Institut de Recherche Mathématique Avancée
7, rue Rene Descartes
68084 Strasbourg Cedex
France

Thèse

Sur deux formes d'approximation

On two forms of approximation

Simon Kristensen

"Life is hard, but mathematics is harder"

— Robert Storm Petersen
Danish cartoonist

Preface

In this thesis, we will discuss two forms of approximation. In the first part, we will discuss metrical Diophantine approximation. In particular, we will discuss Diophantine approximation over the field of Laurent series with coefficients from a finite field.

Diophantine approximation is a closer study of the density of the rationals in the reals. Metrical Diophantine approximation is a subset of this study using measure theory. The field of Laurent series is a characteristic p locally compact field with number theoretical properties remarkably similar to the ones found in the reals. In particular, we have analogues to integers, rationals and irrationals, and the Euclidean algorithm works over the field. These things and more will be discussed in the course of Part 1.

In the first chapter, we begin with an introduction to the topic of Diophantine approximation. Secondly, we introduce the specific setting of this thesis. Namely, we construct the field of Laurent series with coefficients from a finite field. We also introduce the Haar measure and Hausdorff dimension in this setting. Thirdly, we remark on some differences and similarities between our setting and the field of p-adic numbers, where some previous results in metrical number theory exist. Finally we give a brief discussion of some previous results in Diophantine approximation in the field of Laurent series. Most of these are far from the metrical theory.

In the second chapter, we discuss Lüroth series over the field of Laurent series. Lüroth series are a series expansion of any given element in the field, where each summand is a rational. Hence, these expansions give some information on the distribution of the rational elements in the real elements. Using probabilistic methods, we prove a number of results on subsequences of the coefficients of these expansions.

In the third chapter, we prove the first metrical results regarding the set of elements which are approximable by rationals with "small" denominators. Also, this is the first chapter where our setting is expanded from one to several dimensions. We prove the main theorem of this chapter both in one and multiple dimensions in order to illustrate the difficulties encountered in this transition. The result is a zero–one law: Almost all elements are approximable by rational elements with denominators of a certain magnitude, but when a certain threshold is reached, almost none are approximable in this way.

In the fourth and fifth chapter, we discuss two exceptional sets of matrices (or linear forms) arising from the zero–one law of Chapter 3. That is, two sets of measure zero that are interesting from the point of view of a number theorist. In Chapter 4, we discuss linear forms that are so close to rationals that the measure of the set of these is zero. In fact, we study a continuum of such sets, and we are able to calculate the Hausdorff dimension of the sets as a function of the degree of "approximability" required.

In Chapter 5, we discuss the converse notion of a set of elements that are so far from being rational that the measure of the set is zero. In this case, we only discuss one such set. Again, we are able to calculate the Hausdorff dimension of the set, which turns out to be full.

In the final chapter of Part 1, we propose a few directions for further research in the field of metrical Diophantine approximation in the field of Laurent series. Some of these would require a large amount of work, since preciously little is known about the setting in which the work would have to be done. Others are less dramatic, and could probably be worked out over some time given a certain amount of determination. The list of research proposals are by no means complete. There is still much to be worked out in this field.

In Part 2, we discuss Gaussian approximation in ergodic theory. This part consists of discussions about the circumstances under which we can find Gaussian behaviour in a measure preserving system.

In Chapter 7, we begin with a survey of some previous results, related to the Central Limit Theorem and the Almost Sure Central Limit Theorem. The survey contains a result which to some extent builds a bridge between the two parts of this thesis. This survey culminates in the proof of a tool allowing us to produce quite general results in Gaussian approximation. This tool involves the construction of Rokhlin towers. It is applied in Chapter 8 to obtain some such results.

Chapter 9 contains some extensions of the previously known results in this field. In Chapter 9, we discuss weighted partial sums as opposed to the standard partial sums discussed previously. It turns out that when we consider weighted sums instead of the standard partial sums, we may still obtain Gaussian approximation theorems. We give two different approaches to this. In the first approach, we follow the methods of Chapter 8. This turns out to produce quite weak results, so we take another approach involving Abel summation.

In the final chapter of the thesis, we will give some directions for further research in the field of Gaussian approximation in ergodic theory. As was the case in Chapter 6, the list is clearly not exhaustive, but some ideas and pointers are mentioned for a few possible directions of new research.

At this point, some acknowledgements should be made. I would like to express my thanks to a number of people and organisations. First of all, I thank Institut de Recherche Mathématique Avancée, Université Louis Pasteur (Strasbourg I) and Department of Mathematical Sciences, University of Liverpool for providing me with the physical framework for the writing of this thesis. Also, I thank my supervisors M. Weber (Strasbourg) and R. Nair (Liverpool) for the kind suggestions and encouragement. Without either of these gentlemen, this thesis would never have come into existence. The short French abstract of the thesis would have been considerably less readable, had it not

been for the linguistic assistance of S. Racaniere, and for his effort I am deeply grateful. I also thank I. Assani (University of North Carolina – Chapel Hill) for inviting me to spend a month at UNC and present some of the results of this thesis.

I would also like to thank the organisations providing me with my livelihood during the writing of the thesis. First of all, thanks to the European Doctoral College of Strasbourg for providing support for travel and lodging and in particular for providing me with a grant for the first year of my studies. Also, thanks to Rejselegat for Matematikere (Travelling Scholarship for Mathematicians) for providing me with all the financial support I could ever wish for during the final year of my studies.

Finally, I would like to thank some people not directly involved in the mathematical process leading to the present thesis. First of all, I should express my thanks to each and every one of my mathematics teachers since my birth. A number of names come to mind, but if none are mentioned, none are forgotten. Secondly, I express my deepest thanks to my best friend and companion Lykke. Lykke means “good fortune” in Danish, and she has certainly been my good fortune in my time as a Ph.D. student.

Contents

Introduction, résumé et conclusion

Cette thèse concerne deux formes d'approximation. Dans le première partie de la thèse, nous étudions l'approximation diophantienne dans un corps de séries formelles. Dans la deuxième partie de la thèse, nous étudions l'approximation gaussienne dans la théorie ergodique. Dans cette introduction, nous décrivons les méthodes et les résultats de chaque partie de la thèse.

Approximation diophantienne dans un corps des séries formelles

Le sujet de la première partie de la thèse est l'approximation diophantienne dans un corps de séries formelles. Soit $\mathbb{F}$ un corps fini à k éléments, et posons

$$\mathcal{L} = \left\{ \sum_{i=-n}^{\infty} \alpha_{-i} X^{-i} : n \in \mathbb{Z}, \alpha_{-i} \in \mathbb{F}, \alpha_n \neq 0 \right\}.$$

Cet ensemble est un corps. On définit une valeur absolue sur $\mathcal{L}$ en posant

$$\left\| \sum_{i=-n}^{\infty} \alpha_{-i} X^{-i} \right\| = k^n.$$

Avec cette valeur absolue, $(\mathcal{L}, \|\cdot\|)$ est un espace de Banach, c'est à dire qu'il est complet pour la métrique $d(x,y) = \|x-y\|$. La valeur absolue satisfait la propriété ultramétrique :

$$\|x+y\| \leq \max(\|x\|, \|y\|) \quad \text{pour chaque } x, y \in \mathcal{L}.$$

Dans le chapitre 1, après une introduction à l'approximation diophantienne, nous allons faire la construction de $\mathcal{L}$ en détails et démontrer ces propriétés.

A l'aide de la valeur absolue, nous sommes capable de caractériser la mesure de Haar sur $\mathcal{L}$. L'égalité suivante en est une caractérisation complète :

$$\mu(B(c,k^r)) = \mu\{x \in \mathcal{L} : \|x-c\| \leq k^r\} = k^{r+1}.$$

En plus, nous sommes capables de définir la dimension de Hausdorff. Soit $\delta > 0$ et $s \geq 0$. Pour chaque recouvrement $\mathcal{C}_\delta$ d'un ensemble $E \subseteq \mathcal{L}$ par des boules de rayons $< \delta$, posons

$$l^s(\mathcal{C}_\delta) = \sum_{B \in \mathcal{C}_\delta} \rho(B)^s$$

la s-longueur de $\mathcal{C}$. Comme dans le cas réel, on définit la s-mesure de Hausdorff :

$$\mathcal{H}^s(E) = \lim_{\delta \to 0} \inf_{\text{recouvrements } \mathcal{C}_\delta} l^s(\mathcal{C}_\delta)$$

et la dimension de Hausdorff :

$$\dim_{\mathrm{H}}(E) = \inf\{s \geq 0 : \mathcal{H}^s(E) = 0\}$$

Ces définitions sont les instruments principaux dans la première partie de la thèse. Comme dans le cas de la construction de l'espace $\mathcal{L}$, ces définitions seront vues en détail dans le chapitre 1. Aussi, nous allons décrire des résultats déjà connus dans ce domaine à la fin du chapitre 1.

Dans le chapitre 2, nous étudierons des expansions de Lüroth. Nous démontrerons que pour chaque $x \in \mathcal{L}$, on peut écrire

$$x = q_0 + \frac{1}{q_1} + \sum_{i=1}^{\infty} \frac{1}{q_1(q_1-1)\cdots q_{i-1}(q_{i-1}-1)q_i}, \quad q_i \in \mathbb{F}[X].$$

$\mathbb{F}[X]$ est l'anneau des polynômes à coefficients dans $\mathbb{F}$. On appelle ces expansions des expansions de Lüroth. Inversement, chaque expansion de cette forme correspond à un élément de $\mathcal{L}$.

Posons $I = \{x \in \mathcal{L} : \|x\| < 1\}$. Maintenant, $(I, \mathcal{B}, \mu)$ est un espace de probabilité, où $\mathcal{B}$ est la famille des ensembles de Borel et μ est la mesure de Haar. Donc les coefficients dans l'expansion de Lüroth, $q_i(x)$, $x \in I$, sont des variables aléatoires. Dans le chapitre 2, nous allons montrer à l'aide d'une caractérisation dynamique :

Théorème. *Les coefficients de l'expansion de Lüroth q_i sont des variables aléatoires indépendantes et identiquement distribuées.*

A l'aide de ce théorème, nous sommes capable de démontrer quelques résultats dans la théorie métrique de ces expansions. A l'aide de la loi forte des grands nombres, on peut déjà montrer des résultats. En utilisant des théorèmes probabilistes plus fort (le théorème central limite, la loi du logarithme itéré etc...), on peut montrer des théorèmes métriques plus forts. Les démonstrations et les théorèmes métriques se trouvent dans le chapitre 2.

Dans le chapitre 3, nous étudierons des espaces vectoriels et des matrices à coefficients dans $\mathcal{L}$. Définissons encore quelques valeurs absolues. Pour $v = (v_1, \ldots, v_n)$ dans $\mathcal{L}^n$, définissons la hauteur de v,

$$\|v\|_\infty = \max(\|v_1\|, \ldots, \|v_n\|).$$

De plus, définissons la distance au treillis des entiers,

$$|\langle x\rangle| = \min_{p\in\mathbb{F}[X]^n} \|x-p\|_\infty$$

Soient $m,n\in\mathbb{N}$, et soit $\psi:\mathbb{F}[X]^m\to\mathbb{R}^+$ une fonction telle que $\psi(p)=\psi(\|p\|)$. Nous étudions l'ensemble de matrices $A\in\mathcal{L}^{m+n}$,

$$\mathcal{S}(\psi)=\{A\in I^{mn} : |\langle qA\rangle| < \psi(q) \text{ pour une infinité de } q\in\mathbb{F}[X]^m\}.$$

Avec ces définitions, nous sommes capables de démontrer un analogue du théorème de Khintchine et Grošev :

Théorème (Le théorème de Khintchine–Grošev). *Soit $\psi:\mathbb{F}[X]\to\mathbb{R}^+$ une fonction, décroissante en norme, telle que $\psi(p)=\psi(\|p\|)$.*

1.

$$\text{Si} \sum_{q\in\mathbb{F}[X]^m}\psi(q)^n<\infty, \text{ alors } \mu(\mathcal{S}(\psi))=0.$$

2.

$$\text{Si} \sum_{q\in\mathbb{F}[X]^m}\psi(q)^n=\infty, \text{ alors } \mu(\mathcal{S}(\psi))=1.$$

La démonstration est assez compliquée. Il faut utiliser la théorie ergodique, quelques outils de théorie des nombres, des probabilités etc. Les détails se trouvent dans le chapitre 3.

Dans le chapitre 4, nous étudions un ensemble exceptionnel dans le théorème précédent. Soit $v>0$ et soit $\psi(q)=\|q\|_\infty^{-v}$. On définit l'ensemble $\mathcal{S}_v=\mathcal{S}(\psi)$ des formes linéaires bien approximables. On peut facilement démontrer que $\mu(\mathcal{S}_v)=1$ si $v\leq\frac{m}{n}$ et $\mu(\mathcal{S}_v)=0$ si $v>\frac{m}{n}$. Dans le dernier cas, nous sommes capables de calculer la dimension d'Hausdorff de l'ensemble $\mathcal{S}_v$. Dans le chapitre 4, nous allons démontrer l'analogue suivant du théorème de Jarník et Besicovitch :

Théorème (Le théorème de Jarník–Besicovitch). *Soit $v\geq\frac{m}{n}$. Alors,*

$$\dim_{\mathrm{H}}(\mathcal{S}_v)=(m-1)n+\frac{m+n}{v+1}.$$

Dans la démonstration de ce théorème, il y a deux parties. Dans la première partie, on montre que la dimension est inférieure ou égale à $(m-1)n+\frac{m+n}{v+1}$. Cette estimation est une conséquence d'un lemme du type Borel–Cantelli. La deuxième partie de la démonstration est plus compliquée. En utilisant des systèmes doués d'ubiquité, on peut montrer que la dimension d'Hausdorff d'un sous-ensemble de $\mathcal{S}_v$ est supérieure ou égale à $(m-1)n+\frac{m+n}{v+1}$. Cette partie de la démonstration est assez technique.

Dans le chapitre 5, nous étudions un autre ensemble exceptionnel. Cette fois, nous considérons l'ensemble,

$$\mathfrak{B}(m,n)=\left\{A\in\mathcal{L}^{mn} : \exists K>0\,\forall q\in\mathbb{F}[X]^m\ |\langle qA\rangle|^n>\frac{K}{\|q\|_\infty^m}\right\},$$

des formes linéaires mal approximables. On montre facilement que $\mu(\mathfrak{B}(m,n)) = 0$, parce que l'ensemble complémentaire à pour mesure ∞. Dans le chapitre 5, nous allons démontrer un analogue du théorème de Jarník sur les nombres mal approximables.

Théorème (Le théorème de Jarník). *Soit $m,n \in \mathbb{N}$. Alors,*

$$\dim_{\mathrm{H}} \mathfrak{B}(m,n) = mn.$$

Pour la démonstration de ce théorème, on utilise une méthode de Schmidt, en utilisant des «jeux (α,β)». Soient $\alpha,\beta \in (0,1)$. Dans le jeux (α,β), on a deux joueurs, A et B. Le joueur A commence à prendre une boule fermée $A_1 = B(C(A_1),\rho(A_1)) \subseteq \mathcal{L}^{mn}$. Ensuite, B prend une boule $B_1 = B(C(B_1),\rho(B_1)) \subseteq A_1$ telle que $\rho(B_1) = \alpha\rho(A_1)$. Maintenant, A prend une boule $A_2 \subseteq B_1$, telle que $\rho(A_2) = \beta\rho(B_1)$. Le jeux continue à l'infini. B gagne le jeux, si $\cap_{i=1}^{\infty} B_i = \cap_{i=1}^{\infty} A_i \subseteq \mathfrak{B}(m,n)$. Sinon, A a gagné. Nous allons démontrer que pour tout $\beta \in (0,1)$ et tout $\alpha \in (0,\frac{1}{k+1})$, le jouer A peut toujours gagner le jeux (α,β). Ensuite, nous allons montrer que cette propriété implique le théorème.

Dans la dernier chapitre de la première partie, nous ébauchons quelques problèmes d'approximations diophantiennes dans un corps de séries formelles, qui sont dans la continuité de cette thèse. Nous avons choisi trois problèmes, mais il reste sûrement beaucoup de questions ouvertes dans ce domaine de recherche.

Approximation gaussienne en théorie ergodique

Le sujet de la deuxième partie de la thèse est l'approximation gaussienne en théorie ergodique. Nous étudierons quelques aspects gaussiens des systèmes ergodiques.

Dans le chapitre 7, nous allons décrire quelques résultats déjà connus dans ce domaine. On commence avec les similarités entre le théorème ergodique et la loi forte des grandes nombres pour des variables aléatoires indépendantes. Ensuite, nous allons décrire des résultats pour des variables aléatoires non indépendantes, mais satisfaisant quelques propriétés plus faibles.

A l'aide de ces résultats, quelques auteurs ont démontré des théorèmes d'existence de fonctions dans certains systèmes dynamiques, pour lesquels le comportement des sommes partielles imite celui de variables aléatoires indépendantes. En particulier, le théorème central limite est satisfait.

Nous allons décrire des éléments de démonstrations de plusieurs résultats de ce type. Pour la majorité de ces résultats, les outils fondamentaux sont les séries de Fourier et le théorie spectrale. Ces techniques de démonstration s'appliquent donc seulement pour des rotations du tore.

Dans les sections 7.4 et 7.5, nous allons décrire une autre méthode en utilisant des tours de Rokhlin. Cette méthode s'applique pour tout système ergodique et apériodique, mais l'ergodicité est impérative. La méthode des tours de Rokhlin est notre méthode principale dans les chapitres suivants. Dans le chapitre 8 nous allons démontrer quelques résultats récents en approximation gaussienne due a Volný.

Dans le chapitre 9, nous allons généraliser des résultats dans le chapitre 8 pour des sommes partielles avec des poids. Soit (a_n) un suite de réels positifs, bornée par un nombre $M > 0$ et T une automorphisme d'un espace de probabilité non-atomique. Dans chapitre le 9, nous allons démontrer quelques résultats pour des sommes partielles avec des poids définis par

$$A_n f = \sum_{l=0}^{n-1} a_l T^l f.$$

Dans le section 9.1, nous essayons de utiliser la méthode de Volný, sans des outils additionnels. Pour ces méthodes, on a besoin de quelques hypothèses sur la suite des poids. La première hypothèse est

$$\Delta = \sum_{n=1}^{\infty} |a_n - a_{n-1}| < \infty. \tag{H1}$$

La deuxième hypothèse est un peu plus technique :

$$\left|\mathcal{A}_n - K^2\right| = \left| 2 \sum_{\log_3 n < k < \log_2 n} \frac{\left(\sum_{i=n-2^k+1}^{n} a_i\right)^2}{4^k k} - K^2 \right| = O\left(\frac{1}{\log n}\right), \tag{H2}$$

où $K > 0$ est un nombre réel. La troisième hypothèse est la plus compliquée :

$$\frac{1}{n} \sum_{\log_3 n < k < \log_2 n} \sum_{l=\left[n^{\log_3 2}\right]}^{n-1} \frac{\left(\sum_{i=l-2^k+1}^{n-2^k+2} \sum_{h=0}^{2^k-2} (a_{i+h} - a_{i+h-1})\right)^2}{4^k k} = O\left(\frac{\log\log\log\log n}{\log n}\right). \tag{H3}$$

Avec ces trois hypothèses, on peut démontrer trois généralisations de les théorèmes de Volný, mais les hypothèses ne sont pas tout a fait naturelles dans la domaine de probabilités.

Dans le section 9.2, nous allons démontrer une théorème dans l'approximation gaussienne sous une hypothèse plus naturelle. Soit (a_l) une suite décroissante des réels telle que

$$\frac{n \log\log\log\log n}{\log n} = o\left(\sum_{l=0}^{n-1} a_l^2\right). \tag{H4}$$

Sous cette hypothèse, nous allons démontrer :

Théorème. *Soit $(X, \mathcal{B}, \mu)$ un espace de probabilité non-atomique, et soit $T : X \to X$ un automorphisme apériodique. Il existe $f \in L^2(X)$ et des variables aléatoires $Z_l' \sim \mathcal{N}(0, 2(\log\log 3 - \log\log 2))$ tel que*

$$\lim_{n\to\infty} \left\| \frac{A_n f}{\left(\sum_{l=0}^{n-1} a_l^2\right)^{1/2}} - Z_n' \right\| = 0.$$

Dans le dernier chapitre de cette thèse, nous ébaucherons quelques problèmes de recherché dans le domaine des approximations gaussiennes en théorie ergodique. Comme au chapitre 6, la liste des problèmes n'est pas complète. Les possibilités de recherches futures sont nombreuses.

Part I.

Diophantine approximation in a field of formal series

Contents

1. Introduction

In this part of the thesis, we will be concerned with so-called metrical (or measure theoretical) Diophantine approximation. We will be working in a specific setting – namely in the field of Laurent series with coefficients from a finite field.

This first chapter serves as an introduction to this theory. To get acquainted with the topic, we begin with a short introduction to Diophantine approximation over the real numbers. This will include some of the classical theorems like Dirichlet's Theorem and Khintchine's Theorem. This is meant to serve as an informal introduction to the more abstract parts of the subsequent discussions. For a complete treatment on Diophantine approximation over the reals, the reader is referred to [7].

Secondly, we construct the setting, we will be working in. We will discuss the similarities and differences of this setting and the real numbers. Also, we will give a construction of the Haar measure on our field, as well as a definition of Hausdorff dimension.

Thirdly, we give a short discussion of the differences and similarities between our field and another locally compact non-Archimedean field — the p-adics — in which much more work has been done.

In the final part of the introduction, we will be giving a survey of the previous results in Diophantine approximation in this setting. It is worth noting, that almost all of these results are outside of the domain of the metrical theory. The results in the subsequent sections are hence extending the body of knowledge about fields of Laurent series.

1.1. An introduction to Diophantine approximation

Diophantine approximation is a way of studying the density of the rational numbers in the real numbers. It is a well-known fact that the rationals are dense in the reals. That is, you can find a rational number $\frac{p}{q}$ in any neighbourhood of a real number x. A number of questions can be asked to give more precise statements of this density. We state some of these informally:

- Given a real number x, how can we easily construct rationals $\frac{p}{q}$ close to x?

- How large a neighbourhood of x do we need to take to find a rational number $\frac{p}{q}$ with q small?
- How many real numbers x have the property of being close to infinitely many rationals $\frac{p}{q}$ where closeness depends on the magnitude of q?

Quite a few answers have been given to these and similar questions. It is the purpose of this section to give a brief overview of some of the most important of these.

Starting with the first of the questions, a classical way of constructing rationals close to a given real number is the continued fraction expansion (for a reference on the theory of continued fractions, see [29]). Basically, this construction is as follows: Let $x \in \mathbb{R}$. We can write

$$x = a_0 + \frac{1}{r_1},$$

where a_0 denotes the greatest integer less than x, and $\frac{1}{r_1}$ is the fractional part of x. If x is an integer, we let $r_1 = \infty$. Clearly, $r_1 > 1$, so we can do the same thing again,

$$r_1 = a_1 + \frac{1}{r_2}.$$

Continuing in this way, we obtain sequence which can be written as a continued fraction

$$x = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \cdots}}. \tag{1.1}$$

In fact, this sequence converges and is unique (Theorem 14 in [29]).

We can express the n'th fraction in the expansion (1.1) – the so-called n'th convergent of x – as a standard fraction $\frac{p_n}{q_n}$, where p_n and q_n can be calculated by recursive formulae, which we do not deduce here. Using these recursive formulae yields

$$\left| x - \frac{p_n}{q_n} \right| < \frac{1}{q_n^2} \tag{1.2}$$

We have sketched a proof of

Theorem 1.1 (Dirichlet's Theorem). *For any $x \in \mathbb{R}$ and any $N \in \mathbb{N}$, there exist $p, q \in \mathbb{Z}$ with $|q| \leq N$ such that*

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

This provides an answer to the second question, we posed.

Remark. The above version of Dirichlet's Theorem is not the original version. The original version version is slightly stronger, but the above serves just as well as an introduction to the principle.

Another construction of rational numbers approximating real numbers is the so-called Lüroth expansions, discussed in Chapter 2.

Theorem 1.1 provides an answer to the second question, but what about the third question? An answer to this question has also been given by Khintchine (Theorem 32 in [29]).

Theorem 1.2 (Khintchine's Theorem). *Let $\psi : \mathbb{N} \to \mathbb{R}_+$ be a decreasing function such that $n\psi(n)$ is non-increasing. Define the set*

$$\mathcal{K}_\psi = \left\{ x \in \mathbb{R} : \left| x - \frac{p}{q} \right| < \frac{\psi(q)}{q} \text{ for infinitely many } p, q \in \mathbb{Z} \right\}.$$

Then

$$\mu\left(\mathcal{K}_\psi\right) = \begin{cases} 0 & \text{for } \sum_{i=1}^{\infty} \psi(i) < \infty, \\ \infty & \text{for } \sum_{i=1}^{\infty} \psi(i) = \infty, \end{cases}$$

where μ denotes the Lebesgue measure on $\mathbb{R}$.

Khintchine's original proof of this theorem involves the continued fractions expansion and is rather involved. We will discuss newer approaches to the problem in Chapter 3, where we will prove a theorem corresponding to the above in a different setting.

Theorem 1.2 gives us some quantitative information about the density of the rationals in the reals. We see that whenever the approximation function ψ decreases slowly, most reals have many rational neighbours with sufficiently small denominators. However, the theorem does not distinguish between approximation functions $\psi(q) = q^{-2}$ and $\psi'(q) = q^{-3}$. The measures of both $\mathcal{K}_\psi$ and $\mathcal{K}_{\psi'}$ respectively are zero. But surely, $\mathcal{K}_{\psi'}$ is a proper subset of $\mathcal{K}_\psi$? To analyse the difference in the sizes of these sets, we may use Hausdorff dimension. Jarník and Besicovitch independently proved the following theorem ([25] and [8]):

Theorem 1.3 (The Jarník–Besicovitch Theorem). *Let $v \geq 1$. Define the set*

$$\mathcal{K}_v = \left\{ x \in \mathbb{R} : \left| x - \frac{p}{q} \right| < \frac{1}{q^{1+v}} \text{ for infinitely many } p, q \in \mathbb{Z} \right\}.$$

Then

$$\dim_{\mathrm{H}}\left(\mathcal{K}_v\right) = \frac{2}{v+1}.$$

We will return to the discussion of these in Chapter 4, where we will prove corresponding multidimensional results in our setting.

Yet another set of measure zero arises from Theorem 1.2. Namely, we see that the set

$$\mathfrak{B} = \left\{ x \in \mathbb{R} : \text{ there is a } K > 0 \text{ s.t. for any } q \in \mathbb{Z} \text{ there is a } p \in \mathbb{Z} \text{ s.t. } \left| x - \frac{p}{q} \right| > \frac{K}{q^2} \right\}$$

has measure zero, since the complementary set has full measure. Jarník calculated the Hausdorff dimension of this set ([24]):

Theorem 1.4 (Jarník's Theorem).

$$\dim_{\mathrm{H}}(\mathfrak{B}) = 1.$$

Again, we postpone a discussion of the proof of this theorem to a later chapter. We will return to the problem in our new setting in Chapter 5.

The above theorems constitute a basis for the metrical theory of Diophantine approximation. There are plenty of generalisations of the theorems, and it is some of these we will discuss in this part of the thesis. We will, however, be working in another field than the real numbers. In the next section, we construct this field.

1.2. Laurent series, measure and dimension

In this section, we construct the field, we will be working in. Subsequently, we discuss some properties of the different norms and norm-like functions, we will be using. Finally, we will construct an explicit formula for the Haar measure on the field and define Hausdorff dimension.

1.2.1. Construction of $\mathcal{L}$

We will now construct the field that will replace $\mathbb{R}$ in our treatment. First, we will recall the following construction of the reals. Begin with the integers $\mathbb{Z}$. Since $\mathbb{Z}$ is an integral domain, we can construct the field of fractions $\mathbb{Q} = \mathbb{Z} \times (\mathbb{Z} \setminus \{0\})/\sim$, where $\sim$ denotes the equivalence relation $(a,b) \sim (c,d)$ if and only if $ad = cb$.

On $\mathbb{Q}$, we now introduce the standard norm (or absolute value). Now, it makes sense to define the set of Cauchy sequences $\mathcal{C} \subseteq \mathbb{Q}$. It is straightforward to prove, that this is a ring, and that the set of null-sequences $\mathcal{N}$ (that is, sequences converging to zero) is an ideal in $\mathcal{C}$. We can now complete $\mathbb{Q}$ in the norm by letting $\mathbb{R} = \mathcal{C}/\mathcal{N}$. This set is then shown to be a field. This is a standard construction of the real numbers, which is carried out in almost any textbook on undergraduate abstract algebra or for instance, in a more abstract setting, in [39].

Note that the only property needed in order to construct a complete field in this way, is the fact that the initial ring is an integral domain. One such starting point would be the polynomial ring over a finite field. Hence, we let $\mathbb{F}$ denote the finite field of $k = p^l$ elements, where p is a prime and l is a positive integer. Furthermore, we let $\mathbb{F}[X]$ denote the polynomial ring over $\mathbb{F}$. Since this is an integral domain, we can construct the field of fractions denoted $\mathbb{F}(X)$. Elements herein can be expressed as quotients $x = \frac{p}{q}$, where $p, q \in \mathbb{F}[X]$. By formally carrying out the division, we obtain a formal power series

$$x = \sum_{i=-n}^{\infty} \alpha_{-i} X^{-i}, \quad \alpha_{-i} \in \mathbb{F}, \quad \alpha_n \neq 0, \tag{1.3}$$

which is another representation of the elements in $\mathbb{F}(X)$. Note, that not all choices of sequences (α_i) give elements in $\mathbb{F}(X)$. In particular, only eventually periodic (including finite) sequences occur in this way.

In order to complete the second step in the construction, we need to define a norm. For any $p \in \mathbb{F}[X]$, we define the norm of p to be $\|p\| = k^{\deg(p)}$, where $\deg(p)$ denotes the degree of the polynomial p. This induces a norm on $\mathbb{F}(X)$ by $\left\|\frac{p}{q}\right\| = \frac{\|p\|}{\|q\|}$, or alternatively

$$\left\| \sum_{i=-n}^{\infty} \alpha_{-i} X^{-i} \right\| = k^n \tag{1.4}$$

in the above representation (1.3).

Note that this norm is non-Archimedean. That is, it satisfies the following properties:

$$\|x\| \geq 0 \text{ for any } x \in \mathbb{F}(X) \text{ and } \|x\| = 0 \text{ if and only if } x = 0, \tag{1.5a}$$

$$\|xy\| = \|x\|\,\|y\| \text{ for all } x, y \in \mathbb{F}(X), \tag{1.5b}$$

$$\|x+y\| \leq \max\{\|x\|, \|y\|\} \text{ for all } x, y \in \mathbb{F}(X). \tag{1.5c}$$

That is, we have a stronger triangle inequality (1.5c) than we do for the usual norms.

We define a metric d on $\mathcal{L}$ by $d(x,y) = \|x-y\|$. The metric space obtained in this fashion is not complete. Indeed, let $(\alpha_n) \in \{0,1\}^{\mathbb{N}}$ be some sequence which is not eventually periodic. Since distinct elements corresponding to 0 and 1 exist in every finite field, this defines a sequence in each such field. Define the sequence $(x_n) \in \mathcal{L}$ by

$$x_n = \sum_{i=1}^{n} \alpha_i X^{-i}$$

Obviously, this is a Cauchy sequence in $\mathbb{F}(X)$ in the norm defined above. However, the sequence of coefficients is non-periodic, so the limit point does not exist in $\mathbb{F}(X)$.

However, the field can be completed, just as $\mathbb{Q}$ could be completed to obtain $\mathbb{R}$. Hence, we let $\mathcal{L}$ be the ring of Cauchy sequences modulo the ideal of null sequences. Clearly, this field contains $\mathbb{F}(X)$ as a dense subset, and the norm maintains all properties (1.5a), (1.5b) and (1.5c). We have obtained three sets, all corresponding to subsets of the real numbers by construction:

- The polynomial ring $\mathbb{F}[X]$ corresponds to the integers $\mathbb{Z}$.
- The field of fractions $\mathbb{F}(X)$ corresponds to the rational numbers $\mathbb{Q}$.
- The completion of the field of fractions with respect to the norm (1.4) $\mathcal{L}$ corresponds to the real numbers $\mathbb{R}$.

We will now find a suitable representation of the elements in $\mathcal{L}$. It is not surprising that a representation of the space consists of *all* formal power series of the form (1.3).

A proof of this can be found in Chapter XII, § 6 of [39]. We have now obtained the locally compact field,

$$\mathcal{L} = \left\{ \sum_{i=-n}^{\infty} \alpha_{-i} X^{-i} : n \in \mathbb{Z}, \alpha_{-i} \in \mathbb{F}, \alpha_n \neq 0 \right\},$$

with a non-Archimedean norm $\|\cdot\|$. The field is locally compact, since any closed ball with radius k^n for some $n \in \mathbb{Z}$ is clearly compact, so any point has a compact neighbourhood. Since the elements of $\mathcal{L}$ are of form of Laurent series, we call the field $\mathcal{L}$ *the field of Laurent series with coefficients from* $\mathbb{F}$.

Inside this field, we will need a few additional objects. First, we define the sub-ring

$$\mathcal{J} = \{A \in \mathcal{L} : \|A\| \leq 1\}.$$

Traditionally, this is called the *ring of integers in* $\mathcal{L}$. We will further define the *integral part of an element* $A \in \mathcal{L}$ to be

$$\left[\sum_{i=-n}^{\infty} \alpha_{-i} X^{-i} \right] = \begin{cases} \sum_{i=-n}^{0} \alpha_{-i} X^{-i} & \text{for } n \geq 0, \\ 0 & \text{otherwise,} \end{cases}$$

where $\alpha_n \neq 0$ is the leading coefficient of the Laurent polynomial. Now, we take the maximal ideal I in the ring of integers. Clearly, this has the form

$$I = \{A \in \mathcal{L} : \|A\| < 1\} = \{A \in \mathcal{L} : [A] = 0\}. \tag{1.6}$$

This set plays the role of the unit interval in this setting.

1.2.2. Norms and vector spaces

We have already defined the norm in $\mathcal{L}$. An interesting consequence of the ultra-metric property (1.5c), which will be used repeatedly throughout the thesis, is the following:

Proposition 1.5. *Let* $(X, \|\cdot\|)$ *be an ultra-metric space. That is, a normed space with the additional property* (1.5c). *Let* B_1, B_2 *be both open or both closed balls in* X. *Then* $B_1 \subseteq B_2$, $B_2 \subseteq B_1$ *or* $B_1 \cap B_2 = \emptyset$.

Proof. Let c_i and ρ_i, denote the centres and radii of B_i, $i = 1, 2$. Assume without loss of generality that $\rho_1 \geq \rho_2$.

Assume the balls are open and that there is an $x' \in B_2 \cap B_1$. Let $x \in B_2$. We will prove that $x \in B_1$. By (1.5c),

$$\begin{aligned} \|x - c_1\| &= \left\|(x - c_2) + (c_2 - x') + (x' - c_1)\right\| \\ &\leq \max\{\|x - c_2\|, \left\|c_2 - x'\right\|, \left\|x' - c_1\right\|\} < \max\{\rho_2, \rho_1, \rho_1\} = \rho_1. \end{aligned}$$

Hence, $x \in B_1$. The proof is completely analogous for closed balls. □

From Section 3.2 and onwards, we will be working in vector spaces of several dimensions over $\mathcal{L}$. Hence, we will need a multi-dimensional norm.

Definition 1.1. Let $h \in \mathbb{N}$ and let $\mathcal{L}^h$ denote the h-dimensional $\mathcal{L}$-vector space over $\mathcal{L}$. For any $x \in \mathcal{L}^h$, we define the *height of* x to be $\|x\|_\infty = \max_{1\leq i\leq h} \|x_i\|$, where x_i denotes the i'th coordinate of x.

Clearly, the height has properties (1.5a) and (1.5c). In particular, Proposition 1.5 holds. Also, we will need the following proposition:

Proposition 1.6. *Let $h \in \mathbb{N}$ and let $a, b \in \mathcal{L}^h$. Let $a \cdot b$ denote the usual inner product. Then*

$$\|a \cdot b\| \leq \|a\|_\infty \|b\|_\infty .$$

Proof. Let $a = (a_1, \dots, a_h)$ and $b = (b_1, \dots, b_h)$. We have

$$\begin{aligned} \|a \cdot b\| = \|a_1 b_1 + \cdots + a_h b_h\| &\leq \max_{1\leq i\leq h} \{\|a_i\| \|b_i\|\} \\ &\leq \max_{1\leq i\leq h} \|a_i\| \max_{1\leq j\leq h} \|b_j\| = \|a\|_\infty \|b\|_\infty . \end{aligned}$$

This completes the proof. □

Since we are concerned with Diophantine approximation, it is natural to define the distance to the nearest "integer". Recall, that the object replacing the integers in the construction in Section 1.2.1 was the polynomial ring $\mathbb{F}[X]$. This motivates the following definition:

Definition 1.2. Let $n \in \mathbb{N}$ and let $\mathcal{L}^n$ denote the n-dimensional $\mathcal{L}$-vector space over $\mathcal{L}$. For any $x \in \mathcal{L}^n$, we define the *distance from x to the integer lattice* to be

$$|\langle x \rangle| = \min_{p\in\mathbb{F}[X]^n} \|x - p\|_\infty$$

This is the final norm-like function, we will need.

From Section 3.2 and onwards, we will be concerned with linear forms over $\mathcal{L}$. That is, $m \times n$ matrices with entries from $\mathcal{L}$. We will define inner products, matrix products and determinants in our setting exactly as in the real case. Obviously, where we would previously expect a real number or vector to result from one of these operations, we will now obtain an element in $\mathcal{L}$ or a vector of such elements.

Since it will cause no ambiguity, we will throughout the thesis identify the set of $m \times n$ matrices over $\mathcal{L}$ with $\mathcal{L}^{mn}$. This immediately extends the definition of the height of a vector to the height of a matrix.

For lack of a better place for this, we will just introduce a few notational conventions. Given two real quantities x and y, we will write $x \ll y$ if there exists a constant $K > 0$ such that $x \leq Ky$. If $x \ll y$ and $y \ll x$, we will write $x \asymp y$. This notation will be used throughout the thesis.

1.2.3. Measure and dimension

We will now follow a construction due to Sprindžuk ([59]) of an explicit form of the Haar measure on $\mathcal{L}$. Let $n, m \in \mathbb{Z}$, $m \leq n$ and let $\alpha_m^*, \dots, \alpha_n^* \in \mathbb{F}$. We define cylinders

$$B(\alpha_m^*, \dots, \alpha_n^*) = \left\{ \sum_{i=-n}^{\infty} \alpha_{-i} X^{-i} \in \mathcal{L} : \alpha_i = \alpha_i^* \text{ for } i = m, \dots, n \right\} \tag{1.7}$$

That is, sets where the first $n-m$ coefficients of the series are fixed.

It is obvious that

$$B(\alpha_m^*, \dots, \alpha_n^*) = B(\omega, k^m) = \{x \in \mathcal{L} : \|x - \omega\| < k^m\},$$

where $\omega = \alpha_n^* X^n + \cdots + \alpha_n^* X^m$. We define a function μ on the set of these balls by

$$\mu(B(\alpha_m^*, \dots, \alpha_n^*)) = k^m.$$

We now extend this function to all balls by noting, that given a ball $B(c, k^m)$, there exists an element $t \in \mathcal{L}$, such that $c + t$ is of the general form ω. Hence, we define $\mu(B(c, k^m)) = k^m$ for any ball $B(c, k^m)$.

We will now extend this function to the finite unions of balls. Let $I \subseteq \mathbb{N}$ be a finite set, and let B_i be balls for any $i \in I$. By Proposition 1.5, there exists a subset $I' \subseteq I$ such that

$$B = \bigcup_{i \in I} B_i = \bigcup_{i \in I'} B_i$$

where the B_i are disjoint for $i \in I'$. Hence, we define

$$\mu(B) = \mu\left(\bigcup_{i \in I} B_i\right) = \sum_{i \in I'} \mu(B_i).$$

Clearly, the function defined in this way is a σ-additive measure on the σ-algebra generated by the balls in $\mathcal{L}$. That is, the Borel σ-algebra. Since the Haar measure is unique up to scaling, we have proved the following proposition:

Proposition 1.7. *The Haar measure on $\mathcal{L}$ is completely characterised by*

$$\mu(B(c, k^m)) = \mu(\{x \in \mathcal{L} : \|x - c\| < k^m\}) = k^m$$

for any $c \in \mathcal{L}$ and any $m \in \mathbb{Z}$.

Remark. Note, that the above scaling of the Haar measure has the property of assigning measure 1 to the set I. Hence, this a the natural scale of the measure, given the analogy with the real numbers. Note also that we immediately obtain the measure of the corresponding closed ball. Indeed, since the norm only assumes values k^r, where $r \in \mathbb{Z}$,

$$\mu(\{x \in \mathcal{L} : \|x - c\| \leq k^m\}) = \mu\left(\left\{x \in \mathcal{L} : \|x - c\| < k^{m+1}\right\}\right) = k^{m+1}$$

We will now define Hausdorff dimension in $\mathcal{L}$. Let $E \subseteq \mathcal{L}^{mn}$ be some set. For any countable cover $\mathcal{C}$ of E with balls $B_i = B(c_i, \rho_i)$, we define the s-length of $\mathcal{C}$ to be the sum

$$l^s(\mathcal{C}) = \sum_{B \in \mathcal{C}} \rho_i^s$$

for any $s \geq 0$. Let $\delta > 0$ and restrict to covers $\mathcal{C}_\delta$ such that $\rho_i < \delta$ for all $B_i \in \mathcal{C}_\delta$. We can define an outer measure,

$$\mathcal{H}^s(E) = \lim_{\delta \to 0} \inf_{\text{covers } \mathcal{C}_\delta} l^s(\mathcal{C}_\delta). \tag{1.8}$$

We will prove, that this is indeed an outer measure.

Proposition 1.8. *The function $\mathcal{H}^s : 2^{\mathcal{L}} \to \mathbb{R} \cup \{\infty\}$ defined in* (1.8) *is an outer measure for any $s > 0$. That is,*

- $\mathcal{H}^s(\emptyset) = 0$.
- *$\mathcal{H}^s$ is sub-additive.*

Proof. The first property is clearly satisfied. Hence, let $E_1, E_2, \dots$ be an at most countable family of sets and let $\varepsilon, \delta > 0$ be arbitrary, but fixed numbers. Without loss of generality, we may assume that

$$\mathcal{H}^s_\delta(E_j) := \inf_{\mathcal{C}_\delta} l^s(\mathcal{C}_\delta) < \infty \quad \text{for } j \in \mathbb{N}.$$

Hence, for each $j \in \mathbb{N}$ there is a δ-cover $\mathcal{C}^{(j)}$ of E_j such that

$$l^s(\mathcal{C}^{(j)}) < \mathcal{H}^s_\delta(E_j) + \frac{\varepsilon}{2^j} \leq \mathcal{H}^s(E_j) + \frac{\varepsilon}{2^j}$$

The union of these $\mathcal{C} = \cup_{j=1}^\infty \mathcal{C}^{(j)}$ clearly covers $\cup_{j=1}^\infty E_j$; so

$$\mathcal{H}^s_\delta\left(\bigcup_{j=1}^\infty E_j\right) \leq \sum_{j=1}^\infty \sum_{C \in \mathcal{C}^{(j)}} \rho(C)^s \leq \sum_{j=1}^\infty \mathcal{H}^s(E_j) + \varepsilon$$

by the above calculations.

For small δ, we have

$$\mathcal{H}^s_\delta\left(\bigcup_{j=1}^\infty E_j\right) > \mathcal{H}^s\left(\bigcup_{j=1}^\infty E_j\right) - \varepsilon,$$

so in conclusion

$$\mathcal{H}^s\left(\bigcup_{j=1}^\infty E_j\right) \leq \sum_{j=1}^\infty \mathcal{H}^s(E_j) + 2\varepsilon.$$

Since ε was arbitrary, this proves sub-additivity of $\mathcal{H}^s$. ∎

Proposition 1.9. *For any set $E \subseteq \mathcal{L}$, there is an $s' \geq 0$ such that*

$$\mathcal{H}^s(E) = \begin{cases} \infty & \text{for } s < s', \\ 0 & \text{for } s > s'. \end{cases}$$

The value of $\mathcal{H}^{s'}(E)$ may be any number in $\mathbb{R}_+ \cup \{0, \infty\}$.

Proof. Let $E \subseteq \mathcal{L}$ and let $\varepsilon > 0$ be some arbitrary, but fixed number. We may without loss of generality assume that $\mathcal{H}^s(E) < \infty$. First, we prove that $\mathcal{H}^{s+\varepsilon} = 0$ and subsequently that if $\mathcal{H}^s(E) > 0$, then $\mathcal{H}^{s-\varepsilon} = \infty$. This will imply the proposition.

By definition, we may choose a δ-cover $\mathcal{C}_\delta$ of E such that

$$l^s(\mathcal{C}_\delta) \leq \mathcal{H}^s_\delta(E) + 1 \leq \mathcal{H}^s(E) + 1 < \infty.$$

Now, we consider some set $C \in \mathcal{C}_\delta$. By definition $\rho(C) < \delta$, so $\rho(C)^{s+\varepsilon} < \delta^\varepsilon \rho(C)^s$. Hence,

$$\mathcal{H}^{s+\varepsilon}_\delta(E) \leq l^{s+\varepsilon}(\mathcal{C}_\delta) \leq \delta^\varepsilon l^s(\mathcal{C}_\delta) \leq \delta^\varepsilon(\mathcal{H}^s(E) + 1).$$

Letting δ tend to zero in the above, we obtain

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

since $\varepsilon > 0$.

To prove the second claim, let $E \subseteq \mathcal{L}$ be a set such that $\mathcal{H}^s(E) > 0$. We assume that $\mathcal{H}^{s-\varepsilon}(E) < \infty$. The above implies that $\mathcal{H}^s(E) = 0$, which is clearly a contradiction. This completes the proof. ☐

In the light of Proposition 1.9, we make the following definition:

Definition 1.3. Let $E \subseteq \mathcal{L}$. The *Hausdorff dimension of E* is defined as

$$\dim_{\mathrm{H}}(E) = \inf\left\{s \geq 0 : \mathcal{H}^s(E) = 0\right\}.$$

1.3. *p*-adic numbers

The reader may have noticed some similarities between our field $\mathcal{L}$ and the p-adic numbers $\mathbb{Q}_p$. We remind the reader of the construction of the p-adic numbers.

