JOHANNES KEPLER
UNIVERSITY LINZ

Two Contemporary Issues in Number Theory: Sequences with Low Star Discrepancy and the Asymptotic Behaviour of the Sudler Product

Submitted by
DI Mario Neumüller

Submitted at
Institut für Finanzmathematik und angewandte Zahlentheorie

Supervisor and
First Examiner
A.Univ.-Prof. Dr. Friedrich Pillichshammer

Second Examiner
Assoc.Prof. Dr.techn. Christoph Aistleitner

Juni 2019

Doctoral Thesis
to obtain the academic degree of
Doktor der technischen Wissenschaften
in the Doctoral Program
Technische Wissenschaften

**JOHANNES KEPLER
UNIVERSITY LINZ**
Altenbergerstraße 69
4040 Linz, Österreich
www.jku.at
DVR 0093696

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende Dissertation selbstständig und ohne fremde Hilfe verfasst, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt bzw. die wörtlich oder sinngemäß entnommenen Stellen als solche kenntlich gemacht habe. Die vorliegende Dissertation ist mit dem elektronisch übermittelten Textdokument identisch.

Neumüller Mario

Ort, Datum

Acknowledgements

Let me start by expressing my gratitude to my supervisor Friedrich Pillichshammer. First of all I am thankful for all his lectures and tutorials I participated as a student which formed the basis of my mathematical knowledge in number theory and related subjects. Second I would like to thank him for his constant support and advise in my studies and giving me always the appropriate amount of freedom for organizing my work.

Further I would like to thank Roswitha Hofer for employing me at her FWF project for about 2 years. In addition I want to express my gratitude to Sigrid Grepstad for working together with me on half of the material that is part of this thesis. It was a pleasure working with you. But I am even more grateful for the fact that we became good friends.

I would like to thank all my colleagues for this wonderful time. A special thanks goes to all my PhD colleagues Ralph, Florian, Katharina, Simon, Lisa and Wolfgang. Not to forget Melanie who earned a special thanks since she has a solution for every organizational problem that appears. Thank you for all the funny discussions, coffee breaks, TAC-evenings and the good time in general.

Last but not least I would like to thank my parents for supporting me in every situation. But my greatest thanks earns my better half Helene who supported me every single day of my studies and without her this thesis would not be the same. Thank you very much for everything!

The research described in this thesis was supported by the Austrian Science Fund (FWF): Project F5505-N26 and Project F5509-N26, which are both part of the special research Program “Quasi-Monte Carlo Methods: Theory and Applications”.

Abstract

In this thesis we study two interesting topics which both are covered by the mathematical discipline of number theory. On the one hand we will investigate certain problems which are related to numerical integration and discrepancy theory and on the other hand we will analyse the asymptotic behaviour of a special sequence of trigonometric products.

In various application (e.g. finance, physics, economics, computer graphics, ...) it is inevitable to efficiently perform high-dimensional numerical integration. One way to overcome this problem is to use quasi-Monte Carlo methods to approximate the desired integrals. When using this methods the resulting approximation error is intimately linked to certain distribution properties of the underlying integration nodes. Therefore many different methods to efficiently construct finite sets of integration nodes which perform well in the context of quasi-Monte Carlo algorithms already exist. In the first chapter we will transfer two of the above-mentioned methods into a digital setting. Roughly speaking this means we will switch from the usual integer arithmetic to arithmetic with polynomials over finite fields. This change results in new point sets with new properties which can again be used in the setting of quasi-Monte Carlo methods.

For the second part of this thesis the main quantity of interest will be the so-called Sudler product

$$P_N(\alpha) = \prod_{r=1}^{N} |2\sin(\pi r \alpha)|,$$

where $N \in \mathbb{N}$ and $\alpha \in \mathbb{R}$. This sequence of trigonometric products appears in a variety of different fields in both pure and applied mathematics e.g. continued fraction theory, Padé approximation, q-series, KAM theory, strange non-chaotic attractors (SNA), continuation of Dirichlet series, ... We will analyse the asymptotic behaviour of the Sudler product and special subsequences for certain choices of α.

Kurzfassung

In dieser Arbeit betrachten wir verschiedene Fragestellungen aus dem Bereich der Zahlentheorie. Einerseits werden wir uns mit Problemen bezüglich der Diskrepanz einer endlichen Punktmenge bzw. mit numerischer Integration beschäftigen. Andererseits untersuchen wir im zweiten Teil dieser Arbeit eine spezielle Folge von trigonometrischen Produkten.

In einer großen Anzahl von Anwendungen (z.B.: Finanzmathematik, Physik, Wirtschaft, Computergrafik, ...) is eine effiziente Vorgehensweise für numerische Integration unabdingbar. Eine mögliche Lösung für diese Herausforderung bieten sogenannte quasi-Monte Carlo Methoden. Der unter Verwendung dieser Methoden entstehende Approximationsfehler ist eng verknüpft mit bestimmten Verteilungseigenschaften der zugrundeliegenden Integrationspunkten. Aufgrund eben dieser Verbindung gibt es bereits eine Vielzahl an Methoden um effizient geeignete Punktmengen zu konstruieren. Im ersten Kapitel werden wir zwei dieser oben genannten Methoden genauer analysieren und in ein digitales Setting transferieren. Somit entstehen neue Punktmengen mit neuen Eigenschaften, welche sich ebenfalls im Rahmen von quasi-Monte Carlo Methoden verwenden lassen.

Im zweiten Teil widmen wir unsere Aufmerksamkeit dem sogenannten Sudler Produkt

$$P_N(\alpha) = \prod_{r=1}^{N} |2\sin(\pi r \alpha)|,$$

wobei $N \in \mathbb{N}$ und $\alpha \in \mathbb{R}$. Diese Folge von Produkten erscheint in den verschiedensten Gebieten der reinen und angewandten Mathematik z.B.: Kettenbrüche, Padé Approximation, q-series, KAM Theorie, Fortsetzung von Dirichlet Reihen, ... Insbesondere interessieren wir uns für die asymptotischen Eigenschaften des Sudler Produkts und speziellen Teilfolgen davon für gewisse Wahlen von α.

Contents

Abbreviations and notations

The following lists of abbreviations and notations will be used frequently in this thesis:

List of abbreviations

QMC	quasi-Monte Carlo
CBC construction	component-by-component construction
i.i.d.	independent and identically distributed

List of notations

$\{\cdot\}$	fractional part
$\lceil\cdot\rceil$	ceiling function
$\lfloor\cdot\rfloor$	floor function
$\deg(q)$	degree of the polynomial q
$\gcd(a,b)$	greatest common divisor of a and b
$[s]$	the set $\{1,\ldots,s\}$
$\mathbb{N}$	natural numbers
$\mathbb{N}_0$	natural numbers starting from 0
$\mathbb{Z}$	integers
$\mathbb{Q}$	rational numbers
$\mathbb{R}$	real numbers
$\mathbb{C}$	complex numbers
$\mathbb{P}$	set of prime numbers
$\mathbb{F}_p$	finite field of order p, $p\in\mathbb{P}$
$\mathbb{F}_p((x^{-1}))$	field of formal Laurent series over $\mathbb{F}_p$
$\mathcal{L}^2(D)$	$\left\{f: D\to\mathbb{C} \mid f \text{ measurable}, \int_D f^2 \, \mathrm{d}\lambda < \infty\right\}$
$\mathbb{1}_A$	characteristic function of a set A

More notation:

- Usually bold Latin or Greek letters denote either vectors of dimension s or sequences if not explicitly stated otherwise. For example $\boldsymbol{x} = (x_1, \ldots, x_s)$ or $\boldsymbol{\gamma} = (\gamma_n)_{n\in\mathbb{N}}$.
- For $\alpha \in \mathbb{R}$ we denote the continued fraction expansion of α by $[a_0; a_1, a_2, \ldots]$ where $a_0 = \lfloor \alpha \rfloor$ and $(a_n)_{n\in\mathbb{N}} \subseteq \mathbb{N}$ are the continued fraction coefficients of α. Moreover, by $\alpha = [a_0; a_1, \ldots, a_h, \overline{a_{h+1}, \ldots, a_{h+\ell}}]$ we denote that α has a periodic continued fraction expansion with preperiod $a_1, \ldots, a_h$ and period $a_{h+1}, \ldots, a_{h+\ell}$.
- For two functions $f, g : \mathbb{R} \to \mathbb{R}$ we use the following notations:
  - We write $f(x) = \mathcal{O}(g(x))$ as $x \to a \in [-\infty, \infty]$ if there exists a constant $c > 0$ and a neighbourhood $B(a)$ of a such that f and g satisfy $|f(x)| \leq c|g(x)|$ for all $x \in B(a)$. Usually a will be 0 or ∞ and we will not explicitly state that $x \to a$ if it is clear from the context.
  - We use the notation $f(x) = \Theta(g(x))$ as $x \to a \in [-\infty, \infty]$ if there exist constants $c, C > 0$ and a neighbourhood $B(a)$ of a such that $c|g(x)| \leq |f(x)| \leq C|g(x)|$ for all $x \in B(a)$. Again we will omit $x \to a$ if it is clear from the context.
  - We write $f(x) \sim g(x)$ if $\lim_{x\to\infty} f(x)/g(x) = 1$.
- For $\emptyset \neq \mathfrak{u} \subseteq [s]$ and some s-dimensional function $f(\boldsymbol{x})$ with $\boldsymbol{x} = (x_1, \ldots, x_s)$ we use the following notations:
  - $f(\boldsymbol{x}_\mathfrak{u}, \mathbf{1}) = f(y_1, \ldots, y_s)$, where $y_i = x_i$ if $i \in \mathfrak{u}$ and $y_i = 1$ if $i \notin \mathfrak{u}$.
  - For $\mathfrak{u} = \{i_1, \ldots i_r\}$ we denote $\mathrm{d}\boldsymbol{x}_\mathfrak{u} = \mathrm{d}x_{i_1} \ldots \mathrm{d}x_{i_r}$.
  - For $\mathfrak{u} = \{i_1, \ldots i_r\}$ the term $\dfrac{\partial^{|\mathfrak{u}|} f(x_1, \ldots, x_s)}{\partial \boldsymbol{x}_\mathfrak{u}}$ denotes $\dfrac{\partial^{|\mathfrak{u}|} f(x_1, \ldots, x_s)}{\partial x_{i_1} \partial x_{i_2} \cdots \partial x_{i_r}}$.

Preface

In applications of finance, economics, physics, ... it is necessary to numerically solve high-dimensional integrals. One class of algorithms which are able to deal with this kind of problems are quasi-Monte Carlo algorithms. This sort of algorithms heavily depends on the choice of the underlying point set, which should be "well" distributed in the corresponding domain (in this thesis we will only consider the s-dimensional unit cube). More precise, the error, which stems from the usage of such algorithms, depends on certain distribution properties of the underlying integration nodes. Due to this interesting connection of quasi-Monte Carlo algorithms and applications in various fields, this topic has been studied extensively in the last decades. Moreover, there already exist efficient algorithms for constructing finite point sets which perform well in the context of numerical integration with quasi-Monte Carlo algorithms.

In the first chapter we will describe ways to transform two of these algorithms, for creating such finite point sets with the desired properties, into a digital setting. Roughly speaking this means we will switch from the usual integer arithmetic to polynomial arithmetic over finite fields. Performing this transformation results in new point sets with different properties which can again be feed into the machinery of quasi-Monte Carlo methods.

In the second part of this thesis we will focus on a special sequence of trigonometric products

$$P_N(\alpha) = \prod_{r=1}^{N} |2\sin(\pi r \alpha)|,$$

where $N \in \mathbb{N}$ and $\alpha \in \mathbb{R}$. This sequence pops up in many different fields of pure and applied mathematics. For example there are interesting connections of $P_N(\alpha)$ to continued fraction theory, partition theory, q-series, ... In the second chapter we will study the asymptotic behaviour of $P_N(\alpha)$ for certain choices of α. Most of the time we will stick to the cases where α is either the golden ratio or a general quadratic irrational. More precise, we are interested in the properties of the objects $\liminf_{N\to\infty} P_N(\varphi)$ and $\lim_{n\to\infty} P_{q_n}(\alpha)$, where φ is the golden ratio, α a quadratic irrational and $(q_n)_{n\in\mathbb{N}_0}$ is the sequence of best approximation denominators of α.

These two above-mentioned limits are the central objects of questions and conjectures stated by Lubinsky in [76] and Verschueren and Mestel in [104] which will be investigated in Chapter 2.

The main part of this thesis consists of the following four papers:

1. F. Pillichshammer, M. Neumüller: *Metrical star discrepancy bounds for lacunary subsequences of digital Kronecker-sequences and polynomial tractability*, Unif. Distrib. Theory 13 (1) (2018), 65–86.

2. R. Kritzinger, H. Laimer, M. Neumüller: *A reduced fast construction of polynomial lattice point sets with low weighted star discrepancy.* In Monte Carlo and quasi-Monte Carlo methods 2016. Springer, Cham, 2018, pp. 377–394.

3. S. Grepstad, M. Neumüller: *Asymptotic behaviour of the Sudler product of sines for quadratic irrationals*, J. Math. Anal. Appl. 465 (2) (2018), 928–960.

4. S. Grepstad, L. Kaltenböck, M. Neumüller: *A positive lower bound for* $\liminf_{N\to\infty} P_N(\varphi)$ (to appear in Proc. Am. Math. Soc.).

This thesis is structured as follows. First of all we start with a very brief introduction to the mathematical field of number theory. The first chapter deals with the part which is related to discrepancy theory. Therefore we start in Section 1.1 with an introduction to quasi-Monte Carlo methods and discrepancy theory and continue in Sections 1.2 and 1.3 with the second and first paper mentioned in the list above. We close this chapter with Section 1.4, which contains a short conclusion and ideas for further research.

