A. Karatsuba. Trigonometric sums in number theory and analysis. *de Gruyters Exp. Math. 39*, 2004.

D. Badziahin, A. Pollintong, and S. Velani. On a problem in simultaneous diophantine approximation: Schmidt's conjecture. *Annals of Mathematics vol 174*, 2011.

C. J. Bishop and P. W. Jones. Hausdorff dimension and kleinian groups. *Acta Math. 111*, 1997.

É. Borel. Les probabilités dénombrables et leurs applications arithmétiques. *Rend. Circ. Math. Palermo, 27*, 1909.

R. Broderick, L. Fishman, D. Kleinbock, A. Reich, and B. Weiss. The set of badly approximable vectors is strongly c^1 incompressible. *Mathematical Proceedings of the Cambridge Philosophical Society*, 2012.

Y. Bugeaud. *Approximation by algebraic numbers.*, volume 160 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 2004.

Y. Bugeaud. Diophantine approximation and Cantor sets. *Math. Ann.*, 341(3): 677–684, 2008. ISSN 0025-5831.

J.W.S. Cassels. *An introduction to the geometry of numbers*. Springer-Verlag, 1959.

H. Davenport and P. Erdős. Note on normal decimals. *Canad. J. Math. 4, p. 58-63*, 1952.

M. M. Dodson and S. Kristensen. Hausdorff dimension and Diophantine approximation. In *Fractal geometry and applications: a jubilee of Benoît Mandelbrot. Part 1*, volume 72 of *Proc. Sympos. Pure Math.*, pages 305–347. Amer. Math. Soc., Providence, RI, 2004.

M. Drmota and R. F. Tichy. *Sequences, discrepancies and applications.*, volume 1651 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1997.

M. Einsiedler, A. Katok, and E. Lindenstrauss. Invariant measures and the set of exceptions to littlewood's conjecture. *Annals of Mathematics vol 164*, 2006.

R. Esdahl-Schou and S. Kristensen. On badly approximable complex numbers. *Glasgow Mathematical Journal*, 2010.

K. Falconer. *Fractal Geometry - Mathematical Foundtains and Applications.* John Wiley, 2003.

J.-L. Fernandez and M. V. Melián. Bounded geodesics of riemann surfaces and hyperbolic manifolds. *Trans. Amer. Math. Soc. 27*, 1995.

L. Fishman. Schmidt's game on fractals. *Israel J. Math. 171*, 2009.

W. J. Gilbert. Radix representations of quadratic fields. *Journal of mathematical analysis and applications, vol 83.*, 1981.

W. J. Gilbert. Complex bases and fractal similarity. *Ann. sc. math. Québec vol. 11*, 1987.

B. Gittenberger and J. M. Thuswaldner. Asymptotic normality of b-additive functions on polynomial sequences in the Gaussian number field. *J. Number Theory*, 84(2):317–341, 2000. ISSN 0022-314X.

L.-K. Hua. On exponential sums over an algebraic number field. *Canadian J. Math.*, 3:44–51, 1951.

L.-K. Hua. *Additive theory of prime numbers.* Translations of Mathematical Monographs, Vol. 13. American Mathematical Society, Providence, R.I., 1965.

J. E. Hutchinson. Fractals and self-similarity. *Indiana Univ. Math. J. 30*, 1981.

M. N. Huxley. *Area, lattice points, and exponential sums.*, volume 13 of *London Mathematical Society Monographs. New Series.* The Clarendon Press Oxford University Press, New York, 1996. Oxford Science Publications.

K.-H. Indlekofer, I. Kátai, and P. Racskó. Number systems and fractal geometry. *Probability Theory and Applications, Math. Appl. 80, p. 319-334*, 1992.

V. Jarník. Zur metrischen theorie der diophantischen approximationen. *Prace Mat. Fiz. 36*, 1929.

I. Kátai and B. Kovács. Canonical number systems in imaginary quadratic fields. *Acta Math. Acad. Sci. Hungar.*, 37(1-3):159–164, 1981.

I. Kátai and I. Környei. On number systems in algebraic number fields. *Publ. Math. Debrecen vol. 41, p. 289-294*, 1992.

I. Kátai and J. Szabó. Canonical number systems for complex integers. *Acta. Sci. Math. (Szeged) 37, p. 255-260*, 1975.

A. Khintchine. Zur metrischen theorie der diophantischen approximationen. *Math. Z. 24*, 1926.

D. Kleinbock and B. Weiss. Modified schmidt games and diophantine approximation with weights. *Adv. Math. 223*, 2010.

D. Kleinbock, E. Lindenstrauss, and B. Weiss. On fractal measures and Diophantine approximation. *Selecta Math. (N.S.)*, 10(4):479–523, 2004. ISSN 1022-1824.