Let $p \in \mathbb{N}$ be a prime. As in the construction of the reals, we begin with the integers $\mathbb{Z}$ and construct the field of fractions $\mathbb{Q}$. We introduce a norm on this field: First, for any $\frac{a}{b} \in \mathbb{Q}$, we reduce the number $\frac{a}{b} = p^n \frac{a'}{b'}$ such that $p \nmid a'$ and $p \nmid b'$. Now, we define the norm $\left|\frac{a}{b}\right| = p^{-n}$. We take the completion of $\mathbb{Q}$ in the metric induced by this norm and obtain the p-adic numbers $\mathbb{Q}_p$. The p-adic integers $\mathbb{Z}_p$ are the p-adic numbers with norm less than or equal to one. For an extensive treatment on the p-adic numbers, the reader is referred to [52].

Both the field of p-adic numbers and $\mathcal{L}$ are ultra-metric spaces. That is, the norm on each of the spaces has properties (1.5a), (1.5b) and (1.5c). They are, however, not the same objects. For one thing, there are more $\mathcal{L}$-spaces than there are spaces $\mathbb{Q}_p$, and their algebraic structures are different. Still, it is possible to do Diophantine approximation in the field of p-adics, and many of the results, we will prove in this part of the thesis already have analogues in the p-adic setting.

Regarding the well-approximable matrices, Abercrombie proved a partial result towards the most general form of the Jarník-Besicovitch Theorem in [2]. This result was subsequently completed by Dodson, Dickinson and Yuan in [17], where they showed

Theorem 1.10. *Let $m,n \in \mathbb{N}$ and $v > 0$. Define the set*

$$\mathcal{W}(v) = \left\{ X \in \mathbb{Z}_p^{mn} : |qX|_{\infty,p} < |q|_\infty^{-v} \text{ for infinitely many } q \in \mathbb{Z}^m \right\},$$

where $|\cdot|_{\infty,p}$ and $|\cdot|_\infty$ denotes the height in p-adic and standard norm respectively. Then

$$\dim_{\mathrm{H}}(\mathcal{W}(v)) = \begin{cases} (m-1)n + \frac{m}{v} & \text{for } v \geq \frac{m}{n}, \\ mn & \text{otherwise.} \end{cases}$$

We will prove a Laurent analogue of this theorem in Chapter 4, but in our setting we will look at distances to the integer lattice instead of the distance to the origin. This means that the Hausdorff dimension in our setting is different than the one obtained here.

Also, the set of badly approximable matrices has been examined by Abercrombie in [1], albeit only in one dimension. Abercrombie proved that the Hausdorff dimension of this set is one. This is the analogous result to Jarník's Theorem (Theorem 1.4). His method was largely measure theoretic and failed to generalise to multiple dimensions. In Chapter 5, we prove a multi-dimensional analogue of Jarník's Theorem for $\mathcal{L}$. Our method relies mostly on the fact that $\mathcal{L}$ is an ultra-metric space, and it should be possible to adapt this method to the p-adic setting, though some additional difficulties inevitably will arise.

1.4. Previous results in $\mathcal{L}$

Previous results in Diophantine approximation have mainly been about properties of the continued fraction expansion and it's relationship with certain algebraic properties of elements in $\mathcal{L}$. Until now, the only results in metrical Diophantine approximation in $\mathcal{L}$ were analogues of Dirichlet's Theorem and Khintchine's Theorem. A large number of results have been found in the non-metrical setting, but a complete discussion of these would be too extensive to complete here. Instead, the reader is referred to [40], [57] and the works cited therein. We discuss the continued fractions map and some consequences, relevant to the subsequent questions in this thesis.

Development of the continued fraction expansion in $\mathcal{L}$ goes back to Artin ([4]). In § 12 of his paper, Artin defines the continued fractions algorithm for any $x \in \mathcal{L}$:

$$\begin{aligned} x &= [x] + \{x\}, & a_0 &= [x], & x_1 &= \tfrac{1}{\{x\}} \\ x_1 &= [x_1] + \{x_1\}, & a_1 &= [x_1], & x_2 &= \tfrac{1}{\{x_1\}} \\ &\vdots \\ x_n &= [x_n] + \{x_n\}, & a_n &= [x_n], & x_{n+1} &= \tfrac{1}{\{x_n\}} \end{aligned}$$

and so on. He proves, that the continued fractions expansion

$$x = a_0 + \cfrac{1}{a_1 + \frac{1}{a_2 + \cdots}}$$

is convergent and that it terminates if and only if $x \in \mathbb{F}(X)$.

Furthermore, he shows that as in the real case, we get a recursively defined system of convergents

$$p_n = p_{n-1}a_{n-1} + p_{n-2}, \quad q_n = q_{n-1}a_{n-1} + q_{n-2},$$

where $p_0 = 1$, $p_1 = a_0$, $q_0 = 0$ and $q_1 = 1$. For this system,

$$\left\| x - \frac{p_n}{q_n} \right\| = \frac{1}{\|q_n q_{n+1}\|} \leq \frac{1}{\|q_n\|^2} \tag{1.9}$$

for any $n \in \mathbb{N}$. Hence, the analogue of Dirichlet's Theorem (Theorem 1.1) holds in $\mathcal{L}$. With a little more care, it can be shown that $\frac{p}{q}$ is a convergent to x if and only if (1.9) holds with p_n, q_n replaced with p, q.

A certain way of classifying the interesting elements in $\mathcal{L}$ is a result of the continued fractions expansion. For any $x \in \mathcal{L}$ and any $\alpha \in \mathbb{R}$, we define numbers

$$A(x,\alpha) = \liminf_{\|q\| \to \infty} \|q\|^{\alpha} \left\| x - \frac{p}{q} \right\|,$$

where (p,q) runs over all elements in $\mathbb{F}[X] \times (\mathbb{F}[X] \setminus \{0\})$. This makes it possible to define the approximation exponent of x:

$$\nu(x) = \sup\{\alpha \in \mathbb{R} : A(x,\alpha) < \infty\}.$$

These two functions enables us to classify the elements of $\mathcal{L}$:

- $x \in \mathcal{L}$ is *badly approximable* if $\nu(x) = 2$ and $A(x,2) > 0$. Performing deductions from the continued fractions algorithm analogously to the ones leading to Theorem 23 in [29], this can be seen to be equivalent to the property that the partial coefficients are bounded.

- $x \in \mathcal{L}$ is *normally approximable* if $\nu(x) = 2$ and $A(x,2) = 0$.
- $x \in \mathcal{L}$ is *well-approximable* if $\nu(x) > 2$.

A large number of results are known about the relationship between the above classification, the continued fractions map and certain algebraic equations. For a reference on these, see [57] and [40]. In this thesis, we will be interested in "counting" the elements of the above types. That is, finding the Haar measure of the sets of elements of each type, and – whenever the measure is zero – finding the Hausdorff dimension of the sets.

One previous result involving measuring the sets in question is known. Namely, de Mathan proved an adelic version of Khintchine's Theorem in his thesis ([16]). Just stating his theorem requires a bit of work, since the norms, de Mathan was working with, differ from ours.

For any prime ideal $\langle P\rangle \subseteq \mathbb{F}[X]$, we can write any rational polynomial

$$\frac{p(X)}{q(X)} = P^l \frac{p'(X)}{q'(X)}, \quad \text{for some } l \in \mathbb{N}, \text{ where } P \not| \, p'(X), P \not| \, q'(X).$$

We can now define an alternative norm on the rationals by

$$\left\| \frac{p(X)}{q(X)} \right\|_P = \left\| P^l \frac{p'(X)}{q'(X)} \right\|_P = \left(k^{\deg(P)}\right)^l.$$

Completing $\mathbb{F}(X)$ in this norm instead of the norm defined in (1.4), we obtain a new field $\mathcal{L}_P$. This field is related to $\mathcal{L}$ in approximately the same fashion as the p-adic numbers are related to the reals. In this setting, de Mathan proved the following:

Theorem 1.11. *Let I be a finite subset of the prime ideals in $\mathbb{F}[X]$ and let for any $P \in I$, ψ_P be a real function on the set of all non-zero powers of k. Let $\psi = \prod_{P\in I} \psi_P$. Furthermore, let $\mathcal{V}_I = \prod_{P\in I} \mathcal{L}_P$ equipped with the product measure of the normalised Haar measures on each $\mathcal{L}_P$. Consider the inequalities for $x = (x_i)_{i\in I} \in \mathcal{V}_I$,*

$$\left\| x_P - \frac{p}{q} \right\|_P \leq \psi_P(\max\{\|p\|, \|q\|\}) \quad \text{for any } P \in I. \tag{1.10}$$

If the series $\sum_{n=0}^{\infty} k^{2n}\psi(k^n)$ converges, then (1.10) has only finitely many solutions (p,q) for almost all $x \in \mathcal{V}_I$. If the series diverges, the sequence $k^{2n}\psi(k^n)$ decreases and there exists a $K > 0$ such that for any $P \in I$, $\psi(k^n) \leq K\psi(k^m)$ whenever $m \leq n$, then there are infinitely many solutions (p,q) to (1.10) for almost all $x \in \mathcal{V}_I$.

This is de Mathan's version of Khintchine's Theorem (Theorem 1.2) in $\mathcal{L}$. Another version of this theorem will be given in Chapter 3.

We will not give a complete discussion of the constructions involved in de Mathan's proof, since this will lead us in an entirely different direction to the one we wish to follow. The results and constructions are, however, interesting in their own right, and certainly deserve mention in this place.

2. Lüroth expansions

In this chapter, we prove some metric results on approaching any element in $\mathcal{L}$ with rational elements by a certain method. The approximation obtained via the continued fractions expansion has been extensively studied, and in Section 1.4 we gave some of the results and arguments involved in the study of this algorithm. The approximation we examine in this case, is the approximation by the so-called Lüroth expansions of the elements. It is at a first glance slightly less elegant than the continued fractions algorithm, but on the other hand this algorithm has virtues which allow us to obtain quite beautiful results on the coefficients. In particular, it is easy to examine using probability theory.

We begin with the construction of the Lüroth expansions. Then we prove the theorem which allows us to obtain the metrical results of this chapter. Finally, we prove these metrical results. The results in this chapter are published in [36].

2.1. Construction of the Lüroth expansions

We will follow the construction of the Lüroth expansions from the reals used by Perron in [49]. In fact, Perron constructs both the Lüroth series, the Sylvester series and the Engel series here, since these are quite similar. We will only construct the Lüroth series. The metrical theory of these series over the reals has been discussed among other places in [23]. For previous results in $\mathcal{L}$, see [31] and [32].

Lüroth series are a way of approximating irrationals with rationals given by a recursive algorithm. We construct the algorithm. Let $x \in \mathcal{L}$, and let $q_0 = [x]$. If $x - q_0 = 0$, x is already rational in which case the recursion stops. That is, we let $a_i = q_i = \infty$ for $i \geq 1$ as we did in the case of the continued fractions expansion. We can now write

$$x = q_0 + \frac{1}{a_1} \text{ where } \|a_1\| > 1.$$

We now define $q_1 = [a_1] + 1$. If $q_1 = 0$, x is rational and the recursion stops. Otherwise,

$$\left\|\frac{1}{q_1}\right\| = \left\|\frac{1}{a_1}\right\| \leq \left\|\frac{1}{q_1 - 1}\right\| = \left\|\frac{1}{q_1} + \frac{1}{(q_1 - 1)q_1}\right\|. \tag{2.1}$$

If $q_1 - a_1 = 0$, x is rational and the recursion stops. If not, we write

$$\frac{1}{a_1} = \frac{1}{q_1} + \frac{1}{(q_1 - 1)q_1}\frac{1}{a_2} \text{ where } \|a_2\| > 1.$$

Continuing in this fashion, we define (possibly finite) sequences (q_i) and (a_i) such that

$$\frac{1}{a_i} = \frac{1}{q_i} + \frac{1}{(q_i-1)q_i}\frac{1}{a_{i+1}} \quad \text{for } i = 1,2,\ldots. \tag{2.2a}$$

$$q_i \in \mathbb{F}[X], \|q_i\| > 1 \quad \text{for } i = 0,1,\ldots, \tag{2.2b}$$

$$\|a_i\| \geq 1 \quad \text{for } i = 1,2,\ldots, \tag{2.2c}$$

If x is rational, the recursion stops at some point. If x is irrational, the recursion continues ad infinitum. In either case, bearing in mind that $\frac{1}{\infty} = 0$ by convention, this construction gives an expansion of x on the form

$$x = q_0 + \frac{1}{q_1} + \sum_{i=1}^{\infty} \frac{1}{q_1(q_1-1)\cdots q_{i-1}(q_{i-1}-1)q_i}, \quad q_i \in \mathbb{F}[X]. \tag{2.3}$$

This expansion is called *the Lüroth expansion of x*. Note, that the series is convergent by (2.2b), since the radius of convergence of $\sum_{n=0}^{\infty} z^n$ is 1, and this series majorises the Lüroth series for an appropriate z with $|z| < 1$.

Proposition 2.1. *There is a one-to-one correspondence between the series on the form* (2.3) *and the elements of $\mathcal{L}$.*

Proof. We have already seen that all elements in $\mathcal{L}$ have an expansion of the form (2.3). Also, convergence has been proved for each expansion, so each such series is convergent. Hence, we need only prove that the expansions are unique.

Assume that an element $x \in \mathcal{L}$ has two distinct Lüroth series:

$$\begin{aligned} x &= q_0 + \frac{1}{q_1} + \sum_{i=1}^{\infty} \frac{1}{q_1(q_1-1)\cdots q_{i-1}(q_{i-1}-1)q_i} \\ &= q_0' + \frac{1}{q_1'} + \sum_{i=1}^{\infty} \frac{1}{q_1'(q_1'-1)\cdots q_{i-1}'(q_{i-1}'-1)q_i'}. \end{aligned}$$

Now, we define elements γ_j, γ_j' by the equations

$$\begin{aligned} \frac{1}{\gamma_j} &= \frac{1}{q_j} + \sum_{i=1}^{\infty} \frac{1}{q_j(q_j-1)\cdots q_{j+i-1}(q_{j+i-1}-1)q_{j+i}}, \\ \frac{1}{\gamma_j'} &= \frac{1}{q_j'} + \sum_{i=1}^{\infty} \frac{1}{q_j'(q_j'-1)\cdots q_{j+i-1}'(q_{j+i-1}'-1)q_{j+i}'}. \end{aligned}$$

Clearly, $x = q_0 + \frac{1}{\gamma_1} = q_0' + \frac{1}{\gamma_1'}$. By calculations as the ones leading to (2.1),

$$\left\|\frac{1}{q_j}\right\| = \left\|\frac{1}{\gamma_j}\right\| \leq \left\|\frac{1}{q_j-1}\right\| < 1 \quad \left\|\frac{1}{q_j'}\right\| = \left\|\frac{1}{\gamma_j'}\right\| \leq \left\|\frac{1}{q_j'-1}\right\| < 1, \tag{2.4}$$

so we have that $q_0 = q'_0 = [x]$ and $\gamma_1 = \gamma'_1$.

We now see that

$$\frac{1}{\gamma_j} = \frac{1}{q_j} + \frac{1}{q_j(q_j-1)}\frac{1}{\gamma_{j+1}}, \quad \frac{1}{\gamma'_j} = \frac{1}{q'_j} + \frac{1}{q'_j(q'_j-1)}\frac{1}{\gamma'_{j+1}} \tag{2.5}$$

for any $j \in \mathbb{N}$. Hence by (2.4) we get

$$\|q_1 - \gamma_1\| = \left\|\frac{\gamma_1}{q_1 - 1}\right\| \left\|\frac{1}{\gamma_2}\right\| < \frac{\|\gamma_1\|}{\|q_1 - 1\|} = 1.$$

Since $\|q_1\| = \|\gamma_1\|$, we have that $q_1 = [\gamma_1]$. Similarly, $q'_1 = [\gamma'_1] = [\gamma_1]$. Uniqueness follows by induction with the above as the basis, using (2.5) at each step. □

Clearly, the above construction gives us a way of approximating arbitrary elements with rational elements, since the summands in the series expansion are rationals. Hence, it is interesting to examine the coefficients of the expansion to obtain further information on how well the algorithm works. The coefficients of the Lüroth expansions turn out to be particularly easy to examine, since they have nice probabilistic properties. To prove these, we will need a dynamical interpretation of the construction.

We will only consider the dynamics of elements in the ideal I, since our results will generalise to all of $\mathcal{L}$ by translation. Also, the normalised Haar measure induces a probability measure on I by restriction, thus allowing us to use tools from probability theory directly on I. On this ideal, we define the two operators $q : I \setminus \{0\} \to \mathbb{F}[X]$ and $T : I \to I$ by

$$q(x) = \left[\frac{1}{x}\right] + 1, \quad T(x) = \begin{cases} 0 & \text{for } x = 0, \\ (q(x)-1)(xq(x)-1) & \text{otherwise.} \end{cases}$$

Clearly, q maps I into $\mathbb{F}[X]$. We need to check that T maps I into I. Since

$$\|T(x)\| = \|x\|^{-1}\,\|x\| \left\|q(x) - \frac{1}{x}\right\| < 1,$$

this is also true.

We now claim that with the operators above, the coefficients of the Lüroth expansion of an element $x \in I$ are of the form $q_r = q(T^{r-1}(x))$. To prove this, it is sufficient to prove that $T^{r-1}(x) = \frac{1}{a_r}$, where a_r is the element in (2.2a). Since $x \in I$, this certainly holds for $r = 1$. Assume, that the claim holds for some $r > 0$. By (2.2a),

$$T^r(x) = T\left(\frac{1}{a_r}\right) = (q_r - 1)\left(\frac{1}{a_r}q_r - 1\right) = \frac{1}{a_r}(q_r-1)q_r - q_r + 1 = \frac{1}{a_{r+1}}.$$

Hence, this is indeed a dynamical description of the expansion.

With these tools in place, we can prove that the coefficients of the Lüroth expansions are as nice as we could possibly hope for from a probabilistic point of view.

Theorem 2.2. *The coefficients q_i of the Lüroth expansion of an element $A \in \mathcal{L}$ are independent, identically distributed random variables.*

Proof. For any $n \in \mathbb{N}$ and any $p_1, \dots, p_n \in \mathbb{F}[X] \setminus \mathbb{F}$, we define the set

$$I_n = I_n(p_1, \dots, p_n) = \{x \in I : q_1(x) = p_1, \dots, q_n(x) = p_n\}$$

along with the ball $I_0 = I$, where $q_i(x) = q(T^{r-1}(x))$ denotes the i'th coefficient in the Lüroth expansion of x. We call these sets Lüroth cylinders.

Now, let $x \in I_n$ for some Lüroth cylinder. Since $x \in I$, $q_0(x) = 0$. Defining elements

$$d_0 = 1, \quad d_i = \frac{1}{p_1(p_1-1)\cdots p_i(p_i-1)}, \quad c_n = \sum_{i=1}^{n} \frac{d_{i-1}}{p_i},$$

we can even find a general form of any $x \in I_n$. Indeed,

$$x = c_n + d_n \sum_{i=1}^{\infty} \frac{1}{q_{n+1}(q_{n+1}-1)\cdots q_{n+i-1}(q_{n+i-1}-1)q_{n+i}}, \tag{2.6}$$

where the q_{n+i} are the last coefficients from the Lüroth expansion of x.

We recognise the sum appearing in (2.6) as the tail of the Lüroth expansion of x. That is, the sum is equal to $\frac{1}{a_{n+1}}$ from the construction. By our dynamical description of the expansion, we see that the function $\phi_n : I \to I_n$ defined by $\phi_n(x) = c_n + d_n x$ is a left inverse of the restriction of T^n to the Lüroth cylinder I_n. That is, for any $x \in I_n$, $\phi_n(T^n(x)) = x$.

Obviously, ϕ_n is surjective. This gives us an alternative characterisation of I_n,

$$I_n = \phi_n(I) = c_n + d_n I = B(c_n, \|d_n\|).$$

Now, we immediately have the measure of the Lüroth cylinders. Indeed, from the characterisation in Section 1.2.3,

$$\mu(I_n(p_1, \dots, p_n)) = \frac{1}{\|p_1(p_1-1)\cdots p_n(p_n-1)\|} = \frac{1}{\|p_1 \cdots p_n\|^2}.$$

From this, it follows directly that the coefficients are identically distributed and independent. In fact, we have calculated the probability of a given coefficient taking a given value:

$$\mu\{x \in I : q_r = p\} = \frac{1}{\|p\|^2}, \tag{2.7}$$

where $r \in \mathbb{N}$ and $p \in \mathbb{F}[X] \setminus \mathbb{F}$. □

2.2. Metrical subsequence results

Using Theorem 2.2, we may deduce a number of results about the coefficients of the Lüroth expansions from the classical theorems of probability theory. In the following, this is exactly what we will do. Our results are subsequence results. In particular, they imply some results by John and Arnold Knopfmacher ([31] and [32]), who proved similar results for the sequence $n_k = k$.

The first couple of results are simple consequences of the Strong Law of Large Numbers.

Proposition 2.3. *Let $(n_i) \subseteq \mathbb{N}$ be a strictly increasing sequence and let $p \in \mathbb{F}[X] \setminus \mathbb{F}$.*

$$\lim_{i\to\infty} \frac{1}{i} \left|\{r \le i : q_{n_r}(x) = p\}\right| = \|p\|^{-2}$$

for almost every $x \in I$.

Proof. We define random variables $X_i = \mathbb{1}_{\{x\in I: q_{n_i}(x)=p\}}$. By Theorem 2.2, these are independent and identically distributed. By (2.7), $\mathbb{E}(X_1) = \|p\|^{-2}$. The Strong Law of Large Numbers (Theorem 3.30 in [11]) now says,

$$\lim_{i\to\infty} \frac{1}{i} \left|\{r \le i : q_{n_r}(x) = p\}\right| = \lim_{i\to\infty} \frac{\sum_{j=1}^{i} X_j}{i} = \mathbb{E}(X_1) = \|p\|^{-2}$$

for almost every $x \in I$. This completes the proof. □

We are also able to obtain an estimate on the average of the degree of the coefficients.

Proposition 2.4. *Let $(n_i) \subseteq \mathbb{N}$ be a strictly increasing sequence. For almost every $x \in I$,*

$$\lim_{i\to\infty} \frac{1}{i} \sum_{r=1}^{i} \deg(q_{n_r}(x)) = \frac{k}{k-1}.$$

Proof. We define random variables $Y_i = \deg(q_{n_r})$. By Theorem 2.2, these are independent and identically distributed. We calculate the expectation of Y_1.

$$\begin{aligned}\mathbb{E}(Y_1) = \int_I \deg(q_{n_1}(x))\, d\mu(x) &= \sum_{\substack{p\in\mathbb{F}[X] \\ \|p\|>1}} \int_{\{x\in I: q_{n_1}(x)=p\}} \deg(q_{n_1}(x))\, d\mu(x) \\ &= \sum_{\substack{p\in\mathbb{F}[X] \\ \|p\|>1}} \deg(p)\, \|p\|^{-2} = \sum_{r=1}^{\infty} r(k-1)k^r k^{-2r} = \frac{k}{k-1}.\end{aligned}$$

Now, the Strong Law of Large Numbers yields,

$$\lim_{i\to\infty} \frac{1}{i} \sum_{r=1}^{i} \deg(q_{n_r}(x)) = \lim_{i\to\infty} \frac{\sum_{j=1}^{i} Y_j(x)}{i} = \mathbb{E}(Y_1) = \frac{k}{k-1}$$

for almost every $x \in I$. □

We immediately have the following corollary:

Corollary 2.5. *Let $(n_i) \subseteq \mathbb{N}$ be a strictly increasing sequence. For almost every $x \in I$,*

$$\lim_{i\to\infty} \|q_{n_1}(x)\cdots q_{n_i}(x)\|^{1/i} = k^{k/(k-1)}.$$

Proof. Note that $\deg(p) = \log_k \|p\|$. Inserting this in Proposition 2.4, we obtain the corollary. □

These were just a few easy results following from the Strong Law of Large Numbers. Fortunately, there are more probabilistic theorems, stating beautiful results on independent, identically distributed random variables, which allow us to deduce stronger results from Theorem 2.2. We begin with the following proposition, strengthening Proposition 2.3.

Proposition 2.6. *Let $(n_i) \subseteq \mathbb{N}$ be a strictly increasing sequence, let $p \in \mathbb{F}[X] \setminus \mathbb{F}$ and define for all $r \in \mathbb{N}$ random variables $Z_{r,p} = |\{i \le r : q_{n_i} = p\}|$. For almost all $x \in I$,*

$$\limsup_{r\to\infty} \frac{Z_{r,p} - r\|p\|^{-2}}{\sqrt{r\log\log r}} = \sqrt{2\|p\|^{-2}\left(1-\|p\|^{-2}\right)}.$$

Furthermore, for any $s \in \mathbb{R}$,

$$\lim_{r\to\infty} \mu\left\{x \in I : Z_{r,p} - r\|p\|^{-2} < \frac{s}{\|p\|}\sqrt{r\left(1-\|p\|^2\right)}\right\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{s} e^{-u^2/2}du.$$

Proof. We let $p \in \mathbb{F}[X] \setminus \mathbb{F}$ be fixed but arbitrary and consider the random variables $X_i = \mathbb{1}_{\{x\in I: q_{n_i}(x)=p\}}$. By Theorem 2.2, these are independent and identically distributed. We need the variance of these, so we calculate the first and second moments. Since the random variables are indicator functions, $X_i = X_i^2$. Hence by (2.7),

$$\mathbb{E}\left(X_i^2\right) = \mathbb{E}(X_i) = \int_I X_i d\mu = \int_I \mathbb{1}_{\{x\in I: q_{n_i}(x)=p\}} = \|p\|^{-2}.$$

By a standard result in probability theory,

$$\sigma^2(X_i) = \mathbb{E}\left(X_i^2\right) - \mathbb{E}(X_i)^2 = \|p\|^{-2}\left(1-\|p\|^{-2}\right).$$

If we centre the random variables and calculate the partial sums S_r of the centred random variables, we see that

$$S_r = \sum_{i=1}^{r}(X_1 - \mathbb{E}(X_i)) = \sum_{i=1}^{r}\left(\mathbb{1}_{\{x\in I: q_{n_i}(x)=p\}} - \|p\|^{-2}\right) = Z_{r,p} - r\|p\|^{-2}.$$

Clearly, the variance is preserved under the translation, so by the Law of the Iterated Logarithm (Theorem 13.25 in [11]),

$$\limsup_{r\to\infty} \frac{Z_{r,p} - r\|p\|^{-2}}{\sqrt{r\log\log r}} = \limsup_{r\to\infty} \frac{S_r}{\sqrt{r\log\log r}} = \sqrt{2\sigma^2(X_i)} = \sqrt{2\|p\|^{-2}\left(1-\|p\|^{-2}\right)}$$

for almost every $x \in I$. This is the first statement of the proposition.

To prove the second claim, we see that the Central Limit Theorem (Corollary 8.23 in [11]) states that

$$\frac{Z_{r,p} - r\|p\|^{-2}}{\|p\|^{-1}\sqrt{r\left(1-\|p\|^{-2}\right)}} = \frac{S_n}{\sigma(X_i)\sqrt{r}} \xrightarrow[r\to\infty]{\mathcal{D}} \mathcal{N}(0,1),$$

where $\mathcal{N}(0,1)$ denotes the standard normal distribution and the convergence is in distribution. But this is the second statement by definition of convergence in distribution. □

A generalisation of Proposition 2.4 using the Law of the Iterated Logarithm is also possible.

Proposition 2.7. *Let $(n_i) \subseteq \mathbb{N}$ be a strictly increasing sequence. For almost every $x \in I$,*

$$\limsup_{r\to\infty} \frac{\sum_{i=1}^{r} \deg(q_{n_i}(x)) - (k/(k-1))\,r}{\sqrt{r\log\log r}} = \frac{\sqrt{2k}}{k-1}.$$

Proof. We truncate the interesting random variables in order to obtain more precise estimates:

$$Y_i'(x) = \begin{cases} \deg\left(q_{n_i}(x)\right) & \text{for } \|q_{n_i}(x)\| \leq i^2, \\ 0 & \text{otherwise.} \end{cases}$$

Once again, we calculate the first and second moment of these variables. A calculation similar to the one performed in the proof of Proposition 2.4 yields

$$\begin{aligned} \mathbb{E}\left(Y_i'\right) &= \sum_{\substack{p\in\mathbb{F}[X] \\ 1<\|p\|\leq i^2}} \int_{\{x\in I : q_{n_i}(x)=k\}} \deg\left(q_{n_i}(x)\right) d\mu \\ &= \sum_{r:k^r\leq i^2} k^{-2r}k^r(k-1)r = \mathbb{E}(\deg\left(q_{n_i}(x)\right)) + O\left(\frac{\log i}{i^2}\right). \end{aligned}$$

Similarly,

$$\mathbb{E}\left(Y_i'^2\right) = \mathbb{E}\left(\deg\left(q_{n_i}(x)\right)^2\right) + O\left(\frac{\log^2 i}{i^2}\right).$$

Hence,

$$\sigma^2(Y_i') = \mathbb{E}\left(Y_i'^2\right) - \mathbb{E}\left(Y_i'\right)^2 = \sigma^2\left(\deg\left(q_{n_i}(x)\right)\right) + O\left(\frac{\log^2 i}{i^2}\right).$$

Now, we define a new quantity

$$B_r := \sum_{i=1}^{r} \sigma^2(Y_i') = \sum_{i=1}^{r} \sigma^2\left(\deg\left(q_{n_i}(x)\right)\right) + O(1) = \frac{kr}{(k-1)^2} + O(1),$$

where the calculation of $\sigma^2(\deg(q_{n_i}(x)))$ in the last equality is performed completely analogously to the calculation of the expectation in the proof of Proposition 2.4. Also,

$$Y_i'(x) \leq 2\log_k i = O\left(\sqrt{\frac{B_i}{\log\log B_i}}\right).$$

Now, we can use the Law of the Iterated Logarithm, since the Y_i' are independent by Theorem 2.2.

$$\limsup_{r\to\infty} \frac{\sum_{i=1}^r (Y_i' - \mathbb{E}(Y_i'))}{\sqrt{2B_r \log\log B_r}} = 1 \tag{2.8}$$

for almost every $x \in I$.

Asymptotically, we have $B_r \log\log B_r \sim \frac{k}{(k-1)^2} r \log\log r$. Hence, (2.8) implies,

$$\limsup_{r\to\infty} \frac{\sum_{i=1}^r (Y_i' - \mathbb{E}(\deg(q_{n_i}(x))))}{\sqrt{2\frac{k}{(k-1)^2} r\log\log r}} = \limsup_{r\to\infty} \frac{\sum_{i=1}^r Y_i' - k/(k-1)r}{\sqrt{2\frac{k}{(k-1)^2} r\log\log r}} = 1 \tag{2.9}$$

for almost every $x \in I$.

This is almost the required result, except that the random variables Y_i' are not necessarily equal to $\deg(q_{n_i}(x))$. However, this is easily repaired. We define sets $U_i = \{x \in I : Y_i'(x) \neq \deg(q_{n_i}(x))\}$. We calculate the sum of the measures of all these sets:

$$\sum_{i=1}^{\infty} \mu(U_i) = \sum_{i=1}^{\infty} \sum_{\|p\|\geq i^2} \|p\|^{-2} < \sum_{i=1}^{\infty} \frac{1}{i^2} < \infty.$$

The Borel–Cantelli Lemma (Lemma 3.14 in [11]) implies that for almost every $x \in I$, there exists an $i_0 \in \mathbb{N}$ such that for $i \geq i_0$, $Y_i'(x) = \deg(q_{n_i}(x))$. Hence, (2.9) does indeed imply the proposition. □

Just as was the case with Proposition 2.4, where we were able to obtain Corollary 2.5 by a simple substitution, we can obtain a stronger corollary from the stronger Proposition 2.7.

Corollary 2.8. *Let $(n_i) \subseteq \mathbb{N}$ be a strictly increasing sequence. For almost every $x \in I$,*

$$\|q_{n_1}(x)\cdots q_{n_r}(x)\|^{1/r} = k^{k/(k-1)} + O\left(\frac{\sqrt{\log\log r}}{r}\right).$$

Proof. Once again, observe that $\deg(q_{n_r}(x)) = \log_k \|q_{n_r}(x)\|$. Inserting this in Proposition 2.7 yields the corollary. The error term arises when we estimate the multiplicative error from Proposition 2.7 in terms of an additive one. □

The final result of this chapter tells us that the norms of the partial coefficients in the Lüroth expansions are bounded in a certain sense for almost all $x \in I$. Once again, it is a result about convergence in probability.

Proposition 2.9. *Let $(n_i) \subseteq \mathbb{N}$ be a strictly increasing sequence. For any $\varepsilon > 0$,*

$$\lim_{i\to\infty} \mu\left\{x \in I : \frac{1}{i\log_k i}\left|\sum_{r=1}^{i} \|q_{n_r}(x)\| - (k-1)\right| > \varepsilon\right\} = 0.$$

Proof. Let $i \in \mathbb{N}$ be fixed but arbitrary. We split the sum of the norms of the partial coefficients up into two random variables. For $r \leq i$, we define

$$V_r(x) = \begin{cases} \|q_{n_r}(x)\| & \text{when } \|q_{n_r}(x)\| \leq i\log_k i, \\ 0 & \text{otherwise,} \end{cases}$$

$$W_r(x) = \begin{cases} 0 & \text{when } \|q_{n_r}(x)\| \leq i\log_k i, \\ \|q_{n_r}(x)\| & \text{otherwise.} \end{cases}$$

Clearly,

$$\mu\left\{x \in I : \frac{1}{i\log_k i}\left|\sum_{r=1}^{i} \|q_{n_r}(x)\| - (q-1)\right| > \varepsilon\right\}$$
$$\leq \mu\left\{x \in I : \left|\sum_{r=1}^{i} V_r - (k-1)i\log_k i\right| > \varepsilon i\log_k i\right\} + \mu\left\{x \in I : \sum_{r=1}^{i} W_r \neq 0\right\}. \quad (2.10)$$

We will treat the measure of each of the sets on the right hand side separately, proving that the measures tend to zero as i tends to infinity.

By Theorem 2.2, the V_r are independent and identically distributed. Hence by standard probability theory, $\mathbb{E}\left(\sum_{r=1}^{i} V_r\right) = i\mathbb{E}(V_1)$ and $\sigma^2\left(\sum_{r=1}^{i} V_r\right) = i\sigma^2(V_1)$. We calculate estimates of these quantities. First the expectation:

$$\mathbb{E}(V_1) = \int_I V_1 d\mu = \sum_{\substack{p\in\mathbb{F}[X] \\ 1<\|p\|\leq i\log_k i}} \int_{\{x\in I: q_{n_1}(x)=p\}} \|q_{n_1}(x)\|\, d\mu$$
$$= \sum_{\substack{p\in\mathbb{F}[X] \\ 1<\|p\|\leq i\log_k i}} \|p\|^{-1} = \sum_{r:1<k^r\leq i\log_k i} k^{-r}(k-1)k^r = (k-1)\left[\log_k(i\log_k i)\right],$$

where $[\cdot]$ denotes the integral value function. Similarly, for the variance we get,

$$\sigma^2(V_1) < \mathbb{E}\left(V_1^2\right) = \sum_{\substack{p\in\mathbb{F}[X] \\ 1<\|p\|\leq i\log_k i}} 1 = \sum_{r:1<k^r\leq i\log_k i} (k-1)k^r < ki\log_k i.$$

We apply Chebychev's Inequality (Proposition 1.7 in [11]) and the independence and identical distribution of the V_i to these estimates,

$$\mu\left\{x \in I : \left|\sum_{r=1}^{i} V_r - i\mathbb{E}(V_1)\right| > \varepsilon i\mathbb{E}(V_1)\right\} \leq \frac{\mathbb{E}\left(\sum_{r=1}^{i} V_r - i\mathbb{E}(V_1)\right)}{(\varepsilon i\mathbb{E}(V_1))^2}$$
$$= \frac{i\sigma^2(V_1)}{(\varepsilon i\mathbb{E}(V_1))^2} < \frac{ki^2\log_k i}{(\varepsilon(k-1)i\left[\log_k(i\log_k i)\right])^2}. \quad (2.11)$$

This quantity tends to zero as i tends to infinity. Also, $\mathbb{E}(V_1) \sim (k-1)\log_k i$ when i tends to infinity, so the first summand in (2.10) tends to zero.

For the second summand, not quite as much work is needed. It suffices to observe that by Theorem 2.2,

$$\mu\left\{x \in I : \sum_{r=1}^{i} W_r \neq 0\right\} \leq i\mu\{x \in I : \|q_{n_1}(x)\| > i\log_k i\}$$
$$= i \sum_{\substack{p\in\mathbb{F}[X] \\ \|p\|>i\log_k i}} \|p\|^{-2} < \frac{1}{\log_k i}.$$

This also tends to zero as i tends to infinity, so the expression in (2.10) tends to zero. This completes the proof. □

3. Approximation and Haar measure

In this section, we will calculate the Haar measure of certain subsets of $\mathcal{L}$ and subsets of the matrices over $\mathcal{L}$. It turns out that the interesting sets – speaking in the context of metrical Diophantine approximation – have either null or full measure. This is our main result in this section, which also serves to motivate the next two chapters, where we examine the null-sets more closely.

We will begin with a one-dimensional result, which is subsequently generalised to linear forms. Clearly, the multidimensional result implies the one-dimensional result, but the proof of the multidimensional result is much more complicated. The inclusion of both proofs serves to illustrate the kind of difficulties one encounters when passing from one to several dimensions in the examination of these questions. This will be the only chapter of this part of the thesis, where we will have results both in one dimension and several dimensions. For the remaining theorems, even the simple, one-dimensional proofs are quite extensive and since the results are included in the multi-dimensional results and the methods of proof are very similar, the one-dimensional proofs are omitted.

3.1. Approximation and measure in one dimension

The goal of this section is to prove the following theorem, a sketch of which is found in [33]:

Theorem 3.1. *Let $\psi : \mathbb{F}[X] \to \mathbb{R}^+$ be some function decreasing with respect to the norm and only dependent on the norm. Then,*

1. *If $\sum_{q \in \mathbb{F}[X]} \psi(q) < \infty$, then*

$$\mu\{x \in \mathcal{L} : |\langle qx \rangle| < \psi(q) \text{ for infinitely many } q \in \mathbb{F}[X]\} = 0.$$

2. *If* $\sum_{q\in\mathbb{F}[X]}\frac{1}{\deg(q)}\psi(q)=\infty$, *then*

$$\mu\{x\in\mathcal{L}:|\langle qx\rangle|<\psi(q) \textit{ for infinitely many } q\in\mathbb{F}[X]\}=\infty.$$

The reader will notice that this theorem is similar to Khintchine's Theorem (Theorem 1.2) and de Mathan's Theorem (Theorem 1.11), and this is no coincidence. Theorem 3.1 is the first version of Khintchine's Theorem discussed in this thesis. In Section 3.2, we will prove a multi-dimensional version of the theorem. This involves considerably more intricate arguments, but of course the resulting theorem implies Theorem 3.1.

Note, that there is a "hole" in Theorem 3.1. For instance, for $\psi(q)=(\|q\|\deg q)^{-1}$, we get $\sum_{q\in\mathbb{F}[X]}\psi(q)=\infty$, so the first case does not apply. However, we also have $\sum_{q\in\mathbb{F}[X]}\frac{1}{\deg(q)}\psi(q)<\infty$, so neither does the second case. We conjecture that the condition $\sum_{q\in\mathbb{F}[X]}\psi(q)=\infty$ is sufficient to ensure full measure, and that this may be shown with some extra care in the calculations. However, for the purposes of the sets to be examined in this thesis, the above gives the required results.

Proof of Theorem 3.1. First, we note that it suffices to consider the restriction of the set in question to the unit ball, since $\mathcal{L}$ can be written as disjoint translates of I. This has the advantage that we may use tools from probability theory, since the normalised Haar measure induces a probability measure on I. To simplify notation, we define the set

$$\mathcal{S}(\psi)=\left\{x\in I\,\middle|\,|\langle qx\rangle|<\psi(q)\text{ for infinitely many } q\in\mathbb{F}[X]\right\}. \tag{3.1}$$

We wish to prove that

1.
$$\text{If }\sum_{q\in\mathbb{F}[X]}\psi(q)<\infty,\text{ then }\mu(\mathcal{S}(\psi))=0.$$

2.
$$\text{If }\sum_{q\in\mathbb{F}[X]}\psi(q)=\infty,\text{ then }\mu(\mathcal{S}(\psi))=1.$$

We begin with a proof of (1). For any $q\in\mathbb{F}[X]$, we define sets

$$B(q)=\{x\in I:|\langle qx\rangle|<\psi(q)\}. \tag{3.2}$$

Since we have restricted ourselves to considering the unit ball, we immediately see that

$$\begin{aligned} B(q) &= \bigcup_{\substack{p\in\mathbb{F}[X] \\ \|p\|<\|q\|}} \{x\in I : \|qx-p\| < \psi(q)\} \\ &= \bigcup_{\substack{p\in\mathbb{F}[X] \\ \|p\|<\|q\|}} \left\{x\in I : \left\|x-\frac{p}{q}\right\| < \frac{\psi(q)}{\|q\|}\right\} = \bigcup_{\substack{p\in\mathbb{F}[X] \\ \|p\|<\|q\|}} B\left(\frac{p}{q},\frac{\psi(q)}{\|q\|}\right). \end{aligned} \tag{3.3}$$

Since $B(q)$ is the union of balls, we know how to calculate an upper bound on the measure of $B(q)$:

$$\mu(B(q)) = \mu\left(\bigcup_{\substack{p\in\mathbb{F}[X] \\ \|p\|<\|q\|}} B\left(\tfrac{p}{q},\frac{\psi(q)}{\|q\|}\right)\right) \leq \sum_{\substack{p\in\mathbb{F}[X] \\ \|p\|<\|q\|}} k\frac{\psi(q)}{\|q\|} = \psi(q). \tag{3.4}$$

Hence,

$$\sum_{q\in\mathbb{F}[X]} \mu(B(q)) \leq \sum_{q\in\mathbb{F}[X]} \psi(q) < \infty,$$

and since by definition of $B(q)$,

$$\mathcal{S}(\psi) = \left\{x\in I \middle| x\in B(q) \text{ for infinitely many } q\in\mathbb{F}[X]\right\}, \tag{3.5}$$

the Borel–Cantelli Lemma implies (1).

We now prove (2). This requires a bit more finesse. We will use a little ergodic theory. For any $q\in\mathbb{F}[X]$, we define the function $T_q : I\to I$ by $T_q(x) = \{qx\}$, where $\{\cdot\}$ denotes the fractional part, $\{x\} = x-[x]$. We claim that for any $q\in\mathbb{F}[X]$, the transformation T_q is ergodic. That is, any T_q-invariant set has measure either 1 or 0.

Let $q\in\mathbb{F}[X]$ be fixed and let $E\in I$ be some T_q-invariant set. Assume that $\mu(E)>0$, and let x_0 be a point of metric density for E. For any $h\in\mathbb{N}$, we define the open balls $B_h = B(x_0,\|q\|^{-h})$. Using the invariance of E, we see that by substitution,

$$\begin{aligned} \mu(E\cap B_h) &= \int_{B_h} \mathbb{1}_E(x)d\mu(x) = \int_{B_h} \mathbb{1}_E(T_q^h x)d\mu(x) \\ &= \|q\|^{-h}\int_I \mathbb{1}_E(x)d\mu(x) = \mu(B_h)\mu(E). \end{aligned}$$

Since x_0 is a point of metric density for E, we see that $\mu(E)=1$ by letting h tend to infinity, so the transformation is ergodic.

Clearly, for any $q\in\mathbb{F}[X]$, the set $B(q)$ defined in (3.2) is invariant under T_q. Hence, it has measure 0 or 1. From (3.5), we see that $\mathcal{S}(\psi)$ can be written as an intersection of unions of the sets $B(q)$. Hence, $\mathcal{S}(\psi)$ has measure 0 or 1. Thus, it is sufficient to prove that $\mu(\mathcal{S}(\psi))>0$.

We prove this for a suitable subset. Let $I = \{q \in \mathbb{F}[X] : q \text{ irreducible and monic}\}$. We define the set

$$\mathcal{S}_{\text{irr.}}(\psi) = \{x \in I : x \in B(q) \text{ for infinitely many } q \in I\} \subseteq \mathcal{S}(\psi).$$

For $q, q' \in I$, we clearly have

$$\mu(B(q) \cap B(q')) \ll \mu(B(q))\mu(B(q')). \tag{3.6}$$

We now prove that $\sum_{q\in I} \psi(q) = \infty$. For this task, we need the Möbius function $\mu : \mathbb{N} \to \{-1, 0, 1\}$ defined by:

$$\mu(d) = \begin{cases} (-1)^r & \text{for } d \text{ a product of } r \text{ distinct primes,} \\ 0 & \text{otherwise.} \end{cases}$$

It should cause no ambiguity that both the Haar measure and the Möbius function are denoted μ, since they will never be used in the same calculation.

Since we know that the number of monic irreducible polynomials of degree r in $\mathbb{F}[X]$ is $\frac{1}{r}\sum_{d|r}\mu(\frac{r}{d})k^d$, we get

$$\sum_{q\in I} \psi(q) = \sum_{r=1}^{\infty} \left(\frac{1}{r} \sum_{d|r} \mu(\frac{r}{d}) \right) \psi(k^r) \asymp \sum_{q\in\mathbb{F}[X]} \frac{1}{\deg(q)} \psi(q),$$

which diverges.

From (3.3) and (3.4) it follows that $\mu(B(q)) \asymp \psi(q)$. Hence, for any $N \in \mathbb{N}$,

$$0 < \sum_{\substack{q,q'\in I \\ ||q||,||q'||\le N}} \mu(B(q) \cap B(q')) \ll \sum_{\substack{q,q'\in I \\ ||q||,||q'||\le N}} \psi(q)\psi(q')$$
$$< c\left(\sum_{\substack{q\in I \\ ||q||\le N}} \psi(q) \right)^2 \asymp c\left(\sum_{\substack{q\in I \\ ||q||\le N}} \mu(B(q)) \right)^2.$$

A converse to the Borel–Cantelli Lemma (see Lemma 2.3 in [22]) tells us,

$$\mu(\mathcal{S}(\psi))\mu(\mathcal{S}_{\text{irr.}}(\psi)) \ge \limsup_{N\to\infty} \frac{\left(\sum_{\substack{q\in I \\ ||q||\le N}} \mu(B(q)) \right)^2}{\sum_{\substack{q,q'\in I \\ ||q||,||q'||\le N}} \mu(B(q) \cap B(q'))} > c^{-1} > 0.$$

This completes the proof. □

The sets discussed in Theorem 3.1 are a little more general than the ones we will be discussing in the subsequent chapters. The sets, we will be interested in, correspond to the special functions $\psi(q) = ||q||^{-v}$, where $v > 0$. However, we will not prove any results specific to these specialised sets until we have passed to multiple dimensions. In multiple dimensions we will also need to impose extra assumptions on ψ.