The second chapter will be dedicated to the second topic of this thesis, the Sudler product. Again we begin with an introductory part (Section 2.1), followed by the fourth and third paper of the list above in Sections 2.2 and 2.3. In the last part (Section 2.4) we close again the chapter with some concluding remarks and ideas for generalisations and extensions of the main results in Chapter 2.

Chapter 1

Discrepancy theory and quasi-Monte Carlo integration

1.1 Introduction

First of all it should be pointed out that there exists a more detailed and very nicely structured introduction on discrepancy theory and quasi-Monte Carlo integration which was written by Leobacher and Pillichshammer [73]. The introduction of this thesis will follow to some extend the lines of this book and stick to the overall structure of it.

If one would have to explain discrepancy theory in one sentence then the comment of Art Owen at the MCQMC conference in Stanford 2016 sums it up perfectly: "We are counting points in boxes." With this of course very rough comparison in mind let us directly hop into one of the most central definitions of this chapter.

Definition 1.1.1 (Local discrepancy function, star discrepancy)**.**
Let $\mathcal{P}_N = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\} \subset [0,1)^s$ be an N-element point set and let $\boldsymbol{t} = (t_1, \ldots, t_s) \in [0,1]^s$. Then $[\mathbf{0}, \boldsymbol{t}) = \prod_{i=1}^s [0, t_i)$ is a s-dimensional axis parallel box and we define the local discrepancy function of $\mathcal{P}_N$ as

$$\delta(\boldsymbol{t}, \mathcal{P}_N) := \frac{1}{N} \sum_{j=1}^{N} \mathbb{1}_{[\mathbf{0},\boldsymbol{t})}(\boldsymbol{x}_j) - \prod_{i=1}^{s} t_i. \tag{1.1}$$

Moreover, the star discrepancy of $\mathcal{P}_N$ is defined as

$$D_N^*(\mathcal{P}_N) := \sup_{\boldsymbol{t} \in [0,1]^s} |\delta(\boldsymbol{t}, \mathcal{P}_N)|. \tag{1.2}$$

Note that for some s-dimensional box $[\mathbf{0}, \boldsymbol{t})$ the local discrepancy function compares the relative number of points which are contained in the box to the volume of the box. It is exactly this counting process, included in the local discrepancy function, which the comment of Art Owen refers to. Finally by considering the supremum of the local discrepancy function, the star discrepancy reflects the performance of the worst possible box.

If we are interested in the discrepancy of a sequence $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ we will write $D_N^*((\boldsymbol{x}_n)_{n\in\mathbb{N}})$ which denotes the star discrepancy of the first N elements of the sequence $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$.

The motivation for studying the discrepancy of point sets or sequences stems from two different mathematical regions. It is already clear that this definition measures some kind of uniformity of the point set $\mathcal{P}_N$. It will turn out in the next section that the star discrepancy is closely related to the concept of uniform distribution modulo 1. Therefore the motivation for the notion of discrepancy comes from a number theoretic background. Whereas in numerical mathematics the notion of discrepancy is strongly connected to the problem of numerical integration by using quasi-Monte Carlo algorithms.

1.1.1 Motivation

Let us now explain more detailed why the discrepancy of a point set or sequence is a measure of a certain kind of uniformity, which is called uniform distribution modulo 1. This concept was introduced in 1916 by Hermann Weyl in his celebrated work "Über die Gleichverteilung von Zahlen modulo Eins" (cf. [106]) and has developed into a fruitful and active branch of mathematics in the last century. One should again emphasise that exactly this work was the beginning of the up to now well studied branch of uniform distribution theory. The most central definition states the following:

Definition 1.1.2 (Uniform distribution modulo 1)**.**
Let $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ be a sequence in $[0,1)^s$. The sequence $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ is called uniformly distributed modulo 1 (or uniformly distributed) iff for all $\boldsymbol{t} = (t_1, \ldots, t_s) \in [0,1]^s$ we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_{[\mathbf{0},\boldsymbol{t})}(\boldsymbol{x}_n) = \lambda_s([\mathbf{0}, \boldsymbol{t})),$$

where $\lambda_s([\mathbf{0}, \boldsymbol{t}))$ is the s-dimensional Lebesgue measure of $[\mathbf{0}, \boldsymbol{t}) = \prod_{i=1}^{s} [0, t_i)$.

Probably the most famous example for a uniformly distributed sequence is the so-called Kronecker sequence given by $\boldsymbol{x}_n = \{n\boldsymbol{\alpha}\}$, where $\boldsymbol{\alpha} \in \mathbb{R}^s$ and $\{\cdot\}$ denotes the fractional part which has to be read component-wise. It turns out that the condition that $1, \alpha_1, \ldots, \alpha_s$ have to be linearly independent over $\mathbb{Q}$ is necessary and sufficient for the sequence $\{n\boldsymbol{\alpha}\}$ to be uniformly distributed. Although it was already proved by Bohl in 1909 that the above-mentioned sufficient condition is indeed a necessary condition as well, the following criterion established by Weyl provides a very elegant way to prove this fact.

Theorem 1.1.3 (Weyl criterion, [106], [64, Ch. 1, Theorem 2.1])**.** *A sequence* $(\boldsymbol{x}_n)_{n\in\mathbb{N}} \subset [0,1)^s$ *is uniformly distributed modulo 1 if and only if*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{h}\cdot\boldsymbol{x}_n} = 0$$

for all $\boldsymbol{h} \in \mathbb{Z}^s \setminus \{\boldsymbol{0}\}$*. Here* $\cdot$ *denotes the usual Euclidean inner product.*

Let us revisit the definition of the star discrepancy (see Definition 1.1.1). Now it is clear that the star discrepancy is a measure for the property of a point set of being uniformly distributed. This can also be summarized more mathematically in the following well known statement (see [64, Ch. 2, Theorem 1.1])

$$(\boldsymbol{x}_n)_{n\in\mathbb{N}} \text{ is uniformly distributed } \Leftrightarrow \lim_{N\to\infty} D_N^*((\boldsymbol{x}_n)_{n\in\mathbb{N}}) = 0. \tag{1.3}$$

As we already mentioned the second motivation in order to investigate the discrepancy of certain point sets and sequences comes from numerical mathematics. Consider the following situation: Given a function $f : [0,1]^s \to \mathbb{R}$, where f belongs to some suitable function space $\mathcal{F}$ equipped with a norm $\|\cdot\|$. We are interested in the quantity

$$I_s(f) := \int_{[0,1]^s} f(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}. \tag{1.4}$$

A convenient way to approximate $I_s(f)$ is to use an equal weighted quadrature rule of the form

$$\frac{1}{N} \sum_{n=1}^{N} f(\boldsymbol{x}_n),$$

where $\mathcal{P} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\} \subset [0,1]^s$ are the integration nodes. Several questions arise in this context. First of all there is the question of how to choose the point set $\mathcal{P}$, which is indeed the most central question concerning quadrature rules. Second, depending on the choice of the underlying point set, what can we say about the quality of our approximation which is usually measured in terms of the integration error:

$$e_N(f,\mathcal{P}) := \left| I_s(f) - \frac{1}{N}\sum_{n=1}^{N} f(\boldsymbol{x}_n) \right|.$$

A first idea could be to choose the point set $\mathcal{P}$ at random. This approach is called Monte Carlo algorithm and it is well known if $\boldsymbol{x}_1,\dots,\boldsymbol{x}_N$ are i.i.d. random variables in $[0,1]^s$ then

$$\mathbb{E}(e_N(f,\mathcal{P})) \le \frac{\sigma(f)}{\sqrt{N}}, \tag{1.5}$$

where $\sigma(f)$ is the standard deviation of f which is defined as

$$\sigma(f) = \sqrt{\mathbb{E}\left(\left(f - \mathbb{E}(f)\right)^2\right)}$$

and $\mathbb{E}(f) = \int_{[0,1]^d} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = I_s(f)$ if we interpret f as a random variable on the probability space $([0,1]^s, \mathcal{B}, \lambda_s)$, where $\mathcal{B}$ denotes the Borel σ-algebra of $[0,1]^s$ and λ_s is the s-dimensinal Lebesgue measure.

However, the probabilistic error bound and also the convergence rate of $\mathcal{O}(N^{-1/2})$ are not sufficient for some applications. Therefore we will approximate the quantity $I_s(f)$ by a so-called QMC algorithm of the form

$$Q_{N,s}(f,\mathcal{P}_N) := \frac{1}{N}\sum_{n=1}^{N} f(\boldsymbol{x}_n), \tag{1.6}$$

where now in contrast to before the point set $\mathcal{P}_N = \{\boldsymbol{x}_1,\dots,\boldsymbol{x}_N\} \subset [0,1]^s$ is chosen deterministically. A natural way to quantify the error one obtains by using a QMC algorithm is to consider the so-called worst case error which is given by

$$e_N(\mathcal{F},\mathcal{P}_N) := \sup_{\substack{f\in\mathcal{F}\\ \|f\|\le 1}} e_N(f,\mathcal{P}_N). \tag{1.7}$$

We will see that the integration error $e_N(f,\mathcal{P}_N)$ one obtains by using a QMC algorithm is intimately connected to the star discrepancy of the underlying point set $\mathcal{P}_N$ via the famous Koksma-Hlawka inequality. Before actually stating the inequality let us set for some function $f:[0,1]^s \to \mathbb{R}$ where all mixed partial derivatives of f exist and are continuous on $[0,1]^s$

$$\|f\|_{s,1} := f(\mathbf{1}) + \sum_{\emptyset\neq\mathfrak{u}\subseteq[s]} \int_{[0,1]^{|\mathfrak{u}|}} \left| \frac{\partial^{|\mathfrak{u}|}}{\partial \boldsymbol{x}_{\mathfrak{u}}} f(\boldsymbol{x}_{\mathfrak{u}},\mathbf{1}) \right| \mathrm{d}\boldsymbol{x}_{\mathfrak{u}}, \tag{1.8}$$

where we used the corresponding notation declared on page xiv. The subsequent theorem is the combined result of Koksma (cf. [60], $s = 1$) and Hlawka (cf. [50], $s > 1$).

Theorem 1.1.4 (Koksma-Hlawka inequality)**.** *Let $\mathcal{P}_N \subset [0,1)^s$ and let f be a function on $[0,1]^s$ with $\|f\|_{s,1} < \infty$. Then we have*

$$e_N(f,\mathcal{P}_N) = |I_s(f) - Q_{N,s}(f,\mathcal{P}_N)| \leq \|f\|_{s,1} D_N^*(\mathcal{P}_N). \tag{1.9}$$

If we set $\mathcal{F}_{s,1} := \{f \ : \ \|f\|_{s,1} < \infty\}$ we get as a direct consequence of Theorem 1.1.4 that the worst case error in $\mathcal{F}_{s,1}$ is exactly the star discrepancy of the underlying point set, i.e.

$$e_N(\mathcal{F}_{s,1},\mathcal{P}_N) = \sup_{\substack{f\in\mathcal{F}_{s,1} \\ \|f\|_{s,1}\leq 1}} e_N(f,\mathcal{P}_N) = D_N^*(\mathcal{P}_N).$$

Furthermore note that due to the Koksma-Hlawka inequality the integration error when using a QMC algorithm is now split into two parts: One part is depending on the integrand f and the second part is exactly the star discrepancy of the underlying integration nodes $\mathcal{P}_N$. In other words this means that studying point sets with low discrepancy has a direct application in solving numerical integration problems. The inequality (1.9) is still true in a more general case where $\|f\|_{s,1}$ is replaced by the so-called variation of f in the sense of Hardy and Krause (denoted by $V(f)$), i.e.

$$e_N(f,\mathcal{P}_N) = |I_s(f) - Q_{N,s}(f,\mathcal{P}_N)| \leq V(f) D_N^*(\mathcal{P}_N).$$

Roughly speaking $V(f)$ can be interpreted as a measure for the fluctuation of f. For a comprehensive proof of Theorem 1.1.4 and a definition of $V(f)$ we refer the interested reader to [64, Chapter 5]. But observe that if all mixed partial derivatives of f are continuous on $[0,1]^s$ then $V(f) = \|f\|_{s,1} - f(\mathbf{1})$.

For more information on the connection of discrepancy theory and numerical integration we refer to [71, 81, 82, 83, 101]. We will continue the introduction with a very brief survey on some classical results of discrepancy theory.

1.1.2 Discrepancy theory: some classical results

We can naturally extend the definition of the star discrepancy to derive different notions of discrepancies. First of all recall the definition of the star discrepancy (Definition 1.1.1), which is given by the supremum over all axis parallel boxes anchored in the origin of the local discrepancy function defined as

$$\delta(\boldsymbol{t},\mathcal{P}_N) = \frac{1}{N}\sum_{j=1}^{N} \mathbb{1}_{[\mathbf{0},\boldsymbol{t})}(\boldsymbol{x}_j) - \prod_{i=1}^{s} t_i.$$

If we now extend the supremum over all axis parallel boxes (not necessarily anchored in the origin) we arrive at a different notion which is called the extreme discrepancy (denoted by $D_N(\mathcal{P}_N)$). Second, one can interpret the star discrepancy as the L_∞-norm of the local discrepancy function $\delta(\cdot, \mathcal{P}_N)$. By interchanging the L_∞-norm with the L_p-norm for $p \in [1, \infty)$ we arrive at another notion of discrepancy, namely the L_p-discrepancy which is then given by

$$L_{p,N}(\mathcal{P}_N) = \|\delta(\cdot, \mathcal{P}_N)\|_p = \left(\int_{[0,1]^s} \delta(\boldsymbol{t}, \mathcal{P}_N)^p \mathrm{d}\boldsymbol{t} \right)^{1/p}.$$

The subsequent proposition summarizes some basic connections between the different notions of discrepancy.