N. M. Korobov. *Exponential sums and their applications.*, volume 80 of *Mathematics and its Applications (Soviet Series).* Kluwer Academic Publishers Group, Dordrecht, 1992. Translated from the 1989 Russian original by Yu. N. Shakhov.

B. Kóvacs. Canonical number systems in algebraic number fields. *Acta. Math. Hungar. 37*, 1981.

B. Kovács and A. Pethő. On a representation of algebraic integers. *Studia Sci. Math. Hungar.*, 27(1-2):169–172, 1992.

S. Kristensen. Approximating numbers with missing digits by algebraic numbers. *Proc. Edinb. Math. Soc. (2)*, 49(3):657–666, 2006. ISSN 0013-0915.

S. Kristensen, Thorn. R., and S. L. Velani. Diophantine approximation and badly approximable sets. *Advances in Mathematics 203*, 2006.

E. Lánczi. Unique prime factorization in imaginary quadratic number fields. *Acta Math. Acad. Sci. Hungar.*, 16:453–466, 1965.

W. J. LeVeque. Continued fractions and approximations i and ii. *Indig. Math. 14*, 1952.

J. Levesley, C. Salp, and S. L. Velani. On a problem of K. Mahler: Diophantine approximation and Cantor sets. *Math. Ann.*, 338(1):97–118, 2007. ISSN 0025-5831.

M. G. Madritsch. A note on normal numbers in matrix number systems. *Math. Pannon. 18, p. 219-227*, 2007.

M. G. Madritsch. Generating normal numbers over Gaussian integers. *Acta Arith.*, 135(1):63–90, 2008. ISSN 0065-1036.

M. G Madritsch, J. M. Thuswaldner, and R. F. Tichy. Normality of numbers generated by the values of entire functions. *J. Number Theory 128 p. 1127-1145*, 2008.

C. T. McMullen. Winning sets, quasiconformal maps and diophantine approximation. *Geometric And Functional Analysis*, 2010.

W. Müller, J. M. Thuswaldner, and R. F. Tichy. Fractal properties of number systems. *Periodica Mathematica Hungarica Vol. 42*, 2001.

H. L. Montgomery. *Ten lectures on the interface between analytic number theory and harmonic analysis.*, volume 84 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC, 1994.

Y.-N. Nakai and I. Shiokawa. Discrepancy estimates for a class of normal numbers. *Acta Arith. 62 p. 271-284*, 1992.

W. Narkiewicz. *Elementary and analytic theory of algebraic numbers.* PWN—Polish Scientific Publishers, Warsaw, 1974. Monografie Matematyczne, Tom 57.

W. Narkiewicz. *Elementary and analytic theory of algebraic numbers.* Springer Verlag, Warsaw, 1990. Monografie Matematyczne, Tom 57.

M. B. Nathanson. *Elementary methods in number theory.*, volume 195 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, 2000.

A. D. Pollington and S. L. Velani. Metric diophantine approximation and 'absolutely friendly' measures. *Selecta Math. 11*, 2005.

K. Sand-Jensen. How to write consistently boring scientific literature. *Oikos, Vol. 116, No. 5. pp. 723-727.*, 2007.

K. Scheicher and J. M. Thuswaldner. Canonical number systems, counting automata and fractals. *Mathematical Proceedings of the Cambridge Philosophical Society, 133.*, 2002.

J. Schiffer. Discrepancy of normal numbers. *Acta Arith. 47 p. 175-186*, 1986.

W. M. Schmidt. On badly approximable numbers and certain games. *Trans. Amer. Math. Soc. 123*, 1966.

W. M. Schmidt. *Diophantine Approximation.* Springer-Verlag, 1980.

D. Sullivan. disjoint spheres, approximation by imaginary quadratic numbers, and the logarithm law for geodesics. *Acta. Math. vol 149*, 1982.

E.C. Titchmarsh. *The Theory of the Riemann Zeta-Function, 2nd ed.* Oxford University Press, 1986.

P. Turán. *On a new method of analysis and its applications.* Pure and Applied Mathematics (New York). John Wiley & Sons Inc., New York, 1984. With the assistance of G. Halász and J. Pintz, With a foreword by Vera T. Sós, A Wiley-Interscience Publication.