3.2. Approximation and measure of linear forms

In this section, we generalise Theorem 3.1 to several dimensions. In fact, we prove a version of the theorem for $m \times n$-matrices over $\mathcal{L}$. First, we define the sets replacing the one-dimensional sets defined in (3.1). Let $\psi : \mathbb{F}[X]^m \to \mathbb{R}^+$ be some function. We define

$$\mathcal{S}(\psi) = \{A \in I^{mn} : |\langle qA \rangle| < \psi(q) \text{ for infinitely many } q \in \mathbb{F}[X]^m\}. \tag{3.7}$$

That is, the set of matrices A where for all entries a_{ij}, we have $a_{ij} \in I$.

The real analogue of the multi-dimensional theorem was proved by Grošev in [21]. The methods used here to prove the theorem are for the most part the ones used by Dodson ([18]). The multidimensional generalisation of Theorem 3.1, which is to appear in [35], is the following:

Theorem 3.2. *Let $\psi : \mathbb{F}[X]^m \to \mathbb{R}^+$ be a function such that $\psi(q) = \psi(\|q\|_\infty)$, which is decreasing with respect to the height.*

1.

$$\textit{If } \sum_{q \in \mathbb{F}[X]^m} \psi(q)^n < \infty, \textit{ then } \mu(\mathcal{S}(\psi)) = 0.$$

2. a)

$$\textit{If } m > 1 \textit{ and } \sum_{q \in \mathbb{F}[X]^m} \psi(q)^n = \infty, \textit{ then } \mu(\mathcal{S}(\psi)) = 1.$$

b)

$$\textit{If } m = 1 \textit{ and } \sum_{q \in \mathbb{F}[X]} \frac{1}{\deg(q)} \psi(q)^n, \textit{ then } \mu(\mathcal{S}(\psi)) = 1.$$

Note, that the "hole" from the one dimensional case also appears here, but only for $m = 1$. The multidimensional case is much easier from a geometric point of view, so there is no "hole" here. Before proving the theorem, we will prove an easy lemma, which will turn out to be extremely useful in the remainder of Part 1 of the thesis.

Lemma 3.3. *Let $m, r \in \mathbb{N}$.*

$$|\{q \in \mathbb{F}[X]^m : \|q\|_\infty = k^r\}| = m\frac{k-1}{k}k^{rm}.$$

Proof. There are m coordinates, so obviously there are m possibilities for choosing the norm bearing coordinate. The leading coefficient of the norm bearing polynomial can be chosen freely in $\mathbb{F} \setminus \{0\}$, so there are $k-1$ possibilities. For each of the remaining $r-1$ coefficients, there are $|\mathbb{F}| = k$ possible choices, so in total, there are k^{r-1}. The remaining $m-1$ coordinates can be chosen arbitrarily with norms less than or equal to k^r. Hence, there are $k^{r(m-1)}$ possibilities. These calculations give

$$|\{q \in \mathbb{F}[X]^m : \|q\|_\infty = k^r\}| = m(k-1)k^{r-1}k^{r(m-1)} = m\frac{k-1}{k}k^{rm}.$$

□

We now specialise the setting a bit further. Let $\psi : \mathbb{F}[X]^m \to \mathbb{R}_+$ be a decreasing function, taking only values in the set $\{k^r : r \in \mathbb{Z}\}$. We define for any $q \in \mathbb{F}[X]^m$ the resonant set

$$R_q = \{A \in I^{mn} : qA = p \text{ for some } p \in \mathbb{F}[X]^n\}. \tag{3.8}$$

The term *resonant set* comes from physics, where the resonance frequencies of for example a string can be found by solving Diophantine equations. Given our function ψ, we define some neighbourhoods of the R_q:

$$B_{\psi(q)}(R_q) = \{A \in I^{mn} : |\langle qA \rangle| < \psi(q)\}. \tag{3.9}$$

We will prove the following propositions:

Proposition 3.4.

$$\mu\left(B_{\psi(q)}(R_q)\right) = \psi(q)^n$$

Proposition 3.5. *Let $q, q' \in \mathbb{F}[X]^m$ be linearly independent over $\mathcal{L}$. Then*

$$\mu\left(B_{\psi(q)}(R_q) \cap B_{\psi(q')}(R_{q'})\right) = \mu\left(B_{\psi(q)}(R_q)\right)\mu\left(B_{\psi(q')}(R_{q'})\right).$$

In both proofs, we follow the method from [18].

Proof of Proposition 3.4. By the rank equation, the solution curves to the equations $qA = p$ are $(m-1)n$ dimensional affine spaces over $\mathcal{L}$. We begin by calculating the number of affine spaces which pass through the unit ball. First, note that if there is a solution to the equation $qA = p$ with $A \in I^{mn}$, then

$$\|p\|_\infty = \|qA\|_\infty \leq \|q\|_\infty \|A\|_\infty < \|q\|_\infty, \tag{3.10}$$

so certainly, the condition $\|p\|_\infty < \|q\|_\infty$ is necessary. We claim that it is also sufficient.

To see this, it suffices to find a solution $A \in I^{mn}$ which satisfies the equation. Suppose that $\|p\|_\infty < \|q\|_\infty$. Assume without loss of generality that $\|q\|_\infty = \|q_1\|$. Now,

$$qA = q \begin{pmatrix} \frac{p_1}{q_1} & \cdots & \frac{p_n}{q_n} \\ 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots 0 & \end{pmatrix} = p \tag{3.11}$$

and $A \in I^{mn}$.

As in [18], we consider the simplest non-trivial case ($m = 2, n = 1$) and subsequently extend this to the general case. In this case, the solution curves to the equations $qA = p$ define $\|q\|_\infty$ affine 1-dimensional spaces in I^2. These partition I^2 into $\|q\|_\infty$ sets, $\tilde{S}_i$ say. The distance between each affine 1-space is $\frac{1}{\|q\|_\infty}$. The measure of each such strip may be calculated using a characterisation of a translation invariant measure due to Mahler (see [42]).

Using a method more geometric than the one in Chapter 1, Mahler constructs a translation invariant measure on $\mathcal{L}^h$, such that the measure of the parallelepiped spanned by the linearly independent vectors $q_1, \ldots, q_h$ is equal to the reciprocal of the determinant of the matrix having the q_i as it's columns. Since Mahler's measure and the one we constructed in Chapter 1 agree on the set I^h, they must be the same measures by uniqueness of the Haar measure.

Hence, the solution curves partition I^2 into sets of equal size, $\mu(\tilde{S}) = \frac{1}{\|q\|_\infty}$. By the same characterisation, we find that around each solution curve we have a component, S_i say, of the set $B_{\psi(q)}(R_q)$ of measure $\psi(q)/\|q\|_\infty$. Hence

$$\mu\left(B_{\psi(q)}(R_q)\right) = \frac{\mu\left(B_{\psi(q)}(R_q)\right)}{\mu(I^2)} = \frac{\mu\left(\cup S_i\right)}{\mu\left(\cup \tilde{S}_i\right)} = \frac{\mu(S_i)}{\mu(\tilde{S}_i)} = \frac{\frac{\psi(q)}{\|q\|_\infty}}{\frac{1}{\|q\|_\infty}} = \psi(q). \tag{3.12}$$

To obtain the proposition for general $m, n \in \mathbb{N}$, consider n copies of the span of q and apply the above argument to resulting prisms in I^{mn}. This implies the proposition. □

Proof of Proposition 3.5. Again, we consider the simplest non-trivial case, so let $m = 2$ and $n = 1$. Let $q, q' \in \mathbb{F}[X]^2$ be linearly independent over $\mathcal{L}$. We calculate the number of intersections between the solution curves to the equations $qA = p$ and the equations $q'A = p'$, where p, p' run over the possible values. This amounts to solving the system

$$\begin{pmatrix} q_1 & q_2 \\ q_1' & q_2' \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} p \\ p' \end{pmatrix}, \quad \|p\|_\infty < \|q\|_\infty, \left\|p'\right\|_\infty < \left\|q'\right\|_\infty.$$

There are exactly $\left\|\det\begin{pmatrix} q_1 & q_2 \\ q_1' & q_2' \end{pmatrix}\right\|$ such solutions. Indeed, by Cramer's rule (Theorem 4.4 in [39]) we may find at least this number of solutions. To each such solution, we may assign a parallelogram of measure $1/\left\|\det\begin{pmatrix} q_1 & q_2 \\ q_1' & q_2' \end{pmatrix}\right\|$, defined by the four corresponding intersection points of parallel consecutive resonant sets. The measure is calculated by Mahler's method ([42]).

It remains to be shown that each of the parallelograms defined above is a proper subset of I^2. But this must be the case, since any parallelogram may be written as

$$\left\{x \in \mathcal{L}^2 : x = \hat{q}_1 t_1 + \hat{q}_2 t_2 + p, t_1, t_2 \in I\right\}$$

for some $\hat{q}_1, \hat{q}_2 \in \mathcal{L}^2$. Clearly,

$$\left\{x \in \mathcal{L}^2 : x = \hat{q}_1 t_1 + \hat{q}_2 t_2 + p, t_1, t_2 \in I\right\} \subseteq B\left(p, \max\left(\|\hat{q}_1\|_\infty, \|\hat{q}_2\|_\infty\right)\right),$$

so by the ultra-metric property (1.5c), the parallelogram is either fully contained in I^2 or disjoint with I^2. Since the parallelograms bounded by the solution curves are disjoint, there can be no more than the required number.

Furthermore, around each intersection point, there is another parallelogram of measure $\psi(q)\psi(q')/\left\|\det\begin{pmatrix} q_1 & q_2 \\ q_1' & q_2' \end{pmatrix}\right\|$ which forms a part of $B_{\psi(q)}(R_q) \cap B_{\psi(q')}(R_{q'})$ if it is a subset of I^2.

With the above tools, we may apply a proportional argument analogous to (3.12) to obtain the proposition in this case. For the general case, we consider n copies of the span of q and q' and apply the above to the mn dimensional prisms to obtain the proposition. □

We are now ready to embark on the proof of the main theorem of this chapter.

Proof of Theorem 3.2. Again, we consider the solution curves to certain Diophantine equations. From equations (3.8) and (3.9), we see that

$$\mathcal{S}(\psi) = \left\{A \in I^{mn} : A \in B_{\psi(q)}(R_q) \text{ for infinitely many } q \in \mathbb{F}[X]^m\right\}.$$

As in the proof of Theorem 3.1, we use the Borel–Cantelli Lemma for the first part. We no longer have the restriction on the possible values for $\psi(q)$, but clearly Proposition 3.4 implies that $\mu(B_{\psi(q)}(R_q)) \asymp \psi(q)^n$. Just as in the one-dimensional case, the Borel–Cantelli Lemma now yields (1).

As in the proof of Theorem 3.1, proving the second part is more difficult. Furthermore, in the multi-dimensional setting we need to distinguish between the cases $m = 1$ and $m > 1$. For $m = 1$, we use the same method as in the proof of Theorem 3.1. For $m > 1$, this method breaks down and we need a different approach.

Assume first that $m = 1$. As in the proof of Theorem 3.1, we define for any $q \in \mathbb{F}[X]$ an automorphism $T_q : I^n \to I^n$ by $T_q(A) = \{A\}$. Once again, $\{A\}$ denotes the fractional part, but this time the fractional part is taken in each coordinate.

This transformation is ergodic for any $q \in \mathbb{F}[X]$. Indeed, let $q \in \mathbb{F}[X]$, let $E \subseteq I^n$ be some T_q-invariant set with $\mu(E) > 0$ and let A_0 be a point of metric density for E. Furthermore, define for any $h \in \mathbb{N}$, $B_h = B(A_0, \|q\|_\infty^{-h})$. Now, by substitution

$$\begin{aligned}\mu(E \cap B_h) = \int_{B_h} \mathbb{1}_E(x)d\mu(x) \leq \int_{B_h} \mathbb{1}_E(T_q^h x)d\mu(x) \\ = \|q\|_\infty^{-hn} \int_I \mathbb{1}_E(x)d\mu(x) = \mu(B_h)\mu(E).\end{aligned}$$

Once again, letting h tend to infinity reveals that $\mu(E) = 1$ by choice of A_0 as a point of metric density. Hence the transformation is ergodic, and the same argument used in the proof of Theorem 3.1 reveals that since $B_{\psi(q)}(R_q)$ is T_q-invariant for any $q \in \mathbb{F}[X]$, we have $\mu(\mathcal{S}(\psi)) \in \{0,1\}$.

We now apply the same trick used in the one dimensional case. As before, we let I be the set of monic, irreducible polynomials and consider the set $\mathcal{S}_{\text{irr.}}(\psi)$, such that the desired inequality has infinitely many solutions $q \in I$. Again, we easily see that for $q_1, q_2 \in I$,

$$\mu\left(B_{\psi(q_1)}(R_{q_1}) \cap B_{\psi(q_2)}(R_{q_2})\right) \asymp \mu\left(B_{\psi(q_1)}(R_{q_1})\right) \mu\left(B_{\psi(q_2)}(R_{q_2})\right).$$

Divergence of the series $\sum_{q\in I} \psi(q)^n$ follows by the same argument as in the one dimensional case, using the strengthened assumption for this case. Using Proposition 3.4, we obtain

$$\begin{aligned} 0 < \sum_{\substack{q_1,q_2\in I \\ \|q_1\|,\|q_2\|\leq N}} \mu\left(B_{\psi(q_1)}(R_{q_1}) \cap B_{\psi(q_2)}(R_{q_2})\right) &\ll \sum_{\substack{q_1,q_2\in I \\ \|q_1\|,\|q_2\|\leq N}} (\psi(q_1)\psi(q_2))^n \\ &< c\Big(\sum_{\substack{q\in I \\ \|q\|\leq N}} \psi(q)^n\Big)^2 \asymp c\Big(\sum_{\substack{q\in I \\ \|q\|\leq N}} \mu\left(B_{\psi(q)}(R_q)\right)\Big)^2 \end{aligned}$$

for some $c > 0$. Hence, by the converse Borel–Cantelli Lemma used in the proof of Theorem 3.1,

$$\mu(\mathcal{S}(\psi)) \geq \limsup_{N\to\infty} \frac{\left(\sum_{\substack{q\in I \\ \|q\|\leq N}} \mu\left(B_{\psi(q)}(R_q)\right)\right)^2}{\sum_{\substack{q_1,q_2\in I \\ \|q_1\|,\|q_2\|\leq N}} \mu\left(B_{\psi(q_1)}(R_{q_1}) \cap B_{\psi(q_2)}(R_{q_2})\right)} > c^{-1} > 0.$$

Now, (2) follows for $m = 1$.

Now, assume that $m > 1$. The function T_q defined in the proof for $m = 1$ fails to be an automorphism, and the obvious automorphisms which may be constructed on the basis of T_q fail to be ergodic. Hence, another method is needed. Fortunately, the independence of events proved in Proposition 3.5 gives us the possibility of using a stronger version of the converse Borel–Cantelli Lemma (Lemma 3.14, II in [11]).

For any $r \in \mathbb{N}$, we define sets

$$S_r = \{(q_1, \dots, q_m) \in \mathbb{F}[X]^m : \gcd(q_1, \dots, q_m) = 1, \|q\|_\infty = k^r, q_m \text{ monic}\}.$$

Here, gcd denotes the greatest norm of the common divisors of the arguments, so we require the coordinates of each element in S_r to be relatively prime. We also define $P_N = \bigcup_{i=1}^N S_i$ and $P_\infty = \bigcup_{i=1}^\infty S_i$.

Let $N \in \mathbb{N}$ and let $q, q' \in P_N$ with $q \neq q'$. We claim that q and q' are linearly independent over $\mathcal{L}$. Indeed, assume that $\alpha q = \alpha' q'$ for some $\alpha, \alpha' \in \mathcal{L}$. Since $q, q' \in \mathbb{F}[X]^m$, there is no loss of generality in assuming that $\alpha, \alpha' \in \mathbb{F}[X]$ and that they are relative prime. Hence, α divides each coordinate of q' and vice versa. By the condition on the greatest common divisor, $\alpha, \alpha' \in \mathbb{F}$. Since q_m and q'_m are monic, this means that $\alpha = \alpha'$, so $q = q'$. This contradiction proves the claim.

We now count the elements in S_r. We recall that the Möbius function has the property that

$$\sum_{d|k} \mu(d) = \begin{cases} 1 & \text{for } k=1, \\ 0 & \text{otherwise.} \end{cases}$$

Using Lemma 3.3, we get

$$\begin{aligned} |S_r| = \sum_{\substack{\|q\|_\infty = k^r \\ \gcd(q_1,\dots,q_m)=1 \\ q_m \text{ monic}}} 1 &= \sum_{\substack{\|q\|_\infty = k^r \\ \gcd(q_1,\dots,q_m)=p \\ q_m \text{ monic}}} \sum_{d|p} \mu(d) = \sum_{d=0}^{r} \mu(k^d) \sum_{\substack{\|v\|_\infty = k^{r-d} \\ v_m \text{ monic}}} 1 \\ &= \sum_{d=0}^{r} \mu(k^d)(m-1)\frac{k-1}{k} k^{(r-d)(m-1)} \left(\sum_{i=1}^{r-d-1} k^i \right) \\ &= \sum_{d=0}^{r} \mu(k^d)(m-1)\frac{1}{k}\left(k^{m(r-d)} - k^{(r-d)(m-1)+1}\right). \end{aligned}$$

We think of this expression as a polynomial in k. The dominating term of the polynomial is $(m-1)k^{mr-1}$, so asymptotically, $|S_r| \sim (m-1)k^{mr-1}$ for r tending to infinity. Hence, using Lemma 3.3 and the fact that $\psi(q)$ depends solely on $\|q\|_\infty$,

$$\begin{aligned} \sum_{q \in P_\infty} \psi(q)^n = \sum_{r=1}^{\infty} \sum_{q \in S_r} \psi(q)^n \asymp \sum_{r=1}^{\infty} (m-1)k^{mr-1} \psi(k^r)^n = \\ \frac{m-1}{m(k-1)} \left(m \frac{k-1}{k} \sum_{r=1}^{\infty} k^{mr} \psi(k^r)^n \right) = \frac{m-1}{m(k-1)} \sum_{\substack{q \in \mathbb{F}[X]^m \\ \|q\|_\infty \geq k}} \psi(q)^n. \end{aligned}$$

Hence, if the series $\sum_{q \in \mathbb{F}[X]^m} \psi(q)^n$ diverges then the series $\sum_{q \in P_\infty} \psi(q)^n$ diverges. Since also,

$$\left\{ A \in \mathcal{L}^{mn} : A \in B_{\psi(q)}(R_q) \text{ for infinitely many } q \in P_\infty \right\} \subseteq \mathcal{S}(\psi),$$

the strong converse Borel–Cantelli Lemma (Lemma 3.14, II in [11]) implies the theorem for $m > 1$. □

Now, we will define the sets to be discussed in the remainder of Part 1 of this thesis. These are special cases of $\mathcal{S}(\psi)$, but they have the advantage of being technically easier to deal with.

Definition 3.1. Let $v > 0$. The set of matrices

$$\mathcal{S}_v = \left\{ A \in I^{mn} : |\langle qA \rangle| < \|q\|_\infty^{-v} \text{ for infinitely many } q \in \mathbb{F}[X]^m \right\}$$

is called *the set of v-approximable elements in I^{mn}*.

As a corollary to Theorem 3.2, we obtain the following:

Corollary 3.6.

$$\mu(\mathcal{S}_v) = \begin{cases} 1 & \text{for } v \leq \frac{m}{n}, \\ 0 & \text{for } v > \frac{m}{n}. \end{cases}$$

Proof. Let $\psi(q) = \|q\|_\infty^{-v}$. Consider first the case $m > 1$. By Lemma 3.3,

$$\sum_{q \in \mathbb{F}[X]^m} \|q\|_\infty^{-vn} = \sum_{r=0}^{\infty} \sum_{\substack{q \in \mathbb{F}[X]^m \\ \|q\|_\infty = k^r}} k^{-rnv} = m\frac{k-1}{k} \sum_{r=0}^{\infty} \left(k^{m-nv}\right)^r.$$

This series converges if and only if $v > \frac{m}{n}$. Now, the corollary follows in this case from Theorem 3.2.

Now, consider the case $m = 1$. In this case,

$$\sum_{q \in \mathbb{F}[X]} \frac{1}{\deg(q)} \|q\|_\infty^{-vn} = \frac{k-1}{k} \sum_{r=0}^{\infty} \frac{1}{r} \left(k^{1-nv}\right)^r.$$

This certainly converges for $v > \frac{1}{n}$ and diverges for $v < \frac{1}{n}$. For $v = \frac{1}{n}$, it degenerates to $\sum_{r=0}^{\infty} \frac{1}{r}$, which is also divergent. Hence, the corollary follows from Theorem 3.2. □

4. Well-approximable linear forms over $\mathcal{L}$

In this chapter, we calculate the Hausdorff dimension of the first obvious exceptional set resulting from Corollary 3.6. We will first find an upper bound on the Hausdorff dimension, using methods quite similar to the ones used to prove the easy part of Theorem 3.2. In particular, we need to prove a Hausdorff version of the Borel–Cantelli Lemma, which is valid for our definition of Hausdorff dimension in $\mathcal{L}$. The results of this chapter are to appear in [35].

Subsequently, we will use a method due to Dodson, Rynne and Vickers ([19]), which in turn uses a result due to Frostman ([20]) to calculate a lower bound on the Hausdorff dimension. This is by far the most difficult part of the chapter, and indeed of the thesis so far.

4.1. An upper bound on the Hausdorff dimension

The sets we are concerned with in this section are the sets $\mathcal{S}_v$ of v-approximable matrices from Definition 3.1. In fact, we will prove the following theorem:

Theorem 4.1. *Let $v \geq \frac{m}{n}$. Then*

$$\dim_{\mathrm{H}}(\mathcal{S}_v) = (m-1)n + \frac{m+n}{v+1}.$$

The one-dimensional real analogue of this theorem was originally proved by Jarník ([25]). Subsequently, Besicovitch found another proof ([8]) independent of Jarník. The real analogue of Theorem 4.1 is thus known as the general Jarník-Besicovitch Theorem.

We begin with the lemma, which will allow us to compute an upper bound on the Hausdorff dimension.

Lemma 4.2 (The Hausdorff–Cantelli–Laurent Lemma). *Let $E \subseteq \mathcal{L}^{mn}$ be some set and let $\mathcal{C} = (B_i) = B(c_i, \rho_i)$, $i \in \mathbb{N}$ be some sequence of $\|\cdot\|_\infty$-balls in $\mathcal{L}^{mn}$. Assume that*

$$E \subseteq \{A \in \mathcal{L}^{mn} : A \in B_i \text{ for infinitely many } i \in \mathbb{N}\}.$$

If for some $s \geq 0$*,*

$$\sum_{i=1}^{\infty} \rho_i^s < \infty,$$

then $\mathcal{H}^s(E) = 0$ *and* $\dim_{\mathrm{H}}(E) \leq s$*.*

The reader will note some similarity with the Borel–Cantelli Lemma, and indeed, this lemma takes the place of the Borel–Cantelli lemma in the proof of Theorem 3.2 and is the key to obtaining the upper bound. By expressing $\mathcal{S}_v$ as a limsup set of an appropriate cover, we can obtain the required upper bound. First, we prove the Hausdorff–Cantelli–Laurent Lemma.

Proof of Lemma 4.2. By assumption, E is contained in the limsup set of the B_i,

$$E \subseteq \bigcap_{N=1}^{\infty} \bigcup_{i=N}^{\infty} B_i.$$

Hence, for any $N \in \mathbb{N}$, $\mathcal{C}_N = \{B_i : i \geq N\}$ is a cover of E.

Let $\delta > 0$. Since $\sum_{i=1}^{\infty} \rho_i^s < \infty$, we have $\rho_i \to 0$ as $i \to \infty$. Hence, there exists an $N_0 = N_0(\delta) \in \mathbb{N}$ such that $\rho_i < \delta$ for $i \geq N_0$. Furthermore, for any $\varepsilon > 0$, there exists an $N_1 = N_1(\varepsilon) \in \mathbb{N}$ such that for $N \geq N_1$, $\sum_{i=N}^{\infty} \rho_i^s < \varepsilon$. But then, for any $N \geq N_1$:

$$\mathcal{H}_\delta^s(E) \leq l^s(\mathcal{C}_N) = \sum_{i=N}^{\infty} \rho_i^s < \varepsilon.$$

Letting δ tend to zero corresponds by the above to letting N_0 tend to infinity, so we can make the Hausdorff-δ-s-measure arbitrarily small. Hence, $\mathcal{H}^s(E) = 0$ and thus $\dim_{\mathrm{H}}(E) \leq s$. □

Lemma 4.3. *Let* $v \geq \frac{m}{n}$*. Then,*

$$\dim_{\mathrm{H}}(\mathcal{S}_v) \leq (m-1)n + \frac{m+n}{v+1}.$$

Proof. We consider the resonant sets R_q as in (3.8), and neighbourhoods of these, $B_v(R_q) = B_{\psi(q)}(R_q)$ for $\psi(q) = \|q\|_\infty^{-v}$, defined as in (3.9). The $B_v(R_q)$ cover $\mathcal{S}_v$.

We will now construct a cover $\mathcal{C}_q$ of each of the $B_v(R_q)$ with balls. As in the proof of Theorem 3.2 in the last chapter, we see that each resonant set is contained in a union of $(m-1)n$-dimensional affine spaces. Furthermore, since $A \in I^{mn}$, each resonant set R_q will be contained in the affine spaces

$$R_{q,r} = \{A \in I^{mn} | qA = r\} \text{ with } \|r\|_\infty < \|q\|_\infty.$$

There are precisely $\|q\|_\infty^n$ such r, so R_q is contained in this number of affine spaces $R_{q,r}$.

We proceed to cover each of these affine spaces with balls. First, we choose the centres of these at distances of integer multiples of $\|q\|_\infty^{-v-1}$ from each other. In this way, we choose

$$\left(\|q\|_\infty^{-v-1}\right)^{-\dim(R_{q,r})} = \|q\|_\infty^{(1+v)(m-1)n}$$

centres. Now, take balls with these centres and radii $2\|q\|_\infty^{-v-1}$. These define a cover $\mathcal{C}_q$ of $B_v(R_q)$. Calculating the s-length, we see that

$$l^s(\mathcal{C}_q) \leq 2^s \|q\|_\infty^{n+(1+v)(m-1)n-s(1+v)}.$$

Finally, we let $\mathcal{C}$ be the cover of $\mathcal{S}_v$ obtained by taking the union of all the $\mathcal{C}_q$ and calculate an upper bound on the s-length of $\mathcal{C}$ from this estimate and Lemma 3.3,

$$\begin{aligned} l^s(\mathcal{C}) &\ll \sum_{q\in\mathbb{F}[X]^m\setminus\{0\}} l^s(\mathcal{C}_q) \\ &\leq 2^s \sum_{q\in\mathbb{F}[X]^m\setminus\{0\}} \|q\|_\infty^{n+(1+v)(m-1)n-s(1+v)} \\ &= 2^s m\frac{k-1}{k}\sum_{r=0}^{\infty}\left(k^{m+n+(1+v)(m-1)n-s(1+v)}\right)^r, \end{aligned}$$

which converges for any $s > (m-1)n + \frac{m+n}{v+1}$. Now, the lemma follows directly from Lemma 4.2. □

4.2. A lower bound on the Hausdorff dimension

In this section, we will prove that the upper bound for the Hausdorff dimension of $\mathcal{S}_v$ given in Lemma 4.3 is optimal. That is, we will calculate a lower bound for the Hausdorff dimension, which is equal to the upper bound. In order to do this, we use a method due to Dodson, Rynne and Vickers ([19]), using the so-called ubiquitous systems.

We begin with some definitions. Let

$$\rho_N = \left[(1+\tfrac{m}{n})N\right], \tag{4.1}$$

where $[x]$ denotes the integral part of the real number x. We also define the sets

$$B(R_q;k^{-\rho_N}) = \left\{A \in I^{mn} : \mathrm{dist}_\infty(A,R_q) < k^{-\rho_N}\right\}, \tag{4.2}$$

where the R_q are the resonant sets defined in (3.8), and dist_∞ denotes the distance in the height-norm $\|\cdot\|_\infty$. Finally, we define the sets

$$A(N) = \bigcup_{\|q\|_\infty \leq k^N} B(R_q;k^{-\rho_N}). \tag{4.3}$$

This system of sets is an example of an ubiquitous system. We will not define ubiquity in full generality, but in informal terms, ubiquity means that we can find rational elements in $\mathcal{L}$ with "small" denominators "close" to any element in $\mathcal{L}$. Here, the meaning of the term "small" depends on the "closeness" required. Thus, the property contains somewhat more information than density of the rationals in $\mathcal{L}$. The following lemma formalises the above discussion in this particular case.

Lemma 4.4.

$$\lim_{N\to\infty} \mu(I^{mn} \setminus A(N)) = 0.$$

Proof. Let $\varepsilon > 0$ be fixed but arbitrary. By Corollary 3.6, we can choose $N \in \mathbb{N}$ so large that

$$\mu\left(I^{mn} \setminus \bigcup_{1\leq \|q\|_\infty \leq k^N} B_{\|q\|_\infty k^{-(1+m/n)N}}(R_q)\right) < \varepsilon. \tag{4.4}$$

This follows since there are infinitely many solutions to $|\langle qA\rangle| < \|q\|_\infty^{-m/n}$ for almost all $A \in I^{mn}$. Using this, we may fill out the unit ball with sets as above. However, for $\|q\|_\infty \leq k^N$ we also have

$$B_{\|q\|_\infty k^{-(1+m/n)N}}(R_q) \subseteq B(R_q; k^{-(1+m/n)N}) \subseteq B(R_q; k^{-\rho_N}). \tag{4.5}$$

Indeed, let $A \in B_{\|q\|_\infty^{-m/n}}(R_q)$ and $p \in \mathbb{F}[X]^n$ be such that

$$\|qA - p\|_\infty < \|q\|_\infty k^{-(1+m/n)N}$$

Then,

$$\|q\|_\infty \operatorname{dist}_\infty(A, R_q) \leq \inf_{A' \in R_q} \left\|qA - qA'\right\|_\infty \leq \|qA - p\|_\infty < \|q\|_\infty k^{-(1+m/n)N}.$$

Dividing by $\|q\|_\infty$ reveals that $A \in B(R_q; k^{-(1+m/n)N})$. The second inclusion is trivial, since $k^{-(1+m/n)N} \leq k^{-\rho_N}$.

Hence by (4.5),

$$I^{mn} \setminus A(N) \subseteq I^{mn} \setminus \bigcup_{1\leq \|q\|_\infty \leq k^N} B_{\|q\|_\infty^{-m/n}}(R_q).$$

Since ε was arbitrary, the result follows from (4.4). □

We will now use this property to calculate the Hausdorff dimension of a certain subset of $\mathcal{S}_v$. Once again, we consider an arbitrary approximation function. Hence, we take an arbitrary decreasing function $\psi : \mathbb{R}^+ \to \mathbb{R}^+$, such that $\psi(k^N) \leq k^{-\rho_N}$. We define the set

$$\Lambda(\psi) = \left\{A \in I^{mn} : \operatorname{dist}_\infty(A, R_q) < \psi(\|q\|_\infty) \text{ for infinitely many } q \in \mathbb{F}[X]^m\right\}.$$

Also, we define

$$\gamma = \limsup_{N\to\infty} \frac{-\rho_N \log k}{\log \psi(k^N)} \geq 0.$$

We will find a lower bound on the Hausdorff dimension of $\Lambda(\psi)$. This will enable us to compute a lower bound on the Hausdorff dimension of $\mathcal{S}(\psi')$ for appropriate functions $\psi' : \mathbb{F}[X]^m \to \mathbb{R}^+$. In particular, for a clever choice of ψ, we will have that $\Lambda(\psi) \subseteq \mathcal{S}_v$ and thus obtain the required lower bound on the set under examination.

In order to calculate a lower bound on $\dim_{\mathrm{H}}(\Lambda(\psi))$, we will use a version of the easy half of Frostman's Lemma, adapted to the field of Laurent series.

Lemma 4.5 (Frostman's Lemma, Easy Half). *Let $E \subseteq \mathcal{L}^{mn}$ be a Borel set and let $s \geq 0$. If there exists a probability measure ν with support $\mathrm{supp}(\nu) \subseteq E$, which has the property that for any ball $B = B(c,\rho) \subseteq \mathcal{L}^{mn}$, $\nu(B) \ll \rho^s$, then $\mathcal{H}^s(E) > 0$.*

Proof. Assume that ν is a measure as in the statement of the lemma. That is, for any ball $B \subseteq \mathcal{L}^{mn}$ we have $\nu(B) \leq K\rho(B)^s$ for some $K > 0$. Let $\mathcal{C}_\delta$ be a cover of E with balls of radius less than some arbitrary but fixed $\delta > 0$.

$$0 < \nu(E) \leq \nu\left(\bigcup_{B\in\mathcal{C}_\delta} B\right) \leq \sum_{B\in\mathcal{C}_\delta} \nu(B) \leq K \sum_{B\in\mathcal{C}_\delta} \rho(B)^s.$$

Taking the infimum over such covers and letting δ tend to zero yields,

$$0 < \lim_{\delta\to 0} \inf_{\text{covers } \mathcal{C}_\delta} K \sum_{B\in\mathcal{C}_\delta} \rho(B)^s = K\mathcal{H}^s(E).$$

This completes the proof. ☐

Remark. Note, that it suffices to prove that there is a probability measure with support on E such that $\nu(B) \ll \rho(B)^s$ for all sufficiently small balls B. This is the property, we will use.

Lemma 4.6.

$$\dim_{\mathrm{H}}(\Lambda(\psi)) \geq (m-1)n + \gamma n.$$

Proof. This proof falls into four different parts. First, we define some partitions and families of sets in I^{mn}. Secondly, we construct a Cantor set in I^{mn} based on these families and partitions. Thirdly, we construct a measure on I^{mn} with support on our Cantor set. Finally, we prove that this measure is sufficiently nice to allow us to apply Frostman's Lemma.

Step 1: We begin with the construction of the partitions and families. For this, we need to define a family of lattices in $\mathcal{L}^{mn}$. Namely, we define for any $N \in \mathbb{N}$ the lattice $\Gamma(N) = X^{-N}\mathbb{F}[X]^{mn}$. It is a straightforward matter to prove that for $x, y \in \Gamma(N)$, $\|x-y\|_\infty \geq k^{-N}$.

For each $N \in \mathbb{N}$, we define the partition $\mathcal{H}_N$ of I^{mn} to be the family of $\|\cdot\|_\infty$-balls $B(c,k^{-\rho_N})$, where $c \in \Gamma(\rho_N)$. By the Proposition 1.5, these balls are disjoint. By Proposition 1.7, we can calculate the Haar measure of these balls. In our normalisation, each has measure $k^{-\rho_N mn}$. By counting, we see that there are exactly $k^{\rho_N mn}$ of these balls, so since the balls are disjoint, they do indeed define a partition of I^{mn}.

For any ball $B = B(c,\rho)$ and any positive number α, we let αB denote the ball $B(c,\alpha\rho)$. Now, we define the family of bad balls in $\mathcal{H}_N$,

$$\mathcal{E}_N = \left\{H \in \mathcal{H}_N \,|\, \tfrac{1}{k}H \cap A(N) = \emptyset\right\}.$$

These are bad, because they still have not been completely caught by the ubiquitous system $A(N)$. The rest of the balls have been caught, so we define the family of good balls, $\mathcal{G}_N = \mathcal{H}_N \setminus \mathcal{E}_N$.

We see that $\mathcal{E}_N \subseteq I^{mn} \setminus A(N)$, so by Lemma 4.4, $|\mathcal{E}_N| \cdot k^{-\rho_N mn} = \mu(\mathcal{E}_N) \to 0$ as $N \to \infty$. Hence, again by Lemma 4.4,

$$|\mathcal{G}_N| = |\mathcal{H}_N| - |\mathcal{E}_N| \asymp k^{mn\rho_N}$$

for N large enough. For any subset $X \subseteq I^{mn}$, we let $\mathcal{G}_N(X)$ be the family of good balls in X as induced by $\mathcal{G}_N$.

Let $H \in \mathcal{G}_N$ and take a $q \in \mathbb{F}[X]^m$ such that $\|q\|_\infty \le k^N$ and $\frac{1}{k}H \cap B(R_q;k^{-\rho_N}) \neq \emptyset$. We construct a family $\mathcal{D}(H)$ of subsets of H as follows:

1. First, we choose points c_i in R_q such that for $i \neq j$,
$$\left\|c_i - c_j\right\|_\infty \ge k^{mn}\psi\left(k^N\right).$$
2. Choose balls D_i with centres c_i and radii $\frac{\psi(k^N)}{k}$.
3. Remove all points belonging to R_q from the sets obtained in the above fashion. This is to ensure, that we use different $q \in \mathbb{F}[X]^m$ at each step of the construction of the Cantor set.

Note, that for each $u \in D \in \mathcal{D}(H)$, we have

$$0 < \mathrm{dist}_\infty(u,R_q) \le \frac{\psi\left(k^N\right)}{k} < \psi(\|q\|_\infty). \tag{4.6}$$

We proceed with counting the number of sets in $\mathcal{D}(H)$. By the Proposition 1.5, we can find an $A' \in R_q$ such that $\frac{1}{k}H \subseteq B(A',k^{-\rho_N})$. Let $R_{q,r}$ be the $(m-1)n$-dimensional affine space from R_q containing A'. The intersection between $B(A',k^{-\rho_N})$ and this plane is isometric with a ball in $I^{(m-1)n}$ with measure $\left(k^{-\rho_N+1}\right)^{(m-1)n}$. It is not difficult to count the maximal number of elements herein with distances as (1) in the above construction.

Indeed, by translation invariance of the Haar measure, it suffices to consider the ball in $\mathcal{L}^{(m-1)n}$ with centre 0 and radius $k^{-\rho_N}$. Counting points in this ball with distances greater than or equal to $k^{mn}\psi\left(k^N\right)$ corresponds to counting points of a certain lattice in the ball. Let $N' \in \mathbb{Z}$ be maximal such that $\psi\left(k^N\right) \geq k^{N'}$. We scale the whole setting by X^{ρ_N}, and now we note that $|\mathcal{D}(H)|$ is approximately equal to the number of elements in

$$X^{mn+N'+\rho_N}\mathbb{F}[X]^{(m-1)n} \bigcap I^{(m-1)n}.$$

But this is essentially the same calculation as the one performed in the beginning of the proof, so

$$|\mathcal{D}(H)| \asymp k^{(-mn-N'-\rho_N)(m-1)n} \asymp \left(\frac{k^{-\rho_N}}{\psi(k^N)}\right)^{(m-1)n}.$$

Finally, we define the sets

$$T_N = \bigcup_{H\in\mathcal{G}_N} \bigcup_{D\in\mathcal{D}(H)} D,$$

and the numbers

$$t_N = \sum_{H\in\mathcal{G}_N} \sum_{D\in\mathcal{D}(H)} 1 \asymp k^{mn\rho_N} \left(\frac{k^{-\rho_N}}{\psi(k^N)}\right)^{(m-1)n} = \frac{k^{n\rho_N}}{\psi(k^N)^{(m-1)n}}.$$

To complete step one, we need an additional technical lemma.

Lemma 4.7. *Let $X \in I^{mn}$ be a $\|\cdot\|_\infty$-ball. There exists an $N_0 \in \mathbb{N}$ such that for any $N \geq N_0$,*

$$\mu(X\cap T_N) \asymp t_N\mu(X).$$

Proof of Lemma 4.7. Let $\mathcal{G}_X(N) = \{H \in \mathcal{G}_N : H \subseteq X\}$ and $X(N) = \cup_{H\in\mathcal{G}_X(N)}H$. It is immediate that $\mu(X\setminus X(N)) \to 0$ as $N \to \infty$. Hence,

$$\mu(X\cap T_N) \gg \sum_{H\in\mathcal{G}_X(N)} \sum_{D\in\mathcal{D}(H)} \mu(D) \gg t_N\mu(X)$$

for N large enough.

For the converse inequality, we let $\mathcal{G}'_X(N) = \{H \in \mathcal{G}_N : H\cap X \neq \emptyset\}$ and define $X'(N) = \cup_{H\in\mathcal{G}'_X(N)}H$. Again, $\mu(X\setminus X'(N)) \to 0$ as $N \to \infty$. Hence,

$$\mu(X\cap T_N) \ll \sum_{H\in\mathcal{G}'_X(N)} \sum_{D\in\mathcal{D}(H)} \mu(D) \ll t_N\mu(X)$$

for N large enough. □

Step 2: Now, we construct the Cantor set. First, we choose an increasing sequence $(M_r) \subseteq \mathbb{N}$ such that

$$\gamma = \lim_{r\to\infty} \frac{-\rho_{M_r} \log k}{\log \psi(k^{M_r})},$$

and an $N_1 \in (M_r)$ such that Lemma 4.7 holds for $X = I^{mn}$. Let

$$T_{N_1} = \bigcup_{H\in\mathcal{G}_{N_1}} \bigcup_{D\in\mathcal{D}(H)} D, \qquad t_{N_1} \asymp \frac{k^{n\rho_{N_1}}}{\psi(k^{N_1})^{(m-1)n}}.$$

This completes the first step in the construction of the Cantor set.

Let $N_2 \in (M_r)$ be such that $N_2 > N_1$, such that Lemma 4.7 holds for every set in the family T_{N_1} — that is, every $X = D \in T_{N_1}$ — and such that $|\mathcal{G}_{N_2}(D)| \asymp k^{-mn\rho_{N_2}}\mu(D)$. Let

$$T_{N_2} = \bigcup_{H\in\mathcal{G}_{N_2}(T_{N_1})} \bigcup_{D\in\mathcal{D}(H)} D.$$

The choice of N_2 immediately implies that

$$t_{N_2} \asymp \frac{k^{n\rho_{N_2}}}{\psi(k^{N_2})^{(m-1)n}} t_{N_1}\mu(D) \asymp \frac{k^{n\rho_{N_2}}\psi\left(k^{N_1}\right)^{mn}}{\psi(k^{N_2})^{(m-1)n}} t_{N_1}.$$

We repeat this construction to obtain

$$\#\mathcal{G}_{N_r}(D) \gg k^{-mn\rho_{N_r}}\psi\left(k^{N_{r-1}}\right)^{mn} \quad \text{for } D \in T_{N_{r-1}}, \tag{4.7a}$$

$$\psi(k^{N_r}) \leq \prod_{i=1}^{r-1}\left(\psi\left(k^{N_i}\right)k^{\rho_{N_i}}\right)^{rn}. \tag{4.7b}$$

Also, we define

$$T_{N_r} = \bigcup_{H\in\mathcal{G}_{N_r}(T_{N_{r-1}})} \bigcup_{D\in\mathcal{D}(H)} D, \tag{4.7c}$$

$$t_{N_r} = \sum_{H\in\mathcal{G}_{N_r}(T_{N_{r-1}})} \sum_{D\in\mathcal{D}(H)} 1 \asymp k^{-mn\rho_{N_r}} \left(\frac{k^{-\rho_{N_r}}}{\psi(k^{N_r})}\right)^{(m-1)n} t_{N_{r-1}}\psi\left(k^{N_{r-1}}\right)^{mn}. \tag{4.7d}$$

Calculating backwards, since by (4.7d) for any $i \in \mathbb{N}$,

$$t_{N_i}\psi\left(k^{N_i}\right)^{mn} \asymp \left(\frac{k^{-\rho_{N_i}}}{\psi(k^{N_i})}\right)^{-n} t_{N_{i-1}}\psi\left(k^{N_{i-1}}\right)^{mn},$$

we get

$$t_{N_r} \asymp \prod_{i=1}^{r}\left(\frac{k^{-\rho_{N_i}}}{\psi(k^{N_i})}\right)^{-n}\psi\left(k^{N_r}\right)^{-mn}. \tag{4.8}$$

We define our Cantor set to be $T_\infty = \cap_{i=1}^{\infty} T_{N_i}$. By (4.6), $T_\infty \subseteq \Lambda(\psi)$.

Step 3: We now construct a probability measure on I^{mn} with support on T_∞. For any $r \in \mathbb{N}$ and any $D \in T_{N_r}$, we define the auxiliary measure $\nu_r(D) = \frac{1}{t_{N_r}}$, and extend this measure in the natural way to the Borel σ-algebra of I^{mn} (that is, letting it be arbitrary with legal values on subsets of each D and zero on subsets of the complement to T_{N_r}).

This gives us a sequence of probability measures (ν_r) on I^{mn} which correspond to linear functionals on I^{mn}. Hence, we obtain a sequence in the unit ball of the dual space of I^{mn}. This sequence is inside the unit ball in $(I^{mn})^*$, since the measures are probability measures. By Alaoglu's Theorem (Theorem 2.5.2 in [48]), the unit ball in the dual space is weakly compact, so this sequence has a weak*-limit point ν, which itself corresponds to a probability measure. Clearly, this measure has its support on T_∞.

Step 4: The final part of the proof consists in proving that for any $\eta > 0$ and any sufficiently small $\|\cdot\|_\infty$-ball $C \subseteq I^{mn}$ with radius $\rho(C)$, we have $\nu(C) \ll \rho(C)^s$, where $s = (m-1)n + \gamma n - \eta$. By (4.7b) and (4.8), we can calculate $\nu_r(D)$ for $D \in T_{N_r}$:

$$\begin{aligned}
\nu_r(D) = \frac{1}{t_{N_r}} &\asymp \prod_{i=1}^{r} \left(\frac{k^{-\rho_{N_i}}}{\psi\left(k^{N_i}\right)} \right)^n \psi\left(k^{N_r}\right)^{mn} \\
&\asymp k^{-\rho_{N_r} n} \psi\left(k^{N_r}\right)^{-n} \psi\left(k^{N_r}\right)^{nm} \prod_{i=1}^{r-1} \left(\frac{k^{-\rho_{N_i}}}{\psi\left(k^{N_i}\right)} \right)^n \\
&\ll k^{-\rho_{N_r} n} \psi\left(k^{N_r}\right)^{(m-1)n - \frac{1}{r}} \\
&\ll \psi\left(k^{N_r}\right)^{s_r},
\end{aligned}$$

where

$$s_r = \frac{-n\rho_{N_r} \log k + ((m-1)n - \frac{1}{r}) \log \psi\left(k^{N_r}\right)}{\log \psi\left(k^{N_r}\right)} = (m-1)n - \frac{1}{r} + n \frac{-\rho_{N_r} \log k}{\log \psi\left(k^{N_r}\right)}.$$

By choice of N_r, we see that $\lim_{N \to \infty} s_r = (m-1)n + n\gamma$. Hence, we can choose r such that

$$\nu(D) \ll \psi\left(k^{N_r}\right)^s \qquad \text{for any } D \in T_{N_r}. \tag{4.9}$$

Now, let C be a $\|\cdot\|_\infty$-ball. To calculate the ν-measure of C, we must count the number of $D \in T_{N_r}$ such that $D \subseteq C$ for r large enough. We do this in the same way we estimated t_N initially. For $k^{-\rho_{N_r}} < \rho(C)$, we take the maximal number of good balls H in C and estimate the number $|\mathcal{D}(H)|$ to get the total:

$$\nu(C) \ll \left(\frac{\rho(C)}{k^{-\rho_{N_r}}} \right)^{mn} \left(\frac{k^{-\rho_{N_r}}}{\psi\left(k^{N_r}\right)} \right)^{(m-1)n} \nu(D).$$

Otherwise, we only need to be concerned with the maximal number of $\mathcal{D}$-sets inside C, so

$$\nu(C) \ll \left(\frac{\rho(C)}{\psi\left(k^{N_r}\right)} \right)^{(m-1)n} \nu(D).$$

Now, let C be so small that the r for which $\psi\left(k^{N_r}\right) < \rho(C) \leq \psi\left(k^{N_{r-1}}\right)$ is so large that (4.9) holds for r and $r-1$. In the first case,

$$\nu(C) \ll \frac{t_{N_r}}{t_{N_{r-1}}}\left(\frac{\rho(C)}{\psi(k^{N_{r-1}})}\right)^{mn}\nu(D) \ll \frac{1}{t_{N_{r-1}}}\left(\frac{\rho(C)}{\psi(k^{N_{r-1}})}\right)^{mn}$$
$$\ll \rho(C)^{mn}\psi\left(k^{N_{r-1}}\right)^{s-mn} \ll \rho(C)^s$$

by (4.9). In the second case (also by (4.9)),

$$\nu(C) \ll \rho(C)^{(m-1)n}\psi\left(k^{N_r}\right)^{s-(m-1)n} \ll \rho(C)^s.$$

Hence, Frostman's Lemma (Lemma 4.5) implies

$$\dim_{\mathrm{H}}\Lambda(\psi) \geq s = (m-1)n + \gamma n - \eta.$$

Since $\eta > 0$ was arbitrary, this completes the proof. □

This lemma allows us to find a lower bound on the Hausdorff dimension of $\mathcal{S}_v$. We complete the proof of Theorem 4.1.