Proposition 1. *Let $\mathcal{P}$ be an N-element point set in $[0,1)^s$. Then we have*

1. $D_N^*(\mathcal{P}) \leq D_N(\mathcal{P}) \leq 2^s D_N^*(\mathcal{P})$;
2. $L_{p,N}(\mathcal{P}) \leq D_N^*(\mathcal{P}) \leq (L_{p,N}(\mathcal{P}))^{\frac{p}{p+s}}$.

The lower bounds for $D_N(\mathcal{P})$ and $D_N^*(\mathcal{P})$ follow immediately from the definition and the monotonicity of the L_p-norms. The proof for the upper bound of $D_N(\mathcal{P})$ is mainly based on the fact that one can describe an unanchored axis parallel box as the union and exclusion of at most 2^s anchored boxes and the proof for the upper bound of the star discrepancy in terms of the L_p-discrepancy can be found in [31, Theorem 1.8].

In what follows we will for the sake of simplicity just state the parameter dependence for the appearing constants. In other words this means that the constants in the rest of this section are not necessarily the same although they are denoted by the same symbol.

We have already seen in Section 1.1.1 that we want to have point sets and sequences with a small discrepancy. Of course this leads to an interest in upper and lower bounds for the discrepancy of point sets and sequences as well as the corresponding asymptotic behaviour. Let us start our small survey with the famous result by Roth proven in 1954.

Theorem 1.1.5 (Roth, [93])**.** *For every $s \in \mathbb{N}$ there exists a constant $c_s > 0$ such that for every $N \geq 2$ and every N-element point set $\mathcal{P} \subset [0,1]^s$ we have that*

$$D_N(\mathcal{P}) \geq D_N^*(\mathcal{P}) \geq L_{2,N}(\mathcal{P}) \geq c_s \frac{(\log N)^{\frac{s-1}{2}}}{N}. \tag{1.10}$$

This result of Roth is known to be best possible in the sense that for every $s, N \in \mathbb{N}$ there exists a point set $\overline{\mathcal{P}} \subset [0,1)^s$ and some constant $c_s > 0$ such that

$$L_{2,N}(\overline{\mathcal{P}}) \leq c_s \frac{(\log N)^{\frac{s-1}{2}}}{N}. \tag{1.11}$$

The first construction for such point sets in arbitrary dimensions was given by Chen and Skriganov in [19]. Later on Dick and Pillichshammer were able to give explicit constructions in [27].

Remark 1.1.6. Although not explicitly stated it is clear that together with the monotonicity of the L_p norms and Theorem 1.1.5 we get that for all $s \in \mathbb{N}$, $p \in [2, \infty]$ there exists a constant $c_{p,s} \geq 0$ such that for every $\mathcal{P} \subset [0,1)^s$ we have that

$$L_{p,N}(\mathcal{P}) \geq c_{p,s} \frac{(\log N)^{\frac{s-1}{2}}}{N}. \tag{1.12}$$

Moreover, Schmidt [95] could extend Theorem 1.1.5 to the case where $p \in (1,2)$, $s \in \mathbb{N}$ and Halász [41] was able to prove that Theorem 1.1.5 is still true for $p = 1$ and $s = 2$.

That the lower bound in the inequality (1.10) might not be optimal for the star discrepancy is not very surprising if one compares the averaging behaviour of the L_2-norm with the more localized one of the L_∞-norm. Indeed Schmidt was able to prove a stronger lower bound for the star discrepancy in dimension 2.

Theorem 1.1.7 (Schmidt, [94])**.** *There exists a constant $c > 0$ such that for every $N \geq 2$ and every N-element point set $\mathcal{P} \subset [0,1]^2$ we have that*

$$D_N^*(\mathcal{P}) \geq c \frac{\log N}{N}. \tag{1.13}$$

It is known that this lower bound is best possible in N. Some classical examples of point sets with the matching upper star discrepancy bound are provided at the end of this subsection. If we consider the case $s \geq 3$ the currently best lower bound was proven in an outstanding paper by Bilyk, Lacey and Vagharshakyan who were able to improve the bound of Roth for the star discrepancy by an additional term η_s in the $\log N$-exponent.

Theorem 1.1.8 (Bilyk, Lacey, Vagharshakyan, [13])**.** *Let $s \geq 3$ then there exist constant $c_s > 0$ and $\eta_s \in (0, 1/2)$ with the following property: for every $N \geq 2$ and every point set $\mathcal{P} \subset [0,1)^s$ with cardinality N we have that*

$$D_N^*(\mathcal{P}) \geq c_s \frac{\log(N)^{(s-1)/2+\eta_s}}{N} \tag{1.14}$$

Furthermore, the $\log N$*-exponent* η_s *tends to zero (approximately with order* s^{-2}*) for growing* s*.*

This raises the question: Given a point set $\mathcal{P} \subset [0,1)^s$.

$$\text{What is the optimal order of } D_N^*(\mathcal{P}) \text{ for } s \geq 3?$$

It is not surprising that we will not be able to answer this question in this thesis since one definitely can state that this question is the biggest open problem in discrepancy theory. Moreover, the experts in this field do not even agree on what to conjecture. Probably the two most prominent conjectures are the following:

Conjecture 1.1.9. For $s \geq 3$ there exists constants $c_s, \tilde{c}_s > 0$ such that for sufficiently large N and every point set $\mathcal{P} \subset [0,1)^s$ with cardinality N we have that

$$D_N^*(\mathcal{P}) \geq c_s \frac{(\log N)^{s-1}}{N} \text{ or } D_N^*(\mathcal{P}) \geq \tilde{c}_s \frac{(\log N)^{s/2}}{N}.$$

Observe that both cases agree with the lower bound of Schmidt for the case $s = 2$.

Let us now consider sequences instead of point sets. Note that if we analyse the discrepancy of a sequence $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ we investigate the discrepancy of the increasing sets $\{\boldsymbol{x}_1\}$, $\{\boldsymbol{x}_1, \boldsymbol{x}_2\}$, $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3\}, \ldots$. Therefore it is much easier to increase the number of points used in a QMC algorithm if the integration nodes stem from a sequence since we just have to add the desired number of points to the already existing integration nodes. Whereas in the case of point sets we would have to calculate a completely new point set if we would like to increase the number of points. This is exactly the reason why sometimes the two settings are referred to as the dynamic and static setting. Due to this more demanding situation one could expect worse discrepancy bounds for sequences than in the case of point sets which is indeed the case.

Similar to the static case the optimal order of the L_2-discrepancy is known for sequences. The lower bound was given by Proinov in 1986.

Theorem 1.1.10 (Proinov, [89])**.** *Let* $p \in (1, \infty]$*,* $s \geq 2$ *then there exists a constant* $c_{p,s} > 0$ *such that for every sequence* $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ *in* $[0,1)^s$ *we have that*

$$D_N^*((\boldsymbol{x}_n)_{n\in\mathbb{N}}) \geq L_{p,N}((\boldsymbol{x}_n)_{n\in\mathbb{N}}) \geq c_{p,s} \frac{(\log N)^{s/2}}{N} \text{ for infinitely many } N \in \mathbb{N}. \tag{1.15}$$

Similar as before we know that this lower bound is optimal for the L_2-discrepancy since Dick and Pillichshammer gave an explicit construction of a sequence whose L_2-discrepancy is at most $c_s (\log N)^{s/2}/N$ for some constant $c_s > 0$ (see [27]). However, for the star discrepancy the situation is different. There is an improvement for $s = 1$ by Schmidt in [94].

Theorem 1.1.11 (Schmidt). *There exists a constant $c > 0$ such that for every sequence $(x_n)_{n\in\mathbb{N}}$ in $[0,1)$ we have that*

$$D_N^*((x_n)_{n\in\mathbb{N}}) \geq c\frac{\log N}{N}. \tag{1.16}$$

Again this result is known to be best possible. An example for a 2-dimensional sequence which obtains the above order of the star discrepancy is given at the end of this subsection. For sequences the situation is quite similar to before. The correct order of the star discrepancy is not known and there are different opinions and conjectures about the correct exponent of the $\log N$ term. For example s and $(s+1)/2$, which of course both agree with the result of Schmidt for $s = 1$ (Theorem 1.1.11).

Although the optimal order of the star discrepancy is not known for sequences (and point sets) there is a consensus to call a sequence $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ or a point set $\mathcal{P}$ a *low-discrepancy sequence* or *low-discrepancy point set* if

$$D_N^*((\boldsymbol{x}_n)_{n\in\mathbb{N}}) \leq c_s\frac{(\log N)^s}{N} \text{ or } D_N^*(\mathcal{P}) \leq c_s\frac{(\log N)^{s-1}}{N}.$$

Let us conclude this subsection with several famous examples of low-discrepancy sequences and point sets:

- **Van der Corput sequence**:
  The nth sequence element of this one-dimensional sequence is defined as $x_n = \phi_b(n)$, where ϕ_b is the b-adic radical inverse function. More precise, for some base $b \in \mathbb{N}$ we have that $\phi_b : \mathbb{N}_0 \to [0,1)$ and

  $$\phi_b(n) = n_0/b + n_1/b^2 + n_2/b^3 + \cdots , \tag{1.17}$$

  where the n_i are the digits of the b-adic expansion of n, i.e. $n = \sum_{i=0}^r n_i b^i$ and $n_i \in \{0,\dots,b-1\}$. It is well known that the van der Corput sequence in base b satisfies the following star discrepancy bound for some constant $c_b > 0$

  $$D_N^*((\phi_b(n))_{n\in\mathbb{N}}) \leq c_b\frac{\log(N)}{N}.$$

  For $b = 2$ this was first proven by van der Corput in [102, 103].

- **Halton sequence**:
  The Halton sequence can be interpreted as the s-dimensional analogue of the van der Corput sequence. For some base $b \in \mathbb{N}$ let again ϕ_b be the radical inverse function (defined in (1.17)). For $s \in \mathbb{N}$ and integers $b_1,\dots,b_s \geq 2$ the

nth element of the Halton sequence in bases $b_1, \ldots, b_s$, denoted by $\mathcal{S}_{(b_1,\ldots,b_s)} = (\boldsymbol{x}_n)_{n\in\mathbb{N}}$ is defined as

$$\boldsymbol{x}_n = (\phi_{b_1}(n), \phi_{b_2}(n), \ldots, \phi_{b_s}(n)).$$

Halton [42] could show show that for pairwise coprime bases $b_1, \ldots, b_s \geq 2$ the Halton sequence fulfills the following discrepancy bound for some constant $c_{s,\boldsymbol{b}} > 0$ and $\boldsymbol{b} = (b_1, \ldots, b_s)$

$$D_N^*(\mathcal{S}_{(b_1,\ldots,b_s)}) \leq c_{s,\boldsymbol{b}} \frac{(\log N)^s}{N}.$$

- **Hammersley point set**:
  For $s \in \mathbb{N}$ and pairwise coprime integers $b_1, \ldots, b_{s-1} \geq 2$ we define the Hammersley point set $\mathcal{H}_{N,(b_1,\ldots,b_{s-1})} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\}$ with

  $$\boldsymbol{x}_n = \left(\frac{n}{N}, \phi_{b_1}(n), \ldots, \phi_{b_{s-1}}(n)\right).$$

  Again it is known (see [42]) that the Hammersley point set, which was first introduced in [43], is an example for a s-dimensional low-discrepancy point set. In other words we have for some constant $c_{s,\boldsymbol{b}} > 0$ and $\boldsymbol{b} = (b_1, \ldots, b_s)$ that

  $$D_N^*(\mathcal{H}_{N,(b_1,\ldots,b_s)}) \leq c_{s,\boldsymbol{b}} \frac{(\log N)^{s-1}}{N}.$$

For more detailed information in this direction we refer to [26, 73].

1.1.3 Lattice point sets and component-by-component constructions

A well studied example for useful point sets concerning QMC methods are so-called lattice point sets, which were first introduced independently by Hlawka [51] and Korobov [61]. These point sets can be constructed with the help of a generating integer vector $\boldsymbol{g} \in \mathbb{Z}^s \setminus \{\boldsymbol{0}\}$ and the corresponding point set is defined as follows

$$\mathcal{P}_N(\boldsymbol{g}) := \left\{\left\{\frac{n}{N}\boldsymbol{g}\right\} \mid n = 0, \ldots, N-1\right\}, \tag{1.18}$$

where $\{\cdot\}$ denotes the fractional part and is applied component-wise. Observe that this kind of point sets can be interpreted as a rational analogue of the uniformly distributed sequence $(\{n\boldsymbol{\alpha}\})_{n\in\mathbb{N}}$ for $\boldsymbol{\alpha} \in \mathbb{R}^s$, which was mentioned in the beginning of

Subsection 1.1.1. In the analysis of the discrepancy of lattice point sets the so-called dual lattice

$$\mathcal{L}(\boldsymbol{g}, N) := \{\boldsymbol{h} \in \mathbb{Z}^s \mid \boldsymbol{h} \cdot \boldsymbol{g} \equiv 0 \pmod N\} \tag{1.19}$$

plays a crucial role. The following inequality, which was proven by Niederreiter in [80, Chapter 5] is a promising starting point to derive good (star) discrepancy bounds for lattice point sets.