Proof of Theorem 4.1. We let $\psi(N) = N^{-v-1}$. It is easy to see that

$$B(R_q;\psi(\|q\|_\infty)) \subseteq B_{\|q\|_\infty^{-v}}(R_q). \tag{4.10}$$

Indeed, let $A' \in R_q$ be such that

$$\left\|A - A'\right\|_\infty < \|q\|_\infty^{-v-1}.$$

Then for some $p \in \mathbb{F}[X]^n$ we have by Proposition 1.6,

$$\|q\|_\infty^{-v} > \|q\|_\infty \left\|A - A'\right\|_\infty \geq \left\|qA - qA'\right\|_\infty = \|qA - p\|_\infty \geq |\langle qA\rangle|,$$

so $A \in B_{\|q\|_\infty^{-v}}(R_q)$. Hence, $\Lambda(\psi) \subseteq \mathcal{S}_v$. We now calculate a lower bound for the Hausdorff dimension of $\Lambda(\psi)$.

We see, that

$$\limsup_{N\to\infty} \frac{-\rho_N \log k}{\log\left(\left(k^N\right)^{-v-1}\right)} = \limsup_{N\to\infty} \frac{-\left[(1+\frac{m}{n})N\right]\log k}{-(v+1)N\log k} = \left(1+\frac{m}{n}\right)\frac{1}{v+1}.$$

It follows from Lemma 4.6 that the lower bound is the right one. The upper bound was calculated in Lemma 4.3. This completes the proof. □

Remark. Theorem 4.1 implies the classical Jarník–Besicovitch Theorem (Theorem 1.3) for $\mathcal{L}$, when we let $m = n = 1$. This theorem was announced with a sketched proof in [33].

5. Badly approximable linear forms over $\mathcal{L}$

In this chapter, we will be considering a notion complementary (in a certain informal sense) to the well-approximable linear forms, discussed in Chapter 4 – namely the set of badly approximable linear forms.

We will first define this set and prove that it's Haar measure is zero. Also, we will prove some auxiliary results, regarding the geometry of numbers in $\mathcal{L}$. Secondly, we will discuss the so-called (α,β)-games, which is the tool that allows us to calculate the Hausdorff dimension of the set. Thirdly, we will prove that the set of badly approximable linear forms has non-zero winning dimension. In the final section, we prove that this implies that the Hausdorff dimension of the set of badly approximable linear forms is full.

The one-dimensional analogue of the main theorem about the Hausdorff dimension of the set of badly approximable real numbers was proved by Jarník in [24]. The methods used here to calculate the Hausdorff dimension were developed by Schmidt in [55] to prove the analogous real one-dimensional result along with other results in number theory. Subsequently, he used the method to prove the real analogue of the main theorem of this chapter in [56]. The results of this chapter are to appear in [34].

5.1. Definitions and preliminaries

Well-approximable linear forms are the matrices that infinitely often send points in the integer lattice in their domain to points "close" to the integer lattice in the image of the matrices. A converse notion to this one would be the set of linear forms such that all points in the integer lattice of the domain have images bounded away from the integer lattice in the image, for some appropriate notion of being bounded away. In precise terms, the set we will be interested in in this chapter is the following.

Definition 5.1. The set of matrices,

$$\mathfrak{B}(m,n) = \left\{ A \in \mathcal{L}^{mn} : \exists K > 0, \forall q \in \mathbb{F}[X]^m, |\langle qA \rangle|^n > \frac{K}{\|q\|_\infty^m} \right\},$$

is called *the set of badly approximable elements in $\mathcal{L}^{mn}$*.

We begin with a proof of the following consequence of Corollary 3.6:

Theorem 5.1.

$$\mu(\mathfrak{B}(m,n)) = 0.$$

Proof. Clearly, $\mathfrak{B}(m,n)$ is contained in the complementary set of

$$\mathcal{S} = \left\{ A \in \mathcal{L}^{mn} : |\langle qA \rangle| < \|q\|_\infty^{-m/n} \text{ for infinitely many } q \in \mathbb{F}[X]^m \right\}.$$

Hence, it is sufficient to prove, that $\mu(\mathcal{S}) = \infty$. But this follows from Corollary 3.6, since $\mathcal{L}^{mn}$ can be written as the disjoint union of translates of I. ☐

With Theorem 5.1 in place, it is once again natural to enquire about the Hausdorff dimension of the set $\mathfrak{B}(m,n)$. We will prove the following:

Theorem 5.2.

$$\dim_{\mathrm{H}}(\mathfrak{B}(m,n)) = mn.$$

The proof of this is quite lengthy, and uses a variety of different methods. One of these is an extension of parts of the theory known as the geometry of numbers to $\mathcal{L}$. For an extensive treatment on the geometry of numbers over the reals, see [14]. The corresponding theory over $\mathcal{L}$ has been extensively developed by Mahler in [42]. Here, we provide simple proofs of the results we will need in order to obtain the Hausdorff dimension of $\mathfrak{B}(m,n)$.

We consider the possible integer solutions to some inequalities defined by slightly different matrices. Namely, we define for any $A \in \mathcal{L}^{mn}$ the matrices

$$\widetilde{A} = \begin{pmatrix} A & I_m \\ I_n & 0 \end{pmatrix}, \qquad \widetilde{A^*} = \begin{pmatrix} A^T & I_n \\ I_m & 0 \end{pmatrix},$$

where I_m and I_n denotes the $m \times m$ and $n \times n$ identity matrices respectively. Let $A^{(j)}$ denote the j'th column of the matrix A. We note, that $A \in \mathfrak{B}(m,n)$ if and only if there exists a $K > 0$ such that

$$\max_{1 \le j \le n} \left(\left\| q \cdot \widetilde{A}^{(j)} \right\| \right)^n > \frac{K}{\max_{1 \le i \le m} (\|q_i\|)^m} \tag{5.1}$$

for any $q \in \mathbb{F}[X]^{m+n}$ with $q_1, \dots, q_m \neq 0$.

These matrix inequalities enable us to examine the situation in terms of parallelepipeds in $\mathcal{L}^{m+n}$. That is, sets of $x \in \mathcal{L}^{m+n}$ defined by inequalities

$$\|(xA)_i\| < c_i, \qquad A \in \mathcal{L}^{(m+n)^2} \text{ invertible}, c_i > 0, i = 1, \dots, m+n. \tag{5.2}$$

Equivalently, we can define these sets in terms of distance functions

$$F_A(x) := \max_{1 \le j \le m+n} \frac{1}{c_j} \left\| \sum_{i=1}^{m+n} x_i a_{ij} \right\| < 1. \tag{5.3}$$

We define for any $\lambda > 0$, the sets

$$P_A(\lambda) = \left\{x \in \mathcal{L}^{m+n} : F_A(x) < \lambda\right\}.$$

Clearly, $P_A(1)$ is the set defined by (5.2). Also, for $\lambda' \leq \lambda$, $P_A(\lambda') \subseteq P_A(\lambda)$.

Definition 5.2. Let $A \in \mathcal{L}^{(m+n)^2}$. We define *the j'th successive minimum λ_j of F_A* as

$$\lambda_j = \inf\left\{\lambda > 0 : P_A(\lambda) \text{ contains } j \text{ linearly independent } a_1, \dots, a_j \in \mathbb{F}[X]^{m+n}\right\}.$$

First, we prove that if λ_i is the i'th successive minimum of F_A then there exists an $a_i \in \mathbb{F}[X]^{m+n}$ such that $F_A(a_i) = \lambda_i$. Indeed, assume to the contrary that for any $\varepsilon > 0$, $\{x \in \mathcal{L}^{m+n} : F_A(x) < \lambda_i - \varepsilon\}$ does not contain i linearly independent points from $\mathbb{F}[X]^{m+n}$, but that $\{x \in \mathcal{L}^{m+n} : F_A(x) < \lambda_i\}$ does contain such points. By the pigeon hole principle, for any $\varepsilon > 0$, there exists an $a_i \in \mathbb{F}[X]^{m+n}$ such that

$$a_i \in \left\{x \in \mathcal{L}^{m+n} : \lambda_i - \varepsilon \leq F_A(x) < \lambda_i\right\}.$$

Letting ε tend to zero proves the claim.

We will need some bounds on these minima. Clearly,

$$\lambda_{m+n} \gg 1. \tag{5.4}$$

Also, we have

Lemma 5.3. *For any invertible $A \in \mathcal{L}^{(m+n)^2}$,*

$$\lambda_1 \cdots \lambda_{m+n} \ll 1.$$

Proof. Let $e_i \in \mathcal{L}^{m+n}$, $i = 1, \dots, m+n$ denote the i'th unit vector. There is a $K > 0$ such that

$$\max_{1 \leq i \leq m+n} F_A(e_i) \leq K^{\frac{1}{m+n}}.$$

Hence,

$$\lambda_1 \cdots \lambda_{m+n} \leq \lambda_{m+n}^{m+n} \leq \left(\max_{1 \leq i \leq m+n} F_A(e_i)\right)^{m+n} \leq K.$$

□

Lemma 5.4. *Let $A \in \mathcal{L}^{(m+n)^2}$ be invertible and let $\lambda_1, \dots, \lambda_{m+n}$ denote the successive minima of F_A and let $\sigma_1, \dots, \sigma_{m+n}$ denote the successive minima of the function F_A^* defined by*

$$F_A^*(y) = \sup_{x \neq 0} \frac{\|x \cdot y\|}{F_A(x)}.$$

Then,

$$\lambda_m \sigma_{n+1} \gg 1.$$

Proof. We choose $a_i, a'_j \in \mathbb{F}[X]^{m+n}$ such that $F_A(a_i) = \lambda_i$ and $F^*_A(a'_j) = \sigma_j$. By definition of F^*_A,

$$F_A(x)F^*_A(x') \geq \left\|x \cdot x'\right\| \quad \text{for any } x, x' \in \mathcal{L}^{m+n}.$$

Hence, for any $i, j \in \{1, \ldots, m+n\}$,

$$\lambda_i \sigma_j \geq \left\|a_i \cdot a'_j\right\|.$$

We will prove that $\left\|a_m \cdot a'_{n+1}\right\| \neq 0$.

We consider the set

$$S = \left\{x \in \mathcal{L}^{m+n} : x \cdot a'_j = 0, j = 1, \ldots, n+1\right\}.$$

Clearly by the rank equation, this is an $(m-1)$-dimensional subspace, so by the pigeon hole principle, there is an a_i, $i \in \{1, \ldots, m\}$ such that $a_i \notin S$. Hence, for appropriate $i \leq m$, $j \leq n+1$, we have $\left\|a_i \cdot a'_j\right\| \neq 0$. Now, $\lambda_m \geq \lambda_i$ and $\sigma_{n+1} \geq \sigma_j$, so

$$\lambda_m \sigma_{n+1} \geq \lambda_i \sigma_j \geq \left\|a_i \cdot a_j\right\| > 0.$$

Finally, since $a_i, a'_j \in \mathbb{F}[X]^{m+n}$, $\left\|a_i \cdot a'_j\right\| \geq 1$. □

5.2. (α, β)-games

A useful tool for proving that a set has full Hausdorff dimension is the (α, β)-games. We will define these in quite an abstract setting. Let $\Omega = \mathcal{L}^{mn} \times \mathbb{R}^+$. We call Ω *the space of balls in* $\mathcal{L}$, where $\omega = (c, \rho) \in \Omega$ is said to have *centre* c and *radius* ρ. We define the map ϕ from Ω to the subsets of $\mathcal{L}^{mn}$, assigning a closed $\|\cdot\|_\infty$-ball to the abstract one defined above. That is, for $\omega = (c, \rho) \in \Omega$,

$$\phi(\omega) = \{x \in \mathcal{L}^{mn} : \|x - c\|_\infty \leq \rho\}.$$

Remark. Note, that we are now considering closed balls as opposed to the preceding chapters, where all the balls, we considered, were open. The reason for this becomes apparent in a moment. It should not cause much confusion.

Definition 5.3. Let $B_1, B_2 \in \Omega$. We say that $B_1 \subseteq B_2$ if $\rho_1 + d(c_1, c_2) \leq \rho_2$.

Also, we define for every $\gamma \in (0,1)$ and $B \in \Omega$:

$$B^\gamma = \left\{B' \subseteq B : \rho(B') = \gamma \rho(B)\right\}.$$

We can now define the following game:

Definition 5.4. Let $S \subseteq \mathcal{L}^{mn}$, and let $\alpha, \beta \in (0,1)$. Let Black and White be two players. The $(\alpha, \beta; S)$-*game* is played as follows:

- Black chooses a ball $B_1 \in \Omega$.
- White chooses a ball $W_1 \in B_1^{\alpha} \subseteq \Omega$.
- Black chooses a ball $B_2 \in W_1^{\beta} \subseteq \Omega$.
- And so on ad infinitum.

In the end, we let $B_i^* = \phi(B_i)$ and $W_i^* = \phi(W_i)$. If $\bigcap_{i=1}^{\infty} B_i^* = \bigcap_{i=1}^{\infty} W_i^* \subseteq S$, then White wins the game. Otherwise Black wins the game.

Our game can be understood in the following way: Initially, Black chooses a closed ball with radius ρ_1. Then, White chooses a ball inside the first one with radius $\alpha\rho_1$. Now, Black chooses a ball inside the one chosen by White with radius $\beta\alpha\rho_1$, and so on. In the end, the intersection of these balls will be non-empty by a simple corollary of Baire's Category Theorem. White wins the game if this intersection is a subset of S. Otherwise Black wins.

Note that because of the somewhat counter-intuitive metric on $\mathcal{L}^{mn}$, we might have a situation where for instance $(c,\rho) \in \Omega$ and $(c',\alpha\rho) \in \Omega$ maps to the same ball in $\mathcal{L}^{mn}$ under the map ϕ. Therefore, we need to keep track of the formal radii throughout the following sections. Hence the distinction between Ω and the space of balls in $\mathcal{L}^{mn}$.

We will first be looking at the $(\alpha,\beta;S)$-game from the point of view of White player. We need to define strategies:

Definition 5.5. Let $\alpha \in (0,1)$ and define for any $n \in \mathbb{N}$:

$$F_n^{\alpha} = \{f : \Omega^n \to \Omega | f(B_1,\dots,B_n) \in B_n^{\alpha}\}.$$

A sequence $(f_n)_{n\in\mathbb{N}}$ where $f_i \in F_n^{\alpha}$ is said to be a *strategy*. A strategy (f_n) is said to be $(\alpha,\beta;S)$-*winning* if for any set of balls $B_1, B_2, \dots, W_1, W_2, \dots$ where

$$B_n \in W_{n-1}^{\beta}, \qquad n = 2,3,\dots \tag{5.5a}$$

$$W_n = f_n(B_1,\dots,B_n), \quad n = 1,2,\dots \tag{5.5b}$$

we have that $\bigcap_{i=1}^{\infty} B_n^* \subseteq S$.

The sets of particular interest to us, are sets S such that White can always win the $(\alpha,\beta;S)$-game.

Definition 5.6. A set $S \subseteq \mathcal{L}$ is said to be (α,β)-*winning* if White can always win an $(\alpha,\beta;S)$-game, or equivalently, if there exists an $(\alpha,\beta;S)$-winning strategy. S is said to be α-*winning* if S is (α,β)-winning for any $\beta \in (0,1)$.

It is a fairly straightforward matter to see that if S is α-winning for some α and $\alpha' \in (0,\alpha)$, then S is α'-winning.

Proposition 5.5. *Let $S \subseteq \mathcal{L}^{mn}$ be α-winning for some $\alpha \in (0,1)$. Let $\alpha' \in (0,\alpha)$. Then S is α'-winning.*

Proof. First, let α, α' and S be as in the statement of the proposition. Let $\beta \in (0,1)$ be arbitrary and let $\beta' = \frac{\alpha\beta}{\alpha'}$. We will first prove that S is (α', β')-winning.

Let (h_n) be an $(\alpha, \beta; S)$-winning strategy. We will construct an $(\alpha', \beta'; S)$-winning strategy (f_n). Given $B_1, \dots, B_n \in \Omega$, we let $\tilde{W}_n = h(B_1, \dots, B_n)$ and take some ball $W_n \in \tilde{W}^{\alpha'/\alpha}$. We now define a strategy by $f_n(B_1, \dots, B_n) = W_n$. We claim that this strategy is $(\alpha', \beta'; S)$-winning.

Indeed, assume that (5.5a) and (5.5b) hold for α' and β'. We have

$$B_n \in W_{n-1}^{\beta'} = \tilde{W}_{n-1}^{\beta' \frac{\alpha'}{\alpha}} = \tilde{W}_{n-1}^{\beta}.$$

Also, by construction

$$\tilde{W}_n = h(B_1, \dots, B_n).$$

Since (h_n) was an $(\alpha, \beta; S)$-winning strategy, this implies that $\cap_i B_i^* \subseteq S$. Hence, (f_n) is $(\alpha', \beta'; S)$-winning.

To see that this implies the proposition, note that if $\beta' \in (0,1)$ and $\alpha' < \alpha$ then $\frac{\alpha'\beta'}{\alpha} \in (0,1)$. Since S is α-winning, the above implies that S is α'-winning. ☐

In the light of the above proposition, we make another definition.

Definition 5.7. Let $S \subseteq \mathcal{L}^{mn}$ and define $S^* = \{\alpha \in (0,1) : S \text{ is } \alpha\text{-winning}\}$. The *winning dimension of S* is defined as

$$\operatorname{windim} S = \begin{cases} 0 & \text{if } S^* = \emptyset, \\ \sup S^* & \text{otherwise.} \end{cases}$$

Our final tool in this long series of definitions is the concept of chains. Chains are series of legal Black moves given a White strategy.

Definition 5.8. Let (f_n) be an $(\alpha, \beta; S)$-winning strategy. A sequence $(B_i)_{i \in \mathbb{N}} \subseteq \Omega$ is said to be an *(f_n)-chain* if there exists a sequence $(W_i)_{i \in \mathbb{N}}$ such that we have (5.5a) and (5.5b) from Definition 5.5. An ordered set $(B_1, \dots, B_k) \subseteq \Omega$ is said to be a *finite (f_n)-chain* if there exist $(B_{k+i})_{i \in \mathbb{N}} \subseteq \Omega$ such that (B_i) is an (f_n)-chain.

From this point and onwards, we will not make the careful distinction between the formal balls $B = (c, \rho) \in \Omega$ and the images of these

$$B^* = \phi(B) = \{A \in \mathcal{L}^{mn} : \|A - c\|_\infty \leq \rho\} \subseteq \mathcal{L}^{mn}.$$

However, we will take care to use the proper ρ from the formal representation in the set notation of $\phi(B)$. Bearing this in mind, there is no difference between the two representations, and we are in fact considering the formal balls.

5.3. The winning dimension of $\mathfrak{B}(m,n)$

In this section, we will prove that for appropriate $\alpha,\beta\in(0,1)$, White can always win the $(\alpha,\beta;\mathfrak{B}(m,n))$-game. In fact, we will prove:

Theorem 5.6. *Let $\alpha,\beta\in(0,1)$ be such that $\gamma=k^{-1}+\alpha\beta-(k^{-1}+1)\alpha>0$ and let $m,n\in\mathbb{N}$. Then, $\mathfrak{B}(m,n)$ is (α,β)-winning.*

Immediately, we have the following corollary to Theorem 5.6:

Corollary 5.7.

$$\operatorname{windim}(\mathfrak{B}(m,n))\geq\tfrac{1}{k+1}.$$

Proof. Note that

$$\sup\left\{\alpha\in(0,1):k^{-1}+\alpha\beta-(k^{-1}+1)\alpha>0\;\forall\beta\in(0,1)\right\}=\tfrac{1}{k+1}.$$

Now, the Corollary is immediate from Theorem 5.6. □

For the rest of this section, let $n,m\in\mathbb{N}$ be fixed and let $\alpha,\beta\in(0,1)$ be such that $\gamma=k^{-1}+\alpha\beta-(k^{-1}+1)\alpha>0$.

Now, we start the game. Black begins, so let B_1 be the first ball, and let $\rho=\rho(B_1)$. Furthermore, we let $\sigma>0$ be such that for any $A\in B_1$, $\|A\|_\infty\leq\sigma$. The strategy, White will be using, depends on a constant $R>R_0(m,n,\alpha,\beta,\rho,\sigma)\geq 1$ to be chosen later. We will also need constants

$$\delta=R^{-m(m+n)^2},\quad \delta^*=R^{-n(m+n)^2},\quad \lambda=\frac{m}{m+n},\quad \lambda^*=1.$$

For $i,j\in\mathbb{N}$, we let $B_k,B_h\subseteq\mathcal{L}^{mn}$ be balls occurring in the (α,β)-game chosen by Black such that $\rho(B_k)<R^{-(m+n)(\lambda+i)}$ and $\rho(B_h)<R^{-(m+n)(\lambda^*+j)}$. We will show that White can play in such a way that the following properties hold for $i,j\in\mathbb{N}$:

(a) For $A\in B_k$, there are no $q\in\mathbb{F}[X]^{m+n}$ such that

$$0<\max_{1\leq l\leq m}\{\|q_l\|\}<\delta R^{n(\lambda+i)},\tag{5.6a}$$

$$\max_{1\leq l'\leq n}\left\{\left\|q\cdot\widetilde{A}^{(l')}\right\|\right\}<\delta R^{-m(\lambda+i)-n}.\tag{5.6b}$$

(b) For $A\in B_h$, there are no $q\in\mathbb{F}[X]^{m+n}$ such that

$$0<\max_{1\leq l'\leq n}\{\|q_{l'}\|\}<\delta^*R^{m(\lambda^*+j)},\tag{5.7a}$$

$$\max_{1\leq l\leq m}\left\{\left\|q\cdot\widetilde{A^*}^{(l)}\right\|\right\}<\delta^*R^{-n(\lambda^*+j)-m}.\tag{5.7b}$$

Such a strategy will be $(\alpha, \beta; \mathfrak{B}(m,n))$-winning. Indeed, given a $q \in \mathbb{F}[X]^{m+n}$ with $q_1, \ldots, q_m \neq 0$, we can find an $i \in \mathbb{N}$ such that

$$\delta R^{n(\lambda+i-1)} \leq \max_{1 \leq l \leq m} \{\|q_l\|\} < \delta R^{n(\lambda+i)}, \tag{5.8}$$

which immediately implies that (5.6a) holds. Hence by (5.6b),

$$\max_{1 \leq l' \leq n} \left\{ \left\| q \cdot \widetilde{A}^{(l')} \right\| \right\} \geq \delta R^{-m(\lambda+i)-n},$$

so by (5.8),

$$\begin{aligned}
\max_{1 \leq l' \leq n} \left\{ \left\| q \cdot \widetilde{A}^{(l')} \right\| \right\}^n &\geq \frac{\delta^{m+n} R^{-mn(\lambda+i)-n^2+mn(\lambda+i)-mn}}{\max_{1 \leq l \leq m} \{\|q_l\|\}^m} \\
&\geq \frac{\delta^{m+n} R^{-n^2-mn}}{\max_{1 \leq l \leq m} \{\|q_l\|\}^m} \\
&> \frac{K}{\max_{1 \leq l \leq m} \{\|q_l\|\}^m}
\end{aligned}$$

for any $K \in (0, \delta^{m+n} R^{-n^2-mn})$. By (5.1), we are done.

Now, we define for any $i \in \mathbb{N}$:

- B_{k_i} to be the first ball chosen by Black with $\rho(B_{k_i}) < R^{-(m+n)(\lambda+i)}$.
- B_{h_i} to be the first ball chosen by Black with $\rho(B_{h_i}) < R^{-(m+n)(\lambda^*+i)}$.

Since $\lambda < \lambda^*$, these balls occur such that $B_{k_0} \supseteq B_{h_0} \supseteq B_{k_1} \supseteq B_{h_1} \supseteq \cdots$. By choosing R large enough, we can ensure that the inclusions are proper.

Since

$$\delta R^{n\lambda} = R^{-m(m+n)^2+nm(m+n)^{-1}} = R^{-m\left((m+n)^2-\frac{n}{m+n}\right)} < 1,$$

(5.6a) has no solutions for $i = 0$. Hence, White can certainly play in such a way that the system (5.6a) and (5.6b) has no solutions in $\mathbb{F}[X]^{m+n}$ when $A \in B_{k_0}$. We will recursively construct White's strategy in such a way that

1. Given the beginning of a game $B_1 \supseteq W_1 \supseteq \cdots \supseteq B_{k_0} \supseteq \cdots \supseteq B_{k_i}$ such that (5.6a) and (5.6b) have no integer solutions for any $A \in B_{k_i}$, White can play in such a way that (5.7a) and (5.7b) have no solutions in $\mathbb{F}[X]^{m+n}$ for any $A \in B_{h_i}$.
2. Given the beginning of a game $B_1 \supseteq W_1 \supseteq \cdots \supseteq B_{k_0} \supseteq \cdots \supseteq B_{h_i}$ such that (5.7a) and (5.7b) have no integer solutions for any $A \in B_{h_i}$, White can play in such a way that (5.6a) and (5.6b) have no solutions in $\mathbb{F}[X]^{m+n}$ for any $A \in B_{k_{i+1}}$.

Our first lemma guarantees that we need only worry about solutions to the equations in certain subspaces of $\mathcal{L}^{m+n}$.

Lemma 5.8. *Let $B_1 \supseteq W_1 \supseteq \cdots \supseteq B_{k_i}$ be the start of a game such that* (5.6a) *and* (5.6b) *have no solutions in $\mathbb{F}[X]^{m+n}$ for any $A \in B_{k_i}$. Let $A \in B_{k_i}$ be fixed. The set*

$$\left\{q \in \mathbb{F}[X]^{m+n} : (5.7\text{a}) \textit{ and } (5.7\text{b}) \textit{ hold}\right\}$$

contains at most m linearly independent points.

Proof. Assume that $q_1, \ldots, q_{m+1} \in \mathbb{F}[X]^{m+n}$ are linearly independent points such that (5.7a) and (5.7b) hold. Clearly,

$$\|q_u\|_\infty \ll \delta^* R^{m(\lambda^*+i)} \quad \text{for } 1 \le u \le m+1. \tag{5.9}$$

Let C_{k_i} be the centre of B_{k_i}. For any $A \in B_{k_i}$,

$$\left\|\widetilde{A^*}^{(v)} - \widetilde{C^*}^{(v)}\right\|_\infty \le \rho(B_{k_i}) < R^{-(m+n)(\lambda+i)} \quad \text{for } 1 \le v \le n. \tag{5.10}$$

Hence, there is an $A \in B_{k_i}$ such that (5.10) holds and such that (5.7b) holds for each $q_1, \ldots, q_{m+1}$. Now,

$$\begin{aligned} \max_{\substack{1\le l\le m\\ 1\le u\le m+1}} \left\{\left\|q_u \cdot \widetilde{C^*}^{(l)}\right\|\right\} &\le \max_{\substack{1\le l\le m\\ 1\le u\le m+1}} \left\{\left\|q_u \cdot \widetilde{A^*}^{(l)}\right\|, \left\|q_u \cdot \left(\widetilde{C^*}^{(l)} - \widetilde{A^*}^{(l)}\right)\right\|\right\} \\ &\ll \max\left\{\delta^* R^{-n(\lambda^*+i)-m}, \delta^* R^{m(\lambda^*+i)} R^{-(m+n)(\lambda+i)}\right\} \\ &\ll \delta^* R^{-n(\lambda^*+i)}. \end{aligned} \tag{5.11}$$

We define the parallelepipeds

$$P = \left\{y \in \mathcal{L}^{m+n} : \max_{1\le l'\le n} \{\|y_{l'}\|\} < R^{m(\lambda^*+i)}, \max_{1\le l\le m} \left\{\left\|y \cdot \widetilde{C^*}^{(l)}\right\|\right\} < R^{-n(\lambda^*+i)}\right\},$$

with the corresponding distance function F_C and successive minima $\sigma_1, \ldots, \sigma_{m+n}$. By (5.11), $\sigma_{m+1} \ll \delta^*$. For $n = 1$, (5.4) and (5.11) imply that $1 \ll R^{-(m+1)^2}$, which gives a contradiction by choosing R large enough, so in this case we are done.

Now we assume that $n > 1$. Let

$$P^* = \left\{x \in \mathcal{L}^{m+n} : \max_{1\le l\le m} \{\|x_l\|\} < R^{n(\lambda^*+i)}, \max_{1\le l'\le n} \left\{\left\|x \cdot \widetilde{C}^{(l')}\right\|\right\} < R^{-m(\lambda^*+i)}\right\}.$$

This set admits the distance function F_C^* defined in Lemma 5.4. Indeed, if we insert F_C in the definition from Lemma 5.4 and calculate the value of the resulting function under each of the possible assumptions of the form of F_C, when the supremum is attained, we get for each coordinate exactly the distance function for the set P^*. One needs to take into account the specific form of $\widetilde{C}$ and $\widetilde{C^*}$.

Let $\tau_1, \dots, \tau_{m+n}$ denote the successive minima of F_C^*. By Lemma 5.3 and Lemma 5.4,

$$\tau_1 \ll (\tau_1 \cdots \tau_{n-1})^{\frac{1}{n-1}} \ll (\tau_n \cdots \tau_{m+n})^{\frac{-1}{n-1}} \ll \tau_n^{-\frac{m+1}{n-1}} \\ \ll \sigma_{m+1}^{\frac{m+1}{n-1}} \ll (\delta^*)^{\frac{m+1}{n-1}} \ll R^{-(m+n)^2(m+1)} = \delta R^{-(m+n)^2}. \quad (5.12)$$

Hence, there is a $q \in \mathbb{F}[X]^{m+n} \setminus \{0\}$ with

$$\max_{1 \le l \le m} \{\|q_l\|\} < \delta R^{-(m+n)^2} R^{n(\lambda^*+i)}$$

and

$$\max_{1 \le l' \le n} \left\{ \left\| q \cdot \widetilde{C^*}^{(l')} \right\| \right\} < \delta R^{-(m+n)^2} R^{-m(\lambda^*+i)},$$

when we choose R large enough to absorb the implicit constant in (5.12). This gives a contradiction, since a further modification of R yields a solution to (5.6a) and (5.6b). □

In a completely analogous way, we can prove:

Lemma 5.9. *Let $B_1 \supseteq W_1 \supseteq \cdots \supseteq B_{h_i}$ be the start of a game such that* (5.7a) *and* (5.7b) *have no solutions in $\mathbb{F}[X]^{m+n}$ for any $A \in B_{h_i}$. Let $A \in B_{h_i}$ be fixed. The set*

$$\left\{ q \in \mathbb{F}[X]^{m+n} : (5.6\text{a}) \textit{ and } (5.6\text{b}) \textit{ hold} \right\}$$

contains at most n linearly independent points.

We will now reduce the statement that White has a strategy such that Step 1 on page 58 is possible, to the statement that White can win a certain finite game. The converse Step 2 is completely analogous.

Once again, we assume that $B_1 \supseteq W_1 \supseteq \cdots \supseteq B_{k_i}$ is the beginning of a game such that we have avoided solutions in $\mathbb{F}[X]^{m+n}$ to all relevant inequalities so far. Now, it is sufficient for White to avoid solutions $q \in \mathbb{F}[X]^{m+n}$ to (5.7a) and (5.7b) with

$$\delta^* R^{m(\lambda^*+i-1)} \le \max_{1 \le l' \le n} \{\|q_{l'}\|\} < \delta^* R^{m(\lambda^*+i)}.$$

Since all relevant $q \in \mathbb{F}[X]^{mn}$ are required to satisfy (5.7b), we need only consider the situation

$$\delta^* R^{m(\lambda^*+i-1)} \le \|q\|_\infty < c\delta^* R^{m(\lambda^*+i)} \quad (5.13)$$

for some $c > 0$.

By Lemma 5.8, the set of $q \in \mathbb{F}[X]^{m+n}$ satisfying (5.7a) and (5.7b) is contained in some m-dimensional subspace for any fixed $A \in B_{k_i}$. Let $\{y_1, \dots, y_m\}$ be an orthonormal basis for this space and write all $q \in \mathbb{F}[X]^{m+n}$ in this subspace satisfying (5.13) on the form $q = t_1 y_1 + \cdots + t_m y_m$, $t_1, \dots, t_m \in \mathcal{L}$. Immediately,

$$\delta^* R^{m(\lambda^*+i-1)} \le \max_{1 \le l' \le m} \{\|t_{l'}\|\} < c\delta^* R^{m(\lambda^*+i)}. \quad (5.14)$$

White only needs to avoid solutions to the inequalities

$$\max_{1\le l\le m}\left\|\sum_{l'=1}^{m} t_{l'}\left(y_{l'}\cdot\widetilde{A^*}^{(l)}\right)\right\| < \delta^* R^{-n(\lambda^*+i)-m}.$$

This is a matrix inequality and Cramer's Rule (Theorem 4.4 in [39]) yields a solution set for $v = 1,\dots,m$,

$$\|t_v\|\,\|D\| = \|t_v D\| \le m\delta^* R^{-n(\lambda^*+i)-m}\max_{1\le l'\le m}\left\{\left\|D_{l',v}\right\|\right\},$$

where D denotes the determinant of the matrix with entries $y_l\cdot\widetilde{A^*}^{(l')}$ and $D_{i,j}$ denotes the (i,j)'th co-factor of this determinant. By (5.14), it is sufficient to avoid

$$\|D\| \ll R^{-n(\lambda^*+i)-m-m(\lambda^*+i)}\max_{1\le l,l\le m}\left\{\left\|D_{l,l'}\right\|\right\} \ll R^{-(m+n)(\lambda^*+i)}\max_{1\le l,l\le m}\left\{\left\|D_{l,l'}\right\|\right\}. \tag{5.15}$$

We define the following finite game:

Definition 5.9. Let $y_1,\dots,y_m\in\mathcal{L}^{m+n}$ be a set of orthonormal vectors. Let $B\subseteq\mathcal{L}^{mn}$ be a ball with $\rho(B)<1$ such that for any $A\in B$, $\|A\|_\infty\le\sigma$. Let $\psi,\mu>0$ and let α,β be as in the statement of Theorem 5.6. White and Black take turns according to the rules of the game in Definition 5.4, but the game terminates when $\rho(B_t)<\mu\rho(B)$. White wins the game if

$$\|D\| > \psi\rho(B)\mu\max_{1\le l,l\le m}\left\{\left\|D_{l,l'}\right\|\right\}$$

for any $A\in B_t$.

If White can win the game in Definition 5.9 for any $\mu\in(0,\mu^*)$ for some appropriate $\mu^*=\mu^*(m,n,\alpha,\beta,\sigma,\psi)>0$, then White can guarantee that (5.15) does not hold for $A\in B_{h_i}$. To see this, let $B=B_{k_i}$, ψ be the constant implicit in (5.15) and

$$\mu = \frac{R^{-(m+n)(\lambda^*+i)}}{\rho(B)} \le (\alpha\beta)^{-1}R^{-n}.$$

Choosing R large enough, this will be less than μ^*. Inserting in the winning condition in Definition 5.9 proves that White has a strategy such that solutions to 5.15 are avoided. It remains to be shown that such a μ^* exists. We will do this by induction.

Let $A\in\mathcal{L}^{mn}$, $v\in\{1,\dots,m\}$ and let $(y_1,\dots,y_m)$ be the orthonormal system from Definition 5.9. Taking all possible choices of numbers, $1\le i_1<\cdots<i_v\le m$ and $1\le j_1<\cdots<j_v\le m$, we obtain $\binom{m}{v}^2$ matrices

$$\begin{pmatrix} y_{i_1}\cdot\widetilde{A^*}^{(j_1)} & \cdots & y_{i_1}\cdot\widetilde{A^*}^{(j_v)} \\ \vdots & & \vdots \\ y_{i_v}\cdot\widetilde{A^*}^{(j_1)} & \cdots & y_{i_v}\cdot\widetilde{A^*}^{(j_v)} \end{pmatrix} \tag{5.16}$$

For each $v \in \{1,\dots,m\}$, we define the function $M_v : \mathcal{L}^{mn} \to \mathcal{L}^{\binom{m}{v}^2}$ to have as it's coordinates the determinants of the matrices in (5.16) in some arbitrary but fixed order. Furthermore, define

$$M_{-1}(A) = M_0(A) = 1.$$

For $K \subseteq \mathcal{L}^{mn}$, we define

$$M_v(K) = \max_{A \in K} \|M_v(A)\|_\infty.$$

We will prove the following complicated lemma.

Lemma 5.10. *Let $(y_1,\dots,y_m) \subseteq \mathcal{L}^{mn}$ be an orthonormal system. Let $B \subseteq \mathcal{L}^{mn}$ be a ball, $\rho(B) = \rho_0 < 1$, such that for some $\sigma > 0$, $\|A\|_\infty < \sigma$ for any $A \in B$. Let $\psi > 0$, and let $\alpha,\beta \in (0,1)$ with $k^{-1} + \alpha\beta - (k^{-1}+1)\alpha > 0$. Assume that $0 \le v \le m$.*

There exists a $\mu_v = \mu_v(m,n,\alpha,\beta,\sigma,\psi) > 0$ for which White can play the game in Definition 5.9 in such a way that for the first ball B_{i_v} with $\rho(B_{i_v}) < \rho_0\mu_v$,

$$\|M_v(A)\|_\infty > \psi\rho_0\mu_v M_{v-1}(B_{i_v})$$

for any $A \in B_{i_v}$.

Note that this immediately implies:

Corollary 5.11. *Let $\alpha,\beta \in (0,1)$ with $k^{-1} + \alpha\beta - (k^{-1}+1)\alpha > 0$. White can win the game in Definition 5.9 and hence the (α,β)-game.*

Proof. Use Lemma 5.10 with $v = m$. □

Proof of Lemma 5.10. We will prove the lemma by induction. Clearly, the lemma holds for $v = 0$. Hence, we may assume that $v > 0$ and that there exists a μ_{v-1} such that

$$\|M_{v-1}(A)\|_\infty > \psi\rho_0\mu_{v-1}M_{v-2}(B_{i_{v-1}}) \tag{5.17}$$

for all $A \in B_{i_{v-1}}$ for an appropriate $B_{i_{v-1}}$ occurring in the game. We will prove that White has a strategy such that the lemma holds for some appropriate B_{i_v}. This will require a number of lemmas.

Lemma 5.12. *Let $\varepsilon > 0$ and let $B' \subseteq B_{i_{v-1}}$ be a ball, $\rho(B') \le \varepsilon C_1(m,n)\rho(B_{i_{v-1}})$, where $C_1(m,n) > 0$ is to be fixed later. Then*

$$\left\|M_{v-1}(A) - M_{v-1}(A')\right\|_\infty < \varepsilon\rho_0\mu_{v-1}M_{v-2}(B_{i_{v-1}})$$

for any $A, A' \in B'$.

Proof of Lemma 5.12. Consider the norm of each coordinate of the vector on the left hand side. These have the form

$$\left\| \det \begin{pmatrix} y_{i_1}\cdot\widetilde{A^*}^{(j_1)} & \cdots & y_{i_1}\cdot\widetilde{A^*}^{(j_v)} \\ \vdots & & \vdots \\ y_{i_v}\cdot\widetilde{A^*}^{(j_1)} & \cdots & y_{i_v}\cdot\widetilde{A^*}^{(j_v)} \end{pmatrix} - \det \begin{pmatrix} y_{i_1}\cdot\widetilde{A'^*}^{(j_1)} & \cdots & y_{i_1}\cdot\widetilde{A'^*}^{(j_v)} \\ \vdots & & \vdots \\ y_{i_v}\cdot\widetilde{A'^*}^{(j_1)} & \cdots & y_{i_v}\cdot\widetilde{A'^*}^{(j_v)} \end{pmatrix} \right\|$$

Expanding the determinants by the first row, we see that by Proposition 1.6, this expression is

$$\begin{aligned}
&\leq \left\| y_{i_1} \cdot \left(\left(\widetilde{A^*}^{(j_1)} - \widetilde{A'^*}^{(j_1)} \right) + \cdots + \left(\widetilde{A^*}^{(j_v)} - \widetilde{A'^*}^{(j_v)} \right) \right) \right\| M_{v-2}(B_{i_{v-1}}) \\
&\leq \|y_{i_1}\|_\infty \max_{1 \leq h \leq v} \left\{ \left\| \widetilde{A^*}^{(j_v)} - \widetilde{A'^*}^{(j_v)} \right\|_\infty \right\} M_{v-2}(B_{i_{v-1}}) \\
&\ll \varepsilon \rho(B_{i_{v-1}}) M_{v-2}(B_{i_{v-1}}) \ll \varepsilon \rho_0 M_{v-2}(B_{i_{v-1}}).
\end{aligned}$$

By an appropriate choice of $C_1(m,n)$, this completes the proof. □

Corollary 5.13. *Assume that* $B' \subseteq B_{i_{v-1}}$ *is a ball with* $\rho(B') < \frac{1}{2} C_1(m,n) \rho(B_{i_{v-1}})$. *Then,*

$$\left\| M_{v-1}(A') \right\|_\infty \geq \tfrac{1}{2} M_{v-1}(B')$$

for any $A' \in B'$.

Proof of Corollary 5.13. Just apply Lemma 5.12 with $\varepsilon = \frac{1}{2}\psi$. By induction hypothesis (5.17) and this lemma, for any $A, A' \in B'$,

$$\begin{aligned}
\|M_{v-1}(A)\|_\infty &> \psi \rho_0 \mu_{v-1} M_{v-2}(B_{i_{v-1}}) \\
&= 2\varepsilon \rho_0 \mu_{v-1} M_{v-2}(B_{i_{v-1}}) \\
&\geq \tfrac{1}{2} \left\| M_{v-1}(A) - M_{v-1}(A') \right\|_\infty .
\end{aligned}$$

Holding A fixed and maximising over A', we obtain the result. □

Now, we define the simplest co-factor,

$$D_v(A) = \det \begin{pmatrix} y_1 \cdot \widetilde{A^*}^{(1)} & \cdots & y_1 \cdot \widetilde{A^*}^{(v)} \\ \vdots & & \vdots \\ y_v \cdot \widetilde{A^*}^{(1)} & \cdots & y_v \cdot \widetilde{A^*}^{(v)} \end{pmatrix}.$$

Clearly, this is a function of the mv variables $a_{11}, \ldots, a_{m1}, \ldots, a_{mv}$. We define the *discrete gradient of* D_v to be the vector

$$\nabla D_v(A) = \begin{pmatrix} D_v(A + e_{11}) - D_v(A) \\ \vdots \\ D_v(A + e_{mn}) - D_v(A), \end{pmatrix} \in \mathcal{L}^{mn},$$

where $e_{ij} \in \mathcal{L}^{mn}$ denotes the vector having 1 as the ij'th coordinate and 0 elsewhere. Hence, $\nabla D_v(A)$ has at most mv non-zero coordinates.

Corollary 5.14. *With* ε *and* B' *as in Lemma 5.12 and* $A', A'' \in B'$,

$$\left\| \nabla D_v(A') - \nabla D_v(A'') \right\|_\infty \leq C_2(m,n) \left\| M_{v-1}(A') - M_{v-1}(A'') \right\|_\infty$$

for some $C_2(m,n)$.

Proof. Note, that the coordinates of $\nabla D_v(A)$ are linear combinations of the coordinates of $M_{v-1}(A)$ for any A. Indeed, if we expand the determinants at each coordinate along the column containing e_{mn} we obtain exactly this. Now, apply Lemma 5.12 and choose $C_2(m,n)$ as an upper bound on the coefficients from the linear combinations. □

The discrete gradient turns out to be the key ingredient in the proof. We will need the following lemma:

Lemma 5.15. *Let $B' \subseteq B_{i_{v-1}}$ be such that*

$$\rho(B') < \tfrac{1}{2}\psi C_1(m,n)\rho(B_{i_{v-1}}). \tag{5.18}$$

Let $A' \in B'$ be such that

$$\left\|M_v(A')\right\|_\infty < C_3(n,\psi)\psi M_{v-1}(B'), \tag{5.19}$$

where $C_3(n,\psi)$ is to be chosen later. Furthermore, assume that

$$d_v := \left\|M_{v-1}(A')\right\|_\infty = \left\| \det \begin{pmatrix} y_1 \cdot \widetilde{A'^*}^{(1)} & \cdots & y_1 \cdot \widetilde{A'^*}^{(v-1)} \\ \vdots & & \vdots \\ y_{v-1} \cdot \widetilde{A'^*}^{(1)} & \cdots & y_{v-1} \cdot \widetilde{A'^*}^{(v-1)} \end{pmatrix} \right\|.$$

Then

$$\left\|\nabla D_v(A')\right\|_\infty > C_4(m,n,\sigma)M_{v-1}(B')$$

for some $C_4(m,n,\sigma) > 0$.

Proof of Lemma 5.15. We will consider the norm of the discrete directional derivative in the direction of a vector Z. Discrete directional derivatives are defined in a way analogous to the discrete gradient. That is, for $f : \mathcal{L}^{m+n} \to \mathcal{L}$ and $Z \in \mathcal{L}^{m+n}$, we define the discrete directional derivative of f along Z to be

$$\nabla_Z f(A) = f(A+Z) - f(A).$$

For $Z \in \mathcal{L}^{m+n}$, we consider $\|\nabla_Z D_v(A)\|$. This may be written

$$\left\| \det \begin{pmatrix} y_1 \cdot \widetilde{A'^*}^{(1)} & \cdots & y_1 \cdot \left(\widetilde{A'^*}^{(v)} + Z\right) \\ \vdots & & \vdots \\ y_v \cdot \widetilde{A'^*}^{(1)} & \cdots & y_v \cdot \left(\widetilde{A'^*}^{(v)} + Z\right) \end{pmatrix} - \det \begin{pmatrix} y_1 \cdot \widetilde{A'^*}^{(1)} & \cdots & y_1 \cdot \widetilde{A'^*}^{(v)} \\ \vdots & & \vdots \\ y_v \cdot \widetilde{A'^*}^{(1)} & \cdots & y_v \cdot \widetilde{A'^*}^{(v)} \end{pmatrix} \right\|$$

which is equal to $\|(d_1 y_1 + \cdots + d_v y_v) \cdot Z\|$ by expanding the two determinants in the last column. The d_i are coordinates of $M_{v_{i-1}}(A')$.

Let $Z_1 = d'_1 y_1 + \cdots + d'_v y_v$, where $d'_i = d_i$ for $i = 1, \ldots, v-1$ and $d'_v = X d_v$. Orthonormality of the y_i implies,

$$\|\nabla_{Z_1} D_v(A)\| = \left\| \sum_{i=1}^{v} d_i'^2 \right\| = k^2 \|d_v\|^2 = k^2 \left\|M_{v-1}(A')\right\|_\infty^2 \geq \tfrac{k^2}{4} M_{v-1}(B')^2$$

by Corollary 5.13. The second equality is true since for all $i = 1, \ldots, v-1$, we have $\|d'_i\| < \|d'_v\|$. Hence, the function D_v grows rather rapidly in the direction of Z_1. Unfortunately, we cannot guarantee that $\widetilde{A'^*}^{(v)} + Z_1$ is on the required form. That is, it may not be on the form $(v_1, \ldots, v_m, 0, \ldots, 1, 0, \ldots, 0)$. It needs to have this form in order to correspond to an element $Z' \in \mathcal{L}^{mn}$, so further work is needed to obtain an estimate on the discrete gradient.

We now write

$$y_i = y_i^0 + \lambda_{i1} \widetilde{A'^*}^{(1)} + \cdots + \lambda_{in} \widetilde{A'^*}^{(n)},$$

where y_i^0 has zeros on the last n coordinates. This is possible because of the form of the $\widetilde{A'^*}^{(i)}$. Since the y_i are orthonormal,

$$\sum_{j=1}^{n} \left\|\lambda_{ij}\right\|^2 \leq 1, \quad \left\|y_i^0\right\|_\infty \leq C_4(m,n,\sigma)$$

for some $C_4(m,n,\sigma) > 0$. Now, let $Z_2 = d'_1 y_1^0 + \cdots + d'_v y_v^0$. Certainly, $\widetilde{A'^*}^{(v)} + Z_2$ is on the right form. Furthermore,

$$\begin{aligned}\|\nabla_{Z_2} D_v(A)\| &= \|(d_1 y_1 + \cdots + d_v y_v) \cdot Z_2\| \\ &= \|(d_1 y_1 + \cdots + d_v y_v) \cdot (Z_2 - Z_1 + Z_1)\| \\ &\geq \tfrac{k^2}{4} M_{v-1}(K')^2 - \|(d_1 y_1 + \cdots + d_v y_v) \cdot (Z_1 - Z_2)\|.\end{aligned}$$

We know that

$$Z_1 - Z_2 = \sum_{j=1}^{v} \left(\sum_{i=1}^{v} d'_i \lambda_{ij} \right) \widetilde{A'^*}^{(j)}.$$

Furthermore,

$$\begin{aligned}\left\|(d_1 y_1 + \cdots + d_v y_v) \cdot \widetilde{A'^*}^{(v)}\right\| &= \left\| \det \begin{pmatrix} y_1 \cdot \widetilde{A'^*}^{(1)} & \cdots & y_1 \cdot \widetilde{A'^*}^{(v)} \\ \vdots & & \vdots \\ y_v \cdot \widetilde{A'^*}^{(1)} & \cdots & y_v \cdot \widetilde{A'^*}^{(v)} \end{pmatrix} \right\| \\ &\leq \left\|M_v(A')\right\|_\infty < C_3(n,\psi)\psi M_{v-1}(B')\end{aligned}$$

by choice of A'.

Now, $\|d'_i\| \leq k M_{v-1}(B')$, so

$$\|\nabla_{Z_2} D_v(A)\| \geq \tfrac{k^2}{4} M_{v-1}(B')^2 - k v^2 C_3(n,\psi)\psi M_{v-1}(B')^2.$$

We now choose $C_3(n,\psi)$ so small that

$$\|\nabla_{Z_2} D_v(A)\| \geq C_5(n) M_{v-1}(B')^2,$$

for some $C_5(n) > 0$. Since $\|Z_2\|_\infty \ll k M_{v-1}(B')$, this implies that the discrete directional derivative is large enough in this direction. But surely, the norm of the discrete gradient is larger than norm of the discrete directional derivative $\nabla_Z D_v(A)$ along any vector Z of magnitude one. Indeed, if we write a vector of magnitude one as $Z = (z_1, \dots, z_{m+n})$, we see that $\|z_i\| \leq 1$ for all i. This implies that the gradient has greater norm, because of the special form of $D(A)$. This proves the lemma. □

With all of the above in place, we can complete the proof of Lemma 5.10. We let

$$\gamma = k^{-1} + \alpha\beta - (k^{-1}+1)\alpha > 0, \quad \varepsilon = \frac{\gamma}{8} \frac{C_4(m,n,\sigma)}{C_2(m,n)} > 0.$$

Furthermore, we let

$$j_v = \min\left\{ i \in \mathbb{N} : i > i_{v-1}, \rho(B_i) < \min(\tfrac{1}{2}\psi, \varepsilon) C_1(m,n) \rho(B_{i_{v-1}}) \right\}. \tag{5.20}$$

Clearly,

$$\rho(B_{j_v}) \geq C_6(m,n,\alpha,\beta,\sigma,\psi)\rho_0 \tag{5.21}$$

for an appropriate constant $C_6(m,n,\alpha,\beta,\sigma,\psi) > 0$.

By induction hypothesis and Corollary 5.14, for any $A', A'' \in B_{j_v}$,

$$\begin{aligned}\left\| \nabla D_v(A') - \nabla D_v(A'') \right\|_\infty &< C_2(m,n)\varepsilon\rho_0\mu_{v-1} M_{v-2}(B_{i_{v-1}}) \\ &< \tfrac{\gamma}{8} C_4(m,n,\sigma) M_{v-1}(B_{j_v}).\end{aligned}$$

We now let

$$\mu_v < \frac{\min\left\{ C_3(n,\psi), \frac{\gamma}{8}\alpha\beta C_6(m,n,\alpha,\beta,\sigma,\psi), \frac{1}{\psi}\frac{\gamma}{4} C_4(m,n,\sigma) C_6(m,n,\alpha,\beta,\sigma,\psi) \right\}}{C_7(m,n,\sigma)},$$

where $C_7(m,n,\sigma) > 0$ is to be chosen later. Assume that there exists an $A' \in B_{j_v}$ for which the lemma does not hold. That is,

$$\left\| M_v(A') \right\|_\infty \leq \psi\rho_0\mu_v M_{v-1}(B_{j_v}). \tag{5.22}$$

In this case, we will prove that White has a strategy which will eliminate such elements in a finite number of moves.

By choice of j_v, (5.18) holds. Since $\rho_0 < 1$, (5.19) holds. By rearranging the y_i, we can without loss of generality assume that the condition on the determinant in Lemma 5.15 holds. Hence,

$$\left\| \nabla D_v(A') \right\|_\infty > C_4(m,n,\sigma) M_{v-1}(B_{j_v}). \tag{5.23}$$

Let $D' = \nabla D_v(A')$, and let d_i and c_i denote the centres of W_i and B_i respectively. White can play in such a way that

$$\left\|(c_i - d_i)\cdot D'\right\| \geq k^{-1}(1-\alpha)\rho(B_i)\left\|D'\right\|_\infty. \tag{5.24}$$

Indeed, White may certainly choose d_i such that $\|c_i - d_i\|_\infty > k^{-1}(1-\alpha)\rho(B_i)$. By Proposition 1.5, we have $W_i \subseteq B_i$ for such d_i, so the choice is allowed. Examining each coordinate, we see that in fact White may choose d_i such that for any $j \in \{1,\dots,mn\}$, we have $\left\|c_i^{(j)} - d_i^{(j)}\right\| > k^{-1}(1-\alpha)\rho(B_i)$. Now, for any $j \in \{1,\dots,mn\}$,

$$k^{-1}(1-\alpha)\rho(B_i)\left\|D'\right\|_\infty \leq \left\|c_i^{(j)} - d_i^{(j)}\right\|\left\|D'\right\|_\infty = \left\|\left(c_i^{(j)} - d_i^{(j)}\right)D'\right\|_\infty$$

Since this is valid for any j, it guarantees (5.24).

Also, no matter how Black plays

$$\left\|(c_{i+1} - d_i)\cdot D'\right\| \leq (1-\beta)\rho(W_i)\left\|D'\right\|_\infty, \tag{5.25}$$

since $\|(c_{i+1} - d_i)\cdot D'\| \leq \|c_{i+1} - d_i\|_\infty \|D'\|_\infty$ by Proposition 1.6 and since $B_{i+1} \subseteq W_i$ implies that $\|c_{i+1} - d_i\|_\infty \leq (1-\beta)\rho(W_i)$. Hence,

$$\left\|(c_{i+1} - c_i)\cdot D'\right\| \geq \left(k^{-1}(1-\alpha) - \alpha(1-\beta)\right)\rho(B_i)\left\|D'\right\|_\infty = \gamma\rho(B_i)\left\|D'\right\|_\infty > 0.$$

We choose $t_0 \in \mathbb{N}$ such that $\alpha\beta\frac{\gamma}{2} < (\alpha\beta)^{t_0} \leq \frac{\gamma}{2}$. Clearly,

$$\left\|(c_{i+t_0} - c_i)\cdot D'\right\| \geq \gamma\rho(B_i)\left\|D'\right\|_\infty > 0. \tag{5.26}$$

Since furthermore $\rho(B_{i+t_0}) \leq \frac{\gamma}{2}\rho(B_i)$, for any $A \in B_{i+t_0}$,

$$\left\|(A - c_i)\cdot D'\right\| \geq \left\|(A - c_{i+t_0})\cdot D'\right\| + \left\|(c_{i+t_0} - c_i)\cdot D'\right\| \geq \frac{\gamma}{2}\rho(B_i)\left\|D'\right\|_\infty. \tag{5.27}$$

White will play according to such a strategy.

A simple calculation shows that for any $A \in \mathcal{L}^{mn}$,

$$A\cdot\nabla D_v(A) = \sum_{h=1}^{v} D_v(A).$$

Indeed, in calculating the inner product on the right hand side, each coordinate a_{ij} of A contributes with a factor that the sum of v copies of the appropriate cofactor of $D_v(A)$ times a_{ij} times a sign. This is revealed by expanding the determinants in the right column. The sign is the right one in the expansion of $D_v(A)$, so this proves the claim.

Now, repeated use of the ultra-metric property (1.5c) yields:

$$\begin{aligned}
\|D_v(A)\| &\geq C_7(m,n,\sigma)\left\|D_v(A-c_{j_v})\right\| \\
&\geq C_7(m,n,\sigma)\left\|\sum_{h=1}^{v} D_v(A-c_{j_v})\right\| \\
&= C_7(m,n,\sigma)\left\|(A-c_{j_v})\cdot\nabla D_v(A-c_{j_v})\right\| \\
&\geq C_7(m,n,\sigma)\big(\left\|(A-c_{j_v})\cdot D'\right\| \\
&\quad -\left\|(A-c_{j_v})\cdot(\nabla D_v(A-c_{j_v})-\nabla D_v(A'))\right\|\big) \\
&\geq C_7(m,n,\sigma)\left(\tfrac{\gamma}{2}\rho(B_{j_v})\left\|D'\right\|_\infty - 2\rho(B_{j_v})\tfrac{\gamma}{8}C_4(m,n,\sigma)M_{v-1}(B_{j_v})\right) \\
&\geq C_7(m,n,\sigma)\big(\tfrac{\gamma}{2}\rho(B_{j_v})C_4(m,n,\sigma)M_{v-1}(B_{j_v}) \\
&\quad -\tfrac{\gamma}{4}\rho(B_{j_v})C_4(m,n,\sigma)M_{v-1}(B_{j_v})\big) \\
&> \psi\mu_v\rho_0 M_{v-1}(B_{j_v}),
\end{aligned}$$

where $C_7(m,n,\sigma) > 0$ is chosen such that the first inequality holds (which is clearly possible). This completes the proof of Lemma 5.10 and hence the proof of Theorem 5.6. □

5.4. The Hausdorff dimension of $\mathfrak{B}(m,n)$

In this final section of the chapter, we will prove that if $\alpha > 0$ then any α-winning set in $\mathcal{L}^{mn}$ has full Hausdorff dimension. By Corollary 5.7, this will imply Theorem 5.2. To do this, we change our viewpoint to that of Black player. We will need one additional definition:

Definition 5.10. Let $S \subseteq \mathcal{L}^{mn}$ be (α,β)-winning and let (f_n) be a winning strategy. Let $(E_n)_{n\in\mathbb{N}}$ be a sequence of balls in $\mathcal{L}^{mn}$ and $t \in \mathbb{N}$. (E_n) is said to be a t-(f_n)-*chain* if there exists an (f_n)-chain with $E_i = B_{(i-1)t+1}$.

We now state a rather lengthy hypothesis, which is the basis of the following three results:

Hypothesis 5.1. Let $\alpha,\beta \in (0,1)$. Assume that there exist $t,u \in \mathbb{N}$ such that: Given $h_1,\dots,h_t$, $h_i \in F_i^{\alpha}$ and a ball $C_1 \subseteq \mathcal{L}$, there exist $g^{(0)},\dots,g^{(u-1)} \in F_1^{\beta}$, such that for $C_2^{(j)},\dots,C_{t+1}^{(j)}$ and $D_1^{(j)},\dots,D_t^{(j)}$ defined for $j=0,\dots,u-1$ recursively by the relations

$$C_i^{(j)} = g^{(j)}\left(D_{i-1}^{(j)}\right), \qquad i=2,\dots,t+1, \tag{5.28a}$$

$$D_i^{(j)} = h_i\left(C_1,C_2^{(j)},\dots,C_i^{(j)}\right), \qquad i=1,\dots,t, \tag{5.28b}$$

we have that $C_{t+1}^{(j)} \cap C_{t+1}^{(k)} = \emptyset$ for $j \neq k$.

Roughly speaking, this hypothesis says that if Black knows the strategy of White, she can choose a strategy such that after t moves, the game is in one of u entirely different states.

Lemma 5.16. *Under Hypothesis 5.1, if $E_1,\dots,E_k$ is a t-(f_n)-chain, there exist pairwise disjoint $E_{k+1}^{(0)},\dots,E_{k+1}^{(u-1)}$, such that $E_1,\dots,E_k,E_{k+1}^{(i)}$ is a t-(f_n)-chain for every $i=0,\dots,u-1$.*

Proof. Let $B_1,\dots,B_{(k-1)t+1}$ be an (f_n)-chain with $E_i = B_{(i-1)t+1}$. We define functions h_i, $i=1,\dots,t$ by

$$h_i(C_1,\dots,C_i) = f_{(k-1)t+i}\left(B_1,\dots,B_{(k-1)t},C_1,\dots,C_i\right).$$

Let $C_1 = E_k = B_{(k-1)t+1}$ and define $E_{k+1}^{(j)} = C_{t+1}^{(j)}$, where the $C_{t+1}^{(j)}$ are the ones from Hypothesis 5.1. These fulfil the requirements of the lemma. Indeed, the sets form a finite (f_n)-chain for each $i=0,\dots,u-1$,, since the equations (5.28a) and (5.28b) are fulfilled, and hence (5.5a) and (5.5b) are fulfilled. By Hypothesis 5.1, the last sets are disjoint. □

Lemma 5.17. *Under Hypothesis 5.1, there exist balls $C_1(i_1),C_2(i_1,i_2),\dots$ for any choice of $i_j\in\{0,\dots,u-1\}$, $j\in\mathbb{N}$, such that for each sequence (i_j), the corresponding sets $C_1(i_1),C_2(i_1,i_2),\dots$ form a t-(f_n)-chain.*

Furthermore, for any $k\in\mathbb{N}$, the u^k balls $C_k(i_1,\dots,i_k)$ corresponding to different choices of $(i_1,\dots,i_k)$ are pairwise disjoint and have radii $(\alpha\beta)^{kt}$.

Proof. Let $C_1(i_1)$ be u disjoint balls of radius $(\alpha\beta)^t$. Obviously, these are finite (f_n)-chains, so the next u^2 balls can be chosen using Lemma 5.16. In this way, we continue to obtain the disjoint balls at each step of the t-(f_n)-chain. By the construction of Lemma 5.16, these balls will have the required radii, so they fulfil the requirements of the lemma. □

These lemmas allows us to prove a theorem giving a lower bound on the Hausdorff dimension of (α,β)-winning sets.

Theorem 5.18. *Under Hypothesis 5.1, if $S\subseteq\mathcal{L}^{mn}$ is an (α,β)-winning set, then*

$$\dim_{\mathrm{H}}(S) \geq \frac{\log u}{|t\log\alpha\beta|}.$$

Proof. Let $\Lambda = \{0,\dots,u-1\}^{\mathbb{N}}$ and let $(i_j)\in\Lambda$. Clearly, $\bigcap_{j=1}^{\infty} C_j(i_1,\dots,i_j) = \{x\} = \{x(\lambda)\}\subseteq S$. We define

$$S^* = \bigcup_{\lambda\in\Lambda}\{x(\lambda)\}\subseteq S.$$

We define a surjective function (possibly multi-valued) $f: S^*\to[0,1]$ by

$$x\mapsto y = 0,i_1i_2\dots \quad \text{where } x = x(i_1,i_2,\dots)$$

where we have scaled everything by u, so that f is actually onto $[0,1]$. We extend this function to all subsets of $\mathcal{L}^{mn}$ in the following way: For $T \subseteq S^*$, let $f(T) = \bigcup_{x\in T} f(x)$. For $R \subseteq \mathcal{L}^{mn}$, let $f(R) = f(R\cap S^*)$.

Let $\mathcal{C} = (B_l)_{l\in\mathbb{N}}$ be a cover of S with balls, where B_l has radius ρ_l. Clearly, the family $\mathcal{C}^* = (B_l \cap S^*)_{l\in\mathbb{N}}$ is a cover of S^*, whence $f(\mathcal{C}^*) = (f(B_l\cap S^*))_{l\in\mathbb{N}} = (f(B_l))_{l\in\mathbb{N}}$ is a cover of $[0,1]$. Thus, the union of the sets $f(B_l)$ has outer Lebesgue measure $\overline{\mu}$ greater than 1, so by sub-additivity

$$\sum_{l=1}^{\infty} \overline{\mu}(f(B_l)) \geq 1. \tag{5.29}$$

Now, let

$$j_l = \left[\frac{\log 2\rho_l}{t\log\alpha\beta}\right].$$

For ρ_l sufficiently small, we have that $j_l > 0$ and $\rho_l < (\alpha\beta)^{tj_l}$. Hence, by the Proposition 1.5, B_l is contained in at most one ball of the form $C_{j_l}(i_1,\dots,i_{j_l})$ from Lemma 5.17. But such a ball clearly maps into an interval of length u^{-j_l}, since the image of the ball is consists of all numbers of the form $0,i_1\cdots i_{j_v}*$ in base u, where $*$ denotes any sequence of elements in $\{0,\dots,u-1\}$. Hence, $\overline{\mu}(f(B_l)) \leq u^{-j_l}$. By (5.29), we have

$$1 \leq \sum_{l=1}^{\infty} \overline{\mu}(f(B_l)) \leq \sum_{l=1}^{\infty} u^{-j_l} = \sum_{l=1}^{\infty} u^{-\left[\frac{\log 2\rho_l}{t\log\alpha\beta}\right]} \leq 2^{\frac{\log u}{|t\log\alpha\beta|}} \sum_{l=1}^{\infty} \rho_l^{\frac{\log u}{|t\log\alpha\beta|}}.$$

Now, for any such cover $\mathcal{C}$ of S with small enough balls, the s-length $l^s(\mathcal{C})$ is strictly positive for $s = \frac{\log u}{|t\log\alpha\beta|}$. This implies that $\mathcal{H}^{s-\varepsilon}(S) > 0$ for any $\varepsilon > 0$, so

$$\dim_{\mathrm{H}}(S) \geq \frac{\log u}{|t\log\alpha\beta|}.$$

□

Theorem 5.18 allows us to prove that $\dim_{\mathrm{H}}(\mathfrak{B}(m,n)) = mn$. First, we have a corollary:

Corollary 5.19. *Let $\beta \in (0,1)$ and let $N(\beta) \in \mathbb{N}$ be such that for any ball $B \in \Omega$, $\phi(B^\beta)$ contains $N(\beta)$ pairwise disjoint balls. Let $S \subseteq \mathcal{L}$ be (α,β)-winning. Then*

$$\dim_{\mathrm{H}}(S) \geq \frac{\log(N(\beta))}{|\log\alpha\beta|}.$$

Proof. This is just Theorem 5.18 with $t = 1$ and $u = N(\beta)$. □

We can now complete the proof of Theorem 5.2:

Proof of Theorem 5.2. By Corollary 5.19, we need only estimate the number $N(\beta)$ to get a lower bound for the Hausdorff dimension. This is a simple combinatorial problem. By scaling and translation, we note that it suffices to consider the $\|\cdot\|_\infty$-ball $B(0,1) = I^{mn}$.

We choose the number $i \in \mathbb{Z}$ such that $k^{i-1} < \beta \leq k^i$ and consider the family of balls $B(c,\beta) \subseteq I^{mn}$ where $c \in X^{i+1}\mathbb{F}[X]^{mn}$. By choice of i and Proposition 1.5, these are clearly disjoint. Furthermore, counting these balls we see that

$$N(\beta) = \left(k^{-i-1}\right)^{mn} = \frac{1}{k^{mn}} \frac{1}{\left(k^i\right)^{mn}} \asymp \frac{1}{\beta^{mn}}.$$

Hence, by Corollary 5.19,

$$\dim_{\mathrm{H}}(\mathfrak{B}(m,n)) \geq \frac{mn\,|\log\beta|}{|\log\alpha| + |\log\beta|} \xrightarrow[\beta\to 0]{} mn.$$

This completes the proof. □

Remark. Note, that for $m = n = 1$, this is an analogue of Jarník's Theorem (Theorem 1.4) in $\mathcal{L}$. This result was announced in [33] with a sketch of the proof. The Theorem given here is an analogue of Schmidt's Theorem on badly approximable linear forms, as mentioned in the beginning of this chapter.

6. Further research problems

The results of this first part of the thesis contributes to the development of a field previously neglected to a large extent. However, they only form a basis of a potentially larger body of knowledge concerning metrical Diophantine approximation over $\mathcal{L}$. It is the purpose of this chapter to state problems that could form a basis for further research in this field.

6.1. Continued fractions

In Chapter 1, we discussed among other thing the continued fractions algorithm. It was mentioned that the algorithm works over $\mathcal{L}$, and that most previous results in Diophantine approximation over $\mathcal{L}$ are related to this algorithm. In Chapter 2, we discussed another type of expansion of real elements in $\mathcal{L}$, and we obtained a number of metrical results on the coefficients of this expansion. A natural question would be: Is it possible to obtain similar results for the coefficients of the continued fractions expansion?

This question can be further motivated by the analogous question in the reals. In this setting, a lot of estimates have been found, and the methods used to obtain these results were quite similar to the ones we used in Chapter 2. Defining maps $T : [0,1) \to [0,1)$ and $a : [0,1) \to \mathbb{N}$ by

$$T(x) = \begin{cases} \left\{\frac{1}{x}\right\} & \text{for } x \neq 0, \\ 0 & \text{for } x = 0, \end{cases} \qquad a(x) = \begin{cases} \left[\frac{1}{x}\right] & \text{for } x \neq 0, \\ \infty & \text{for } x = 0. \end{cases}$$

Now, the partial coefficients a_n of the continued fractions expansion of any number $x \in [0,1)$ are easily seen to be $a_n(x) = a(T^{n-1}(x))$. So far, so good. This also works for $I \subseteq \mathcal{L}$, where it is just an easy consequence of Artin's algorithm (see Section 1.4).

This is where the trouble starts, because unlike the Lüroth expansions, the coefficients of the continued fractions expansion are not independent and identically distributed with respect to the Lebesgue measure (or the Haar measure in the $\mathcal{L}$-case). However, in the real case, it has been shown (see for example [10]) that T is ergodic with respect to the Gauss measure ν defined by

$$\nu(A) = \frac{1}{\log 2}\int_A \frac{1}{1+x}dx \quad \text{for } A \in \mathcal{B}([0,1)). \tag{6.1}$$

Birkhoff's Ergodic Theorem (Theorem 1.3 in [10]) now allows us to derive results of the type given in Chapter 2 for Lüroth expansions. For instance,

$$\lim_{i\to\infty} (a_1(x)\dots a_i(x))^{1/i} = \prod_{j=1}^{\infty} \left(1+\frac{1}{j^2+2j}\right)^{\log j/\log 2}.$$

To derive subsequence results, one would have to use subsequence ergodic theory — for instance by using Proposition 4 in [43]. This would impose additional restrictions on the permissible subsequences.

In $\mathcal{L}$, much less is known. In three notes ([45], [46], [47]), Paysant-Leroux and Dubois study the Jacobi–Perron algorithm in $\mathcal{L}$. This is a multidimensional analogue of the continued fractions algorithm, and hence, results about this algorithm implies results about the continued fractions algorithm.

Paysant-Leroux and Dubois introduce (in [47]) a dynamical system, which can be used to calculate the partial coefficients of this algorithm. Also, they prove that there exists a measure ν for which this system is ergodic, and that this measure is absolutely continuous with respect to the Haar measure. Hence, by the Radon–Nikodym Theorem (Theorem 6.5.4 in [48]), it is possible to express this measure as $\nu = \int f d\mu$ for some Borel function f. However, a closed form of f is not given, and this author has so far been unable to find one. Finding such a function would immediately put all the tools from ergodic theory at our disposal.

6.2. Algebraic elements

In the Chapter 4, we were examining approximation of real elements by rational elements. In particular, we were looking at the sets of real elements that were well-approximable by rationals. Another large area of research is concerned with approximating real elements with algebraic elements of degree $n \in \mathbb{N}$. That is, elements that are roots in polynomials of degree n with integer coefficients.

We define the following:

$$A_n = \{y \in \mathcal{L} : P(y) = 0 \text{ for some polynomial } P \in \mathbb{F}[X][Y], \deg P = n\}.$$

Here, $\mathbb{F}[X][Y]$ denotes the ring of polynomials in Y with coefficients from $\mathbb{F}[X]$. It can be shown that for each element a in the set A_n, there exists a primitive polynomial $P_a = c_n X^n + \cdots c_1 X + C_0$ with a as a root. Define the height $H(a)$ of $a \in A_n$ to be the unique number

$$H(a) = H(P_a) := \max_{1\le i\le n} \|c_i\|.$$

Let $\psi : \mathbb{N} \to \mathbb{R}^+$ be a decreasing function. We define the sets

$$\mathcal{K}_n(\psi) = \{x \in \mathcal{L} : \|x-a\| < \psi(H(a)) \text{ for infinitely many } a \in A_n\}.$$

Sprindžuk examined this set, as well as the analogous real and p-adic sets in [59] for a special ψ. He showed that for $\lambda > 1$ and $\psi(N) = N^{-(n+1)\lambda}$, the Haar measure (Lebesgue measure in the real case) of $\mathcal{K}_\alpha(\psi)$ is zero.

In Chapter 3, we proved a similar result, and proceeded to calculate the Hausdorff dimension of the sets. This has been done in the real case by Baker and Schmidt ([5]) for the same function ψ, but no result exists for $\mathcal{L}$.

Also, we have results for arbitrary functions ψ over the reals. In particular, if $\sum_{m=1}^{\infty} m^n \psi(m) < \infty$, Beresnevich showed ([6]) that the measure of $\mathcal{K}_\alpha(\psi)$ is zero. He also proved that the measure is full whenever the series diverges. Jarník calculated the Hausdorff measure of this set in the case when $n = 1$ ([26]), and recently Bugeaud extended this result to arbitrary $n \in \mathbb{N}$ ([12]). Clearly, a calculation of the Hausdorff measure yields the Hausdorff dimension.

None of these results have analogues in $\mathcal{L}$, but it seems reasonable to conjecture that similar results hold.

6.3. Inhomogeneous linear forms

This final research proposal is somewhat more ambitious than the previous ones. Indeed, the obvious way to start looking for a solution involves extending deep theory from various parts of mathematics to $\mathcal{L}$, which at least the present author has no idea of how to go about, save sheer will, determination and effort.

The problem itself looks innocent enough. We have already seen that the set of badly approximable linear forms in $\mathcal{L}^{mn}$ has full Hausdorff dimension (Theorem 5.2). How about affine forms? We first define the space of these. For any set of points $(A, b) \in \mathcal{L}^{mn} \times \mathcal{L}^n$, we obtain an affine function $q \mapsto qA + b$, $q \in \mathbb{F}[X]^m$. An affine form is badly approximable if

$$\liminf_{\substack{p \in \mathbb{F}[X]^n \\ q \in \mathbb{F}[X]^n \\ \|q\|_\infty \to \infty}} \|qA + b + p\|_\infty^n \|q\|_\infty^m > 0.$$

For $b = 0$, this is equivalent to Definition 5.1. The property is defined analogously for the reals.

We make another definition. Let $h \in \mathbb{N}$. A set $E \in \mathbb{R}^h$ is said to be thick if for any open subset $W \subseteq \mathbb{R}^h$, $\dim_{\mathrm{H}}(W \cap E) = \dim_{\mathrm{H}}(E)$. Clearly, this implies full Hausdorff dimension of E. In fact, it implies full Hausdorff dimension at any point in $\mathbb{R}^h$. Since thickness is defined solely in terms of Hausdorff dimension, the notion extends naturally to $\mathcal{L}^h$.

Kleinbock examined the set of badly approximable affine forms in [30], where he proved that the set is thick. This leads to the question of whether or not the same is true for $\mathcal{L}$. The obvious way of proving this would be to extend Kleinbock's method to $\mathcal{L}$. However, Kleinbock uses methods involving flows on real Lie groups, and the corresponding tools would be very difficult to extend to $\mathcal{L}$. It might be possible, but it would surely be a lot of work.

Part II.

Gaussian approximation in ergodic theory

Contents

7. Introduction

In this part of the thesis, we discuss Gaussian approximation in ergodic theory. We begin with a discussion of the setting itself. An important element in the discussion is the links between probability theory and ergodic theory. In particular, we will emphasise the possible uses of Gaussian random variables in ergodic theory and some of the possible forms of the Central Limit Theorem in the setting of ergodic theory. The aim of this discussion is to clarify the topic of this part of the thesis.

Following this preliminary discussion, we will discuss the several approaches to Gaussian approximation in ergodic theory. We will discuss the methods in turn. Along the way, we will also give some results obtained by these methods. In most cases, we spare the reader a number of technical details of the proofs of the results. The full details are given in cited papers. The only cases where we do in fact complete the proofs are the ones where the original paper left some doubt — at least in the mind of the present author — as to the validity of the method.

We begin with some results relating to sequences of random variables, that are not independent, but which do fulfil the Central Limit Theorem. Namely lacunary trigonometric series. Here, we will discuss two different approaches to the problem. One is solely related to the Central Limit Theorem, and the other to the more general setting of Gaussian approximation.

The first approach related directly to dynamical systems involves spectral theory. This method forms the basis of a large number of the previous results in the field. Here, one construct functions such that the appropriate form of the Central Limit Theorem is fulfilled by it's partial sums solely by considering the Fourier coefficients of the function. This gives good control over the moments and correlations of the partial sums, and since Fourier series are trigonometric, the results in the preceding section can be applied to this setting. However, only a limited number of dynamical systems may be examined in this way.

In Section 7.4, we will discuss one way of moving from the limited number of possible systems that can be treated by the method from spectral theory to more general systems. This construction involves the so-called Rokhlin towers, which is a classical construction in ergodic theory.

Using Rokhlin towers, it is possible to obtain more general results than the ones obtained solely by spectral theory. We introduce some fundamental tools in Section 7.5, which we will need in Chapter 8.

7.1. The Central Limit Theorem and Gaussian random variables

We will begin with a brief discussion on the similarities between ergodic theory and the theory of independent and identically distributed random variables. We state two classical theorems from each setting.

Theorem 7.1 (Birkhoff's Ergodic Theorem). *Let $(X,\mathcal{B},\mu)$ be a probability space, let $T : X \to X$ be an ergodic measure preserving transformation and let $f \in L^1(X,\mathcal{B},\mu)$.*

$$\lim_{m\to\infty} \frac{\sum_{j=0}^{m-1} f(T^j x)}{m} = \int f d\mu \quad a.e.$$

Theorem 7.2 (The Strong Law of Large Numbers). *Let $(X_j)_{j=0}^{\infty}$ be a sequence of independent and identically distributed random variables on some probability space $(X,\mathcal{B},\mu)$.*

$$\lim_{m\to\infty} \frac{\sum_{j=0}^{m-1} X_j}{m} = \mathbb{E}(X_0) \quad a.e.$$

We note that even though we do not necessarily have independence in the ergodic case, the Strong Law of Large Numbers does indeed hold for the sequence of random variables $(f \circ T^j)_{j=0}^{\infty}$ in an ergodic system for any L^1-function f. This poses the natural question: Do other classical results from probability theory transfer to (ergodic) dynamical systems?

The probabilistic theorem, we are especially interested in, is the Central Limit Theorem. We state the following form:

Theorem 7.3 (The Central Limit Theorem). *Let $(X_j)_{j=0}^{\infty}$ be a sequence of independent and identically distributed random variables on some probability space $(X,\mathcal{B},\mu)$ with $\mathbb{E}(X_0) = 0$, $\sigma^2 = \mathbb{E}\left(X_0^2\right) < \infty$.*

$$\frac{\sum_{j=0}^{m-1} X_j}{\sigma\sqrt{m}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1),$$

where the convergence is in distribution as m tends to infinity and $\mathcal{N}(0,1)$ denotes the standard normal (Gaussian) distribution.

A related theorem is the Almost Sure Central Limit Theorem:

Theorem 7.4 (The Almost Sure Central Limit Theorem). *Let $(X_j)_{j=1}^{\infty}$ be a sequence of independent and identically distributed random variables on some probability space $(X,\mathcal{B},\mu)$ with $\mathbb{E}(X_1) = 0$, $\sigma^2 = \mathbb{E}\left(X_1^2\right) = 1$. Let $S_m = X_1 + \cdots + X_m$. Then*

$$\frac{1}{\log m} \sum_{j=1}^{m} \frac{1}{j} \delta_{\{S_j/\sqrt{j}\}} \xrightarrow[m\to\infty]{\mathcal{D}} \mathcal{N}(0,1) \quad a.s.$$

where δ_z is the Dirac point-measure at the point z.

Whereas the Central Limit Theorem is a theorem about the overall Gaussian behaviour of random variables, the Almost Sure Central Limit Theorem is about the Gaussian behaviour of the orbits of points in the space on which the random variables are defined.

We do not prove these theorems for random variables. It is sufficient for the purposes of this thesis to make the reader aware of their existence and validity. What we are interested in, is the analogous results in measure preserving systems.

Throughout this part of the thesis, we will denote the distribution function of the standard normal distribution by $\Phi(u)$.

A few questions are natural following Theorem 7.3. First of all, can the independence condition be weakened or omitted? This question has been extensively examined by Philipp and Stout ([51]), who developed a method allowing them to make conclusions of this form for a number of sequences of random variables in terms of Brownian motions. The results obtained by Philipp and Stout concerns a number of specific cases of dependent random variables for which these results hold. Another question along the same lines would be to ask, whether or not similar results hold in dynamical systems, and if so, what conditions are needed on the dynamical system (ergodicity, mixing, aperiodicity etc.).

7.2. Lacunary trigonometric series

One of the first interesting papers in the context of Gaussian approximation and the Central Limit Theorem for non-independent random variables is a paper in two parts by Salem and Zygmund ([53], [54]). In this paper, the authors consider the particular random variables on the unit interval $\mathbb{T} = [0,1)$ which are trigonometric series:

$$Y_k = a_k \cos 2\pi l_k x + b_k \sin 2\pi l_k x, \quad \text{where } k \in \mathbb{N}. \tag{7.1}$$

Not surprisingly, some assumptions on the sequences involved are needed in order to prove the Central Limit Theorem for these random variables. In particular, the case that Salem and Zygmund examined was the case where the sequence (l_i) is lacunary.

The partial sums of the functions in (7.1) are Fourier series, and hence they cover quite a lot of territory. This is the reason why their theorem has been used a lot in the subsequent work on the Central Limit Theorem in dynamical systems. We will return to some applications of their theorem in the next section. Also, their methods are based on calculations on Fourier series. This is in analogy with one of the methods used in subsequent sections.

We will begin with noting that for the examination of the partial sums of random variables on the form (7.1), it is sufficient to consider random variables that only depend on cosine. Indeed,

$$a_k \cos 2\pi l_k x + b_k \sin 2\pi l_k x = \sqrt{a_k^2 + b_k^2} \cos\left(2\pi l_k x + \omega_k\right)$$

for appropriate choice of ω_k. Hence, if we can we prove the Central Limit Theorem for the random variables

$$X_k = a_k \cos 2\pi l_k x \tag{7.2}$$

with certain assumptions on the l_k and a_k, we will obtain an analogous theorem for the Y_k as defined in (7.1). Salem and Zygmund proved the following:

Theorem 7.5 (The Central Limit Theorem for Lacunary Trigonometric Series). *Let $(l_k) \subseteq \mathbb{R}^+$ be a sequence such that $\frac{l_{k+1}}{l_k} > q > 1$ for some q. Let $(a_k) \subseteq \mathbb{R} \setminus \{0\}$ be some sequence and define for $m \in \mathbb{N}$*

$$S_m(x) = \sum_{k=1}^{m} a_k \cos 2\pi l_k x, \quad A_m = \left(\frac{1}{2} \sum_{k=1}^{m} a_k^2 \right)^{1/2}.$$

Suppose that $A_m \to \infty$ and $|a_m|/A_m \to 0$ as m tends to infinity. Then

$$\frac{S_m}{A_m} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1).$$

Sketch of proof. First, we reduce the theorem to a simpler statement. It is sufficient to prove that for $E \in \mathcal{B}([0,1))$ with $\mu(E) > 0$,

$$\mu(E)^{-1} \int_E e^{i\lambda \frac{S_m(x)}{A_m}} dx \xrightarrow[m\to\infty]{} e^{\lambda^2/2} \tag{7.3}$$

uniformly for a finite range of λ. That is, we may assume that $\lambda = O(1)$. Since the intervals generate $\mathcal{B}([0,1))$, we assume without loss of generality that $E = (a,b)$.

Now, let $\eta > 0$ and define the function

$$K_\eta(x) = \begin{cases} \frac{x-a}{\eta} & \text{for } x \in (a, a+\eta), \\ 1 & \text{for } x \in [a+\eta, b-\eta], \\ \frac{b-x}{\eta} & \text{for } x \in (b-\eta, b), \\ 0 & \text{otherwise.} \end{cases}$$

Letting $k_\eta = \int K_\eta(x)dx$, we see that it is sufficient to prove

$$k_\eta^{-1} \int_a^b K_\eta(x) e^{i\lambda \frac{S_m(x)}{A_m}} dx \xrightarrow[m\to\infty]{} e^{\lambda^2/2} \tag{7.4}$$

uniformly, since letting η tend to zero yields (7.3).

Let $f_{k,m}(x) = i\lambda A_m^{-1} a_k \cos(2\pi l_k x)$ for $1 \le k \le m$ and note that since $|a_m|/A_m \to 0$, $|f_k(x)| \to 0$ as m tends to infinity. Since $e^z = (1+z)e^{z^2/2+o(|z|^2)}$ as $z \to 0$, we can write

$$\begin{aligned} \int_a^b K_\eta(x) e^{i\lambda \frac{S_m(x)}{A_m}} dx &= \int_a^b K_\eta(x) \prod_{k=1}^{m} e^{f_k(x)} dx \\ &= \int_a^b e^{o(1)} K_\eta(x) \left(\prod_{k=1}^{m} \left(1 + \frac{i\lambda a_k}{A_m} \cos(2\pi l_k x) \right) \right) e^{\left(-\frac{\lambda^2}{2A_m^2} \sum_{k=1}^{m} \frac{a_k^2}{A_m^2} \cos^2(2\pi l_k x) \right)} dx. \end{aligned} \tag{7.5}$$

By the various assumptions, $e^{o(1)}$ tends to one uniformly. Consider now the argument of the exponential function in (7.5). A double angle formula gives

$$\sum_{k=1}^{m} \frac{a_k^2}{A_m^2} \cos^2(2\pi l_k x) = 1 + \sum_{k=1}^{m} \frac{a_k^2}{2A_m^2} \cos(4\pi l_k x) = 1 + \xi_m(x).$$

Using Chebychev's inequality and lacunarity of (l_k),

$$\mu\{x \in (a,b) : |\xi_m(x)| \geq \delta\} = o(1)$$

for any $\delta > 0$. Hence, inserting the above in (7.5) we get for $m \to \infty$,

$$\int_a^b K_\eta(x) e^{i\lambda \frac{S_m(x)}{A_m}} dx \sim e^{-\lambda^2/2} \int_a^b K_\eta(x) \prod_{k=1}^{m} \left(1 + \frac{i\lambda a_k}{A_m} \cos(2\pi l_k x)\right). \tag{7.6}$$

We write the product above as a Fourier series $\sum_\nu c_\nu \cos 2\pi \nu x$. The lacunarity of (l_k) implies certain bounds on the solution to some Diophantine inequalities, which in turn implies nice bounds on the Fourier coefficients a_ν. In this way, one can prove that

$$\int_a^b K_\eta(x) \prod_{k=1}^{m} \left(1 + \frac{i\lambda a_k}{A_m} \cos(2\pi l_k x)\right) \to k_\eta$$

uniformly as $m \to \infty$. Inserting in (7.6), we get (7.4). □

In later sections of this part of the thesis, we will be looking at dynamical systems, where we approximate each partial sum with Gaussian random variables. Such a result also exists for lacunary trigonometric series (see [51], Chapter 6). We include the result here, since it is an early example of a method, which — albeit in a different disguise — we will use to prove results in ergodic theory later on. Philipp and Stout prove theorems similar to the following for the partial sums of a number of sequences of different types of weakly dependent random variables.

Theorem 7.6 (The Gaussian Approximation Theorem for Lacunary Trigonometric Series). *Let (l_k), (a_k) and A_m be as in the statement of Theorem 7.5. Let*

$$S(t) = S(t,x) = \sum_{k=1}^{m} a_k \cos 2\pi l_k x \quad \text{for } t \in [A_m^2, A_{m+1}^2).$$

Assume that $A_m \to \infty$ as m tends to infinity and that there exists a $\delta \in (0,1]$ such that $a_m \ll A_m^{1-\delta}$. Then, we can refine the process $S(t)$ on a richer probability space without changing the distribution such that for every $\lambda < \delta/32$,

$$S(t) - X(t) \ll t^{1/2-\lambda} \quad a.s.$$

where $X(t)$ is the standard Brownian motion.

In fact, Theorem 7.6 is (almost) a generalisation of Theorem 7.5. To see this, let $S_m(t) = S(A_m^2 t)$ for $0 \le t \le 1$. By Theorem 7.6,

$$S_m(t) - X(A_m^2 t) \ll A_m A_m^{-2\lambda} t^{1/2-\lambda}.$$

Dividing by A_m and using standard facts about the standard Brownian motion, we see that

$$\frac{S_m(t)}{A_m} \xrightarrow{\mathcal{D}} X(t) \quad \text{for } t \in [0,1].$$

Inserting $t = 1$ and using Proposition 12.4 in [11], we get the convergence claimed in Theorem 7.5.

Note that $a_m \ll A_m^{1-\delta}$ implies that $|a_m| / A_m \to 0$, but not vice versa. Hence Theorem 7.6 generalises Theorem 7.5 with the weaker assumption on the coefficients.

Sketch of proof of Theorem 7.6. In all the cases examined by Philipp and Stout in [51], they use the same method. We only give a sketch of the method, since the technical details get quite involves. Their main idea is to use the fact that under appropriate circumstances, every element in a sequence of random variables can be written as a martingale plus a co-boundary. In fact, one has the following easy lemma:

Lemma 7.7. *Let $(X_j)_{j\in\mathbb{N}}$ be some sequence of random variables and let $(\mathcal{B}_j)_{j=0}^{\infty}$ be a non-decreasing sequence of σ-fields such that X_j is $\mathcal{B}_j$-measurable for all $j \in \mathbb{N}$, and $\mathcal{B}_0$ is the trivial σ-field. Let*

$$u_j = \sum_{k=1}^{\infty} \mathbb{E}\left(X_{j+k} | \mathcal{B}_{j-1}\right) \quad \textit{for } j \in \mathbb{N}$$

and assume that

$$\sum_{k=0}^{\infty} \mathbb{E}\left(\left|\mathbb{E}\left(X_{j+k} | \mathcal{B}_j\right)\right|\right) < \infty \quad \textit{for } j \in \mathbb{N}. \tag{7.7}$$

Then for any $j \in \mathbb{N}$,

$$X_j = Y_j + (u_j - u_{j+1}),$$

where $\{Y_j, \mathcal{B}_j\}_{j\in\mathbb{N}}$ is a martingale.

Proof of Lemma 7.7. The sequence

$$Y_j = \sum_{k=1}^{\infty} \left(\mathbb{E}\left(X_{j+k} | \mathcal{B}_j\right) - \mathbb{E}\left(X_{j+k} | \mathcal{B}_{j-1}\right)\right)$$

fulfils the requirements of the lemma. □

With the lemma in place, one splits the partial sums up into blocks $S_m = \sum_{j=1}^{M} X_j$, where $X_j = \sum_{k\in I_j} a_k \cos(2\pi l_k x)$. The blocks I_j may be chosen such that for large j, the transfer function u_j of the co-boundary is small in comparison with X_j and such that

(7.7) holds, where the $\mathcal{B}_j$ are the σ-fields generated by $X_1,\ldots,X_j$. With some small errors, Lemma 7.7 implies,

$$X_j \sim Y_j \quad \text{and} \quad S_m = \sum_{j=1}^{M} X_j \sim \sum_{j=1}^{M} Y_j.$$

Hence, it suffices to prove the theorem for the partial sums of the martingale Y_j. For this, one can use Skorohkod's Representation Theorem (Chapter 7.2 in [58]) to find non-negative random variables T_j such that

$$\sum_{j=1}^{M} Y_j = X\left(\sum_{j=1}^{M} T_j\right) \quad a.s.,$$

$$\mathbb{E}\left(T_j | \mathcal{B}_{j-1}\right) = \mathbb{E}\left(Y_j^2 | \mathcal{B}_{j-1}\right) \quad a.s.,$$

$$\mathbb{E}\left(T_j^p\right) \ll \mathbb{E}\left(\left|Y_j\right|^{2p}\right) \quad \text{for } p > 1,$$

where X is the standard Brownian motion. With these tools, it is possible to prove the theorem. We refer to [51] for the technical details. □

7.3. Spectral theory

Having seen that theorems of the type in which we are interested exist for certain trigonometric series, we will now consider a particular example of an ergodic dynamical system. Namely the case of an irrational rotation of the circle.