Theorem 1.1.12. *Let $N \geq 2$, $\boldsymbol{g} \in \mathbb{Z}^s$ and $\mathcal{P}_N(\boldsymbol{g})$ be the corresponding lattice point set in $[0,1)^s$. Then we have*

$$D_N(\mathcal{P}_N(\boldsymbol{g})) \leq 1 - \left(1 - \frac{1}{N}\right)^s + R_N(\boldsymbol{g}), \tag{1.20}$$

where

$$\begin{aligned} R_N(\boldsymbol{g}) &:= \sum_{\boldsymbol{h} \in C_s^* \cap \mathcal{L}(\boldsymbol{g}, N)} \frac{1}{r(\boldsymbol{h}, N)}, \\ C_s^* &= ((-N/2, N/2] \cap \mathbb{Z})^s \setminus \{\mathbf{0}\}, \\ r(\boldsymbol{h}, N) &= \prod_{i=1}^s r(h_i, N), \end{aligned}$$

and for $h \in \mathbb{Z}$

$$r(h, N) = \begin{cases} N \sin(\pi |h|/N) & \text{if } h \neq 0, \\ 1 & \text{if } h = 0. \end{cases}$$

Observe that by definition a lattice point set is fully determined by the generating vector $\boldsymbol{g}$ and that it is enough to search for $\boldsymbol{g}$ in the set $\{0, \ldots, N-1\}^s$. Since there are no explicit constructions for good lattice point sets in dimension $s \geq 3$ (see [14] for explicit constructions for $s \in \{1, 2\}$) one employes computer search algorithms to find generating vectors which construct useful lattice point sets. In consideration of the N^s possible choices for a generating vector an exhaustive search is not feasible even for moderate values of N and s. In order to handle this problem one switches to a greedy algorithm. The scheme of the algorithm described in the following lines was first invented by Korobov in the 1960s in [62] and then rediscovered by Sloan and Reztsov in 2002 in [97]. These so-called component-by-component constructions (CBC) start by setting the first coordinate g_1 of $\boldsymbol{g}$ to some sufficiently good starting value. Then we choose the next component g_2 such that g_2 minimizes a certain quantity. The optimal case would be if g_2 minimizes the star discrepancy of the lattice point set but again this is computationally not feasible. Therefore the figure of merit will be $R_N(\boldsymbol{g})$, i.e. g_2 is chosen such that $R_N(\boldsymbol{g})$ is minimized. The third component is again chosen in a way such that it minimizes the quantity $R_N(\boldsymbol{g})$. One follows this procedure until we obtain a generating vector of full size s.

Observe that we have $g_i \in \{0, \dots, N-1\}$ for all $i \in [s] := \{1, \dots, s\}$, i.e. in each step of the algorithm described above there are N possible choices for each component. Hence the size of the search space for a generating vector $\boldsymbol{g}$ reduces to sN. To be more precise we actually exclude 0 since the choice of $g_i = 0$ would lead to bad distributional properties for the generated point set and in order to improve the distributional properties of the resulting lattice point set even further we restrict ourselves to search sets of the form $G_N = \{n \in \{1, \dots, N-1\} : \gcd(n, N) = 1\}$. Note that we have $|G_N| = \varphi(N)$, where φ is Euler's totient function and the size of the search space for $\boldsymbol{g}$ equals $s\varphi(N)$. Now the algorithm described above reads as follows:

Algorithm 1.1.13 (CBC algorithm)**.** *Let $s, N \in \mathbb{N}$.*

1. *Choose $g_1 = 1$.*
2. *For $d = 2, \dots, s$ do: choose $z = g_d \in G_N$ such that $R_N((g_1, g_2, \dots, g_{d-1}, z), N)$ is minimized as a function of z.*

The guarantee that the output of this algorithm is indeed a generating vector which constructs a lattice point set with small (star) discrepancy is given by the next theorem.

Theorem 1.1.14. *Let N be a prime number. If the generating vector $\boldsymbol{g} = (g_1, \dots, g_s)$ is constructed with the help of Algorithm 1.1.13 then we have for $d \in [s]$ that*

$$D_N(\mathcal{P}_N(\boldsymbol{g}_d)) \leq \frac{d}{N} + \frac{2^d}{N}(\log N + 1)^d, \tag{1.21}$$

where $\boldsymbol{g}_d = (g_1, \dots, g_d)$.

For a proof of Theorem 1.1.14 we refer to [73, Chapter 4]. Observe that the discrepancy bound in Theorem 1.1.14 implies

$$D_N(\mathcal{P}_N(\boldsymbol{g})) = \mathcal{O}\left(\frac{(\log N)^s}{N}\right), \tag{1.22}$$

if $\boldsymbol{g}$ is the output vector of Algorithm 1.1.13. Let us conclude this section with a pure existence result of Bykovskii [18] in 2012 which uses a different approach and achieves a slightly better order of magnitude. For $s \geq 2$ and $N \geq 3$ there exist $\boldsymbol{g} \in \mathbb{Z}^s$ such that

$$D_N(\mathcal{P}_N(\boldsymbol{g})) = \mathcal{O}\left(\frac{(\log N)^{s-1} \log \log N}{N}\right), \tag{1.23}$$

where the implied constant depends on the dimension s.

The construction cost of the standard CBC construction (Algorithm 1.1.13) is of order $\mathcal{O}(sN^2)$ (see [71]), where s is the dimension and N the cardinality of the resulting lattice point set $\mathcal{P}_N(\boldsymbol{g})$. The order of magnitude in N and in s of the construction cost of Algorithm 1.1.13 can be significantly reduced by combining different ideas and concept of various authors. This topic will be covered in more detail in Section 1.2.

1.1.4 Formal Laurent series: basic notations

In this thesis we will encounter in several sections objects and definitions which frequently make use of finite fields, polynomial arithmetic over finite fields and formal Laurent series. Therefore it is beneficial to introduce some basic notations and provide some elementary information concerning these notations, which is exactly the aim of this subsection.

For a prime number p, let $\mathbb{F}_p$ be the finite field of prime order p. We identify $\mathbb{F}_p$ with the set $\{0, 1, \ldots, p-1\}$ equipped with arithmetic modulo p. Moreover, we denote by $\mathbb{F}_p[x]$ the set of polynomials over $\mathbb{F}_p$. In certain situations we want to identify each $n \in \mathbb{N}$ with a polynomial in $\mathbb{F}_p[x]$. This is done in the following way: Each $n \in \mathbb{N}$ has a unique p-adic expansion of the form $n = n_0 + n_1 p + \cdots + n_{m-1} p^{m-1}$ with digits $n_r \in \{0, 1, \ldots, p-1\}$ for $r \in \{0, 1, \ldots, m-1\}$. We can therefore associate to each integer n the uniquely determined polynomial $n(x) = \sum_{r=0}^{p-1} n_r x^r \in \mathbb{F}_p[x]$.

The field of formal Laurent series over $\mathbb{F}_p$ will be denoted by $\mathbb{F}_p((x^{-1}))$, i.e.

$$\mathbb{F}_p((x^{-1})) = \left\{ \sum_{l=\omega}^{\infty} t_l x^{-l} \ : \ \omega \in \mathbb{Z}, t_l \in \mathbb{F}_p \right\}.$$

Note that we have $\mathbb{F}_p \subset \mathbb{F}_p[x] \subset \mathbb{F}_p((x^{-1}))$. The fractional part of a formal Laurent series $g = \sum_{l=\omega}^{\infty} t_l x^{-l}$ is denoted in the same way as in the real case and defined by

$$\{g\} := \sum_{l=\max\{1,\omega\}}^{\infty} t_l x^{-l}.$$

Further, we set

$$\overline{\mathbb{F}}_p((x^{-1})) := \big\{ \{g\} \ : \ g \in \mathbb{F}_p((x^{-1})) \big\} \tag{1.24}$$

$$= \left\{ \sum_{l=\omega}^{\infty} t_l x^{-l} \,:\, \omega \in \mathbb{N}, t_l \in \mathbb{F}_p \right\}.$$

Finally, for $m \in \mathbb{N}$ we introduce the following functions

$$\phi_m : \overline{\mathbb{F}}_p((x^{-1})) \to [0,1),\ \phi_m \left(\sum_{l=\omega}^{\infty} t_l x^{-l} \right) = \sum_{l=\omega}^{m} t_l p^{-l}, \tag{1.25}$$

$$\phi : \overline{\mathbb{F}}_p((x^{-1})) \to [0,1],\ \phi \left(\sum_{l=\omega}^{\infty} t_l x^{-l} \right) = \sum_{l=\omega}^{\infty} t_l p^{-l}. \tag{1.26}$$

Observe that the function ϕ is surjective on $[0,1]$ but not injective since for $x \in \mathbb{Q} \cap [0,1)$ with finite p-adic expansion of the form $x = \sum_{l=1}^{L} t_l p^{-l}$ and $t_L \neq 0$ we can also write $x = \sum_{l=1}^{\infty} u_l p^{-l}$, where $u_l = t_l$ for $l \in \{1, \ldots, L-1\}$, $u_L = t_L - 1$ and $u_l = p-1$ for $l > L$. To overcome this problem let us set

$$\mathcal{C} := \{ g \in \overline{\mathbb{F}}_p((x^{-1})) \,:\, t_k = p-1 \text{ for all but finitely many } k \geq 1 \}. \tag{1.27}$$

Note that $\mathcal{C}$ is a countable set. Further, we define $\overline{\mathbb{F}}_p^*((x^{-1})) := \overline{\mathbb{F}}_p((x^{-1})) \setminus \mathcal{C}$. Then one can check that the map

$$\phi : \overline{\mathbb{F}}_p^*((x^{-1})) \to [0,1),\ \phi \left(\sum_{l=\omega}^{\infty} t_l x^{-l} \right) = \sum_{l=\omega}^{\infty} t_l p^{-l}$$

is a bijection.

1.1.5 A digital analogue: polynomial lattice point sets

In Subsection 1.1.3 we studied lattice point sets and algorithms to construct them. Now we are interested in a certain analogue of lattice point sets which are called polynomial lattice point sets. Roughly speaking we are switching from integer arithmetic to arithmetic with polynomials over a finite field.

For a given dimension $s \geq 2$ and some integer $m \geq 1$ we choose a modulus $f \in \mathbb{F}_p[x]$ with $\deg(f) = m$ as well as polynomials $g_1, \ldots, g_s \in \mathbb{F}_p[x]$. The vector of polynomials $\boldsymbol{g} = (g_1, \ldots, g_s)$ is again called the generating vector of the polynomial lattice point set. Moreover, we associate to each $n \in \{0, 1, \ldots, p^m - 1\}$ the polynomial $n(x)$ as described in Subsection 1.1.4. With this notation the polynomial lattice point set $\mathcal{P}_N(\boldsymbol{g}, f)$ is defined as the set of $N := p^m$ points of the form

$$\boldsymbol{x}_n = \left(\phi_m \left(\left\{ \frac{n(x) g_1(x)}{f(x)} \right\} \right), \ldots, \phi_m \left(\left\{ \frac{n(x) g_s(x)}{f(x)} \right\} \right) \right) \in [0,1)^s, \tag{1.28}$$

where ϕ_m is defined as in (1.25) and for (see also [26, Chap. 10]). Observe that the definition which results in (1.28) is of a similar structure as the definition for lattice point sets. At first sight the main difference seems to be the switch from integer arithmetic to arithmetic with polynomials over finite fields. Indeed the development of lattice point sets and polynomial lattice point sets follows a parallel track but nevertheless there are certain situations where one is superior to the other (e.g. in terms of error bounds or the function classes where they yield good results for numerical integration). For a more detailed comparison see [88].

Niederreiter [80] proved the existence of polynomial lattice point sets with low star discrepancy by averaging arguments. Generating vectors of good polynomial lattice point sets can be constructed analogous to generating vectors of lattice point sets, i.e. by a CBC construction. This can be done by finding suitable analogues for the quantity $R_N(\boldsymbol{g})$ (see (1.20)) and the search set $G_N = \{1, \ldots, N-1\}$. Following this strategy CBC constructions for generating vectors of polynomial lattice point sets for an irreducible modulus f were provided in [25] and for a reducible f in [22]. In these papers, the authors considered the star discrepancy as well as its weighted version (see Subsection 1.1.7). In Section 1.2 we will deal with this constructions in more detail and focus especially on the speed up of them.

Remark 1.1.15. Polynomial lattice point sets are a special case of a more general class of point sets introduced by Niederreiter in [77] (see also [79, 80]). Later the name digital nets (see for example [26]) was introduced for this class of point sets. We will see that there exists an equivalent definition for polynomial lattice point sets which fits into the framework of digital nets. One of the main characteristic features of this kind of point sets is their construction. An s-dimensional digital net can be constructed with the help of s generating matrices.

Definition 1.1.16 (Digital net)**.** *Let $p \in \mathbb{P}$, $m \in \mathbb{N}$ and $C_1, \ldots, C_s \in \mathbb{F}_p^{m \times m}$. We call the point set $\mathcal{P} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\}$ with $N = p^m$ a digital net over $\mathbb{F}_p$ with generating matrices $C_1, \ldots, C_s$ if the elements of $\mathcal{P}$ are constructed in the following way:*

In order to construct the nth element $\boldsymbol{x}_n$ of $\mathcal{P}$ we first compute the vector $\boldsymbol{n}$ which consists of the p-adic digits of n, i.e. $n = \sum_{k=0}^{m-1} n_k p^k$ and $\boldsymbol{n} = (n_0, \ldots, n_{m-1})^\top$. Note that we construct a point set with $N = p^m$ elements. Second, for each $i \in [s]$ we set $\boldsymbol{y}_n^{(i)} = (y_{n,1}^{(i)}, \ldots, y_{n,s}^{(i)})^\top$ as

$$\boldsymbol{y}_n^{(i)} = C_i \boldsymbol{n}.$$

Finally the entries of $\boldsymbol{y}_n^{(i)}$ are the digits of the ith component of $\boldsymbol{x}_n$. This means we set $\boldsymbol{x}_n = (x_{n,1}, \ldots, x_{n,s})$ and

$$x_{n,i} = \frac{y_{n,1}^{(i)}}{p} + \frac{y_{n,2}^{(i)}}{p^2} + \cdots + \frac{y_{n,m}^{(i)}}{p^m} \text{ for } i \in [s].$$

The construction scheme described above can also be generalised to the cases where p is a prime power or an integer, see [26] and [67, 69, 80], respectively. For a detailed description of the concept of digital nets we refer the interested reader to [26] or [80].

Let us now have a more detailed look on the digital construction of polynomial lattice point sets. Let therefore be $s, m \in \mathbb{N}$ and $p \in \mathbb{P}$. Additionally let $\boldsymbol{g} = (g_1, \ldots, g_s) \in (\mathbb{F}_p[x])^s$ and fix some $f \in \mathbb{F}_p[x]$ with $\deg(f) = m$. For $i \in [s]$ consider the formal Laurent series expansion

$$\frac{g_i(x)}{f(x)} = \sum_{l=w_i}^{\infty} v_l^{(i)} x^{-l} \in \mathbb{F}_p((x^{-1})),$$

where $w_i \in \mathbb{Z}$ with $w_i \leq 1$ and $v_l^{(i)} \in \mathbb{F}_p$. Now we define the s generating matrices $C_1, \ldots, C_s$ of size $m \times m$ as

$$c_{j,r+1}^{(i)} = v_{r+j}^{(i)}, \tag{1.29}$$

where $j \in [m]$, $r \in \{0, \ldots, m-1\}$ and $i \in [s]$. Now the digital net over $\mathbb{F}_p$ with generating matrices $C_1, \ldots, C_s$ given in (1.29) yields exactly the same point set as defined in (1.28). For a proof of this statement see, for example, [26, Theorem 10.5]. Moreover, the digital construction for polynomial lattice point sets described above can be generalised to the case where p is a prime power (see [79]).

1.1.6 Inverse of the star discrepancy

In certain situations and especially in the context of QMC-methods it is useful to study the following question:

How many points N do we need to ensure a discrepancy smaller than some ε?

Let us put this question on a more mathematical basis. Therefore we define for $N, s \in \mathbb{N}$ the minimal star discrepancy

$$\operatorname{disc}^*(N, s) := \inf_{\substack{\mathcal{P} \subset [0,1)^s \\ |\mathcal{P}| = N}} D_N^*(\mathcal{P}). \tag{1.30}$$

In other words $\mathrm{disc}^*(N,s)$ denotes the smallest star discrepancy which can be achieved by a point set with N elements. Now we can reformulate the question from before as the following quantity

$$N^*(\varepsilon, s) := \min\{N \in \mathbb{N} : \mathrm{disc}^*(N,s) \leq \varepsilon\}. \tag{1.31}$$

Often $N^*(\varepsilon, s)$ is referred to as the *inverse of the star discrepancy*.

Remark 1.1.17. The quantity defined in (1.31) is actually just a special example which is part of a more sophisticated and general theory, called information based complexity (IBC). In the language of IBC the inverse of the star discrepancy is a special instance of the so-called information complexity. Here one studies the behaviour of the information complexity for multivariate continuous problems over some suitable function class $\mathcal{F}$ consisting of s variate functions. It was already mentioned in Section 1.1.1 that according to the Koksma-Hlawka inequality the star discrepancy is related to the problem of numerical integration. In IBC one is interested in the dependence of $N^*(\varepsilon, s)$ on the error demand ε and on the dimension s. If the information complexity depends exponentially on the dimension s or on ε^{-1} then the corresponding problem is called intractable, otherwise it is called tractable. In order to classify and categorize the tractable problems, several notions have been established (see [81, 82]). For example we say that the inverse of the star discrepancy is *polynomially tractable* if there exists $C > 0, p, q > 0$ such that

$$N^*(\varepsilon, s) \leq C s^p \varepsilon^{-q}.$$

If $p = 0$ (i.e. no dependence on the dimension s) we say that the inverse of the star discrepancy is strongly polynomially tractable. For a detailed and comprehensive description of tractability theory see [81, 82, 83].

Since the last one and a half decades a lot of effort has been put into the analysis of the star discrepancy with respect to dimensions s tending to infinity. In a seminal work by Heinrich, Novak, Wasilkowski and Woźniakowski [45] it has been shown that there exists an absolute constant $C > 0$ such that

$$\mathrm{disc}^*(N,s) \leq C\sqrt{\frac{s}{N}} \quad \text{for all } s, N \in \mathbb{N} \tag{1.32}$$

(see [45, Theorem 3]). Later Aistleitner [1, Theorem 1] showed that the constant C can be chosen as $C = 10$ and recently Gnewuch and Hebbinghaus (to appear) could improve this result further such that one can choose $C = 2.528\ldots$. From this and (1.32) we obtain

$$N^*(\varepsilon, s) \leq cs\varepsilon^{-2} \text{ with } c = 6.394\ldots.$$

This means that the star discrepancy is *polynomially tractable*, see [81, 82]. On the other hand Hinrichs [48] showed that $N^*(\varepsilon, s) \geq cs\varepsilon^{-1}$ for all $s \in \mathbb{N}$ and for sufficiently small $\varepsilon > 0$. Hence *the inverse of the star discrepancy depends linearly on the dimension*, which is also the programmatic title of [45]. However, the determination of the exact exponent of ε^{-1} is still an open problem.

1.1.7 Weighted discrepancy

We have already seen several notions of discrepancy in the previous sections. The motivation for another concept of discrepancy comes from the aim to classify problems with respect to their tractability properties.

It has been observed in several applications of QMC methods that integrands show different behaviour concerning different variables in the sense that some variables are less important than others for the corresponding integration problem. In order to mathematically model this observations and also gain an advantage from this information Sloan and Woźniakowski [98] introduced the concept of the weighted star discrepancy. Therefore weights have been introduced to reflect the influence of different coordinates on the integration error. Consider weights $\boldsymbol{\gamma} = (\gamma_{\mathfrak{u}})_{\mathfrak{u} \subseteq [s]}$ of nonnegative real numbers, i.e., every group of variables $(x_i)_{i \in \mathfrak{u}}$ is equipped with a weight $\gamma_{\mathfrak{u}}$. Roughly speaking, a small weight indicates that the corresponding variables contribute little to the integration problem. Now the weighted star discrepancy is defined as follows.

Definition 1.1.18 (Weighted star discrepancy)**.**
Let $\boldsymbol{\gamma} = (\gamma_{\mathfrak{u}})_{\mathfrak{u} \subseteq [s]}$ be given weights, $\mathcal{P}_N = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\} \subseteq [0,1)^s$ be an N-element point set and $\delta(\boldsymbol{t}, \mathcal{P}_N)$ be the local discrepancy function of $\mathcal{P}_N$ (see (1.1)*). Then the weighted star discrepancy of $\mathcal{P}_N$ is defined as*

$$D^*_{N,\boldsymbol{\gamma}}(\mathcal{P}_N) := \sup_{\boldsymbol{t} \in (0,1]^s} \max_{\emptyset \neq \mathfrak{u} \subseteq [s]} \gamma_{\mathfrak{u}} |\delta((\boldsymbol{t}_{\mathfrak{u}}, \mathbf{1}), \mathcal{P}_N)|,$$

where $(\boldsymbol{t}_{\mathfrak{u}}, \mathbf{1})$ denotes the vector $(\tilde{t}_1, \ldots, \tilde{t}_s)$ with $\tilde{t}_j = t_j$ if $j \in \mathfrak{u}$ and $\tilde{t}_j = 1$ if $j \notin \mathfrak{u}$.

First of all note that this concept is indeed a generalisation of the star discrepancy defined in Definition 1.1.1 since $D^*_{N,\boldsymbol{\gamma}} = D^*_N$ if $\gamma_{\mathfrak{u}} = 1$ for all $\mathfrak{u} \subseteq [s]$. Secondly, we would like to point out that there also exists a weighted version of the L_p-discrepancy (see [26, Section 3.6]).

The most common examples of weights studied in the literature are:

- Product weights: The importance of the jth variable is reflected by the jth sequence element of a non-increasing sequence of positive real numbers $(\gamma_j)_{j \geq 1}$ with $\gamma_j \leq 1$. Then we set $\gamma_{\mathfrak{u}} := \prod_{j \in \mathfrak{u}} \gamma_j$ and $\gamma_\emptyset := 1$.

- Finite order weights: Let $k \in \mathbb{N}$ be the order of the weights. Then $\gamma_{\mathfrak{u}} = 0$ for $|\mathfrak{u}| > k$ for all $\mathfrak{u} \subseteq [s]$.

As already indicated in the beginning of this section introducing weights can lead to improvements in the high-dimensional behaviour of the corresponding discrepancy. Analogous as in Subsection 1.1.6 we define the inverse of the weighted star discrepancy:

$$\operatorname{disc}^*_{\boldsymbol{\gamma}}(N,s) := \inf_{\substack{\mathcal{P} \subset [0,1)^s \\ |\mathcal{P}|=N}} D^*_{N,\boldsymbol{\gamma}}(\mathcal{P}), \tag{1.33}$$

$$N^*_{\boldsymbol{\gamma}}(\varepsilon,s) := \min\{N \in \mathbb{N} \,:\, \operatorname{disc}^*_{\boldsymbol{\gamma}}(N,s) \le \varepsilon\}. \tag{1.34}$$

Recall that the dependence of the inverse of the star discrepancy on the dimension s was linear. If we consider the weighted version one can show the following theorem which was first shown in [49].

Theorem 1.1.19. *If the weights $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_{\mathfrak{u}})_{\mathfrak{u} \subseteq [s]}$ fulfill the condition*

$$\sup_{s \in \mathbb{N}} \max_{\mathfrak{u} \subseteq [s]} \gamma_{\mathfrak{u}} \sqrt{|\mathfrak{u}|} < \infty$$

then for all $\delta \in (0,1)$ there exists a constant $c_\delta > 0$ such that

$$N^*_{\boldsymbol{\gamma}}(\varepsilon,s) \le \left\lceil c_\delta (\log s + 1)^{1/(1-\delta)} \varepsilon^{-2/(1-\delta)} \right\rceil. \tag{1.35}$$

A proof of this theorem can also be found in [26, Theorem 3.65] as well as a version of this theorem for the weighted L_p-discrepancy.

One can use the notion of the weighted star discrepancy as starting point to develop a concept for numerical integration in weighted spaces with the help of QMC rules, where one uses a weighted version of the Koksma-Hlawka inequality in order to connect the integration error of the QMC rule to the weighted star discrepancy of the underlying point set. For detailed information in this direction see [26, Sections 2.5 and 3.6] or [98].

We will briefly discuss this concept for functions belonging to the function space $W_2^{(1,1,\ldots,1)}([0,1]^s) = \bigotimes_{d=1}^s W_2^1([0,1])$ with the norm $\|f\|_{s,\boldsymbol{\gamma}}$, where $W_2^1([0,1])$ is the set of all absolutely continues functions where the first derivatives belong to $\mathcal{L}_2([0,1])$. The norm is defined as follows:

$$\|f\|_{s,\boldsymbol{\gamma}} := \sum_{\emptyset \neq \mathfrak{u} \subseteq [s]} \gamma_{\mathfrak{u}}^{-1/2} \int_{[0,1]^{|\mathfrak{u}|}} \left| \frac{\partial^{|\mathfrak{u}|}}{\partial \boldsymbol{x}_{\mathfrak{u}}} f(\boldsymbol{x}_{\mathfrak{u}}, 1) \right| \mathrm{d}\boldsymbol{x}_{\mathfrak{u}}, \tag{1.36}$$

where $\mathrm{d}\boldsymbol{x}_{\mathfrak{u}} = \prod_{j\in\mathfrak{u}} x_j$, $\frac{\partial^{|\mathfrak{u}|}}{\partial \boldsymbol{x}_{\mathfrak{u}}} = \frac{\partial^{|\mathfrak{u}|}}{\partial x_{u_1}\cdots\partial x_{u_k}}$ and $f(\boldsymbol{x}_{\mathfrak{u}}, \mathbf{1}) = f(y_1, \ldots, y_s)$ with $y_i = x_i$ if $i \in \mathfrak{u}$ and $y_i = 1$ else. Moreover, we set

$$F_{s,\boldsymbol{\gamma}} = \{f \in W_2^{(1,1,\ldots,1)}([0,1]^s) \;:\; \|f\|_{d,\boldsymbol{\gamma}} < \infty\}. \tag{1.37}$$

Then the weighted Koksma-Hlawka inequality states (see [98]) that for $f \in F_{s,\boldsymbol{\gamma}}$ we have for the QMC integration error that

$$\left| \int_{[0,1]^s} f(\boldsymbol{x})\mathrm{d}\boldsymbol{x} - \frac{1}{N}\sum_{i=1}^{N} f(\boldsymbol{x}_i) \right| \leq \|f\|_{s,\boldsymbol{\gamma}} D^*_{N,\boldsymbol{\gamma}}(\mathcal{P}), \tag{1.38}$$

where $\mathcal{P} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\}$.