For the remainder of this section, we let $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, $\mathbb{T} = [0,1)$ and let $T : \mathbb{T} \to \mathbb{T}$ be defined by $x \mapsto \alpha + x \pmod 1$. This defines a measure preserving system $(\mathbb{T}, \mathcal{B}, \mu, T)$, where μ is the Lebesgue measure. Equivalently, we can map $\mathbb{T}$ to the unit circle in the complex plane and consider the map $T(e^{2\pi i x}) = e^{2\pi i(\alpha + x)}$. This is another description of the same dynamical system.

We are interested in finding functions $f \in L^2(\mathbb{T})$ such that the Central Limit Theorem holds for the normalised partial sums $S_m f / \|S_m f\|_2$, where $S_m f = \sum_{j=0}^{m-1} f \circ T^j$. Since real functions on $\mathbb{T}$ are periodic functions with period 1, we may use Fourier analysis to describe these. In particular, we may assign a Fourier series to each function in $L^2(\mathbb{T})$ and — under certain assumptions on the Fourier coefficients — vice versa. Since Fourier series are trigonometric series, we may in under appropriate circumstances use the results in Section 7.2 to obtain similar results for irrational rotations. A number of different approaches have been taken to this question. We discuss them chronologically.

A general fact about partial sums of Fourier series is that it is particularly easy to calculate their partial sums under irrational rotations. Note that the eigenfunctions of

T are $e_j(x) = e^{2\pi i j x}$ with eigenvalues $Te_j = e^{2\pi i j\alpha} e_j$. Hence, if we let f be defined in terms of it's Fourier series,

$$f(x) = \sum_{n=-\infty}^{\infty} a_n e^{2\pi i n x} = \sum_{n=-\infty}^{\infty} a_n e_n(x),$$

we obtain

$$S_m f(x) = \sum_{j=0}^{m-1} T^j \sum_{n=-\infty}^{\infty} a_n e_n(x) = \sum_{n=-\infty}^{\infty} a_n \left(\sum_{j=0}^{m-1} e^{2\pi i j\alpha} \right) e_n(x)$$
$$= \sum_{n=-\infty}^{\infty} a_n \frac{1 - e^{2\pi i n m\alpha}}{1 - e^{2\pi i n\alpha}} e_n(x).$$

This is also a Fourier series, where the coefficients from the original series have been scaled by certain fractions. Since the e_n are orthogonal in $L^2(\mathbb{T})$, this implies

$$\|S_m f\|^2 = \sum_{n=-\infty}^{\infty} a_n^2 \left| \frac{1 - e^{2\pi i n m\alpha}}{1 - e^{2\pi i n\alpha}} \right|^2.$$

To simplify notation, we define the functions

$$V_m(\theta) = \frac{1 - e^{2\pi i m\theta}}{1 - e^{2\pi i\theta}}. \tag{7.8}$$

These functions are known as the spectral kernels.

7.3.1. Irrational rotations

The first Central Limit Theorem for irrational rotations was proved by Burton and Denker in [13]. They subsequently used this theorem to deduce a general Central Limit Theorem for aperiodic dynamical systems, but unfortunately there is a mistake in this second proof. We will discuss this in more detail in Section 7.4.

Theorem 7.8 ([13], Theorem 1a). *There exists a function $f \in L^2(\mathbb{T})$ with $\mathbb{E}(f) = 0$ such that*

$$\frac{S_m f}{\|S_m f\|} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1). \tag{7.9}$$

Proof. We follow the proof given by Burton and Denker. The strategy is to reduce the theorem to a consequence of Theorem 7.5 by clever choice of f in terms of it's Fourier coefficients. We take sequences $\{l_k\} \subseteq \mathbb{N}$, $\varepsilon_k \to 0$, $\bar{\varepsilon}_k \to 0$ and $\{a_k\} \subseteq \mathbb{R}$ such that

$$\sum_{k=1}^{\infty} a_k^2 l_k < \infty, \tag{7.10}$$

$$\forall k \in \mathbb{N} : \varepsilon_k > \varepsilon_k - \bar{\varepsilon}_k > \tfrac{\varepsilon_k}{2}. \tag{7.11}$$

Let $\beta = e^{2\pi i\alpha}$. By Weyl's Equidistribution Theorem, there are infinitely many $j \in \mathbb{N}$ such that $\left|1-\beta^j\right| \in (\varepsilon_k - \bar{\varepsilon}_k, \varepsilon_k)$ and β^j is in the first quadrant of the complex plane. Hence, we may choose a lacunary sequence of integers (j_k) such that these two properties hold. We split the sequence up into consecutive blocks J_k of size l_k and define a sequence (b_j) by

$$b_j = \begin{cases} a_k & \text{for } j \in J_k \text{ or } -j \in J_k, \\ 0 & \text{otherwise,} \end{cases}$$

and a function $f \in L^2(\mathbb{T})$ by

$$f(x) = \sum_{n=-\infty}^{\infty} b_n e^{2\pi i n x} = \sum_{n=1}^{\infty} 2b_n \cos(2\pi n x).$$

By (7.10), f is indeed in L^2.

As in the beginning of this section, we obtain

$$\|S_m f\|^2 = \sum_{n=1}^{\infty} 2b_n^2 |V_m(n\alpha)|^2 = \sum_{n=1}^{\infty} 2a_n^2 \sum_{j\in J_n} |V_m(n\alpha)|^2.$$

Now, let $j \in J_n$ and $m \in \mathbb{N}$ be such that $m\varepsilon_n \leq 1$. We have the following inequality:

$$0 \leq m(1-m\varepsilon_n) < |V_m(n\alpha)| \leq m. \tag{7.12}$$

Indeed, the last inequality follows since

$$\left|\frac{1-\beta^{mj}}{1-\beta^j}\right| = \left|\sum_{i=0}^{m-1} \left(\beta^j\right)^i\right| \leq m$$

by the triangle inequality. The first inequality is most easily explained by a picture. The denominator of the fraction is strictly less than ε_n, so since $m\varepsilon_n \leq 1$, the fraction is strictly greater than $m\left|1-\beta^{mj}\right|$. Since β^j is in the first quadrant, $\left|1-\beta^j\right| < \varepsilon_n$ and $m\varepsilon_n \leq 1$, β^{jm} does not go full circle. In fact, we have the situation of Figure 7.1 on the following page. Note, that we must have $l_2 > (1-m\varepsilon_n)$. Hence the inequality.

Note that $4\left|1-\beta^j\right|^{-2} \geq |V_m(j\alpha)|^2$. Letting $L(n) = 2a_n^2 l_n$, we obtain by (7.12),

$$L(n_0)m^2(1-m\varepsilon_{n_0}) \leq \|S_m f\|^2 \leq \sum_{n<n_0} 16L(n)\varepsilon_n^{-2} + m^2 L(n_0) + m^2 \sum_{n>n_0} L(n), \tag{7.13}$$

for any n_0 and any m with $m\varepsilon_{n_0} \leq 1$. We can choose the function $L(n)$ freely, as long as (7.10) remains valid. Hence, we choose $L(n) = 2^{-\gamma n^2}$, where $\gamma \in (0,2)$ is arbitrary.

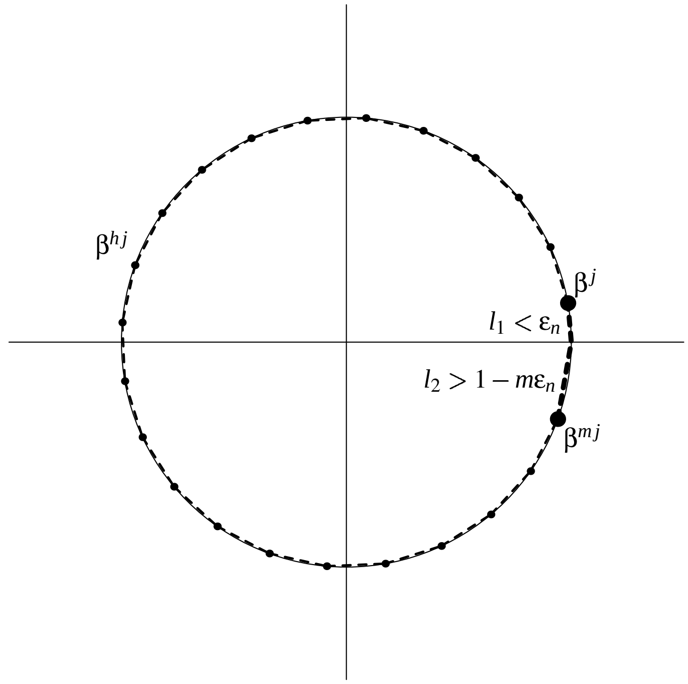

Figure 7.1.: Rotation of β^j (dashed lines are straight).

Now, we take the set $N_0 = \{2^{n^2} : n \in \mathbb{N}\}$ and the numbers $\varepsilon_n = 2^{-n^2-n}$ and $m = 2^{n_0^2}$, where n_0 is chosen so large that

$$\begin{aligned} m^{-\gamma} m^2 (1 - 2^{-n_0}) &\leq \|S_m f\|^2 \\ &\leq 4 \sum_{n<n_0} \left(2^{-\gamma n^2}\right)\left(2^{2n^2+2n}\right) + m^{2-\gamma} + m^2 \sum_{n>n_0} 2^{-\gamma n^2} \\ &\leq 8\left(2^{(2-\gamma)(n_0-1)^2}\right) + m^{2-\gamma} + 2m^2 2^{-\gamma(n_0+1)^2}. \end{aligned}$$

This is possible by (7.13). Hence, $\|S_m f\|^2 \sim m^{2-\gamma}$, and $S_m f$ is well approximated in $L^2(\mathbb{T})$ by the Fourier series restricted to block J_{n_0} for $m = 2^{n_0^2}$, as is seen by calculating the L^2-distance between the two. Hence,

$$S_m f \sim a_{n_0} \sum_{|j| \in J_{n_0}} V_m(j\alpha) e_j$$

when $m = 2^{n_0^2} \to \infty$. For any $\varepsilon' > 0$, we may choose n_0 so large that $\bar{\varepsilon}_n < \varepsilon'$ for any

$n > n_0$. Hence, with arbitrarily small L^2-error,

$$V_m(j\alpha)e_j(x) + V_m(-j\alpha)e_{-j}(x) \\ = \frac{1-\beta^{jm}}{1-\beta^j}e_j(x) + \frac{1-\beta^{-jm}}{1-\beta^{-j}}e_{-j}(x) \sim C_{n_0}\cos(2\pi jx) + D_0\sin(2\pi jx),$$

where C_{n_0} and D_{n_0} depend on ε_{n_0}. Normalising, we see that for $m = 2^{n_0^2}$,

$$\frac{S_m f}{\|S_m f\|} = \frac{1}{A_{n_0}\sqrt{k_{n_0}}} \sum_{j\in J_{n_0}} C_{n_0}\cos(2\pi jx) + D_0\sin(2\pi jx),$$

where $A_{n_0} = \frac{1}{2}\sqrt{C_{n_0}^2 + D_{n_0}^2}$. Since we are still free to choose l_{n_0}, Theorem 7.5 (and the remarks preceding it) implies that the theorem holds along the sequence N_0.

We will now extend the theorem to all of $\mathbb{N}$. For large m, we define the number $n_0 = \sup\{n : 2^{n^2} \leq m\}$. The variance of $S_m f$ is concentrated on the blocks J_{n_0} and J_{n_0+1}. Hence,

$$\frac{S_m f}{\|S_m f\|} = \frac{1}{\|S_m f\|} \sum_{n=1}^{\infty} 2a_k \sum_{|j|\in J_n} V_m(j\alpha)\cos(2\pi jx), \tag{7.14a}$$

and

$$\frac{a_{n_0}\sum_{|j|\in J_{n_0}} V_m(j\alpha)e_j + a_{n_0+1}\sum_{|j|\in J_{n_0+1}} V_m(j\alpha)e_j}{2a_{n_0}^2\sum_{j\in J_{n_0}} |V_m(j\alpha)|^2 + 2a_{n_0+1}^2\sum_{j\in J_{n_0+1}} |V_m(j\alpha)|^2}, \tag{7.14b}$$

have the same limit distribution as $m \to \infty$. By the argument used above to obtain the Central Limit Theorem along a subsequence, we see that for $2^{n^2} \leq m < 2^{(n+1)^2}$, the expression

$$A_m^{-1}\left(a_n \sum_{|j|\in J_n} V_m(j\alpha)e_j + a_{n+1}\sum_{|j|\in J_{n+1}} V_m(j\alpha)e_j\right)$$

tends to $\Phi(u)$. For $2^{(n-1)^2} \leq m < 2^{n^2}$, the expression

$$A_m^{-1}\left(a_{n-1} \sum_{|j|\in J_{n-1}} V_m(j\alpha)e_j + a_n\sum_{|j|\in J_n} V_m(j\alpha)e_j\right)$$

also tends to $\Phi(u)$, where

$$A_m^2 = \frac{1}{2}\left(a_n^2 \sum_{j\in J_n} (2\mathfrak{Re}\,(V_m(j\alpha)))^2 + (2\mathfrak{Im}\,(V_m(j\alpha)))^2\right.$$

$$\left. + a_{n\pm1} \sum_{j\in J_{n\pm1}} (2\mathfrak{Re}\,(V_m(j\alpha)))^2 + (2\mathfrak{Im}\,(V_m(j\alpha)))^2\right)$$

$$= 2a_n^2 \sum_{j\in J_n} |V_m(j\alpha)|^2 + 2a_{n\pm1}^2 \sum_{j\in J_{n\pm1}} |V_m(j\alpha)|^2$$

and $\Phi(u)$ is the standard normal distribution. This completes the proof. □

In fact, if we define the set

$$\mathrm{CLT}(\sigma(m)) = \left\{ f \in L^2(\mathbb{T}) : \text{Theorem 7.8 holds for } f, \|S_m f\| \asymp \sigma(m) \right\}, \tag{7.15}$$

the proof of Theorem 7.8 implies the following corollary:

Corollary 7.9. *For any* $\gamma \in (0,2)$, $\mathrm{CLT}(m^{-\gamma})$ *is dense in* $L^2(\mathbb{T})$.

Proof. By (7.13) and choice of $L(n)$, it is clearly possible to choose elements in $\mathrm{CLT}(n^{-\gamma})$ for each value of $\gamma \in (0,2)$. A closer examination on our freedom of choice of the Fourier coefficients and Parseval's equality shows, that one such element exists arbitrarily close to any $g \in L^2(\mathbb{T})$. This completes the proof. □

Kato made the definition of the set $\mathrm{CLT}(\sigma(m))$ in [28], where he also proved the following:

Theorem 7.10. *Let* $\gamma > 0$. $\mathrm{CLT}(m^2(\log m)^{-\gamma})$ *is dense in* $L^2(\mathbb{T})$.

In the beginning of his proof, Kato defines series $(l_k) \subseteq \mathbb{N}$ and $(a_k) \subseteq \mathbb{R}$, where (l_k) is lacunary and $a_k = \frac{1}{k^{(1+\gamma)/2}}$. Kato's proof consists of splitting the partial sums up into three,

$$S_m f(x) = \sum_{j=0}^{m-1} T^j f(x) = \sum_{|n|\le n_0(m)} a_n V_{l_n}(p\alpha) e_{l_n}(x)$$

$$+ m \sum_{|n|>n_0(m)} a_n e_{l_n}(x) + \sum_{|n|>n_0(m)} a_n \left(V_{l_n}(p\alpha) - m\right) e_{l_n}(x),$$

where $n_0(m) = \min\{n > 0 : \{n\alpha\} \le \frac{1}{2m(\log m)^{1/6}}\}$. Now, he divides by $\|S_m f\|$ and obtains three summands of the normalised partial sums, Σ_1, Σ_2 and Σ_3. Using classical estimates, he proves that Σ_1 and Σ_3 converges to zero in probability as m tends to infinity. For the last summand, Theorem 7.5 applies.

In fact, the estimates made on the three summands are not obtained quite as easily as the above might suggest. Rather intricate arguments from spectral theory are used along with some classical estimates on the $V_n(\theta)$. The details are omitted, but the result deserves mention here.

To obtain density of the set, one notes that the estimates in the proof of the particular case only depends on Fourier coefficients corresponding to arbitrarily small sets having given values. Thus, the density is obtained.

We will not go further into the method of dividing the partial sums up into blocks at this point. There will be plenty of block-dividing arguments in what follows.

7.3.2. Diophantine criteria

The above result by Burton and Denker was subsequently generalised by Lacey (see [38]). Lacey proved a very beautiful result, connecting Diophantine approximation and Gaussian approximation. He showed that the modulus of continuity of a function, such that the Central Limit Theorem holds for it's partial sums under an irrational rotation of the circle, depends on the Diophantine type of the rotation number $\alpha \in \mathbb{R} \setminus \mathbb{Q}$. We make some definitions.

Definition 7.1. Let $\alpha \in \mathbb{R} \setminus \mathbb{Q}$. The *Diophantine type* β of α is defined to be

$$\beta = \inf\left\{ b > 0 : \left|\alpha - \tfrac{p}{q}\right| > C_b q^{-1-b} \text{ for any } p,q \in \mathbb{Z} \text{ for some } C_b > 0 \right\}$$

Relating this to the material in Part I, we note that the Diophantine type of an irrational number is a measure of how badly approximable the number is. In particular, we see that if $\beta \geq 1$, then α is badly approximable (the real one-dimensional analogue of Definition 5.1).

Definition 7.2. Let $f : \mathbb{T} \to \mathbb{R}$ be continuous and $0 < a < 1$. We say that $f \in \text{lip}(a)$ if

$$|f(x) - f(y)| < C\,|x-y|^a$$

for some constant C.

Lacey proved the following theorem:

Theorem 7.11. *Let α be of Diophantine type β. If $a < \frac{1}{2\beta}$ there is an $f \in \text{lip}(a)$,*

$$\frac{S_m f}{\sqrt{m}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1). \tag{7.16}$$

If $a > \frac{1}{2\beta}$ there does not *exist an $f \in \text{lip}(a)$ such that* (7.16) *holds.*

Remark. In fact, Lacey proved a stronger result (Theorem 1.4 in [38]), involving the convergence of the partial sums (in appropriate normalisation) to a standard fractional Brownian motion. For simplicity, we stay with the above form of the Central Limit Theorem.

Lacey needed a few auxiliary results in order to prove Theorem 7.11. We give these here. First of all, equation (1.2) from our sketch of the proof of Dirichlet's Theorem (Theorem 1.1) along with Definition 7.1 immediately implies:

Proposition 7.12. *Let α be irrational of Diophantine type β and define the sequence $d_n = |\alpha q_n - p_n|$. Then for any $\varepsilon > 0$ we have for large n, that $d_n > 1/q_n^{\beta+\varepsilon}$, and there exists an infinite subsequence (n_j) such that $d_{n_j} < 1/q_{n_j}^{\beta-\varepsilon}$.*

As in the previous section, we will be looking at a function defined as a cosine series with Fourier coefficients a_k. This time, we also include a sequence (r_k) of Rademacher functions in the series:

$$f(x) = \sum_{k=1}^{\infty} 2a_k r_k \cos(2\pi kx). \tag{7.17}$$

For such functions, we have (see [27], Theorem 3, page 68):

Theorem 7.13. *If there exists an a, $0 < a < 1$ such that*

$$s_j = \left(\sum_{2^j \le k < 2^{j+1}} a_k^2 \right)^{1/2} = O(2^{-aj})$$

then $f \in \operatorname{lip}(a)$ almost surely.

With the tools in place, we embark on a sketch of the proof of Theorem 7.11.

Sketch of the proof of Theorem 7.11. We prove the negative part in detail. Let $a > \frac{1}{2\beta}$ and assume that $f \in \operatorname{lip}(a)$ is such that for any $u > 0$,

$$\mu\left\{ x \in \mathbb{T} : \left| \frac{S_m f}{\sqrt{m}} \right| > u \right\} \to \phi(u),$$

where $\phi \not\equiv 0$ is some non-trivial distribution function. We wish to arrive at a contradiction. Let $\varepsilon > 0$ be arbitrary and take by Proposition 7.12 a subsequence $(n_i) \subseteq \mathbb{N}$ such that $|\{q_{n_i}\alpha\}| = d_{n_i} < q_{n_i}^{-\beta+\varepsilon}$. Using this and the fact that $f \in \operatorname{lip}(a)$, we get

$$\begin{aligned}
\left| S_{2q_{n_i}} f(x) - 2S_{q_{n_i}} f(x) \right| &\le \sum_{j=0}^{q_{n_i}-1} |f(x + q_{n_i}\alpha + j\alpha) - f(x + j\alpha)| \\
&\le \sum_{j=0}^{q_{n_i}-1} C_f |\{q_{n_i}\alpha\}|^a \le C_f q_{n_i}^{1-(\beta-\varepsilon)a}
\end{aligned}$$

for any $i \in \mathbb{N}$ and any $x \in \mathbb{T}$. Now choose ε so small that $a(\beta - \varepsilon) > \frac{1}{2}$. Then for some $\varepsilon' > 0$, the above implies

$$\left| \frac{S_{2q_{n_i}} f(x)}{\sqrt{q_{n_i}}} - 2\frac{S_{q_{n_i}} f(x)}{\sqrt{q_{n_i}}} \right| \le C_f q_{n_i}^{-\varepsilon'}.$$

Since q_{n_i} tends to infinity with i, this implies

$$\phi(u) = \lim_{i\to\infty}\left|\left\{x\in\mathbb{T} : \left|2\frac{S_{q_{n_i}}f(x)}{\sqrt{q_{n_i}}}\right| > 2u\right\}\right| \le \lim_{i\to\infty}\left|\left\{x\in\mathbb{T} : \left|\frac{S_{2q_{n_i}}f(x)}{\sqrt{q_{n_i}}}\right| > u\right\}\right|$$
$$\le \lim_{i\to\infty}\left|\left\{x\in\mathbb{T} : \left|\frac{S_{2q_{n_i}}f(x)}{\sqrt{2q_{n_i}}}\right| > \tfrac{u}{2}\right\}\right| = \phi(\tfrac{u}{2}).$$

Since ϕ is a distribution function, it is a non-increasing function, which tends to 0 as u tends to infinity. This implies the negative half.

For the positive half of the theorem, we need the second and fourth moment of the partial sums. Using tools from random measures — in particular using Gaussian white noise — Lacey proves that we can choose the Fourier coefficients a_k in (7.17) such that for some $\delta > 0$,

$$\|S_m f\|^2 = m + O\left(m^{1-\delta}\right), \tag{7.18}$$

$$m^{-2}\mathbb{E}\left((S_m f)^4\right) = O\left(m^{-\delta}\right). \tag{7.19}$$

The construction also ensures that $f \in \mathrm{lip}(a)$ by Theorem 7.13.

With these estimates, we may reduce Theorem 7.11 to a consequence of a slightly adapted version of Theorem 6.4 in [64]. We see that

$$\frac{S_m f(x)}{\sqrt{m}} = \sum_{k=1}^{\infty} a_{m,k} r_k \cos(2\pi k x) \quad \text{where } a_{m,k} = \frac{2a_k V_m(k\alpha)}{\sqrt{m}}.$$

It is enough to prove that this converges in distribution to a standard normal distribution. This would follow from the before mentioned adapted theorem, if we can prove that $\frac{1}{2}\sum_k a_{m,k}^2 \sim 1$ and $\sum_m \sum_k a_{m,k}^4 < \infty$. We see by (7.18),

$$\frac{1}{2}\sum_{k=1}^{\infty} a_{m,n}^2 = \frac{1}{2}\sum_{k=1}^{\infty}\frac{4a_k^2\,|V_m(k\alpha)|^2}{m} = \frac{\|S_m(f)\|^2}{m} = 1 + O\left(m^{-\delta}\right).$$

We will now prove that $\sum_m \sum_k a_{m,k}^4 < \infty$. To do this, let $\theta_j = [\theta^j]$ for some arbitrary $\theta > 1$. We split $\mathbb{N}$ up into blocks $I_j = (\theta_j, \theta_{j-1}]$. By (7.19), along this sequence we get

$$\sum_{\theta_j}\sum_k a_{\theta_j,n}^4 \le K\sum_{\theta_j}\theta_j^{-\delta} \le K\sum_{j=1}^{\infty}\left(\frac{1}{\theta^\delta}\right)^j < \infty.$$

Hence, we have the required property along the subsequence (θ_j). We now control the oscillations of $S_m f$ on each block I_j. By (7.18),

$$\left\|S_m f - S_{\theta_j} f\right\|^2 = O((\theta-1)\sqrt{\theta_j}),$$

so letting $\theta \to 1$, we get the required convergence. This concludes our sketch. □

7.3.3. Gaussian randomisation

Another method for proving theorems of the Central Limit Theorem type involving spectral theory was introduced by Weber in [62] and [63] to prove the Almost Sure Central Limit Theorem for irrational rotations. We have so far mainly been interested in the Central Limit Theorem, and it would appear that Weber's method could also work for this theorem. Also, in none of the partial methods, ergodicity is required, so it might be possible to use this method to treat non-ergodic system. Hence, we include it in this survey.

We state Weber's Almost Sure Central Limit Theorem.

Theorem 7.14. *There exists a sequence (σ_j) such that $\sigma_j \asymp \sqrt{j}$ and an $f \in L^2(\mathbb{T})$ such that*

$$\frac{1}{\log m} \sum_{j=1}^{m} \frac{1}{j} \delta_{\{S_j f/\sigma_j\}} \xrightarrow[m\to\infty]{\mathcal{D}} \mathcal{N}(0,1) \quad a.s.$$

Note the resemblance with Theorem 7.4. The main difference is that the normalising factor is only approximately $\sqrt{j}$ and not necessarily equal to $\sqrt{j}$. This is the result of a limitation of the method of proof, which implies that we can not obtain an exact value for the variance of the partial sums. In the following, we sketch the method used by Weber.

The method of proof is based on two principles. One is the almost sure convergence of certain series over a quasi-orthogonal system. The other is the concept of Gaussian randomisation.

A sequence of vectors $(f_n) \subseteq L^2(\mathbb{T})$ (or indeed any Hilbert space) is said to be a *quasi-orthogonal system* if the quadratic form $\sum \langle f_j, f_k \rangle x_j x_k$ is bounded, where $\langle \cdot, \cdot \rangle$ denotes the inner product in $L^2(\mathbb{T})$. It can be shown that if $\sup_{j\geq 1} \sum_k \left|\langle f_j, f_k \rangle\right| < \infty$, then the sequence (f_n) is a quasi-orthogonal system. Also, one can show that for any quasi-orthogonal system (f_n),

$$\sum_{n=1}^{\infty} \frac{1}{\sqrt{n}(\log n)^b} f_n < \infty \quad a.s. \tag{7.20}$$

whenever $b > 3/2$. In fact, this is a corollary of a theorem due to Menchov and Rademacher (see [44], Lemma 1, page 42).

The second element of the proof is the Gaussian randomisation. We fix some number $\Delta > 0$ and construct a function $f = \sum a_k e_{l_k}$ in terms of it's Fourier expansion, where

$$a_k = \begin{cases} \Delta^{k/2} & \text{for } k < 0, \\ 0 & \text{for } k = 0, \\ \Delta^{-k/2} & \text{for } k > 0, \end{cases}$$

and l_k is such that $\{l_k \alpha\} \in [\Delta^{-k-1}, \Delta^{-k})$. As before, we can calculate the variance of the partial sums and — again using the orthonormality of the e_n — we get an upper bound on the correlation of the partial sums. In fact, with the aid of some classical estimates on the $V_n(\theta)$,

$$\|S_m f\|^2 \asymp n, \quad \forall n < m : |\langle S_n f, S_m f\rangle| \leq C(\Delta) n \log\left(\frac{m}{n}\right). \tag{7.21}$$

We will include a random element in the definition of our function. Hence, let $(Y, \mathcal{C}, \nu)$ be another probability space, and let $(g_n)_{n\in\mathbb{Z}}$ and $(g'_n)_{n\in\mathbb{Z}}$ be two isonormal, independent sequences on this space. Let $\gamma_n = g_n + ig'_n$ and define

$$\xi(x) = \sum_{k=-\infty}^{\infty} a_k \mathfrak{Re}\left(\overline{\gamma_k} e_{l_k}(x)\right).$$

For each $x \in \mathbb{T}$, these are random variables on $(Y, \mathcal{C}, \nu)$, and for any $x \in \mathbb{T}$, $m, n \in \mathbb{N}$,

$$\langle S_n(\xi(x)), S_m(\xi(x))\rangle = \langle S_n f, S_m f\rangle. \tag{7.22}$$

Calculating the characteristic function of $S_m\xi(x,\omega)$ in the product probability space $(\mathbb{T} \times Y, \mathcal{B} \otimes \mathcal{C}, \mu \times \nu)$, we see that this is a centred Gaussian sequence. We may use a nice property of Gaussian vectors in $\mathbb{R}^2$ along with the estimates (7.21) and (7.22) to construct an appropriate quasi-orthogonal system, which gives an almost sure convergence statement by (7.20). This all takes place in the extension $(\mathbb{T} \times Y, \mathcal{B} \otimes \mathcal{C}, \mu \times \nu)$. Using Fubini's Theorem, one may deduce that we also have an almost sure convergence in the factor $(\mathbb{T}, \mathcal{B}, \mu)$ for ν-almost all functions in an appropriate subset of $L^2(\mathbb{T})$. Using Kronecker's Lemma ([11], Lemma 3.28), one concludes that the theorem in fact holds for a large number of functions.

7.4. Tower constructions

Irrational rotations of the circle are interesting in their own right, but they hardly make up the whole spectrum of ergodic dynamical systems. Hence, it is a natural question to ask, whether or not the results in Section 7.3 extend to other types of dynamical systems. It turns out that under appropriate circumstances, some of the results do indeed extend to other dynamical systems. A method of passing from rotations to general ergodic, aperiodic systems is by using Rokhlin towers.

A first attempt to prove a Central Limit Theorem for general aperiodic dynamical systems was due to Burton and Denker ([13]), who emulated the behaviour of rotations by constructing a Rokhlin tower inside the new dynamical system. Unfortunately, their proof was flawed. It was later corrected by de la Rue, Ladouceur, Peškir and Weber in [15]. The theorem stated in both articles is the following:

Theorem 7.15. *Let $(X, \mathcal{B}, \mu)$ be a non-atomic probability space and let $T : X \to X$ be an aperiodic automorphism. There exists a function $f \in L^2(\mu)$ with $\mathbb{E}(f) = 0$ such that*

$$\frac{S_m f}{\|S_m f\|} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1).$$

A key element in all known proofs of Central Limit Theorems in general measure preserving systems is Rokhlin's Lemma:

Lemma 7.16 (Rokhlin's Lemma ([50], Lemma 4.7)). *Let $(X,\mathcal{B},\mu)$ be a non-atomic probability space, let $T: X \to X$ be an ergodic measure preserving transformation and let $\varepsilon > 0$, $n \in \mathbb{N}$. There exists a set $A \in \mathcal{B}$ such that*

1. *The sets $\{T^iA\}_{i=0}^{n-1}$ are pairwise disjoint.*
2. $$\mu\left(\bigcup_{i=0}^{n-1} T^iA\right) = \sum_{i=0}^{n-1} \mu\left(T^iA\right) > 1-\varepsilon.$$

The set A is called an (n,ε)-Rokhlin set.

This lemma allows us to trace large parts of the orbits of a substantial subset of points in X. This in turn allows us to emulate the behaviour of irrational rotation, where this information is implied in the bounds on the Fourier coefficients.

The method of using Rokhlin sets to deduce the Central Limit Theorem was first used by Burton and Denker in [13] to generalise Theorem 7.8 to arbitrary systems. The fundamental property needed in order to obtain the general theorem was the existence of functions $g^i : E \to \{-1,1\}$, $i = 0,\dots,N-1$, where E is an (LN,ε)-Rokhlin set for some $L,N \in \mathbb{N}$, such that the sequence $(g^{2n})_{n=0}^{N/2-1}$ as well as the sequence $(g^{2n+1})_{n=0}^{N/2-1}$ each consists of independent and identically distributed random variables with respect to the measure $\mu(E)^{-1}\operatorname{tr}(\mu,E)$, where $\operatorname{tr}(\mu,E)(A) = \mu(A\cap E)$. Also, one needs the property that

$$\mu(E)^{-1}\operatorname{tr}(\mu,E)\left\{x \in X : g^i(x) = \pm 1\right\} = \frac{1}{2} \quad \text{for } i = 0,\dots,N-1.$$

Burton and Denker made a mistake in their proof of the existence of such functions. This mistake was corrected by de la Rue, Ladouceur, Peškir and Weber in [15]. They still needed the sequences of functions. In their paper, the construction is based on choosing independent partitions of the space $(X,\mathcal{B},\mu)$. We will not go into this construction at this point, since independent partitions will be constructed in abundance in the next section. Instead, we will focus on the application of these functions in the construction of functions for which the Central Limit Theorem holds.

One defines the function

$$g(x) = \begin{cases} g^l\left(T^{jN+l}(x)\right) & \text{for } x \in T^{jN+l} \text{ where } j \in \{0,\dots,L-1\}, l \in \{0,\dots,N-1\}, \\ 0 & \text{for } x \notin \bigcup_{i=0}^{NL-1} T^l(F). \end{cases}$$

The information we have on the Rokhlin construction along with the definition of the g^i gives us quite a lot of information about the distribution of the values of the first iterates of the function. In particular, we can get nice bounds on the variance of the partial sums. In fact, if $K(\varepsilon + L^{-1}) \to 0$ as $K, L \to \infty$, $\varepsilon \to 0$,

$$(1-\delta)m \le \|S_m f\|^2 \le (1+\delta)m \qquad \text{for } 1 \le m \le K,$$

$$(1-\delta)K \le \|S_m f\|^2 \le (1+\delta)K \qquad \text{for } K \le m \le N-K,$$

when $\delta \to 0$. Also, we obtain nice independence properties of the partial sums.

Now, by a particularly clever choice of sequences (N_k), (L_k), (K_k) and (ε_k), we may use the above to construct a sequence (g_k) of functions of the above form. Furthermore, we may choose a sequence of reals (a_k) and integers (n_k) — again in a particularly clever way — and define

$$f = \sum_{k=1}^{\infty} a_k g_k, \quad f_k = a_{n_k} g_{n_k} + a_{n_k+1} g_{n_k+1} \quad \text{and} \quad A_k = K_{n_k} |a_{n_k}|^2 + k |a_{n_k+1}|^2.$$

One can prove that the partial sums of f are well approximated in $L^2(\mu)$ by the partial sums of f_k as k tends to infinity. This is similar to the last part of the proof of Theorem 7.8, with the g_k replacing the cosines. Furthermore, the estimates on the variance implies the theorem.

This was a short sketch of the ideas involved in extending the theorems known from irrational rotations to general ergodic, aperiodic dynamical systems. The full details of the constructions involved are quite lengthy and the detail given here is far from complete. We refer the reader to [15] for a detailed account. In that paper, a few results extending Theorem 7.15 are also given.

7.5. A general result on Rokhlin towers

In [61], Volný proved a surprising result on Gaussian approximation in general aperiodic dynamical systems. His construction involved the Rokhlin sets introduced in the preceding section along with a construction similar to the one used in the proof of Theorem 7.6. We prove Volný's general result on Rokhlin towers in this section. The proof of this result given in [61] is essentially a sketch. In a subsequent private communication between this author, M. Weber and D. Volný ([60]), the ideas of the proof became more apparent. The aim of the present exposition of the proof is to present it in a consistent manner.

A key property in Volný's proof is the fact that we can introduce the beginning of any strictly stationary sequence of random variables with arbitrary distribution on the probability space in our aperiodic dynamical systems along with a function for which the iterates are close (in L^2) to the random variables. In fact, we have:

Proposition 7.17. *Let $(\Omega,\mathcal{B},\mu,T)$ be an ergodic, aperiodic measure preserving system, $(X_i)_{i\in\mathbb{Z}}$ be an ergodic, strictly stationary process (defined on a different probability space), let $\xi_1=\{A_1,\ldots,A_K\}$ be a partition of Ω, $n\in\mathbb{N}$, and let $\varepsilon>0$. Then, there exist a measurable finite valued function f such that:*

1. *There exists a random vector $(X_0',\ldots,X_n')$ distributed as $(X_0,\ldots,X_n)$ on Ω,*
$$\mu\{\exists 0\le i\le n, \left|f\circ T^i - X_i'\right| > \varepsilon\} < \varepsilon. \tag{7.23}$$
2. *The partition η_1 generated by $f, f\circ T,\ldots, f\circ T^n$ is ε-independent of ξ_1 in the sense that*
$$\sum_{A\in\eta_1,B\in\xi_1,\mu(B)>0} |\mu(A)-\mu(A|B)| < \varepsilon. \tag{7.24}$$

Proof. To begin with, we impose the assumption, that the X_i be finitely valued. We will see later that this causes no loss of generality.

Let $(\bar{\Omega},\mathcal{C},\nu,S)$ be a representation of the process (X_i). It poses no problem to choose this representation in such a way that there exist S-invariant, independent σ-algebras $\mathcal{C}',\mathcal{C}''\subseteq\mathcal{C}$, such that the corresponding factors are aperiodic and such that X is $\mathcal{C}''$-measurable. Indeed, take the product space of a representation with any other measure preserving system with $\mathbb{Z}$-action and extend $(\overline{\Omega},\mathcal{C},S)$ and ν accordingly.

Now, we use Rokhlin's Lemma to chop things up in a nice and orderly fashion. Let $N\in\mathbb{N}$ be some number to be fixed at a later point and let $F,TF,\ldots,T^NF$ be a Rokhlin tower in Ω. Further, we let $E,SE,\ldots,S^NE$ be a $\mathcal{C}'$-measurable Rokhlin tower in $\bar{\Omega}$ with $\mu(F)=\nu(E)$ and let π_1 and π_2 denote the partitions generated by $F,TF,\ldots,T^NF$ and $E,SE,\ldots,S^NE$ respectively. Clearly, these two partitions are identically distributed.

We now choose a family of measure preserving bijections,
$$\phi_i : T^iF\to S^iE,\quad i\in\{0,\ldots,N\}.$$

For the exceptional set, we choose another measure preserving bijection,
$$\phi_\varepsilon : X\setminus\bigcup_{i=0}^{N} T^iF \to X'\setminus\bigcup_{i=0}^{N} S^iE.$$

We combine the whole thing to a measure preserving bijection which maps π_1 to π_2,
$$\phi(x)=\begin{cases}\phi_i(x) & \text{for } x\in T^iF, i\in\{0,\ldots,N\},\\ \phi_\varepsilon(x) & \text{otherwise.}\end{cases}$$

We map the partition ξ_1 to $\bar{\Omega}$ through this function, that is, $\xi_2=\phi(\xi_1)$.

Now, let η_2 be the partition of $\bar{\Omega}$ generated by $X,SX,\ldots,S^n$. We map this partition to Ω. That is, $\eta_1=\phi^{-1}(\eta_2)$. Since $\mathcal{C}'$ is independent of $\mathcal{C}''$, the partition $\pi_2\vee\xi_2$ is independent of η_2. Hence, η_1 is independent of $\pi_1\vee\xi_1$.

Since all the functions we are interested in are constant on each cell of the partition, we may define the function

$$f(x) = \begin{cases} X(\phi(x)) & \text{for } x \in T^i F, i \in \{0, 1, \dots, N\} \\ 0 & \text{otherwise} \end{cases}$$

and the random variables

$$X_i'(x) = S^i(X(\phi(x)))$$

The random variables are clearly distributed as the X_i, as everything is constant on each cell.

Now, we can get the results. For $x \in \bigcup_{i=0}^{N-n} T^i F$, there is full control of the future of x for the next n steps of the process. Furthermore, by definition, the value of $T^i \circ f(x)$ is the same as the value of $X_i'(x)$ for such x. Hence, we have the first property when we choose N so large that $\mu(\bigcup_{i=0}^{N-n} T^i F) > 1 - \varepsilon$. We also see that any dependence between the partitions defined in the statement of property 2 must occur outside of this set. Hence, this guarantees the second property as well.

A note missing in Volný's original paper is that he can indeed pass from considering finitely valued random variables to any random variable. This turns out to be possible, since we can approach any random variable with finitely valued random variables. This means that we can get a sequence of refinements of partitions of each of our two spaces in such a way that they have the same distributions at each step, and such that the step-functions are constant on each cell in the representation space, where the functions converge to the required random variable. Hence, there is no choice for the mirrored step-functions in the aperiodic system but to converge to a random variable of the same distribution. □

We note, that this proposition implies Proposition 2 in [61] by applying the previous Proposition for each $k \in \mathbb{N}$:

Proposition 7.18. *Let n_k be positive integers, $(X_{k,i})_{i \in \mathbf{Z}}$ be ergodic, strictly stationary processes (defined on different probability spaces), $\varepsilon_k > 0, k = 1, 2, \dots$. Then there exist measurable finite valued functions f_k such that:*

1. *For every $k = 1, 2, \dots$ there exists a random vector $(X'_{k,0}, \dots, X'_{k,n_k})$ with the same distribution as $(X_{k,0}, \dots, X_{k,n_k})$ on Ω,*

$$\mu\left\{\exists 0 \le i \le n_k, \left| f_k \circ T^i - X'_{k,i} \right| > \varepsilon_k \right\} < \varepsilon_k.$$

2. *The partition ξ generated by $f_j \circ t^i, 0 \le i \le n_j, 1 \le j \le k-1$, and η generated by $f_k \circ T^i, 0 \le i \le n_k$, are ε-independent in the sense that*

$$\sum_{A \in \eta, B \in \xi, \mu(B) > 0} |\mu(A) - \mu(A|B)| < \varepsilon_k.$$

A straightforward application of this Proposition gives us Proposition 1 in [61]:

Proposition 7.19. *Let* $(\varepsilon_i)_{i\in\mathbb{N}} \subseteq \mathbb{R}$, $(\alpha_i)_{i\in\mathbb{N}} \subseteq \mathbb{R}$ *and* $(d_i)_{i\in\mathbb{N}} \subseteq \mathbb{N}$. *Then there exist* $(\overline{f}_i)_{i\in\mathbb{N}} \in L^2(\Omega, \mathcal{A}, \mu)$, $\mathbb{E}f_i = 0$ *for any* $i \in \mathbb{N}$*; and independent random variables* $X_{k,i}$, $i = 0, \ldots, 2d_k$, $k \in \mathbb{N}$ *where any given* $X_{k,i}$ *is* $N(0, \alpha_k^2)$*-distributed, such that*

$$\forall k \in \mathbb{N}, i = 0, \ldots, 2d_k : \left\| X_{k,i} - \overline{f}_k \circ T^i \right\| < \varepsilon_k. \tag{7.25}$$

In the following chapter, we will go through the details of Volný's proof of some theorems in Gaussian approximation.

8. Gaussian approximation

In this section, we prove some of Volný's results in Gaussian approximation. Volný's results are the basis of the results of the next chapter, where we use these results to obtain results in weighted Gaussian approximation.

In the first section, we prove results on Gaussian approximation in L^2. We begin by proving the existence of a function for which the n'th partial sum is close to the sum of independent random variables, where the variables are dependent on n. This is the essence of Lemma 8.6 below. To obtain the correct distribution of the random variables, we simply normalise these appropriately, and prove that the non-normalised random variables are asymptotically close to the normalised ones. To also get rid of the dependence on n of the random variables, we consider the asymptotic behaviour of the triangular array of random variables when we move towards the diagonal. This has an impact on the speed of convergence obtained.

In the second section, we prove that the partial sums of the function found in the preceding section also satisfy the strong invariance principle. In Volný's paper there is also a proof of the weak invariance principle. While we do mention this result, we do not prove it, since this would involve additional technique, and since the result is not generalised in the subsequent chapter, where the additional methods fail to work.

8.1. Approximation in L^2

In this section, we will prove the following theorem:

Theorem 8.1. *Let $(X, \mathcal{B}, \mu)$ be a non-atomic probability space. Let $T: X \to X$ be an ergodic, aperiodic automorphism. There exist a function $f \in L^2(X, \mathcal{B}, \mu)$ and independent random variables $Z_j \sim \mathcal{N}(0, 2(\log\log 3 - \log\log 2))$, such that for n sufficiently large,*

$$\max_{1 \le l \le n} \frac{1}{\sqrt{n}} \left\| S_l f - \sum_{j=0}^{l-1} Z_j \right\| = O\left(\frac{\log\log\log\log n}{\log n}\right)^{1/2}.$$

We follow Volný's general approach from [61], but since we do not aim to obtain the weak invariance principle simultaneously with Theorem 8.1, we avoid some technicalities. We do go into other technicalities somewhat deeper that the original proof.

It is worth noting that a version of the Central Limit Theorem follows from Theorem 8.1, just as a version followed from Theorem 7.6. Hence, Volný's Theorem is a generalisation of the Central Limit Theorem for aperiodic dynamical systems (for example in the form of Theorem 7.15).

A converse of Volný's Theorem due to Akcoglu, Baxter, Ha and Jones exists ([3]). In their paper, they prove that given an aperiodic, non-atomic measure preserving system $(X, \mathcal{B}, \mu, T)$, an L^2-function f and a natural number K, one can find another L^2-function g, such that the $T^i g$ are "almost" Gaussian and such that the correlations $\langle T^i g, T^j g\rangle$ are "close to" the correlations $\langle T^i f, T^j f\rangle$ for $i, j \in \{1, \dots, K\}$. For the appropriate definitions of closeness as well as the full results, the reader is referred to [3].

As in Volný's proof, we define numbers

$$d_k = 3^k, \quad p_k = 2^k, \quad \alpha_k = \frac{1}{p_k\sqrt{k}}, \quad \varepsilon_k = 6^{-3k},$$

and take functions $\overline{f}_k$'s and random variables $X_{k,i}$'s as in Proposition 7.19. Again, we follow Volný and define functions

$$f_k = \sum_{i=0}^{p_k-1} T^i \overline{f}_k - T^{d_k} \sum_{i=0}^{p_k-1} T^i \overline{f}_k,$$
$$f = \sum_{k=1}^{\infty} f_k.$$

We will split the function up into three parts,

$$f = f' + f'' + f''' = \sum_{k:d_k \leq n} f_k + \sum_{k:p_k<n<d_k} f_k + \sum_{k:n\leq p_k} f_k.$$

We are considering the partial sums of f, defined by

$$S_j f = \sum_{l=0}^{j-1} T^i f.$$

Clearly, the operator S_j is linear, so

$$S_j f = S_j f' + S_j f'' + S_j f'''.$$

The three partial sums are treated separately. In order to estimate the sums, we need an auxiliary lemma.