1.2 A reduced fast construction of polynomial lattice point sets

As already mentioned in Section 1.1.5 there exist constructions for polynomial lattice point sets which perform well in terms of the resulting star discrepancy or the weighted star discrepancy. It is the aim of the present section to speed up these constructions (in the case of the weighted star discrepancy) by reducing the search sets for the components of the generating vector $\boldsymbol{g}$ of a polynomial lattice point set according to each component's importance. We will consider the weighted star discrepancy as a quality measure for our point sets (see Section 1.1.7). For the weighted star discrepancy of a polynomial lattice point set with modulus f and generating vector $\boldsymbol{g}$ we simply write $D^*_{N,\boldsymbol{\gamma}}(\boldsymbol{g}, f)$. Moreover we have to consider weights $\boldsymbol{\gamma} = (\gamma_{\mathfrak{u}})_{\mathfrak{u} \subseteq [s]}$ of non-negative real numbers. Then each group of variables $\boldsymbol{x}_{\mathfrak{u}} = (x_i)_{i \in \mathfrak{u}}$ is equipped with the weight $\gamma_{\mathfrak{u}}$. For the sake of simplicity we will stick to the case of product weights, i.e. for a non-increasing sequence of positive real numbers $(\gamma_j)_{j \geq 1}$ with $\gamma_j \leq 1$. We set $\gamma_{\mathfrak{u}} := \prod_{j \in \mathfrak{u}} \gamma_j$ and $\gamma_{\emptyset} := 1$. (Note that if we consider a weight for a single variable x_j we have by definition $\gamma_{\{j\}} = \gamma_j$ and therefore in what follows we we will not distinguish between the weight $\gamma_{\{j\}}$ and the sequence element γ_j.)

For the rest of this section let $p \in \mathbb{P}$, $m \in \mathbb{N}$ and $f \in \mathbb{F}_p[x]$ with $\deg(f) = m$. Additionally, by $G_{p,m}$ we denote the set of all polynomials g over $\mathbb{F}_p$ with $\deg(g) < m$. Further we define

$$G_{p,m}(f) := \{g \in G_{p,m} \mid \gcd(g, f) = 1\}. \tag{1.39}$$

Observe that $G_{p,m}(f)$ denotes the search set of the standard CBC construction for polynomial lattice point sets described for example in [23, 25]. It is the nature of product weighted spaces that the components g_j of the generating vector have less and less influence on the quality of the corresponding polynomial lattice point set as j increases. Roughly speaking this is due to the weights $(\gamma_j)_{j \geq 1}$ which are assumed to become ever smaller with increasing index j. We want to exploit this property in the following way. As the components' influence is decreasing with the growth of their indices we want to use less and less time and computational cost to choose these components. To achieve this we choose them from even smaller search sets, which are defined as follows. Let $w_1 \leq w_2 \leq \ldots$ be a non-decreasing sequence of nonnegative integers. This sequence of w_j's is determined in accordance with the weight sequence $(\gamma_j)_{j \geq 1}$. Loosely speaking, the smaller γ_j, the bigger w_j is chosen. For $w \in \mathbb{N}_0$ with $w < m$ we define $G_{p,m-w}$ and $G_{p,m-w}(f)$ analogously to $G_{p,m}$ and $G_{p,m}(f)$, respectively. Further we set $G_{p,m-w}(f) := \{1\} \subset \mathbb{F}_p[x]$ for $w \geq m$. For $w < m$ these sets have cardinality $p^{m-w} - 1$ in the case of an irreducible modulus f and $p^{m-w-1}(p-1)$ for the special case $f : \mathbb{F}_p \to \mathbb{F}_p, x \mapsto x^m$. We will consider these two cases in what follows. Finally, for $d \in [s]$, we define $G^d_{p,m-\boldsymbol{w}}(f) := G_{p,m-w_1}(f) \times \cdots \times G_{p,m-w_d}(f)$. The idea is to choose the ith component of $\boldsymbol{g}$ of the form $x^{w_i} g_i$, where $g_i \in G_{p,m-w_i}(f)$, i.e., the search set for the ith component is reduced by a factor $p^{-\min\{w_i,m\}}$ in comparison to the standard CBC construction. We will show that under certain conditions on the weight sequence $(\gamma_j)_{\geq 1}$ and the parameters w_i a polynomial lattice point set constructed according to the reduced CBC construction has a low weighted star discrepancy of order $N^{-1+\delta}$ for all $\delta > 0$.

The standard CBC construction (cf. [97]) can be done in $\mathcal{O}(sN^2)$ operations. To speed up the construction, in a first step, making use of ideas from Nuyens and Cools [84, 85] on fast Fourier transformation (FFT), the construction cost can be reduced to $\mathcal{O}(sN \log N)$, as for example done in [25]. Combining this with the reduced search sets, which have been described above, we obtain a computational cost that is independent of the dimension eventually. Reduced CBC constructions have been introduced first by Dick et al. in [23] for lattice and polynomial lattice point sets with a small worst case integration error in Korobov and Walsh spaces, respectively, and have also been investigated in [63] for lattice point sets with small weighted star discrepancy.

Recall that $N^*_{\boldsymbol{\gamma}}(s, \varepsilon)$ is the minimal number of points required to achieve a weighted star discrepancy of at most ε (see Section 1.1.7). To keep the construction cost of the generating vector low, it is, of course, beneficial to have a small information complexity and thus to stand a chance to have a polynomial lattice point set of small size. We will show that our reduced fast CBC algorithm finds a generating vector $\boldsymbol{g}$ of a polynomial lattice point set that achieves strong polynomial tractability provided that $\sum_{j=1}^{\infty} \gamma_j p^{w_j} < \infty$ with a construction cost of

$$\mathcal{O}\left(N \log N + \min\{s,t\}N + N \sum_{d=1}^{\min\{s,t\}} (m - w_d) p^{-w_d}\right)$$

operations, where $t = \max\{j \in \mathbb{N} \mid w_j < m\}$.

Before stating our main results we would like to discuss a motivating example. Consider first the standard CBC construction as treated in [22, 25], where $w_j = 0$ for all $j \geq 0$. In this case, a sufficient condition for strong polynomial tractability is $\sum_{j=1}^{\infty} \gamma_j < \infty$, which for instance is satisfied for the special choices $\gamma_j = j^{-2}$ and $\gamma_j = j^{-1000}$. However, in the second example the weights decay much faster than in the first but without any further advantage for the standard CBC construction. We are able to exploit this decay by introducing the sequence $\boldsymbol{w} = (w_j)_{j\geq 0}$ such that the condition $\sum_{j=1}^{\infty} \gamma_j p^{w_j} < \infty$ holds, while still achieving strong polynomial tractability (see Corollary 1.2.4). This way, we can reduce the size of the search sets for the components of the generating vector if the weights γ_j decay very fast. Consider for example the weight sequence $\gamma_j = j^{-k}$ for some $k > 1$. For $w_j = \lfloor (k-\alpha)\log_p j \rfloor$ with arbitrary $1 < \alpha < k$ we find

$$\sum_{j=1}^{\infty} \gamma_j p^{w_j} \le \sum_{j=1}^{\infty} j^{-k} j^{k-\alpha} = \sum_{j=1}^{\infty} j^{-\alpha} = \zeta(\alpha) < \infty,$$

where ζ denotes the Riemann Zeta function. Observe that for large k, i.e., fast decaying weights, we may choose smaller search sets and thereby speed up the CBC algorithm.

1.2.1 A reduced CBC construction

In this section we present a CBC construction for the vector $(x^{w_1}g_1, \ldots, x^{w_s}g_s)$ and an upper bound for the weighted star discrepancy of the corresponding polynomial lattice point set.

First note that if $\boldsymbol{g} \in G_{p,m}^s$, then it is known (see [25]) that

$$D^*_{N,\boldsymbol{\gamma}}(\boldsymbol{g}, f) \le \sum_{\substack{\mathfrak{u} \subseteq [s] \\ \mathfrak{u} \neq \emptyset}} \gamma_{\mathfrak{u}} \left(1 - \left(1 - \frac{1}{N}\right)^{|\mathfrak{u}|}\right) + R^s_{\boldsymbol{\gamma}}(\boldsymbol{g}, f), \tag{1.40}$$

where in the case of product weights we have

$$R^s_{\boldsymbol{\gamma}}(\boldsymbol{g}, f) = \sum_{\substack{\boldsymbol{h} \in G^s_{p,m} \setminus \{\boldsymbol{0}\} \\ \boldsymbol{h} \cdot \boldsymbol{g} \equiv 0 \bmod f}} \prod_{i=1}^{s} r_p(h_i, \gamma_i). \tag{1.41}$$

Here, for elements $\boldsymbol{h} = (h_1, \ldots, h_s)$ and $\boldsymbol{g} = (g_1, \ldots, g_s)$ in $G^s_{p,m}$ we define the scalar product by $\boldsymbol{h} \cdot \boldsymbol{g} := h_1 g_1 + \cdots + h_s g_s$. The numbers $r_p(h, \gamma)$ for $h \in G_{p,m}$ and $\gamma \in \mathbb{R}$ are defined as

$$r_p(h, \gamma) = \begin{cases} 1 + \gamma & \text{if } h = 0, \\ \gamma r_p(h) & \text{otherwise,} \end{cases}$$

where for $h = h_0 + h_1 x + \cdots + h_a x^a$ with $h_a \neq 0$ we set $r_p(h) = \frac{1}{p^{a+1} \sin^2\left(\frac{\pi}{p} h_a\right)}$. Thus, in order to analyse the weighted star discrepancy of a polynomial lattice point set it suffices to investigate the quantity $R^s_{\boldsymbol{\gamma}}(\boldsymbol{g}, f)$. This is due to the result of Joe [55], who proved that for any summable weight sequence $(\gamma_j)_{j\geq 1}$ we have

$$\sum_{\substack{\mathfrak{u}\subseteq[s]\\ \mathfrak{u}\neq\emptyset}} \gamma_{\mathfrak{u}} \left(1-\left(1-\frac{1}{N}\right)^{|\mathfrak{u}|}\right) \leq \frac{\max(1,\Gamma)\mathrm{e}^{\sum_{i=1}^{\infty}\gamma_i}}{N} \quad \text{with } \Gamma := \sum_{i=1}^{\infty} \frac{\gamma_i}{1+\gamma_i}. \tag{1.42}$$

Algorithm 1.2.1. *Let $p \in \mathbb{P}$, $m \in \mathbb{N}$, $f \in \mathbb{F}_p[x]$ with $\deg(f) = m$ and let $(w_j)_{j\geq 1}$ be a non-decreasing sequence of nonnegative integers and consider product weights $(\gamma_j)_{j\geq 1}$. Construct $(g_1, \ldots, g_s) \in G^s_{p,m-\boldsymbol{w}}(f)$ as follows:*

1. *Set $g_1 = 1$.*
2. *For $d \in [s-1]$ assume that $(g_1, \ldots, g_d) \in G^d_{p,m-\boldsymbol{w}}(f)$ is already found. Choose $g_{d+1} \in G_{p,m-w_{d+1}}(f)$ such that $R^{d+1}_{\boldsymbol{\gamma}}((x^{w_1}g_1, \ldots, x^{w_d}g_d, x^{w_{d+1}}g_{d+1}), f)$ is minimized as a function of g_{d+1}.*

In the algorithm above, the search set is reduced for each coordinate of $(g_1, \ldots, g_s)$ according to the weight γ_j, since with increasing w_j the search set becomes smaller, as the weight γ_j and thus the corresponding component's influence on the quality of the generating vector decreases. For this reason we call Algorithm 1.2.1 a reduced CBC algorithm. We will now study Algorithm 1.2.1 for different choices of f.

1.2.2 Polynomial lattice point sets for $f(x) = x^m$

Let us shift our attention to the interesting case where $f\colon \mathbb{F}_p \to \mathbb{F}_p, x \mapsto x^m$. Throughout the rest of this section we write x^m instead of f to emphasise our special choice of f. Note that for $g \in \mathbb{F}_p((x^{-1}))$ the Laurent series g/f can be easily computed in this case by shifting the coefficients of g m times to the left. This is why the choice x^m for the modulus is the most frequently used in practice. Furthermore, the mathematical analysis of the reduced CBC algorithm is slightly less technical in this case, since the proof of the following discrepancy bound requires to compute a sum over all divisors of the modulus f. This is much easier for the special case $f(x) = x^m$ than for a general modulus f. It is the aim of this section to prove the following theorem:

Theorem 1.2.2. *Let $\boldsymbol{\gamma} = (\gamma_j)_{j\geq 1}$ be positive real numbers and $\boldsymbol{w} = (w_j)_{j\geq 1}$ be nonnegative real numbers with $0 = w_1 \leq w_2 \leq \ldots$. Let further $(g_1, \ldots, g_s) \in G^s_{p,m-\boldsymbol{w}}(x^m)$ be constructed using Algorithm 1.2.1. Then we have for every $d \in [s]$*

$$R^d_{\boldsymbol{\gamma}}((x^{w_1}g_1, \ldots, x^{w_d}g_d), x^m) \leq \frac{1}{p^m} \prod_{i=1}^{d} \left(1 + \gamma_i + \gamma_i 2p^{\min\{w_i,m\}} m \frac{p^2-1}{3p}\right).$$

As a direct consequence we obtain the following discrepancy estimate.

Corollary 1.2.3. *Let $N = p^m$ and $\boldsymbol{\gamma}$, $\boldsymbol{w}$ and $(g_1, \ldots, g_s)$ be as in Theorem 1.2.2. Then the polynomial lattice point set $\mathcal{P}((x^{w_1}g_1, \ldots, x^{w_s}g_s), x^m)$ has a weighted star discrepancy*

$$D^*_{N,\boldsymbol{\gamma}}((x^{w_1}g_1, \ldots, x^{w_s}g_s), x^m) \leq \sum_{\substack{\mathfrak{u}\subseteq[s]\\ \mathfrak{u}\neq\emptyset}} \gamma_{\mathfrak{u}} \left(1 - \left(1 - \frac{1}{N}\right)^{|\mathfrak{u}|}\right) + \frac{1}{N}\prod_{i=1}^{s}\left(1 + \gamma_i + \gamma_i 2p^{\min\{w_i,m\}} m\frac{p^2-1}{3p}\right).$$

Knowing the above discrepancy bound, we are now ready to ask about the size of the polynomial lattice point set required to achieve a weighted star discrepancy not exceeding some ε threshold. In particular, we would like to know how this size depends on the dimension s and on ε.

Corollary 1.2.4. *Let $N = p^m$, $\boldsymbol{\gamma}$, and $\boldsymbol{w}$ be as in Theorem 1.2.2. Assume that $(g_1, \ldots, g_s)$ is constructed according to Algorithm 1.2.1. Then $\sum_{j=1}^{\infty}\gamma_j p^{w_j} < \infty$ implies*

$$D^*_{N,\boldsymbol{\gamma}}((x^{w_1}g_1, \ldots, x^{w_s}g_s), x^m) = \mathcal{O}(N^{-1+\delta}),$$

with the implied constant independent of s, for any $\delta > 0$.

Proof. Construct a generating vector $(g_1, \ldots, g_s) \in G^s_{p,m-\boldsymbol{w}}(x^m)$ by applying Algorithm 1.2.1 and consider its weighted star discrepancy, which is bounded in the following way due to Corollary 1.2.3:

$$D^*_{N,\boldsymbol{\gamma}}((x^{w_1}g_1, \ldots, x^{w_s}g_s), x^m) \leq \sum_{\substack{\mathfrak{u}\subseteq[s]\\ \mathfrak{u}\neq\emptyset}} \gamma_{\mathfrak{u}} \left(1 - \left(1 - \frac{1}{N}\right)^{|\mathfrak{u}|}\right) + \frac{1}{N}\prod_{i=1}^{s}\left(1 + \gamma_i + \gamma_i 2p^{\min\{w_i,m\}} m\frac{p^2-1}{3p}\right). \quad (1.43)$$

Recall that by (1.42) we already know that

$$\sum_{\substack{\mathfrak{u}\subseteq[s]\\ \mathfrak{u}\neq\emptyset}} \gamma_{\mathfrak{u}} \left(1 - \left(1 - \frac{1}{N}\right)^{|\mathfrak{u}|}\right) = \mathcal{O}(N^{-1}), \quad (1.44)$$

where the implied constant is independent of N and s. For the second summand in (1.43) we get

$$\begin{aligned}\prod_{i=1}^{s}\left(1+\gamma_i+\gamma_i 2p^{\min\{w_i,m\}}m\frac{p^2-1}{3p}\right) &\leq \prod_{i=1}^{s}\left(1+3\gamma_i p^{\min\{w_i,m\}}m\frac{p^2-1}{3p}\right)\\ &\leq \prod_{i=1}^{s}\left(1+c_p\gamma_i p^{w_i}\log_p N\right),\end{aligned}$$

where $c_p=(p^2-1)/p$. If we define for $d\in\mathbb{N}_0$ the quantity $\sigma_d := c_p\sum_{i=d+1}^{\infty}\gamma_i p^{w_i}$ then it follows by [47, Lemma 3] that

$$\prod_{i=1}^{s}\left(1+c_p\gamma_i p^{w_i}\log_p N\right) \leq (1+\sigma_d^{-1})^d N^{(1+\sigma_0)\sigma_d}.$$

Now let $\delta\in(0,1)$ and choose $d\in\mathbb{N}_0$ sufficiently large such that $\sigma_d\leq\delta(1+\sigma_0)^{-1}$. Then we obtain

$$\prod_{i=1}^{s}\left(1+\gamma_i+\gamma_i 2p^{\min\{w_i,m\}}m\frac{p^2-1}{3p}\right) \leq \prod_{i=1}^{s}\left(1+c_p\gamma_i p^{w_i}\log_p N\right) \leq c_{p,\boldsymbol{\gamma},\delta}N^{\delta}, \quad (1.45)$$

where $c_{p,\boldsymbol{\gamma},\delta}$ is independent of N and s. Applying the estimates (1.44) and (1.45) to (1.43) we get that

$$D^*_{N,\boldsymbol{\gamma}}\left((x^{w_1}g_1,\ldots,x^{w_s}g_s),x^m\right)=\mathcal{O}(N^{-1+\delta}),\quad \text{for } \delta>0$$

and this finishes the proof. □

In order to show Theorem 1.2.2 we need several auxiliary results.

Lemma 1.2.5. *Let $a\in\mathbb{F}_p[x]$ be monic. Then we have*

$$\sum_{\substack{h\in G_{p,m}\setminus\{0\}\\ a|h}} r_p(h) = (m-\deg(a))\,\frac{p^2-1}{3p}p^{-\deg(a)}.$$

In particular, for $a=1$ this formula yields $\sum_{h\in G_{p,m}\setminus\{0\}} r_p(h)=m\frac{p^2-1}{3p}$.

Proof. This fact follows from [22, p. 1055] (by setting $\gamma_{d+1}=1$). The special case $a=1$ also follows from [25, Lemma 2.2] by setting $s=1$. □

For our purposes, it is convenient to write $R^s_{\boldsymbol{\gamma}}(\boldsymbol{g}, f)$ from (1.41) in an alternative way. To this end, we introduce some notation. For a Laurent series $L \in \mathbb{F}_p((x^{-1}))$ we denote by $c_{-1}(L)$ its coefficient of x^{-1}, i.e., its residuum. Further, we set $X_p(L) := \chi_p(c_{-1}(L))$, where χ_p is a non-trivial additive character of $\mathbb{F}_p$. One could for instance choose $\chi_p(n) = \mathrm{e}^{(2\pi\mathrm{i}/p)n}$ for $n \in \mathbb{F}_p$ (see, e.g., [74]). It is clear that $X_p(L) = 1$ if L is a polynomial and that $X_p(L_1 + L_2) = X_p(L_1)X_p(L_2)$ for $L_1, L_2 \in \mathbb{F}_p((x^{-1}))$. From [80, p. 78] we know that for some $q \in \mathbb{F}_p[x]$ we have

$$\sum_{v \in G_{p,m}} X_p\left(\frac{v}{f}q\right) = \begin{cases} p^m & \text{if } f \mid q, \\ 0 & \text{otherwise.} \end{cases} \tag{1.46}$$

With this, it is an easy task to show the following formula.

Lemma 1.2.6. *We have*

$$R^s_{\boldsymbol{\gamma}}(\boldsymbol{g}, f) = -\prod_{i=1}^{s}(1+\gamma_i) + \frac{1}{p^m}\sum_{v \in G_{p,m}}\prod_{i=1}^{s}\left(1 + \gamma_i + \gamma_i \sum_{h \in G_{p,m}\setminus\{0\}} r_p(h) X_p\left(\frac{v}{f}hg_i\right)\right).$$

Proof. We start with (1.41) and employ the properties of X_p as stated above to obtain

$$\begin{aligned} R^s_{\boldsymbol{\gamma}}(\boldsymbol{g}, f) &= -\prod_{i=1}^{s}(1+\gamma_i) + \frac{1}{p^m}\sum_{\boldsymbol{h} \in G^s_{p,m}}\left(\prod_{i=1}^{s} r_p(h_i,\gamma_i)\right)\sum_{v \in G_{p,m}} X_p\left(\frac{v}{f}\boldsymbol{h}\cdot\boldsymbol{g}\right) \\ &= -\prod_{i=1}^{s}(1+\gamma_i) + \frac{1}{p^m}\sum_{v \in G_{p,m}}\prod_{i=1}^{s}\left(\sum_{h_i \in G_{p,m}} r_p(h_i,\gamma_i) X_p\left(\frac{v}{f}h_i g_i\right)\right) \\ &= -\prod_{i=1}^{s}(1+\gamma_i) + \frac{1}{p^m}\sum_{v \in G_{p,m}}\prod_{i=1}^{s}\left(1 + \gamma_i + \gamma_i \sum_{h \in G_{p,m}\setminus\{0\}} r_p(h) X_p\left(\frac{v}{f}hg_i\right)\right), \end{aligned}$$

and the claimed formula is verified. □

Now we study a sum which will appear later in the proof of Theorem 1.2.2 and show an upper bound for it.

Lemma 1.2.7. *Let $w \in \mathbb{N}_0$ and $v \in G_{p,m}$. Let*

$$Y_{p^m,w}(v, x^m) := \sum_{g \in G_{p,m-w}(x^m)}\sum_{h \in G_{p,m}\setminus\{0\}} r_p(h) X_p\left(\frac{v}{x^m}hx^w g\right).$$

Then we have

$$\frac{1}{\#G_{p,m-w}(x^m)} \sum_{v \in G_{p,m}} |Y_{p^m,w}(v,x^m)| \leq 2p^{\min\{w,m\}} m \frac{p^2-1}{3p}.$$

Proof. Let us first assume that $w \geq m$. Then we have $G_{p,m-w}(x^m) = \{1\}$ and therefore

$$Y_{p^m,w}(v,x^m) = \sum_{h \in G_{p,m} \setminus \{0\}} r_p(h) X_p(vhx^{w-m}) = \sum_{h \in G_{p,m} \setminus \{0\}} r_p(h) = m\frac{p^2-1}{3p}$$

with Lemma 1.2.5. Hence, in the case $w \geq m$ we obtain

$$\frac{1}{\#G_{p,m-w}(x^m)} \sum_{v \in G_{p,m}} |Y_{p^m,w}(v,x^m)| = p^m m \frac{p^2-1}{3p} \leq 2p^{\min\{w,m\}} m \frac{p^2-1}{3p}.$$

For the rest of the proof let $w < m$. We abbreviate $\#G_{p,m-w}(x^m)$ by $\#G$ and write

$$\frac{1}{\#G} \sum_{v \in G_{p,m}} |Y_{p^m,w}(v,x^m)| = \frac{1}{\#G} \sum_{\substack{v \in G_{p,m} \\ x^{m-w} | v}} |Y_{p^m,w}(v,x^m)| + \frac{1}{\#G} \sum_{\substack{v \in G_{p,m} \\ x^{m-w} \nmid v}} |Y_{p^m,w}(v,x^m)|.$$

In what follows, we refer to the latter sums as

$$S_1 := \frac{1}{\#G} \sum_{\substack{v \in G_{p,m} \\ x^{m-w} | v}} |Y_{p^m,w}(v,x^m)| \quad \text{and} \quad S_2 := \frac{1}{\#G} \sum_{\substack{v \in G_{p,m} \\ x^{m-w} \nmid v}} |Y_{p^m,w}(v,x^m)|.$$

We may uniquely write any $v \in G_{p,m} \setminus \{0\}$ in the form $v = qx^{m-w} + \ell$, where $q, \ell \in \mathbb{F}_p[x]$ with $\deg(q) < w$ and $\deg(\ell) < m - w$. Using the properties of X_p it is clear that $Y_{p^m,w}(v,x^m) = Y_{p^m,w}(\ell,x^m)$ and hence

$$\begin{aligned} S_1 &= \frac{1}{\#G} \sum_{\substack{v \in G_{p,m} \\ x^{m-w} | v}} |Y_{p^m,w}(0,x^m)| = \sum_{\substack{v \in G_{p,m} \\ x^{m-w} | v}} \frac{1}{\#G} \sum_{g \in G_{p,m-w}(x^m)} \sum_{h \in G_{p,m} \setminus \{0\}} r_p(h) \\ &= \sum_{\substack{v \in G_{p,m} \\ x^{m-w} | v}} m \frac{p^2-1}{3p} = p^{\min\{w,m\}} m \frac{p^2-1}{3p}. \end{aligned}$$

We move on to S_2. Let $e(\ell) := \max\{k \in \{0,1,\dots,m-w-1\} : x^k \mid \ell\}$. With this definition we may display S_2 as

$$S_2 = \frac{p^w}{\#G} \sum_{k=0}^{m-w-1} \sum_{\substack{\ell \in G_{p,m-w}\setminus\{0\} \\ e(\ell)=k}} |Y_{p^m,w}(\ell, x^m)|. \tag{1.47}$$

We compute $Y_{p^m,w}(\ell, x^m)$ for $\ell \in G_{p,m-w} \setminus \{0\}$ with $e(\ell) = k$. Let μ_p be the Möbius function on the set of monic polynomials over $\mathbb{F}_p$, i.e., $\mu_p : \mathbb{F}_p[x] \to \{-1, 0, 1\}$ and

$$\mu_p(h) = \begin{cases} (-1)^\nu & \text{if } h \text{ is squarefree and has } \nu \text{ irreducible factors,} \\ 0 & \text{else.} \end{cases}$$

We call h squarefree if there is no irreducible polynomial $q \in \mathbb{F}_p[x]$ with $\deg(q) \geq 1$ such that $q^2 \mid h$. The fact that $\mu_p(1) = 1$, $\mu_p(x) = -1$ and $\mu_p(x^i) = 0$ for $i \in \mathbb{N}$, $i \geq 2$, yields the equivalence of $\sum_{t|\gcd(x^{m-w},g)} \mu_p(t) = 1$ and $\gcd(x^{m-w}, g) = 1$. Therefore we can write

$$\begin{aligned}
Y_{p^m,w}(\ell, x^m) &= \sum_{h \in G_{p,m}\setminus\{0\}} r_p(h) \sum_{g \in G_{p,m-w}} X_p\left(\frac{\ell}{x^{m-w}} hg\right) \sum_{t|\gcd(x^{m-w},g)} \mu_p(t) \\
&= \sum_{h \in G_{p,m}\setminus\{0\}} r_p(h) \sum_{t|x^{m-w}} \mu_p(t) \sum_{\substack{g \in G_{p,m-w} \\ t|g}} X_p\left(\frac{\ell}{x^{m-w}} hg\right) \\
&= \sum_{h \in G_{p,m}\setminus\{0\}} r_p(h) \sum_{t|x^{m-w}} \mu_p(t) \sum_{a \in G_{p,m-w-\deg(t)}} X_p\left(\frac{\ell}{x^{m-w}} hat\right) \\
&= \sum_{h \in G_{p,m}\setminus\{0\}} r_p(h) \sum_{t|x^{m-w}} \mu_p\left(\frac{x^{m-w}}{t}\right) \sum_{a \in G_{p,\deg(t)}} X_p\left(\frac{a}{t} h\ell\right) \\
&= \sum_{h \in G_{p,m}\setminus\{0\}} r_p(h) \sum_{\substack{t|x^{m-w} \\ t|h\ell}} \mu_p\left(\frac{x^{m-w}}{t}\right) p^{\deg(t)} \\
&= \sum_{t|x^{m-w}} \mu_p\left(\frac{x^{m-w}}{t}\right) p^{\deg(t)} \sum_{\substack{h \in G_{p,m}\setminus\{0\} \\ t|h\ell}} r_p(h).
\end{aligned}$$