Lemma 8.2. *Let $a > 1$, $p > 0$ and $c > 0$. There exists a $K < \infty$ such that*

$$\sum_{k:a^{ck}\leq n} \frac{a^k}{k^p} \leq K \frac{n^{\frac{1}{c}}}{(\frac{1}{2c}\log n)^p}.$$

Proof. The proof is a straightforward calculation:

$$\begin{aligned}
\sum_{k:a^{ck}\le n}\frac{a^k}{k^p} &\le \sum_{k=1}^{[\frac{1}{2c}\log_a n]} a^k + \sum_{k=[\frac{1}{2c}\log_a n]+1}^{[\frac{1}{c}\log_a n]} \frac{a^k}{(\frac{1}{2c}\log_a n)^p}\\
&\le \frac{a^{1+[\frac{1}{2c}\log_a n]}-1}{a-1} + \frac{a^{1+[\frac{1}{c}\log_a n]}-1}{a-1}\frac{1}{(\frac{1}{2c}\log_a n)^p}\\
&\le \frac{a(n^{1/c})^{1/2}-1}{a-1} + \frac{an^{1/c}-1}{a-1}\frac{1}{(\frac{1}{2c}\log_a n)^p}\\
&\le K\frac{n^{\frac{1}{c}}}{(\frac{1}{2c}\log n)^p}.
\end{aligned}$$

□

We can now estimate the first weighted sum.

Lemma 8.3. *Let $n\in\mathbb{N}$. For any $j\in\{0,\dots,n\}$,*

$$\left\|S_j f'\right\| = O\left(\frac{n}{\log n}\right)^{1/2}.$$

Proof. For any $k\in\mathbb{N}$, we define the function

$$g_k = \sum_{j=0}^{d_k-1}\sum_{i=0}^{p_k-1} T^{i+j}\overline{f}_k. \tag{8.1}$$

We immediately note that $f_k = g_k - Tg_k$. That is, f_k is a co-boundary. Hence,

$$S_j f' = \sum_{l=0}^{j-1} T^l \sum_{k:d_k\le n}(g-Tg) = \sum_{l=0}^{j-1}\sum_{k:d_k\le n}\left(T^l g_k - T^{l+1}g_k\right) = \sum_{k:d_k\le n} g_k - T^j \sum_{k:d_k\le n} g_k.$$

Using the triangle inequality and the fact that T is measure preserving, the above implies

$$\left\|S_j f'\right\| \le \sum_{k:d_k\le n}\|g_k\| - \sum_{k:d_k\le n}\|g_k\| \ll \sum_{k:d_k\le n}\|g_k\|.$$

Hence, it suffices to consider the last sum.

We define new functions,

$$\hat{g}_k = \sum_{j=0}^{d_k-1}\sum_{i=0}^{p_k-1} X_{k,i+j}, \quad \tilde{g}_k = g_k - \hat{g}_k. \tag{8.2}$$

We consider $\tilde{g}_k$ first. By the triangle inequality and Proposition 7.19, we have

$$\|\tilde{g}_k\| = \left\| \sum_{j=0}^{d_k-1} \sum_{i=0}^{p_k-1} T^{i+j}\overline{f}_k - \sum_{j=0}^{d_k-1} \sum_{i=0}^{p_k-1} X_{k,i+j} \right\| \leq \sum_{j=0}^{d_k-1} \sum_{i=0}^{p_k-1} \left\| T^{i+j}\overline{f}_k - X_{k,i+j} \right\| \leq d_k p_k \varepsilon_k.$$

Since

$$\sum_{k:d_k \leq n} d_k p_k \varepsilon_k = \sum_{k:d_k \leq n} 3^k 2^k 6^{-3k} = \sum_{k=1}^{[\log_3 n]} \frac{1}{6^{2k}} < 1,$$

we only need to consider the sum of the $\hat{g}_k$.

We know that the $X_{k,i+j}$ are independent. Hence, since $p_k < d_k$,

$$\left\| \sum_{j=0}^{d_k-1} \sum_{i=0}^{p_k-1} X_{k,i+j} \right\|^2 \leq \alpha_k^2 p_k^2 (p_k + d_k) = \frac{d_k + p_k}{k}$$

by choice of sequences. But now we may use Lemma 8.2 to obtain

$$\begin{aligned} \sum_{k:d_k \leq n} \|\hat{g}\| &\leq \sum_{k:d_k \leq n} \sqrt{\frac{d_k + p_k}{k}} = \sum_{k:3^k \leq n} \frac{(3^k + 2^k)^{1/2}}{k^{1/2}} \\ &\leq \sqrt{2} \sum_{k:\sqrt{3}^{2k} \leq n} \frac{\sqrt{3}^k}{k^{1/2}} \leq \sqrt{2} K \frac{n^{1/2}}{(\log n)^{1/2}} = O\left(\frac{n}{\log n}\right)^{1/2}. \end{aligned}$$

This completes the proof. ☐

We now consider the third weighted sum:

Lemma 8.4. *Let $n \in \mathbb{N}$. For any $j \in \{0, \ldots, n\}$,*

$$\left\| S_j f''' \right\| = O\left(\frac{n}{\log n}\right)^{1/2}.$$

Proof. We need a new splitting of the sums. In fact, since

$$S_j f''' = S_j \left(\sum_{k:n \leq p_k} f_k \right) = \sum_{k:n \leq p_k} S_j f_k, \tag{8.3}$$

we will look at the $S_j f_k$. For these,

$$S_j f_k = \sum_{l=0}^{j-1} T^l \left(\sum_{i=0}^{p_k-1} T^i \overline{f}_k - T^{d_k} \sum_{i=0}^{p_k-1} T^i \overline{f}_k \right) = \sum_{l=0}^{j-1} \sum_{i=0}^{p_k-1} T^{l+i} \left(\overline{f}_k - T^{d_k} \overline{f}_k \right).$$

As in the proof of Lemma 8.3, we define

$$\hat{S}_{k,j} = \sum_{l=0}^{j-1} \sum_{i=0}^{p_k-1} \left(X_{k,l+i} - X_{k,l+i+d_k} \right), \quad \tilde{S}_{k,j} = S_j f_k - \hat{S}_{k,j}. \tag{8.4}$$

We may calculate an upper bound on the norm of $\tilde{S}_{k,j}$ using Proposition 7.19 and the triangle inequality:

$$\|\tilde{S}_{k,j}\| = \left\| \sum_{l=0}^{j-1} \sum_{i=0}^{p_k-1} \left(T^{l+i}\overline{f}_k - X_{k,l+i} \right) - \sum_{l=0}^{j-1} \sum_{i=0}^{p_k-1} \left(T^{l+i+d_k}\overline{f}_k - X_{k,l+i+d_k} \right) \right\| \\ \leq 2jp_k\varepsilon_k \leq 2n2^k 6^{-3k}.$$

But since

$$\sum_{k\geq \log_2 n} 2^k 6^{-3k} = O\left(\frac{1}{n^3}\right),$$

we see that

$$\left\| \sum_{k:n\leq p_k} \tilde{S}_{k,j} \right\| = O\left(\frac{1}{n^2}\right). \tag{8.5}$$

The estimate on the norm of $\hat{S}_{k,j}$ is not quite as easy. Once again, we calculate the square of the norm.

$$\begin{aligned} \|\hat{S}_{k,j}\|^2 &= \left\langle \sum_{l=0}^{j-1} \sum_{i=0}^{p_k-1} \left(X_{k,l+i} - X_{k,l+i+d_k} \right), \sum_{l'=0}^{j-1} \sum_{i'=0}^{p_k-1} \left(X_{k,l'+i'} - X_{k,l'+i'+d_k} \right) \right\rangle \\ &\leq 2 \left\langle \sum_{l=0}^{j-1} \sum_{i=0}^{p_k-1} X_{k,l+i}, \sum_{l'=0}^{j-1} \sum_{i'=0}^{p_k-1} X_{k,l'+i'} \right\rangle. \end{aligned} \tag{8.6}$$

The last inequality holds since $\sum_{l=0}^{j-1} \sum_{i=0}^{p_k-1} X_{k,l+i}$ and $\sum_{l=0}^{j-1} \sum_{i=0}^{p_k-1} X_{k,l+i+d_k}$ are equally distributed and uncorrelated, due to the gap between the indices. and independence of the summands.

To calculate the final inner product, we split the sum up into three uncorrelated parts with indices corresponding to Figure 8.1 on the next page. The set Σ_2 contains the diagonals on it's boundary. Note that along each diagonal, the second index of the corresponding random variables is constant. Summing along the diagonals and using the fact that $\|X\|^2 = \langle X, X\rangle$, we may write

$$\left\langle \sum_{l=0}^{j-1} \sum_{i=0}^{p_k-1} X_{k,l+i}, \sum_{l'=0}^{j-1} \sum_{i'=0}^{p_k-1} X_{k,l'+i'} \right\rangle \\ = \left\| \sum_{i=0}^{j-2} (i+1) X_{k,i} \right\|^2 + \left\| j \sum_{i=j-1}^{p_k-1} X_{k,i} \right\|^2 + \left\| \sum_{i=0}^{j-2} (i+1) X_{k,p_k+j-2-i} \right\|^2.$$

These three norms are easily estimated using the independence of the $X_{k,i}$. For the first summand, we get

$$\left\| \sum_{i=0}^{j-2} (i+1) X_{k,i} \right\|^2 \leq (j-1) \left\| \sum_{i=0}^{j-2} X_{k,i} \right\|^2 \leq j^3 \alpha_k^2 \leq n^3 \frac{1}{2^{2k}k}. \tag{8.7}$$

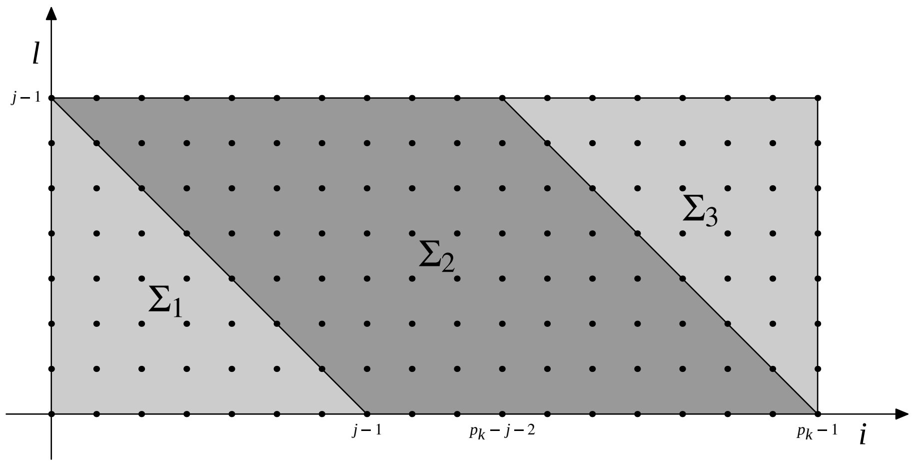

Figure 8.1.: Splitting the sum in $\left\|\hat{S}_{k,j}\right\|$

The same estimate holds for the last summand. For the middle summand,

$$\left\| j \sum_{i=j-1}^{p_k-1} X_{k,i} \right\|^2 \leq j^2 p_k \alpha_k^2 \leq n^2 \frac{1}{2^k k}. \tag{8.8}$$

Now, since

$$\sum_{k \geq \log_2 n} \frac{1}{2^{2k} k} = O\left(\frac{1}{n^2 \log n}\right) \quad \text{and} \quad \sum_{k \geq \log_2 n} \frac{1}{2^k k} = O\left(\frac{1}{n \log n}\right),$$

(8.7) and (8.8) imply that

$$\left\| \sum_{k: n \leq p_k} \hat{S}_{k,j} \right\|^2 = \left(\frac{n}{\log n}\right). \tag{8.9}$$

Putting it all together, we see that (8.3) along with (8.5) and (8.9) implies the lemma. □

We now consider the middle term. This is where the weighted partial sums are close to the sums of Gaussian random variables. First, we define these variables. For $n \in \mathbb{N}$, $0 \leq l \leq n$, we let

$$Y_{n,l} = \sum_{k: p_k < n < d_k} p_k \left(X_{k,l} - X_{k,l+d_k}\right). \tag{8.10}$$

Note that for each $n \in \mathbb{N}$, these are independent Gaussians, since they are themselves the sum of independent Gaussians. Furthermore, for each fixed $n \in \mathbb{N}$, the random variables $Y_{n,0}, \ldots, Y_{n,n-1}$ are identically distributed. This fact will be used to deduce the main results of this chapter.

Lemma 8.5. *Let $n \in \mathbb{N}$.*

$$\max_{1 \le j \le n} \frac{1}{\sqrt{n}} \left\| S_j f'' - \sum_{l=0}^{j-1} Y_{n,l} \right\| = O\left(\frac{1}{\log n}\right)^{1/2}.$$

Proof. As in the proofs of the two previous lemmas, we will use Proposition 7.19 to reduce the statement of the lemma to a statement about random variables. Hence, we define

$$\hat{S}''_j = \sum_{k: p_k < n < d_k} \sum_{l=0}^{j-1} \sum_{i=0}^{p_k - 1} \left(X_{k,i+l} - X_{k,i+l+d_k}\right), \quad \tilde{S}''_j = S_j f'' - \hat{S}''_j. \tag{8.11}$$

Applying Proposition 7.19 and the triangle inequality in the usual way, we obtain

$$\begin{aligned} \|\tilde{S}''_j\| = \Bigg\| & \sum_{k: p_k < n < d_k} \sum_{l=0}^{j-1} \sum_{i=0}^{p_k - 1} \left(T^{i+l} \overline{f}_k - X_{k,i+l}\right) \\ & - \sum_{k: p_k < n < d_k} \sum_{l=0}^{j-1} \sum_{i=0}^{p_k - 1} \left(T^{i+l+d_k} \overline{f}_k - X_{k,i+l+d_k}\right) \Bigg\| \\ \le 2n & \sum_{k: p_k < n < d_k} p_k \varepsilon_k \le 2n \sum_{k: p_k < n < d_k} 2^k 6^{-3k} = O\left(\frac{1}{n^2}\right). \end{aligned} \tag{8.12}$$

Hence, by the triangle inequality, the lemma is reduced to a statement about $\hat{S}''_j$.

Unfortunately, to estimate the norm of $\hat{S}''_j$, we need to split the expression up into a number of parts. First, we split it into two parts,

$$\begin{aligned} \hat{S}''_j = & \sum_{l=0}^{j-1} \sum_{\substack{k: p_k < n < d_k \\ p_k \ge j}} \sum_{i=0}^{p_k - 1} \left(X_{k,i+l} - X_{k,i+l+d_k}\right) \\ & + \sum_{l=0}^{j-1} \sum_{\substack{k: p_k < n < d_k \\ p_k < j}} \sum_{i=0}^{p_k - 1} \left(X_{k,i+l} - X_{k,i+l+d_k}\right) = \Sigma_1 + \Sigma_2. \end{aligned} \tag{8.13}$$

The first of these sums may be estimated immediately. For the second sum, we may use a re-ordering similar to the one performed in the proof of Lemma 8.4 to obtain a splitting into three parts, which may in turn be estimated individually. We begin with the first sum.

First of all, we note that for $j \le p_k$, Σ_1 has the same distribution as

$$\hat{\Sigma}_1 = \sum_{l=0}^{j-1} \sum_{\substack{k: p_k < n < d_k \\ p_k \ge j}} \sum_{i=0}^{p_k - 1} \left(X_{k,i+l} - X_{k,i+l+2p_k}\right). \tag{8.14}$$

Indeed, the gap between the random variables $X_{k,i+l}$ and $X_{k,i+l+2p_k}$ is preserved and the variables are identically distributed, so Σ_1 must have the same distribution as $\hat{\Sigma}_1$. Hence, it suffices to estimate the norm of $\hat{\Sigma}_1$.

We let U be the step-up operator on the second index of the $X_{k,j}$. That is, whenever $X_{k,j}$ and $X_{k,j+1}$ are defined, we let $UX_{k,j} = X_{k,j+1}$. Under this operator, the the expression (8.14) may be written as the weighted sum of a co-boundary. Indeed, if we define

$$g_j'' = \sum_{\substack{k:p_k<n<d_k \\ p_k \geq j}} \sum_{h=0}^{2p_k-1} \sum_{i=0}^{p_k-1} X_{k,i+h}, \tag{8.15}$$

we see that

$$g_j'' - Ug_j'' = \sum_{\substack{k:p_k<n<d_k \\ p_k \geq j}} \sum_{i=0}^{p_k-1} \left(X_{k,i} - X_{k,i+2p_k}\right).$$

Hence,

$$\hat{\Sigma}_1 = \sum_{l=0}^{j-1} U^l (g_j'' - Ug_j'') = g_j'' - U^j g_j''.$$

Clearly, U preserves L^2-norm. Hence, by the triangle inequality,

$$\|\hat{\Sigma}_1\| \leq \|g_j''\| + \|U^j g_j''\| \ll \|g_j''\|.$$

But by the triangle inequality and Lemma 8.2 we can estimate this norm:

$$\|g_j''\|^2 \leq \sum_{\substack{k:p_k<n<d_k \\ p_k \geq j}} 4p_k^3 \alpha_k^2 = 4 \sum_{\substack{k:p_k<n<d_k \\ p_k \geq j}} \frac{2^k}{k} \leq \sum_{k:2^k \leq n} \frac{2^k}{k} = O\left(\frac{n}{\log n}\right).$$

Hence,

$$\|\hat{\Sigma}_1\| = O\left(\frac{n}{\log n}\right)^{1/2}. \tag{8.16}$$

We now consider Σ_2. First, we look at the sum over i and l,

$$\sum_{l=0}^{j-1} \sum_{i=0}^{p_k-1} \left(X_{k,i+l} - X_{k,i+l+d_k}\right).$$

As in the proof of Lemma 8.4, we will re-arrange the terms in this sum according to Figure 8.2.

Again, $i+l$ is constant along each diagonal. Performing the splitting, we get

$$\begin{aligned}
\sum_{l=0}^{j-1} \sum_{i=0}^{p_k-1} \left(X_{k,i+l} - X_{k,i+l+d_k}\right) &= \sum_{i=0}^{p_k-2} (i+1) \left(X_{k,i} - X_{k,i+d_k}\right) \\
&\quad + \sum_{i=p_k-1}^{j-1} p_k \left(X_{k,i} - X_{k,i+d_k}\right) \\
&\quad + \sum_{i=0}^{p_k-2} (i+1) \left(X_{k,p_k+j-2-i} - X_{k,p_k+j-2-i+d_k}\right) \\
&= S_1 + S_2 + S_3.
\end{aligned}$$

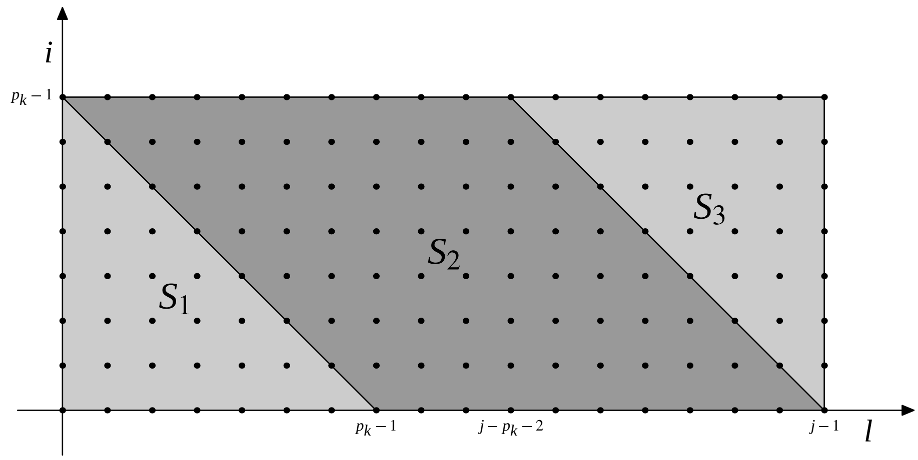

Figure 8.2.: Re-arrangement of the terms.

We examine the terms individually. For S_1, following the argument used to obtain (8.6),

$$\|S_1\|^2 \leq 2\left\|\sum_{i=0}^{p_k-2} iX_{k,i}\right\|^2 \leq 2p_k^3\alpha_k^2 = 2\frac{2^k}{k}.$$

Hence, by Lemma 8.2,

$$\left\|\sum_{\substack{k:p_k<n<d_k\\p_k<j}} S_1\right\|^2 \leq 2\sum_{k:2^k\leq n}\frac{2^k}{k} = O\left(\frac{n}{\log n}\right). \tag{8.17}$$

Analogously,

$$\left\|\sum_{\substack{k:p_k<n<d_k\\p_k<j}} S_3\right\|^2 = O\left(\frac{n}{\log n}\right). \tag{8.18}$$

We now consider the final term. The relevant norm splits up into three sums. To simplify our notation, we introduce new random variables,

$$V_{k,l} = X_{k,l} - X_{k,l+d_k}. \tag{8.19}$$

With this convention,

$$\begin{aligned}\left\|\sum_{l=0}^{j-1} Y_{n,l} - \sum_{\substack{k:p_k<n<d_k\\ p_k<j}} S_2\right\|^2 &= \left\|\sum_{l=0}^{j-1}\sum_{k:p_k<n<d_k} p_k V_{k,l} - \sum_{\substack{k:p_k<n<d_k\\ p_k<j}}\sum_{l=p_k-1}^{j-1} p_k V_{k,l}\right\|^2\\ &= \left\|\sum_{l=0}^{p_k-2}\sum_{\substack{k:p_k<n<d_k\\ p_k<j}} p_k V_{k,l} + \sum_{\substack{k:p_k<n<d_k\\ p_k\geq j}}\sum_{l=0}^{j-1} p_k V_{k,l}\right\|^2\\ &\leq 2\sum_{\substack{k:p_k<n<d_k\\ p_k<j}} p_k^2(p_k-1)\alpha_k^2 + 2\sum_{\substack{k:p_k<n<d_k\\ p_k\geq j}} p_k^2 j\alpha_k^2\\ &= 2\left(\sum_{\substack{k:p_k<n<d_k\\ p_k<j}}\frac{p_k-1}{k} + \sum_{\substack{k:p_k<n<d_k\\ p_k\geq j}}\frac{j}{k}\right).\end{aligned} \tag{8.20}$$

We have used independence of the $X_{k,i}$ and the definition of $V_{k,i}$ to get the inequality, when getting rid of the norm. The first sum may be estimated from Lemma 8.2. Indeed,

$$\sum_{\substack{k:p_k<n<d_k\\ p_k<j}}\frac{p_k-1}{k} \leq \sum_{k:2^k\leq n}\frac{2^k}{k} = O\left(\frac{n}{\log n}\right). \tag{8.21}$$

We only need to estimate the last sum in (8.20). This is where the maximum in the statement of the lemma comes into play. Consider the function

$$\max_{1\leq j\leq n}\frac{1}{n}\sum_{\substack{k:p_k<n<d_k\\ p_k\geq j}}\frac{j}{k} = \max_{1\leq j\leq n}\frac{1}{n}\sum_{\substack{k:\log_3 n<k<\log_2 k\\ \log_2 j\leq k}}\frac{j}{k}.$$

Since $\log_2 j \leq \log_3 n$ if and only if $j \leq n^{\log_3 2}$, the function under the maximum is constant for such j. Hence, the maximum is assumed for some $j \geq n^{\log_3 2}$. We may estimate the function using integrals,

$$\frac{1}{n}\sum_{\substack{k:\log_3 n<k<\log_2 k\\ \log_2 j\leq k}}\frac{j}{k} \leq \frac{j}{n}\sum_{k:\log_2 j<k<\log_2 k}\frac{1}{k} \leq \frac{j}{n}\int_{\log_2 j}^{\log_2 n}\frac{1}{x}dx = \frac{j}{n}\left(\log\log_2 n - \log\log_2 j\right).$$

To estimate the last expression, we use the Mean Value Theorem,

$$\begin{aligned}\log\log_2 n - \log\log_2 j &= \frac{\log\log_2 n - \log\log_2 j}{\log_2 n - \log_2 j}\frac{\log_2 n - \log_2 j}{n-j}(n-j)\\ &\ll \frac{1}{\log_2 j}\frac{1}{j}(n-j).\end{aligned}$$

Since the maximum is assumed for some $j \in [n^{\log_3 2}, n]$, we get

$$\max_{1 \le j \le n} \frac{1}{n} \sum_{\substack{k: p_k < n < d_k \\ p_k \ge j}} \frac{j}{k} \ll \frac{j}{n} \frac{1}{\log_2 j} \frac{1}{j} (n-j) \le \frac{n}{n \log_2 n^{\log_3 2}} = O\left(\frac{1}{\log_2 n}\right). \tag{8.22}$$

Now, by (8.17), (8.18), (8.20), (8.21), and (8.22),

$$\max_{1 \le j \le n} \frac{1}{\sqrt{n}} \left\| \Sigma_2 - \sum_{l=0}^{j-1} Y_{n,l} \right\| = O\left(\frac{1}{\log n}\right)^{1/2}.$$

Along with (8.13), (8.16) and the triangle inequality, this implies the lemma. □

Combining Lemma 8.3, Lemma 8.4 and Lemma 8.5, we get the key lemma of this section.

Lemma 8.6.

$$\max_{1 \le j \le n} \frac{1}{\sqrt{n}} \left\| S_j f - \sum_{l=0}^{j-1} Y_{n,l} \right\| = O\left(\frac{1}{\log n}\right)^{1/2}.$$

This is the first part of the way to Theorem 8.1. However, there are two problems. First of all, the random variables are not properly normalised. We take care of this problem in the next theorem. Secondly, the random variables depend on n. This problem will be solved subsequently.

Theorem 8.7. *Let $(X, \mathcal{B}, \mu)$ be a non-atomic probability space and let $T : X \to X$ be an aperiodic automorphism. For any $n \in \mathbb{N}$ there exists an $f \in L^2(X)$ and independent $\mathcal{N}(0, 2(\log\log 3 - \log\log 2))$ distributed random variables $Z_{n,1}, \dots, Z_{n,n-1}$ such that*

$$\max_{1 \le j \le n} \frac{1}{\sqrt{n}} \left\| S_j f - \sum_{l=0}^{j-1} Z_{n,l} \right\| = O\left(\frac{1}{\log n}\right)^{1/2}.$$

Proof. We may use the same $f \in L^2(X)$ as above. Let $Y_{n,l}$ be as in (8.10) and define

$$Z_{n,l} = \frac{Y_{n,l}}{\|Y_{n,l}\|} \sqrt{2(\log\log 3 - \log\log 2)}. \tag{8.23}$$

For ease of notation, we denote $\sqrt{2(\log\log 3 - \log\log 2)}$ by K in this proof.

These random variables clearly have the right distribution, since the $Y_{n,l}$ are centred. By Lemma 8.6 and the triangle inequality, it suffices to prove that

$$\frac{1}{\sqrt{n}} \left\| \sum_{l=0}^{j-1} \left(Y_{n,l} - Z_{n,l}\right) \right\| = O\left(\frac{1}{\log n}\right)^{1/2} \tag{8.24}$$

for any $j = 1, \dots, n-1$.

Let n be fixed. For any $l = 0, \dots, n-1$,

$$\left\|Y_{n,l}\right\|^2 = 2 \sum_{\log_3 n < k < \log_2 n} \frac{p_k^2}{p_k^2 k} = 2 \sum_{\log_3 n < k < \log_2 n} \frac{1}{k},$$

We need to prove that

$$\left| \left\|Y_{n,l}\right\|^2 - K^2 \right| = O\left(\frac{1}{\log n}\right). \tag{8.25}$$

This amounts to showing that

$$\left| \sum_{\log n/\log 3 < k < \log n/\log 2} \frac{1}{k} - \log\log 3 - \log\log 2 \right| = O\left(\frac{1}{\log n}\right). \tag{8.26}$$

But this follows since for any $a, b \in \mathbb{N}$, $a \leq b$ we have

$$\left| \sum_{i=a}^{b} \frac{1}{k} - \int_a^b \frac{1}{x} dx \right| \leq \sum_{i=a}^{b} \left(\frac{1}{k} - \frac{1}{k+1} \right) \leq \frac{1}{a}.$$

Now,

$$\int_{\log n/\log 3}^{\log n/\log 2} \frac{1}{x} dx = \log\left(\frac{\log n}{\log 2}\right) - \log\left(\frac{\log n}{\log 3}\right) = \log\log 3 - \log\log 2.$$

Even when taking possible error terms into account, coming from the fact that the bounds $\log n/\log 3$ and $\log n/\log 2$ on k may not be integers, we still get (8.26).

Now, let $j \in \{1, \dots, n\}$. By (8.25), since the $Y_{n,l}$ are independent and identically distributed,

$$\begin{aligned} \frac{1}{n} \left\| \sum_{l=0}^{j-1} \left(Y_{n,l} - Z_{n,l}\right) \right\|^2 &\leq \frac{1}{j} \left\| \sum_{l=0}^{j-1} Y_{n,l} \left(1 - \frac{K}{\left\|Y_{n,l}\right\|}\right) \right\|^2 = \left\|Y_{n,l}\right\|^2 \left| 1 - \frac{K}{\left\|Y_{n,l}\right\|} \right|^2 \\ &= \left| \left\|Y_{n,l}\right\| - K \right|^2 \leq \left| \left\|Y_{n,l}\right\|^2 - K^2 \right| = O\left(\frac{1}{\log n}\right). \end{aligned}$$

Taking the square root of this and maximising over j implies the theorem. □

Note that this theorem is a triangular one. That is, for each n there exists an array of random variables $Z_{n,j}$ for which the theorem holds, but the arrays need not be equal on the first entries. We would like a theorem which states that this is the case. That is, a similar theorem in which the dependence on n is removed from the array of random variables. This is essentially Theorem 8.1.

First, we make some definitions of random variables. For any $n \in \mathbb{N}$, we define

$$Y_n = \sum_{k: p_k < n < d_k} p_k \left(X_{k,n} - X_{k,n+d_k}\right). \tag{8.27}$$

Note that if we extend the definition of $Y_{n,l}$ to $1 \leq l \leq n$ (instead of $l < n$), this is the diagonal random variables $Y_{n,n}$. As before, we take the appropriate normalisations,

$$Z_n = \frac{Y_n}{\|Y_n\|}\sqrt{2(\log\log 3 - \log\log 2)} = \frac{Y_n}{\|Y_n\|}K. \tag{8.28}$$

These are the random variables of the theorem.

Proof of Theorem 8.1. The main part of the proof consists in estimating the size of certain sums. Some of these are potentially very close to the bounds, we are looking for. Hence, the following is quite technical. By Lemma 8.6 and the triangle inequality, it is sufficient to prove,

$$\max_{1\leq j\leq n} \frac{1}{\sqrt{n}} \left\| \sum_{l=0}^{j-1} \left(Z_j - Y_{n,j}\right) \right\| = O\left(\frac{\log\log\log\log n}{\log n}\right)^{1/2}. \tag{8.29}$$

We first consider $\left\|\sum_{l=0}^{j-1}(Y_l - Y_{n,l})\right\|$. We use the $V_{k,l}$ defined in (8.19) to simplify notation.

$$\begin{aligned}\sum_{l=0}^{j-1}\left(Y_l - Y_{n,l}\right) &= \sum_{l=0}^{j-1}\left(\sum_{k:p_k<l<d_k} p_k V_{k,l} - \sum_{k:p_k<n<d_k} p_k V_{k,l}\right) \\ &= \sum_{l=0}^{j-1}\sum_{k:p_k<l<d_k\leq n} p_k V_{k,l} - \sum_{l=0}^{j-1}\sum_{k:l\leq p_k<n<d_k} p_k V_{k,l} \\ &= \Sigma_1 - \Sigma_2.\end{aligned}$$

Hence, to obtain an estimate of the norm, it suffices to obtain an estimate of the norm of Σ_1 and Σ_2 individually. Immediately, by independence of the $V_{k,l}$ and definitions of constants,

$$\|\Sigma_1\|^2 \leq 2\sum_{l=0}^{n-1}\sum_{k:p_k<l<d_k\leq n}\frac{1}{k}.$$

Unfortunately, it is not trivial to estimate the double sum in this expression. We split the sum up into two parts,

$$\sum_{l=0}^{\left[n^{\log_3 2}\right]}\sum_{\log_3 l<k<\log_2 l}\frac{1}{k} + \sum_{l=\left[n^{\log_3 2}\right]+1}^{n-1}\sum_{\log_3 l<k<\log_3 n}\frac{1}{k} = S_1 + S_2. \tag{8.30}$$

To obtain estimates of the inner sums, we will use Riemann integrals. From the definition of the Riemann integral, we get for any $a, b \in \mathbb{N}$, $a \leq b$

$$\sum_{k=a}^{b}\frac{1}{k} - \int_a^b \frac{1}{x}dx \leq \sum_{k=a}^{b}\left(\frac{1}{k} - \frac{1}{k+1}\right) \leq \frac{1}{a}. \tag{8.31}$$

From (8.31),

$$\sum_{\log_3 l<k<\log_2 l} \frac{1}{k} \leq \log\log_2 l - \log\log_3 l + \frac{C_1}{\log l} = \log\log 3 - \log\log 2 + \frac{C_1}{\log l}$$

for some $C_1 > 0$. Since the sum on the left hand side is equal to zero for $l \in \{0,1\}$ and the expression on the right hand side is decreasing as a function of l, we may majorise the sum by a constant $C_2 > 0$. Hence,

$$S_1 = \sum_{l=0}^{\left[n^{\log_3 2}\right]} \sum_{\log_3 l<k<\log_2 l} \frac{1}{k} \leq C_2\left(1+n^{[\log_3 2]}\right) = O\left(\frac{n\log\log\log\log n}{\log n}\right). \qquad (8.32)$$

We now estimate S_2. First, by (8.31),

$$\sum_{\log_3 l<k<\log_3 n} \frac{1}{k} \leq \log\left(\frac{\log n}{\log j}\right) + \frac{C_3}{\log n}$$

for some $C_3 > 0$. Hence,

$$\begin{aligned} S_2 &\leq \sum_{l=\left[n^{\log_3 2}\right]+1}^{n-1} \left(\log\left(\frac{\log n}{\log j}\right) + \frac{C_3}{\log n}\right) \\ &= \sum_{l=\left[n^{\log_3 2}\right]+1}^{[n/\log n]} \log\left(\frac{\log n}{\log j}\right) + \sum_{l=[n/\log n]+1}^{n-1} \log\left(\frac{\log n}{\log j}\right) + \frac{C_3 n}{\log n} \\ &= S_2' + S_2'' + S_2'''. \qquad (8.33) \end{aligned}$$

Clearly,

$$S_2''' = O\left(\frac{n\log\log\log\log n}{\log n}\right). \qquad (8.34)$$

Consider S_2'. Here,

$$\begin{aligned} S_2' &= \sum_{l=\left[n^{\log_3 2}\right]+1}^{[n/\log n]} \log\left(\frac{\log n}{\log j}\right) \leq \frac{n}{\log n}\log\left(\frac{\log n}{\log\left(n^{\log_3 2}\right)}\right) \\ &= \frac{n}{\log n}(\log\log 3 - \log\log 2) = O\left(\frac{n\log\log\log\log n}{\log n}\right). \qquad (8.35) \end{aligned}$$

We now estimate S_2''. Let $L = \log\log n$. Since $j \leq n$, we may take a $C_4 > 0$ such that

$$\log\left(\frac{\log n}{\log j}\right) \leq C_4\left(\frac{\log n}{\log j} - 1\right).$$

Now, we split the sum once again,

$$S_2'' = \sum_{l=[n/\log n]+1}^{n-1} \log\left(\frac{\log n}{\log j}\right)$$
$$\leq \sum_{l=[n/\log n]+1}^{[n/L]} C_4\left(\frac{\log n}{\log j} - 1\right) + \sum_{l=[n/L]+1}^{n-1} C_4\left(\frac{\log n}{\log j} - 1\right) = \underline{S}_2'' + \overline{S}_2''. \quad (8.36)$$

We estimate $\underline{S}_2''$.

$$\underline{S}_2'' = \sum_{l=[n/\log n]+1}^{[n/L]} C_4\left(\frac{\log n}{\log j} - 1\right) \leq \frac{nC_4}{L}\left(\frac{\log n}{\log n - \log\log n} - 1\right)$$
$$\leq \frac{2C_4 n}{L}\frac{\log\log n}{\log n} = O\left(\frac{n\log\log\log\log n}{\log n}\right). \quad (8.37)$$

For $\overline{S}_2''$ we get, since $\log n \geq \log\log n = L$,

$$\overline{S}_2'' = \sum_{l=[n/L]+1}^{n-1} C_4\left(\frac{\log n}{\log j} - 1\right) \leq C_4 n\left(\frac{\log n}{\log n - \log L} - 1\right)$$
$$\leq C_4 n\left(\frac{2L}{\log n - \log L} - 1\right) \leq 2nC_4\frac{2\log\log L}{\log n} = O\left(\frac{n\log\log\log\log n}{\log n}\right), \quad (8.38)$$

whenever $n \geq e^{e^e}$. Assume for the remainder of the proof that this is indeed the case. This is what it means for n to be sufficiently large. By (8.36), (8.37) and (8.38),

$$S_2'' = O\left(\frac{n\log\log\log\log n}{\log n}\right). \quad (8.39)$$

Hence, by (8.33), (8.34), (8.35) and (8.39), the same holds for S_2, so by (8.30) and (8.32),

$$\|\Sigma_1\|^2 = O\left(\frac{n\log\log\log\log n}{\log n}\right). \quad (8.40)$$

For the last term,

$$\|\Sigma_2\|^2 \leq 2\sum_{l=0}^{n-1}\sum_{k:l\leq p_k<n<d_k}\frac{1}{k}.$$

This expression may be estimated in a fashion analogous to the previous calculations, so

$$\|\Sigma_2\|^2 = O\left(\frac{n\log\log\log\log n}{\log n}\right). \quad (8.41)$$

Hence we get by (8.40) and (8.41),

$$\left\|\sum_{l=0}^{j-1}(Y_l - Y_{n,l})\right\|^2 = O\left(\frac{n\log\log\log\log n}{\log n}\right). \quad (8.42)$$

With this estimate in place, it is easy to prove the theorem. Once again, K denotes $\sqrt{2(\log\log 3 - \log\log 2)}$. If we can find a good bound on $\left\|\sum_{l=0}^{j-1}(Y_l - Z_l)\right\|^2$, the triangle inequality implies (8.29). Using the independence of the Y_l, we get

$$\begin{aligned}\left\|\sum_{l=0}^{j-1}(Y_l - Z_l)\right\|^2 &= \left\|\sum_{l=0}^{j-1} Y_l\left(1 - \frac{K}{\|Y_l\|}\right)\right\|^2 = \sum_{l=0}^{j-1}(\|Y_l\| - K)^2 \\ &\leq \sum_{l=0}^{j-1}\left\|Y_l - Y_{n,l}\right\| + \sum_{l=0}^{j-1}\left(\left\|Y_{n,l}\right\| - K\right)^2 = O\left(\frac{n\log\log\log\log n}{\log n}\right) + O\left(\frac{n}{\log n}\right),\end{aligned} \tag{8.43}$$

by (8.25) and (8.42). This completes the proof. □

8.2. Invariance principles

In Volný's paper [61], we also find proofs of the weak and the strong invariance principle. Here, we prove the strong invariance principle. We subsequently make a few comments on Volný's proof of the weak invariance principle, and the reasons why it is omitted here. We use the same splitting of the weighted sums as in Section 8.1, so a large amount of references are given in the following.

Theorem 8.8. *There is a Brownian motion $X(n)$, such that*

$$\frac{|S_n f - X(n)|}{\sqrt{n\log\log n}} \to 0 \quad a.s.$$

Proof. Clearly, $\sum_{l=0}^{n-1} Z_l$ has the same distribution as a Brownian motion $X(n)$. We may write

$$S_n f - \sum_{l=0}^{n-1} Z_l = \sum_{l=0}^{n-1}\left(Y_{n,l} - Z_l\right) + \left(S_n f - \sum_{l=0}^{n-1} Y_{n,l}\right). \tag{8.44}$$

By (8.42) and (8.43), we know that

$$\left\|\sum_{l=0}^{n-1}\left(Y_{n,l} - Z_l\right)\right\|^2 = O\left(\frac{n\log\log\log\log n}{\log n}\right), \tag{8.45}$$

where the sum of random variables is itself a Gaussian random variable.

We now split the last term in (8.44) up into eight terms. These are terms, we treated in the proofs of Lemma 8.3, Lemma 8.4 and Lemma 8.5. In particular by (8.2), (8.4) and (8.11),

$$S_n f - \sum_{l=0}^{n-1} Y_{n,l} = \sum_{k:d_k\le n} \hat{g}_k + \sum_{k:d_k\le n} \tilde{g}_k + U^n \sum_{k:d_k\le n} \hat{g}_k + T^n \sum_{k:d_k\le n} \tilde{g}_k + \sum_{k:n\le p_k} \hat{S}_{k,n} + \sum_{k:n\le p_k} \tilde{S}_{k,n} + \left(\hat{S}''_n - \sum_{l=0}^{n-1} Y_{n,l}\right) + \tilde{S}''_n.$$

The first, third, fifth and seventh summand on the right hand side are all Gaussian, and the variance of each one has already been seen to be

$$O\left(\frac{n}{\log n}\right) \le O\left(\frac{n \log\log\log\log n}{\log n}\right).$$

Hence, by the above and (8.44), we may write

$$S_n f - \sum_{l=0}^{n-1} Z_l = G_n + \sum_{k:d_k\le n} \tilde{g}_k + T^n \sum_{k:d_k\le n} \tilde{g}_k + \sum_{k:n\le p_k} \tilde{S}_{k,n} + \tilde{S}''_n, \tag{8.46}$$

where G_n is a Gaussian. Hence, $\hat{G}_n = G_n/\sqrt{n\log\log n}$ is also a Gaussian, and by the above,

$$\left\|\hat{G}_n\right\|^2 = \left\|\frac{G_n}{\sqrt{n\log\log n}}\right\|^2 = O\left(\frac{\log\log\log\log n}{\log n \log\log n}\right). \tag{8.47}$$

Hence, for any $\varepsilon > 0$ we have

$$\mu\left\{x\in X : \left|\hat{G}_n\right| > \varepsilon\right\} = \sqrt{\frac{2}{\pi}} \int_{\varepsilon\sqrt{\frac{\log n \log\log n}{C_1 \log\log\log\log n}}}^{\infty} e^{-x^2/2} dx < n^{-2},$$

where $C_1 > 0$ is the constant implicit in the O in (8.47). By the Borel–Cantelli Lemma, since $\sum_{n=0}^{\infty} n^{-2} < \infty$,

$$\lim_{n\to\infty} \frac{G_n}{\sqrt{n\log\log n}} = \lim_{n\to\infty} \hat{G}_n = 0 \quad a.s. \tag{8.48}$$

We now consider the remaining terms in (8.46). First, let $\tilde{g} = \sum_{k=1}^{\infty} |\tilde{g}_k|$. Then, $\left|\sum_{k:d_k\le n} \tilde{g}_k\right| \le \tilde{g}$ for any $n \in \mathbb{N}$. As in the proof of Lemma 8.3, $\|\tilde{g}_k\| \le d_k p_k \varepsilon_k = 6^{-2k}$. Hence, $\tilde{g} \in L^2(X)$, so for any $\varepsilon > 0$,

$$\sum_{n=1}^{\infty} \mu\left\{x\in X : \left|T^n \sum_{k:d_k\le n} \tilde{g}_k(x)\right| > \varepsilon\sqrt{n}\right\} \le \sum_{n=1}^{\infty} \mu\left\{x\in X : (\tilde{g}(x))^2 > \varepsilon^2 n\right\} < \infty.$$

Hence, the Borel–Cantelli Lemma implies,

$$\lim_{n\to\infty} \frac{\left|T^n \sum_{k:d_k\le n} \tilde{g}_k(x)\right|}{\sqrt{n\log\log n}} \le \lim_{n\to\infty} \frac{\left|T^n \sum_{k:d_k\le n} \tilde{g}_k(x)\right|}{\sqrt{n}} = 0 \quad a.s. \tag{8.49}$$

In exactly the same manner,

$$\lim_{n\to\infty} \frac{\left|\sum_{k:d_k\leq n} \tilde{g}_k(x)\right|}{\sqrt{n\log\log n}} = 0 \quad a.s. \tag{8.50}$$

Now, we consider the remaining random variable. That is, we define

$$R_n = \sum_{k:n\leq p_k} \tilde{S}_{k,n} + \tilde{S}''_n.$$

By (8.5) and (8.12), $\|R_n\| = O\left(\frac{1}{n^2}\right)$. By Chebychev's inequality,

$$\mu\left\{x \in X : |R_n(x)| > \varepsilon\sqrt{n}\right\} \leq \frac{C_2}{\varepsilon^2 n^3}$$

for some $C_2 > 0$ implicit in the O. Since $\sum_{n=0}^{\infty} n^{-3} < \infty$, the Borel–Cantelli Lemma implies that

$$\lim_{n\to\infty} \frac{|R_n|}{\sqrt{n\log\log n}} \leq \lim_{n\to\infty} \frac{|R_n|}{\sqrt{n}} = 0 \quad a.s. \tag{8.51}$$

Now, the theorem follows from (8.46), (8.48), (8.49), (8.50) and (8.51). □

Volný proves both a weak and a strong invariance principle. To state the weak form, we define another more general form of partial sum. For $t \in [0,1]$, we define

$$S_n(t) = \frac{S_{[tn]}f}{\sqrt{n}} + \frac{\{tn\}}{\sqrt{n}} T^{[tn]} f. \tag{8.52}$$

The weak invariance principle states:

Theorem 8.9 (The Weak Invariance Principle).

$$S_n(t) \xrightarrow{\mathcal{D}} X(t),$$

where $X(t)$ is some Brownian motion.

The proof of the weak invariance principle is based on the fact that Theorem 8.9 is implied by two properties. The first is the convergence of all finite dimensional distributions of the $S_n(t)$ to the Brownian motion. That is, for any $k \in \mathbb{N}$ and any vector $(t_1,\ldots,t_k) \in [0,1]^k$,

$$(S_n(t_1),\ldots,S_n(t_k)) \xrightarrow{\mathcal{D}} (X(t_1),\ldots,X(t_k)).$$

This follows directly from Theorem 8.7

The second property is tightness of the sequence $(S_n(t))$ in $C([0,1])$. Tightness is again implied by the fact that for any $\varepsilon > 0$ there is a $\lambda > 1$ such that

$$\mu\left\{x \in X : \max_{1\leq j\leq n} \left|S_j f(x)\right| > \lambda\sqrt{n}\right\} < \frac{\varepsilon}{\lambda^2}. \tag{8.53}$$

This is a consequence of Theorem 8.4 in [9]. Volný proves this last property using a more refined splitting than the one which lead to Theorem 8.7. For each component, he proves that (8.53) holds. This immediately implies that (8.53) holds for the partial sums.

To see that these two properties together implies the weak invariance principle, we need only note that the sequence is in $C([0,1])$. By Theorem 8.1 in [9], the sequence converges to the Brownian motion from the first of the two properties.

We do not go deeper into the proof of the Weak Invariance Principle. The first reason is the increased technicalities involved in a further splitting of the sums. The second reason is that (8.53) is highly dependent on the stationarity of the sequence $S_j f$. Since we will be dealing with weighted sums in the next chapter, the sequence of random variables involved is no longer stationary. Hence the omission.

9. Weighted Gaussian approximation

In this chapter, we generalise the results in the previous chapter to weighted partial sums. We take two approaches to the problem.

The first approach, which we discuss in Section 9.1 is to take Volný's methods and apply them directly to weighted sums instead of the non-weighted sums treated by himself. As we go along, we impose hypotheses on the sequence of weights. This approach does in fact yield some results, but the hypotheses we need on the sequence of weights are both rather strange and rather restrictive. This illustrated the limitations of Volný's methods, when no new elements are added.

In Section 9.2, we add an additional element to Volný's approach and with little effort obtain a weighted result, where the hypothesis on the weights is less restrictive and much more natural.

9.1. The direct approach

In this section, we will discuss the approach of proving a generalisation of Volný's theorem on Gaussian approximation in L^2 (Theorem 8.1) to weighted sums instead of the partial sums studied in Chapter 8 using only Volný's methods. For the remainder of this section, let $(a_n)_{n=0}^{\infty}$ be a sequence of positive reals bounded from above by some $M > 0$. We now state three hypotheses on the weights, which we will need to arrive at a generalisation. We will return to a discussion of these at a later point. The first one is

$$\Delta = \sum_{n=1}^{\infty} |a_n - a_{n-1}| < \infty. \tag{H1}$$

Our second hypothesis is more technical. Namely,

$$\left|\mathcal{A}_n - K^2\right| = \left| 2 \sum_{\log_3 n < k < \log_2 n} \frac{\left(\sum_{i=n-2^k+1}^{n} a_i\right)^2}{4^k k} - K^2 \right| = O\left(\frac{1}{\log n}\right), \tag{H2}$$

where $K > 0$ is some constant. Our third and final hypothesis looks even more complicated.

$$\frac{1}{n} \sum_{\log_3 n < k < \log_2 n} \sum_{l=\left[n^{\log_3 2}\right]}^{n-1} \frac{\left(\sum_{i=l-2^k+1}^{n-2^k+2} \sum_{h=0}^{2^k-2} (a_{i+h} - a_{i+h-1})\right)^2}{4^k k} = O\left(\frac{\log\log\log\log n}{\log n}\right). \tag{H3}$$

The hypotheses may seem restrictive, but they certainly include the standard partial sums. Clearly, (H1) and (H3) holds, since the involved sums are equal to zero. (H2) also holds, since $a_n = C$ implies that $K = C\sqrt{2(\log\log 3 - \log\log 2)}$ as we saw in (8.25). In the same way, we see that the hypotheses are all true if the sequence (a_n) is eventually constant. In fact, we always have that $0 < K < M\sqrt{2(\log\log 3 - \log\log 2)}$, as is seen by majorising the terms of the sum in (H2).

Example. To prove that our Hypotheses does not exclude all non-trivial sequences, we produce an example. From Hypothesis (H1), it is clear that the sequence must be convergent, and from Hypothesis (H2) it follows that the limit must be different from zero. Indeed, otherwise K would be zero. Let $\alpha \in (0,1)$ and let $a_n = 1 + \alpha^n$ for $n = 0, 1, \ldots$. We claim that the Hypotheses (H1), (H2) and (H3) are true for this sequence.

We prove that (H1) holds.

$$\sum_{n=1}^{\infty} |a_n - a_{n-1}| = (1-\alpha) \sum_{n=1}^{\infty} \alpha^{n-1} = 1.$$

Hence, the (H1) certainly holds. To prove Hypothesis (H2), we note that

$$\left(\sum_{i=n-2^k+1}^{n} a_i\right)^2 = \left(2^k + \sum_{i=n-2^k+1}^{n} \alpha^i\right)^2 = 4^k + 2\cdot 2^k \sum_{i=n-2^k+1}^{n} \alpha^i + \left(\sum_{i=n-2^k+1}^{n} \alpha^i\right)^2.$$

Hence, $\mathcal{A}_n$ splits up into three sums,

$$2 \sum_{\log_3 n < k < \log_2 n} \frac{1}{k} + 4 \sum_{\log_3 n < k < \log_2 n} \frac{\sum_{i=n-2^k+1}^{n} \alpha^i}{2^k k} + 2 \sum_{\log_3 n < k < \log_2 n} \frac{\left(\sum_{i=n-2^k+1}^{n} \alpha^i\right)^2}{4^k k}.$$

For the first sum, we know from (8.25),

$$\left| 2 \sum_{\log_3 n < k < \log_2 n} \frac{1}{k} - 2(\log\log 3 - \log\log 2) \right| = O\left(\frac{1}{\log n}\right).$$

Hence, it suffices to prove that the two remaining terms converge to zero rapidly enough. We let $A = 1/(1-\alpha) = \sum_{i=0}^{\infty} \alpha^i$. Then,

$$4 \sum_{\log_3 n < k < \log_2 n} \frac{\sum_{i=n-2^k+1}^{n} \alpha^i}{2^k k} \leq 4A \sum_{\log_3 n < k} \frac{1}{2^k k} = O\left(\frac{1}{n^{\log_3 2} \log n}\right) \leq O\left(\frac{1}{\log n}\right).$$

Similarly,

$$2\sum_{\log_3 n<k<\log_2 n}\frac{\left(\sum_{i=n-2^k+1}^{n}\alpha^i\right)^2}{4^k k}\leq 2A^2\sum_{\log_3 n<k}\frac{1}{4^k k}=O\left(\frac{1}{\log n}\right). \tag{9.1}$$

This completes the proof of Hypothesis (H2).

The third Hypothesis (H3) is in fact even easier to check. First, we see that

$$\begin{aligned}\left(\sum_{i=l-2^k+1}^{n-2^k+2}\sum_{h=0}^{2^k-2}(a_{i+h}-a_{i+h-1})\right)^2 &= \left(\sum_{i=l-2^k+1}^{n-2^k+2}\sum_{h=0}^{2^k-2}\left(\alpha^{i+h}-\alpha^{i+h-1}\right)\right)\\ &= \left(\sum_{i=l-2^k+1}^{n-2^k+2}\alpha^{i+2^k-2}-\sum_{i=l-2^k+1}^{n-2^k+2}\alpha^{i-2}\right)^2\leq 2A^2.\end{aligned}$$

Hence, the left hand side in Hypothesis (H3) is less than or equal to

$$2A^2\frac{1}{n}\sum_{\log_3 n<k<\log_2 n}\sum_{l=\lceil n^{\log_3 2}\rceil}^{n-1}\frac{1}{4^k k}\leq 2A^2\sum_{\log_3 n<k<\log_2 n}\frac{1}{4^k k}=O\left(\frac{1}{\log n}\right),$$

as in (9.1). Hence, Hypothesis (H3) also holds.

Now, let $(X,\mathcal{B},\mu)$ be a non-atomic probability space and $T:X\to X$ be an aperiodic, ergodic automorphism. For any $f\in L^2(X)$ and any $j\in\mathbb{N}$, we define the n'th weighted partial sum,

$$A_j f=\sum_{l=0}^{j-1}a_l T^l f. \tag{9.2}$$

Following Volný's methods, we may prove the following theorem:

Theorem 9.1. *Under the hypotheses* (H1), (H2) *and* (H3), *there exists an* $f\in L^2(X)$ *and independent random variables* $Z_j\sim\mathcal{N}(0,K^2)$, *where* K *is defined by* (H2), *such that for* n *sufficiently large*

$$\max_{1\leq j\leq n}\frac{1}{\sqrt{n}}\left\|A_j f-\sum_{l=0}^{j-1}Z_l\right\|=O\left(\frac{\log\log\log\log n}{\log n}\right)^{1/2}.$$

Note that for $a_n=1$, this is Theorem 8.1.

We will briefly discuss the method of proof. The proof follows the same method as the one we used in Chapter 8. However, occasional problems present themselves, since the sums are weighted. The first problem occurs in the proof of a weighted analogue of Lemma 8.3. In the weighted case, the partial sums of a co-boundary are no longer telescoping. Hence, there is an additional term in the proof, which needs to be controlled. This is the function of Hypothesis (H1), which tells us that the non-telescoping part of the weighted sum of a co-boundary may be controlled.

Hypothesis (H1) gives us sufficient mileage to arrive at a weighted analogue of Lemma 8.6. However, as was the case in the non-weighted setting, the random variables are not identically distributed. In this first weighted setting, the appropriate definitions are

$$Y_{n,l} = \sum_{k:p_k<n<d_k} \left(\sum_{i=n-(p_k-1)}^{n} a_i \right) \left(X_{k,l} - X_{k,l+d_k} \right), \quad Z_{n,l} = \frac{Y_{n,l}}{\left\| Y_{n,l} \right\|} K, \tag{9.3}$$

where K is the constant defined in (H2). This illustrates the function of Hypothesis (H2), which is needed to ensure that we may find the variance of the limiting distribution of the random variables, to which the weighted partial sums are close. This takes us as far as an analogue of the triangular Theorem 8.7:

Theorem 9.2. *Let (a_n) be a bounded sequence of positive reals such that* (H1) *and* (H2) *hold. Let $(X, \mathcal{B}, \mu)$ be a non-atomic probability space and let $T : X \to X$ be an aperiodic automorphism. For any $n \in \mathbb{N}$ there exists an $f \in L^2(X)$ and independent $\mathcal{N}(0, K^2)$ distributed random variables $Z_{n,1}, \ldots, Z_{n,n-1}$, where K is defined by* (H2) *such that*

$$\max_{1 \le j \le n} \frac{1}{\sqrt{n}} \left\| A_j f - \sum_{l=0}^{j-1} Z_{n,l} \right\| = O\left(\frac{1}{\log n} \right)^{1/2}.$$

To obtain the full non-triangular result, we let the random variables of the triangular array tend to the diagonal. That is, we consider the difference between the partial sums of the $Y_{n,l}$ and the Y_n defined by

$$Y_n = \sum_{k:p_k<n<d_k} \left(\sum_{i=n-(p_k-1)}^{n} a_i \right) \left(X_{k,n} - X_{k,n+d_k} \right).$$

The essence of Hypothesis (H3) is that the difference between the partial sums of the off-diagonal $Y_{n,l}$ and the partial sums of the diagonal Y_n may be controlled. With that assumption, we may obtain Theorem 9.1.

In fact, Volný's methods will also allow us to obtain a Strong Invariance Principle for weighted sums as in Theorem 8.8. However, it is not immediately possible to obtain the Weak Invariance Principle by his methods. This is because his approach depends critically on the fact that the random process $\{T^i f\}$ is stationary. This is not generally the case for the weighted process $\{a_i T^i f\}$.

9.2. Abel summation

The results in the preceding section are somewhat artificial. Though the result is indeed a result in Gaussian approximation, the weighted sums behave more or less as the standard partial sums, since we are only allowed weights that oscillate very little around a constant sequence. Such sequences do not occur naturally in probability theory. Therefore, we take another approach, which will give us a Gaussian approximation result under weaker assumptions. This work is joint work between this author and M. Weber ([37]).

Let (a_l) be a decreasing sequence of positive numbers and assume that

$$\frac{n\log\log\log\log n}{\log n} = o\left(\sum_{l=0}^{n-1} a_l^2\right). \tag{H4}$$

Under this assumption, we prove the following theorem:

Theorem 9.3. *Let (a_l) be a decreasing sequence satisfying Hypothesis* (H4). *There exist independent random variables $Z_l \sim \mathcal{N}(0, 2(\log\log 3 - \log\log 2))$ and a function $f \in L^2(X, \mathcal{B}, \mu)$ such that*

$$\lim_{n\to\infty} \left\| \frac{A_n f}{\left(\sum_{l=0}^{n-1} a_l^2\right)^{1/2}} - \frac{\sum_{i=0}^{n-1} a_i Z_i}{\left(\sum_{l=0}^{n-1} a_l^2\right)^{1/2}} \right\| = 0.$$

Furthermore,

$$\lim_{n\to\infty} \frac{\|A_n f\|}{\left(\sum_{l=0}^{n-1} a_l^2\right)^{1/2}} = \sqrt{2(\log\log 3 - \log\log 2)}.$$

This theorem generalises Theorem 8.1. Indeed, for the constant sequence $a_i = 1$ it is exactly this theorem. Also, the theorem gives us control over the variance of $A_n f$. There are several other sequences of weights for which the theorem holds. We give examples at the end of this chapter.

Proof of Theorem 9.3. Let $f \in L^2(X, \mathcal{B}, \mu)$. For $l \in \mathbb{N}$, we consider the sums

$$S_l f = \sum_{i=0}^{l-1} T^i f, \quad A_l f = \sum_{i=0}^{l-1} a_i T^i f \quad \text{and} \quad \overline{A}_l f = \sum_{i=0}^{l-1} (a_i - a_{i+1}) S_i f.$$

Note that

$$A_l f = \overline{A}_l f + a_l S_l. \tag{9.4}$$

This is the application of Abel summation.

Let $f \in L^2(X, \mathcal{B}, \mu)$ and Z_l be the function and the random variables from Theorem 8.1. That is, there is a $K > 0$ for which

$$\max_{0\le l\le n} \left\| S_l f - \sum_{i=0}^{l-1} Z_i \right\| \le K \left(\frac{n\log\log\log\log n}{\log n}\right)^{1/2}.$$

Hence, for any $l \le n$, we have

$$\left\| \overline{A}_l f - \sum_{i=0}^{l-1} (a_i - a_{i+1}) \left(\sum_{j=0}^{i-1} Z_j \right) \right\| \le \sum_{i=0}^{l-1} |a_i - a_{i+1}| \left\| S_l f - \sum_{i=0}^{l-1} Z_i \right\|$$
$$\le K \sum_{i=0}^{l-1} |a_i - a_{i+1}| \left(\frac{n \log\log\log\log n}{\log n} \right)^{1/2} \quad (9.5)$$

Furthermore,

$$\left\| a_l S_l f - a_l \sum_{i=0}^{l-1} Z_i \right\| \le K |a_l| \left(\frac{n \log\log\log\log n}{\log n} \right)^{1/2}. \quad (9.6)$$

Since

$$\sum_{i=0}^{l-1} (a_i - a_{i+1}) \left(\sum_{j=0}^{i-1} Z_j \right) = \sum_{i=0}^{l-1} a_i Z_i - a_l \sum_{i=0}^{l-1} Z_i,$$

we have by (9.4), (9.5) and (9.6),

$$\max_{0 \le l \le n} \left\| A_l f - \sum_{i=0}^{l-1} a_i Z_i \right\| = \max_{0 \le l \le n} \left\| \overline{A}_l f + a_l S_l - \sum_{i=0}^{l-1} a_i Z_i - a_l \sum_{i=0}^{l-1} Z_i + a_l \sum_{i=0}^{l-1} Z_i \right\|$$
$$\le K \max_{0 \le l \le n} \left(\sum_{i=0}^{l-1} |a_i - a_{i+1}| + |a_l| \right) \left(\frac{n \log\log\log\log n}{\log n} \right)^{1/2}$$
$$= K a_0 \left(\frac{n \log\log\log\log n}{\log n} \right)^{1/2}, \quad (9.7)$$

since the sequence of weights is assumed to be decreasing. Since we also have assumed Hypothesis (H4), we have

$$\lim_{n \to \infty} \left\| \frac{A_n f}{\left(\sum_{i=0}^{n-1} a_i^2 \right)^{1/2}} - \frac{\sum_{i=0}^{n-1} a_i Z_i}{\left(\sum_{i=0}^{n-1} a_i^2 \right)^{1/2}} \right\| = 0. \quad (9.8)$$

Hence, to complete the proof of the first part of the theorem, it suffices to prove that the random variables have the right distribution. But this follows since

$$\left\| \frac{\sum_{i=0}^{n-1} a_i Z_i}{\left(\sum_{i=0}^{n-1} a_i^2 \right)^{1/2}} \right\|^2 = \frac{\sum_{i=0}^{n-1} a_i^2 \|Z_i\|^2}{\sum_{i=0}^{n-1} a_i^2} = \|Z_0\|^2 = 2(\log\log 3 - \log\log 2), \quad (9.9)$$

since the Z_i are independent and identically distributed. Also, (9.8) and (9.9) immediately imply the second part of the theorem. □

Slightly strengthening the hypothesis on the weights proves useful in applications of Theorem 9.3. The strengthened hypothesis we are interested in states that for some $\eta > 0$,

$$\frac{n \log\log\log\log n}{\log n} (\log n)^{\eta} = o\left(\sum_{l=0}^{n-1} a_l^2\right). \tag{H5}$$

With this strengthened hypothesis, we get the following corollary of the proof of Theorem 9.3:

Corollary 9.4. *Let (a_l) be a decreasing sequence satisfying Hypothesis* (H5). *There exist independent random variables $Z_l \sim \mathcal{N}(0, 2(\log\log 3 - \log\log 2))$ and a $\delta > 0$ such that*

$$\left\| \frac{A_n f}{\left(\sum_{l=0}^{n-1} a_l^2\right)^{1/2}} - \frac{\sum_{i=0}^{n-1} a_i Z_i}{\left(\sum_{l=0}^{n-1} a_l^2\right)^{1/2}} \right\| = O\left(\frac{1}{(\log n)^{\delta}}\right).$$

Proof. This is immediate from (9.7) and (H5), with $\delta = \eta/2$. □

Theorem 9.3 and Corollary 9.4 may be applied to find functions for which the Central Limit Theorem, the Almost Sure Central Limit Theorem and the Law of the Iterated Logarithm hold for the weighted partial sums under mild assumptions on the weights. These results will be published elsewhere ([37]). For now, we give two examples of sequences satisfying (H4) and (H5). The first one is a trivial one, which implies that our results has Volný's as a consequence. The second one illustrates that there is in fact new information in the results of this chapter.

Example. (i) Let $c > 0$ be some number and let $a_i = c$. This sequence fulfils both (H4) and (H5). Indeed, since

$$\sum_{i=0}^{n-1} a_i^2 = nc^2,$$

we have (H4). To see that (H5) is satisfied, note that any $\eta \in (0,1)$ will do.

(ii) Let $a \in [0, \frac{1}{2})$ and define $a_i = 1/(\log i)^a$ for $i > 1$. This certainly satisfies (H4), since

$$\frac{n \log\log\log\log n}{\left(\sum_{i=0}^{n-1} a_n^2\right) \log n} \ll \frac{n \log\log\log\log n}{n \frac{1}{(\log n)^{2a}} \log n} = \frac{\log\log\log\log n}{(\log n)^{1-2a}},$$

which tends to zero as n tends to infinity, since $1 - 2a > 0$. Since we may choose $\eta > 0$ such that $1 - 2a - \eta > 0$, this sequence also satisfies (H5).

10. Further research problems

In this chapter, we describe some ideas for further research in the field of Gaussian approximation in ergodic theory. We discuss three such ideas, for all of which we describe some ideas which may lead to results in the field. Some of these ideas are quite loose, mainly due to the fact that they have not been pursued so far.

The first problem we describe is a multi-dimensional generalisation of the material in the preceding two chapters. The question is, if there exists an L^2 function in a non-atomic probability space $(X, \mathcal{B}, \mu)$, such that for d commuting, aperiodic, measure preserving transformations $T_1, \ldots, T_d$, the partial sums of f under each such transformation is well approximated by the sum of Gaussian random variables. We give some partial results in this direction and describe where this method fails.

The second problem is the question of, whether or not the approximation theorems remain valid in other norms than the L^2-norm. In the preceding chapters, we have been concerned with approximation in L^2. However, a number of other Banach spaces are contained in L^2, and it is approximation inside these, we will discuss in Section 10.2. Only a few ideas in this direction are given.

In Section 10.3, we discuss a final research problem. Namely, the problem of extending the material in this part of the thesis to non-ergodic dynamical systems. Our methods so far have been highly dependent on the ergodicity of the systems, and only loose ideas for overcoming this requirement are given in this thesis.

10.1. Gaussian approximation for $\mathbb{Z}^d$-actions

In this section, let $(X, \mathcal{B}, \mu)$ be a non-atomic probability space, let $d \in \mathbb{N}$ and let $\mathbb{Z}^d$ act on X by measure preserving transformations T_v, $v \in \mathbb{Z}^d$. We assume that the action is ergodic. For such a system, we may define the partial sums along the vector $v \in \mathbb{Z}^d$ of the function $f \in L^2(X)$,

$$S_j^{(v)} f = \sum_{l=0}^{j-1} T_v^j f. \tag{10.1}$$

Can we find an $f \in L^2(X)$ such that Theorem 8.1 holds for all $v \in \mathbb{Z}^d$ or at least all v in some subset $V \subseteq \mathbb{Z}^d$? For $V = \{e_1, \ldots, e_d\}$, this is the question asked in the beginning of this chapter.

To this author, it seems that the key to answering the above question must lie in a generalisation of Proposition 7.19. Such a generalisation must depend on a multi-dimensional version of Rokhlin's Lemma. In fact, it is possible to obtain such a generalisation. In the following, we do just that.

Theorem 10.1. *Let $(X, \mathcal{B}, \mu)$ be a non-atomic probability space. Let $d \in \mathbb{N}$ and let $(T_v)_{v \in \mathbb{Z}^d}$ be a $\mathbb{Z}^d$-action by aperiodic measure preserving transformations on X. Let $n \in \mathbb{N}$, let $\varepsilon > 0$ and let $I_n = \{0, \ldots, n-1\}^d \subseteq \mathbb{Z}^d$. There exists an $E \in \mathcal{B}$ such that*

$$T_v E \cap T_{v'} E = \emptyset \quad \textit{for any } v, v' \in I_n, v \neq v', \tag{10.2a}$$

$$\mu\left(\bigcup_{v \in I_n} T_v E\right) > 1 - \varepsilon. \tag{10.2b}$$

Proof. Let $\preceq_l$ be the lexicographical ordering on $\mathbb{Z}^d$. We define the following ordering on $(\mathbb{N} \cup \{0\})^d$: Let $v, w \in (\mathbb{N} \cup \{0\})^d$. $v \preceq w$ if and only if

$$\max(v_1, \ldots, v_d) < \max(w_1, \ldots, w_d) \text{ or} \tag{10.3a}$$

$$\max(v_1, \ldots, v_d) = \max(w_1, \ldots, w_d) \text{ and } v \preceq_l w. \tag{10.3b}$$

It is easy to see, that this is a total ordering. For the remainder of this proof, the ordering used (explicitly or implicitly) is $\preceq$. Now, for some $B \in \mathcal{B}$ with $\mu(B) < n^{-d}\varepsilon$, we define the function

$$v_B(x) = \inf_{\preceq} \left\{ v \in (\mathbb{N} \cup \{0\})^d : T_v x \in B \right\}.$$

Furthermore, for $u \in \mathbb{Z}_+^d \setminus \{0\}$, we define sets,

$$B_u = \{x \in B : v_B(x) = u\}.$$

Since the action is ergodic, the sets $T_v B_u$, $u \in \mathbb{Z}_+^d \setminus \ \{0\}, 0 \preceq v \prec u$ cover X up to a set of measure zero. Indeed, otherwise the complementary set of the union of the B_v would be invariant under the action with measure between zero and one, contradicting the ergodicity. Clearly, the sets are also disjoint.

We will construct the set E as the union of some of the disjoint sets defined above. Define

$$E = \bigcup_{u \succeq (n, \ldots, n)} \bigcup_{v=0}^{\left(\left[\frac{u_1 - n - 1}{n}\right], \ldots, \left[\frac{u_d - n - 1}{n}\right]\right)} T_{n \cdot v} B_u.$$

Clearly, $\{T_v E\}_{v \in I_n}$ is a family of pairwise disjoint sets. Furthermore, the construction yields

$$\mu\left(X \setminus \bigcup_{v \in I_n} T_v E\right) \leq n^d \sum_{v \in \mathbb{Z}_+^d \setminus \{0\}} \mu(B_k) \leq n^d \mu(B) < \varepsilon.$$

This completes the proof. □

With the above theorem, it is relatively simple to extend Proposition 7.19 to $\mathbb{Z}^d$-actions.

Proposition 10.2. *Let $(X,\mathcal{B},\mu)$ be a non-atomic probability space. Let $d \in \mathbb{N}$ and let $(T_v)_{v\in\mathbb{Z}^d}$ be a $\mathbb{Z}^d$-action by aperiodic measure preserving transformations on X. Let $(X_v)_{i\in\mathbb{Z}^d}$ be an ergodic strictly stationary process (defined on a different probability space), let $\xi_1 = \{A_1,\dots,A_K\}$ be a partition of X, $n \in \mathbb{N}$, and let $\varepsilon > 0$. Let I_n be as in Theorem 10.1. Then, there exist a measurable finite valued function f such that:*

1. *There exists a random array $(X'_v)_{v\in I_n}$ distributed as $(X_v)_{v\in I_n}$ on X,*

$$\mu\{\exists v \in I_n, \left|T_v f - X'_v\right| > \varepsilon\} < \varepsilon. \tag{10.4a}$$

2. *The partition η_1 generated by $(T_v f)_{v\in I_n}$ is ε-independent of ξ_1 in the sense that*

$$\sum_{A\in\eta_1, B\in\xi_1, \mu(B)>0} |\mu(A)-\mu(A|B)| < \varepsilon. \tag{10.4b}$$

Proof. As before, we begin with the assumption that X_v is finitely valued. We may choose a representation of the process as an ergodic dynamical system with $\mathbb{Z}^d$-action. Let $(X',\mathcal{C},\nu,S_v)$ where $v \in \mathbb{Z}^d$ be an ergodic dynamical system with some integrable function Y such that $S_vY \sim X_v$. This is such a representation.

As in the proof of Proposition 7.19, we may without loss of generality choose the representation in such a way, that there exists sub-σ-algebras $\mathcal{C}', \mathcal{C}'' \subseteq \mathcal{C}$ with the properties that

1. $\mathcal{C}'$ and $\mathcal{C}''$ are S_v invariant.
2. Both $(X',\mathcal{C}',\nu,S_v)$ and $(X',\mathcal{C}'',\nu,S_v)$ are ergodic.
3. Y is $\mathcal{C}''$ measurable.

Now, let $N \in \mathbb{N}$ and $\varepsilon' > 0$ to be fixed later. Let $F \in \mathcal{B}$ be the set corresponding to these values from Theorem 10.1. That is, $\{T_vF\}$ fulfils (10.2a) and (10.2b) with ε replaced by ε' and for $v, v' \in I_N$. By the same theorem, we may choose a family of sets $\{S_vE\} \subseteq \mathcal{C}'$, $v \in I_N$ with the same properties. Furthermore, the base sets E and F may be chosen such that $\mu(F) = \nu(E)$. Let π_1 be the partition of X generated by $\{T_vF\}$, $v \in I_N$ and π_2 be the partition of X' generated by $\{S_vE\}$, $v \in I_N$. Clearly, these are identically distributed.

We now choose a family of measure preserving bijections,

$$\phi_v : T_vF \to S_vE, \quad v \in I_N.$$

For the exceptional set, we choose another measure preserving bijection,

$$\phi_\varepsilon : X \setminus \bigcup_{v\in I_N} T_vF \to X' \setminus \bigcup_{v\in I_N} S_vE.$$

We combine the whole thing to a measure preserving bijection which maps π_1 to π_2.

$$\phi(x) = \begin{cases} \phi_v(x) & \text{for } x \in T_v F, \\ \phi_\varepsilon(x) & \text{otherwise.} \end{cases}$$

We map the partition ξ_1 to X' through this function, that is, $\xi_2 = \phi(\xi_1)$.

Now, let η_2 be the partition of X' induced by $\{S_v Y\}$, $v \in I_n$. We map this partition to X. That is, $\eta_1 = \phi^{-1}(\eta_2)$. Since $\mathcal{C}'$ is independent of $\mathcal{C}''$, the partition $\pi_2 \vee \xi_2$ is independent of η_2. Hence, η_1 is independent of $\pi_1 \vee \xi_1$.

We know, that all images of Y under S_v, $v \in I_n$ are constant on the cells of the partition $\pi_2 \vee \xi_2 \vee \eta_2$, except possibly on the exceptional set in π_2. Hence, we may define a finitely valued function on X by

$$f(x) = \begin{cases} Y(\phi(x)) & \text{for } x \in T_v F, v \in I_N, \\ 0 & \text{otherwise,} \end{cases} \tag{10.5}$$

and random variables,

$$X'_v(x) = S_v\left(Y\left(\phi(x)\right)\right). \tag{10.6}$$

Clearly,

$$(X_v)_{v \in I_n} \sim \left(X'_v\right)_{v \in I_n}.$$

We now consider

$$x \in \bigcup_{v \in I_{N-n}} T_v F.$$

Following from the construction, we know the values of $T_w f(x)$ for any $w \in I_n$, since we know which cells of the partition $\pi_1 \vee \xi_1 \vee \eta_1$, $T_w x$ belongs to. In fact, by (10.5) and (10.6),

$$T_w f(x) = X'_w(x) \quad \text{for } x \in \bigcup_{v \in I_{N-n}} T_v F, w \in I_n.$$

By Theorem 10.1, we may choose F, N and ε' such that

$$\mu\left(\bigcup_{v \in I_{N-n}} T_v F\right) > 1 - \varepsilon.$$

This implies (10.4a). To prove (10.4b), we note that any dependence between the partitions must come from the exceptional set. That is, the set of measure $< \varepsilon$.

It only remains to be shown that the proposition remains true without the assumption that X_v be finitely valued. To see this, we note that we may approximate any X_v by finitely valued $\hat{X}_v^{(h)}$ for each $h \in \mathbb{N}$. Thus, we obtain partitions $\eta_1^{(h)}$ and $\eta_2^{(h)}$ in the construction above for each $h \in \mathbb{N}$. We may choose this approximation such that $\eta_2^{(h+1)}$ is a refinement of $\eta_2^{(h)}$. In this way, we ensure that the convergence takes place in both spaces. □

Proposition 10.3. *Let n_k be positive integers, $(X_{k,v})_{v\in\mathbb{Z}^\nu}$ be ergodic strictly stationary processes (defined on different probability spaces), $\varepsilon_k > 0, k = 1,2,\ldots$. Then there exist measurable finite valued functions f_k such that*

1. *for every $k = 1,2,\ldots$ there exists a random matrix $(X'_{k,v})_{v\in I_{n_k}}$ with the same distribution as $(X_{k,v})_{v\in I_{n_k}}$ on X,*

$$\mu\left\{\exists v\in I_{n_k}, \left|T_v f_k - X'_{k,v}\right| > \varepsilon_k\right\} < \varepsilon_k.$$

2. *the partition ξ generated by $f_j\circ T_v, 1\le j\le k-1, v\in I_{n_j}$, and η generated by $f_k\circ T_w, w\in I_{n_k}$, are ε-independent in the sense that*

$$\sum_{A\in\eta, B\in\xi, \mu(B)>0} |\mu(A)-\mu(A|B)| < \varepsilon_k.$$

Proof. Apply Proposition 10.2 for each $k\in\mathbb{N}$ inductively. □

Proposition 10.4. *Let $(\varepsilon_i)_{i\in\mathbb{N}}\subseteq\mathbb{R}$, $(\alpha_i)_{i\in\mathbb{N}}\subseteq\mathbb{R}$ and $(d_i)_{i\in\mathbb{N}}\subseteq\mathbb{N}$. Then there exist $(\overline{f}_i)_{i\in\mathbb{N}}\in L^2(\Omega,\mathcal{A},\mu)$, $\mathbb{E}f_i = 0$ for any $i\in\mathbb{N}$; and independent random variables $X_{k,v}$, $i\in I_{2d_k}$, $k\in\mathbb{N}$ where any given $X_{k,v}$ is $\mathcal{N}(0,\alpha_k^2)$-distributed, such that*

$$\forall k\in\mathbb{N}, v\in I_{2d_k} : \left\|X_{k,v} - T_v\overline{f}_k\right\| < \varepsilon_k. \tag{10.7}$$

Proof. Apply the above proposition to stationary processes $(X_{k,v})$ consisting of random variables with $\mathcal{N}(0,\alpha_k^2)$-distribution. This yields the proposition. □

Clearly, Proposition 10.4 implies that we may find functions such that their iterates behave much like Gaussian random variables. However, putting this together to a function for which Theorem 8.1 holds for all $v\in\{e_1,\ldots,e_d\}$ is a lot more difficult than it might seem. The attempts made by this author to construct such a function all failed because the partial sums failed to fulfil the analogue of Lemma 8.3.

The proof of this lemma in one dimension was based on the fact that we could re-write the function as the sum of co-boundaries with respect to the automorphism, we iterated. To prove the corresponding lemma for $\mathbb{Z}^d$-actions by the same techniques, we must find a function which is a co-boundary with respect to all the transformations T_v, $v\in V\subseteq\mathbb{Z}^d$. Failing to find such a function, we could find a function which is a co-boundary plus something controllable with respect to transformation T_v, $v\in V\subseteq\mathbb{Z}^d$. It has not been possible for the present author to construct any functions with the required properties. However, this does not mean that such a function does not exist.

10.2. Approximation in L^p-norm

The second problem, we choose to describe, concerns some properties of the function found in Theorem 8.1. In the construction of this function, we found that the function was in L^2. Our question is now: Is f contained in L^p for $2 < p \leq \infty$? Since, for $p, q \in (2, \infty]$, $p \leq q$, it is well known that $L^p(X) \supseteq L^q(X)$, it is clearly sufficient to prove that f as constructed in Section 7.5 is uniformly bounded.

In fact, this is implied by Theorem 8.8. Indeed, the Brownian motion in uniformly bounded, so by Theorem 8.8, $S_n f \in L^\infty(X)$ for n large enough. For such n,

$$T^{n+1} f = S_{n+1} f - S_n f \in L^\infty(X). \tag{10.8}$$

Since T is measure preserving, this must mean that $f \in L^\infty(X)$. Indeed, otherwise for any $M > 0$ there would be a set of positive measure, such that for any x in this set, $f(x) > M$. Since T is measure preserving, this would also be the case for $T^{n+1} f$, contradicting (10.8).

The next obvious question is, whether or not the theorem corresponding to Theorem 8.1 holds, when we replace the L^2-norm $\|\cdot\|$ with the L^p-norm $\|\cdot\|_p$. As in the previous section, one also encounters problems when trying to generalise the method of proof from Volný's paper to this problem. The method of proof used by Volný in [61] and the present author in Chapter 8 was highly dependent on the fact that independence of random variables corresponds to orthogonality in L^2. This is not the case in L^p for $p > 2$, where we do not even have a sensible inner product.

Other problems are likely to present themselves in an attempt to solve this problem. At any rate, the present author has not been able to come up with really useful ideas for solving this problem.

10.3. Non-ergodic systems

The third problem we will describe, is the question of weakening the independence condition implied by the ergodicity of the systems, we have been looking at so far.

Ergodicity can be viewed as a generalisation of independence. Indeed, we may view ergodicity as a generalisation of the strong mixing property for measure preserving systems. Since strong mixing states

$$\mu\left(A \cap T^{-n} A\right) \xrightarrow[n\to\infty]{} \mu(A)^2 \quad \text{for any } A \in \mathcal{B},$$

we may read this as "Any given event is asymptotically independent of it's history". This clearly generalises independence of events to a wider class of random processes. It is well-known that strong mixing implies ergodicity, so with the above interpretation, we may consider ergodicity to be an weakened form of independence of random processes.

In the beginning of this part of the thesis, we posed the question of whether or not the independence criterion could be weakened in the Central Limit Theorem and the Almost Sure Central Limit Theorem. We have answered this question affirmatively for some functions in $L_0^2(X)$ under the assumption that the process could be described as an ergodic measure preserving system. However, we are still left with an assumption stating a sort of generalised independence. The next obvious question is: "Under which circumstances may the independence assumption be weakened further?"

Clearly, this is a very difficult question. In all of the methods used in the preceding, we had to assume ergodicity in some sort of disguise. Whenever we were discussing irrational rotations of the torus, we needed Weyl's Equidistribution Theorem, which is in fact a consequence of ergodicity. In the general measure preserving systems case, the main instrument of our approach was Rokhlin's Lemma, which is again a consequence of ergodicity.

It seems clear to this author that if the results are to be extended to non-ergodic systems, we must revert to the methods known from irrational rotations of the circle. It would appear that we must find some substitute for Weyl's Theorem, which applies in non-ergodic systems. However, we need to go beyond rotations of the circle, since all irrational rotations are ergodic, and hence already treated. Also, it is easy to construct counterexamples to any statement resembling Weyl's Theorem, that would be useful for our purposes, whenever the rotation is not irrational. Hence, it would seem that some definition of equidistribution in other measure spaces along with a replacement of Weyl's Theorem would a the way to overcome the barrier. This author has so far not been able to take these ideas any further.

Bibliography

[1] A. G. Abercrombie, *Badly approximable p-adic integers*, Proc. Indian Acad. Sci. Math. Sci. **105** (1995), no. 2, 123–134.

[2] ______, *The Hausdorff dimension of some exceptional sets of p-adic integer matrices*, J. Number Theory **53** (1995), no. 2, 311–341.

[3] M. A. Akcoglu, J. R. Baxter, D. M. Ha, and R. L. Jones, *Approximation of l_2-processes by Gaussian processes*, New York J. Math. **4** (1998), 75–82.

[4] E. Artin, *Quadratische Körper im Gebiete der höheren Kongruenzen. I. (Arithmetischer Teil.)*, Math. Z. **19** (1924), 153–206.

[5] A. Baker and W. M. Schmidt, *Diophantine approximation and Hausdorff dimension*, Proc. London Math. Soc. (3) **21** (1970), 1–11.

[6] V. Beresnevich, *On approximation of real numbers by real algebraic numbers*, Acta Arith. **90** (1999), no. 2, 97–112.

[7] V. I. Bernik and M. M. Dodson, *Metric Diophantine approximation on manifolds*, Cambridge University Press, Cambridge, 1999.

[8] A. S. Besicovitch, *Sets of fractional dimension (IV): on rational approximation to real numbers*, J. Lond. Math. Soc. **9** (1934), 126–131.

[9] P. Billingsley, *Convergence of probability measures*, John Wiley & Sons Inc., New York, 1968.

[10] ______, *Ergodic theory and information*, Robert E. Krieger Publishing Co., Huntington, N.Y., 1978, Reprint of the 1965 original.

[11] L. Breiman, *Probability*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1992, Corrected reprint of the 1968 original.

[12] Y. Bugeaud, *Approximation par des nombres algébriques de degré borné et dimension de hausdorff*, To appear.

[13] R. Burton and M. Denker, *On the central limit theorem for dynamical systems*, Trans. Amer. Math. Soc. **302** (1987), no. 2, 715–726.

[14] J. W. S. Cassels, *An introduction to the geometry of numbers*, Springer-Verlag, Berlin, 1997, Corrected reprint of the 1971 edition.

[15] T. de la Rue, S. Ladouceur, G. Peškir, and M. Weber, *On the central limit theorem for aperiodic dynamical systems and applications*, Teor. Ĭmovīr. Mat. Stat. (1997), no. 57, 140–159.

[16] B. de Mathan, *Approximations diophantiennes dans un corps local*, Bull. Soc. Math. France Suppl. Mém. **21** (1970), 93.

[17] H. Dickinson, M. M. Dodson, and J. Yuan, *Hausdorff dimension and p-adic Diophantine approximation*, Indag. Math. (N.S.) **10** (1999), no. 3, 337–347.

[18] M. M. Dodson, *Geometric and probabilistic ideas in the metric theory of Diophantine approximations*, Uspekhi Mat. Nauk **48** (1993), no. 5(293), 77–106.

[19] M. M. Dodson, B. P. Rynne, and J. A. G. Vickers, *Diophantine approximation and a lower bound for Hausdorff dimension*, Mathematika **37** (1990), 59–73.

[20] O. Frostmann, *Potentiel d'équilibre et capacité des ensembles avec quelques applications à la théorie des fonctions*, Meddel. Lunds Univ. Math. Sem **3** (1935), 1–118.

[21] A. V. Grošev, *Une théorème sur les systemes des formes linéaires*, Dolk. Akad. Nauk. SSSR **9** (1938), 151–152.

[22] G. Harman, *Metric number theory*, The Clarendon Press Oxford University Press, New York, 1998.

[23] H. Jager and C. de Vroedt, *Lüroth series and their ergodic properties*, Nederl. Akad. Wetensch. Proc. Ser. A 72=Indag. Math. **31** (1969), 31–42.

[24] V. Jarník, *Zür metrischen Theorie der diophantischen Approximationen*, Prace Mat.–Fiz. (1928–1929), 91–106.

[25] ______, *Diophantischen Approximationen und Hausdorffsches Mass*, Mat. Sbornik. **36** (1929), 371–382.

[26] ______, *Über die simultanen Diophantische Approximationen*, Math. Z. **33** (1931), 505–543.

[27] J.-P. Kahane, *Some random series of functions*, D. C. Heath and Co. Raytheon Education Co., Lexington, Mass., 1968.

[28] Y. Kato, *Central limit theorem for Weyl automorphism*, Rep. Statist. Appl. Res. Un. Japan. Sci. Engrs. **34** (1987), no. 3, 1–10.

[29] A. Ya. Khintchine, *Continued fractions*, The University of Chicago Press, Chicago and London, 1964.

[30] D. Kleinbock, *Badly approximable systems of affine forms*, J. Number Theory **79** (1999), no. 1, 83–102.

[31] J. Knopfmacher, *Ergodic properties of some inverse polynomial series expansions of Laurent series*, Acta Math. Hungar. **60** (1992), no. 3-4, 241–246.

[32] J. Knopfmacher and A. Knopfmacher, *Metric properties of algorithms inducing Lüroth series expansions of Laurent series*, Astérisque (1992), no. 209, 15, 237–246, Journées Arithmétiques, 1991 (Geneva).

[33] S. Kristensen, *Approximation diophantienne dans un corps des séries formelles*, Submitted to C. R. Acad. Sci. Paris Sér. I Math.

[34] ———, *Badly approximable linear forms over a field of formal series*, Submitted to Monatsh. Math.

[35] ———, *On the well-approximable matrices over a field of formal series*, To appear in Math. Proc. Cambridge Philos. Soc.

[36] ———, *Some metric properties of Lüroth expansions over the field of Laurent series*, Bull. Austral. Math. Soc. **64** (2001), no. 2, 345–351.

[37] S. Kristensen and M Weber, *Weighted gaussian approximation in ergodic systems*, In preparation.

[38] M. T. Lacey, *On central limit theorems, modulus of continuity and Diophantine type for irrational rotations*, J. Anal. Math. **61** (1993), 47–59.

[39] S. Lang, *Algebra*, second ed., Addison-Wesley Publishing Co., Reading, Mass., 1984.

[40] A. Lasjaunias, *A survey of Diophantine approximation in fields of power series*, Monatsh. Math. **130** (2000), no. 3, 211–229.

[41] E. Lesigne, *Almost sure central limit theorem for strictly stationary processes*, Proc. Amer. Math. Soc. **128** (2000), no. 6, 1751–1759.

[42] Kurt Mahler, *An analogue to Minkowski's geometry of numbers in a field of series*, Ann. of Math. (2) **42** (1941), 488–522.

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

[44] A. M. Olevskiĭ, *Fourier series with respect to general orthogonal systems*, Springer-Verlag, New York, 1975, Translated from the Russian by B. P. Marshall and H. J. Christoffers, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 86.

[45] R. Paysant-Leroux and E. Dubois, *Algorithme de Jacobi-Perron dans un corps de séries formelles*, C. R. Acad. Sci. Paris Sér. A-B **272** (1971), A564–A566.

[46] ______, *Approximations simultanées dans un corps de séries formelles*, C. R. Acad. Sci. Paris Sér. A-B **274** (1972), A437–A440.

[47] ______, *Étude métrique de l'algorithme de Jacobi-Perron dans un corps de séries formelles*, C. R. Acad. Sci. Paris Sér. A-B **275** (1972), A683–A686.

[48] G. K. Pedersen, *Analysis now*, Springer-Verlag, New York, 1989.

[49] O. Perron, *Irrationalzahlen*, Walter de Gruyter & Co., Berlin, 1960.

[50] K. Petersen, *Ergodic theory*, Cambridge University Press, Cambridge, 1983.

[51] W. Philipp and W. Stout, *Almost sure invariance principles for partial sums of weakly dependent random variables*, Mem. Amer. Math. Soc. 2 (1975), no. issue 2, 161, iv+140.

[52] A. M. Robert, *A course in p-adic analysis*, Springer-Verlag, New York, 2000.

[53] R. Salem and A. Zygmund, *On lacunary trigonometric series*, Proc. Nat. Acad. Sci. U. S. A. **33** (1947), 333–338.

[54] ______, *On lacunary trigonometric series. II*, Proc. Nat. Acad. Sci. U. S. A. **34** (1948), 54–62.

[55] W. M. Schmidt, *On badly approximable numbers and certain games*, Trans. Amer. Math. Soc. **123** (1966), 178–199.

[56] ______, *Badly approximable systems of linear forms*, J. Number Theory **1** (1969), 139–154.

[57] ______, *On continued fractions and Diophantine approximation in power series fields*, Acta Arith. **95** (2000), no. 2, 139–166.

[58] A. V. Skorokhod, *Studies in the theory of random processes*, Addison-Wesley Publishing Co., Inc., Reading, Mass., 1965.

[59] V. G. Sprindžuk, *Mahler's problem in metric number theory*, American Mathematical Society, Providence, R.I., 1969, Translated from the Russian by B. Volkmann. Translations of Mathematical Monographs, Vol. 25.

[60] D. Volný, *Private communication*, Letters to M. Weber and S. Kristensen.

[61] ______, *Invariance principles and Gaussian approximation for strictly stationary processes*, Trans. Amer. Math. Soc. **351** (1999), no. 8, 3351–3371.

[62] M. Weber, *Systèmes quasi-orthogonaux et théorème central limite presque sûr*, C. R. Acad. Sci. Paris Sér. I Math. **328** (1999), no. 9, 795–798.

[63] ______, *Un théorème central limite presque sûr à moments généralisés pour les rotations irrationnelles*, Manuscripta Math. **101** (2000), no. 2, 175–190.

[64] A. Zygmund, *Trigonometric series. 2nd ed. Vols. I, II*, Cambridge University Press, New York, 1959.