The equivalence of the conditions $t \mid h\ell$ and $\frac{t}{\gcd(t,\ell)} \mid h$ yields

$$Y_{p^m,w}(\ell, x^m) = \sum_{t|x^{m-w}} \mu_p\left(\frac{x^{m-w}}{t}\right) p^{\deg(t)} \sum_{\substack{h \in G_{p,m}\setminus\{0\} \\ \frac{t}{\gcd(t,\ell)}|h}} r_p(h).$$

We investigate the inner sum and use Lemma 1.2.5 with $a = \frac{t}{\gcd(t,\ell)}$ to find

$$\sum_{\substack{h \in G_{p,m} \setminus \{0\} \\ \frac{t}{\gcd(t,\ell)} | h}} r_p(h) = \left(m - \deg \left(\frac{t}{\gcd(t, \ell)} \right) \right) \frac{p^2 - 1}{3p} p^{-\deg\left(\frac{t}{\gcd(t,\ell)}\right)}.$$

Now we have

$$\begin{aligned} Y_{p^m,w}(\ell, x^m) =& \frac{p^2 - 1}{3p} \sum_{t | x^{m-w}} \mu_p \left(\frac{x^{m-w}}{t} \right) \left(m - \deg \left(\frac{t}{\gcd(t, \ell)} \right) \right) p^{\deg(\gcd(t,\ell))} \\ =& \frac{p^2 - 1}{3p} m \sum_{t | x^{m-w}} \mu_p \left(\frac{x^{m-w}}{t} \right) p^{\deg(\gcd(t,\ell))} \\ & - \frac{p^2 - 1}{3p} \sum_{t | x^{m-w}} \mu_p \left(\frac{x^{m-w}}{t} \right) \deg \left(\frac{t}{\gcd(t, \ell)} \right) p^{\deg(\gcd(t,\ell))}. \end{aligned}$$

From the fact that $e(\ell) = k \leq m - w - 1$ we obtain $\gcd(x^{m-w}, \ell) = \gcd(x^{m-w-1}, \ell) = x^k$. This observation leads to

$$\sum_{t | x^{m-w}} \mu_p \left(\frac{x^{m-w}}{t} \right) p^{\deg(\gcd(t,\ell))} = p^{\deg(\gcd(x^{m-w},\ell))} - p^{\deg(\gcd(x^{m-w-1},\ell))} = 0$$

and

$$\begin{aligned} & \sum_{t | x^{m-w}} \mu_p \left(\frac{x^{m-w}}{t} \right) \deg \left(\frac{t}{\gcd(t, \ell)} \right) p^{\deg(\gcd(t,\ell))} \\ & = \deg\left(\frac{x^{m-w}}{\gcd(x^{m-w}, \ell)} \right) p^{\deg(\gcd(x^{m-w},\ell))} - \deg\left(\frac{x^{m-w-1}}{\gcd(x^{m-w-1}, \ell)} \right) p^{\deg(\gcd(x^{m-w-1},\ell))} \\ & = (m - w - k)p^k - (m - w - k - 1)p^k = p^k. \end{aligned}$$

Altogether we have $Y_{p^m,w}(\ell, x^m) = -\frac{p^2-1}{3p} p^k$. Inserting this result into (1.47) yields

$$S_2 = \frac{p^w}{\#G} \frac{p^2 - 1}{3p} \sum_{k=0}^{m-w-1} p^k \sum_{\substack{\ell \in G_{p,m-w} \setminus \{0\} \\ e(\ell) = k}} 1.$$

Since

$$\#\{\ell \in G_{p,m-w} \setminus \{0\} : e(\ell) = k\}$$

$$
\begin{aligned}
&=\#\{\ell \in G_{p,m-w} \setminus \{0\} : x^k \mid \ell\} - \#\{\ell \in G_{p,m-w} \setminus \{0\} : x^{k+1} \mid \ell\} \\
&=p^{m-w-k} - 1 - (p^{m-w-k-1} - 1) = p^{m-w-k-1}(p-1),
\end{aligned}
$$

we have

$$
\begin{aligned}
S_2 =& \frac{p^w}{p^{m-w-1}(p-1)} \frac{p^2-1}{3p} \sum_{k=0}^{m-w-1} p^k p^{m-w-k-1}(p-1) \\
=& p^w \frac{p^2-1}{3p}(m-w) \leq p^{\min\{w,m\}} m \frac{p^2-1}{3p}.
\end{aligned}
$$

Summarizing, we have shown

$$
\frac{1}{\#G} \sum_{v \in G_{p,m}} |Y_{p^m,w}(v, x^m)| = S_1 + S_2 \leq 2p^{\min\{w,m\}} m \frac{p^2-1}{3p},
$$

which completes the proof.

$\square$

Now we are ready to prove Theorem 1.2.2 using induction on d.

Proof. We show the result for $d = 1$. From Lemma 1.2.6 we have

$$
\begin{aligned}
R^1_{\boldsymbol{\gamma}}((x^{w_1}), x^m) =& -(1+\gamma_1) + \frac{1}{p^m} \sum_{v \in G_{p,m}} \left(1 + \gamma_1 + \gamma_1 \sum_{h \in G_{p,m} \setminus \{0\}} r_p(h) X_p\left(\frac{v}{x^m} h x^{w_1}\right)\right) \\
=& \frac{\gamma_1}{p^m} \sum_{v \in G_{p,m}} \sum_{h \in G_{p,m} \setminus \{0\}} r_p(h) X_p\left(\frac{v}{x^m} h x^{w_1}\right).
\end{aligned}
$$

If $w_1 \geq m$, then

$$
\begin{aligned}
R^1_{\boldsymbol{\gamma}}((x^{w_1}), x^m) =& \frac{\gamma_1}{p^m} \sum_{v \in G_{p,m}} \sum_{h \in G_{p,m} \setminus \{0\}} r_p(h) = \frac{\gamma_1}{p^m} p^{\min\{w_1,m\}} m \frac{p^2-1}{3p} \\
&\leq \frac{1}{p^m}\left(1 + \gamma_1 + \gamma_1 2p^{\min\{w_1,m\}} m \frac{p^2-1}{3p}\right).
\end{aligned}
$$

If $w_1 < m$, then we can write

$$
R^1_{\boldsymbol{\gamma}}((x^{w_1}), x^m) = \frac{\gamma_1}{p^m} \sum_{v \in G_{p,m}} \sum_{h \in G_{p,m} \setminus \{0\}} r_p(h) X_p\left(\frac{v}{x^m} h x^{w_1}\right)
$$

$$= \frac{\gamma_1}{p^m} \sum_{\substack{h \in G_{p,m} \setminus \{0\} \\ x^{m-w_1} | h}} r_p(h) \sum_{v \in G_{p,m}} X_p\left(\frac{v}{x^m} h x^{w_1}\right) + \frac{\gamma_1}{p^m} \sum_{\substack{h \in G_{p,m} \setminus \{0\} \\ x^{m-w_1} \nmid h}} r_p(h) \sum_{v \in G_{p,m}} X_p\left(\frac{v}{x^m} h x^{w_1}\right)$$

$$= \gamma_1 \sum_{\substack{h \in G_{p,m} \setminus \{0\} \\ x^{m-w_1} | h}} r_p(h),$$

where we used (1.46) in the latter step. We use Lemma 1.2.5 with $a = x^{m-w_1}$ to compute

$$\sum_{\substack{h \in G_{p,m} \setminus \{0\} \\ x^{m-w_1} | h}} r_p(h) = \frac{1}{p^m} p^{w_1} w_1 \frac{p^2-1}{3p} \leq \frac{1}{p^m} p^{\min\{w_1,m\}} m \frac{p^2-1}{3p},$$

which leads to the desired result also in this case.
Now let $d \in [s-1]$. Assume that we have some $(g_1, \ldots, g_d) \in G^d_{p,m-\boldsymbol{w}}(x^m)$ such that

$$R^d_{\boldsymbol{\gamma}}((x^{w_1} g_1, \ldots, x^{w_d} g_d), x^m) \leq \frac{1}{p^m} \prod_{i=1}^d \left(1 + \gamma_i + \gamma_i 2 p^{\min\{w_i,m\}} m \frac{p^2-1}{3p}\right).$$

Let $g^* \in G_{p,m-w_{d+1}}(x^m)$ be such that $R^{d+1}_{\boldsymbol{\gamma}}((x^{w_1} g_1, \ldots, x^{w_d} g_d, x^{w_{d+1}} g_{d+1}), x^m)$ is minimized as a function of g_{d+1} for $g_{d+1} = g^*$. Then we have

$$\begin{aligned}
R^{d+1}_{\boldsymbol{\gamma}}((x^{w_1} g_1, \ldots, x^{w_d} g_d, x^{w_{d+1}} g^*), x^m) &= -(1+\gamma_{d+1}) \prod_{i=1}^d (1+\gamma_i) \\
+ \frac{1}{p^m} \sum_{v \in G_{p,m}} \prod_{i=1}^d &\left(1 + \gamma_i + \gamma_i \sum_{h \in G_{p,m} \setminus \{0\}} r_p(h) X_p\left(\frac{v}{x^m} h x^{w_i} g_i\right)\right) \\
\times &\left(1 + \gamma_{d+1} + \gamma_{d+1} \sum_{h \in G_{p,m} \setminus \{0\}} r_p(h) X_p\left(\frac{v}{x^m} h x^{w_{d+1}} g^*\right)\right) \\
= (1+\gamma_{d+1}) R^d_{\boldsymbol{\gamma}}((x^{w_1} g_1, \ldots, x^{w_d} g_d), x^m) &+ L(g^*),
\end{aligned} \tag{1.48}$$

where

$$\begin{aligned}
L(g^*) = &\frac{\gamma_{d+1}}{p^m} \sum_{v \in G_{p,m}} \sum_{h \in G_{p,m} \setminus \{0\}} r_p(h) X_p\left(\frac{v}{x^m} h x^{w_{d+1}} g^*\right) \\
&\times \prod_{i=1}^d \left(1 + \gamma_i + \gamma_i \sum_{h \in G_{p,m} \setminus \{0\}} r_p(h) X_p\left(\frac{v}{x^m} h x^{w_i} g_i\right)\right).
\end{aligned}$$

A minimizer g^* of $R_{\boldsymbol{\gamma}}^{d+1}((x^{w_1}g_1,\ldots,x^{w_d}g_d,x^{w_{d+1}}g_{d+1}),x^m)$ is also a minimizer of $L(g_{d+1})$. With the ideas in the proof of [25, Theorem 2.7], we see that $L(g)\in\mathbb{R}^+$ for all $g\in G_{p,m-w_{d+1}}(x^m)$. Thus we may bound $L(g^*)$ by the mean over all $g\in G_{p,m-w_{d+1}}(x^m)$:

$$
\begin{aligned}
L(g^*)\leq&\frac{1}{\#G_{p,m-w_{d+1}}(x^m)}\sum_{g_{d+1}\in G_{p,m-w_{d+1}}(x^m)}L(g_{d+1})\\
\leq&\frac{\gamma_{d+1}}{p^m}\sum_{v\in G_{p,m}}\frac{1}{\#G_{p,m-w_{d+1}}(x^m)}\\
&\times\left|\sum_{g_{d+1}\in G_{p,m-w_{d+1}}(x^m)}\sum_{h\in G_{p,m}\setminus\{0\}}r_p(h)X_p\left(\frac{v}{x^m}hx^{w_{d+1}}g_{d+1}\right)\right|\\
&\times\prod_{i=1}^{d}\left(1+\gamma_i+\gamma_i\sum_{h\in G_{p,m}\setminus\{0\}}r_p(h)\left|X_p\left(\frac{v}{x^m}hx^{w_i}g_i\right)\right|\right)\\
\leq&\frac{\gamma_{d+1}}{p^m}\prod_{i=1}^{d}\left(1+\gamma_i+\gamma_i m\frac{p^2-1}{3p}\right)\sum_{v\in G_{p,m}}\frac{|Y_{p^m,w_{d+1}}(v,x^m)|}{\#G_{p,m-w_{d+1}}(x^m)},
\end{aligned}
$$

where we used the estimate $\left|X_p\left(\frac{v}{x^m}hx^{w_i}g_i\right)\right|\leq 1$ in the last step. With the induction hypothesis and Lemma 1.2.7 this leads to

$$
\begin{aligned}
&R_{\boldsymbol{\gamma}}^{d+1}((x^{w_1}g_1,\ldots,x^{w_d}g_d,x^{w_{d+1}}g^*),x^m)\\
&\leq(1+\gamma_{d+1})\frac{1}{p^m}\prod_{i=1}^{d}\left(1+\gamma_i+\gamma_i 2p^{\min\{w_i,m\}}m\frac{p^2-1}{3p}\right)\\
&\quad+\frac{\gamma_{d+1}}{p^m}\prod_{i=1}^{d}\left(1+\gamma_i+\gamma_i m\frac{p^2-1}{3p}\right)2p^{\min\{w_{d+1},m\}}m\frac{p^2-1}{3p}\\
&\leq\frac{1}{p^m}\prod_{i=1}^{d}\left(1+\gamma_i+\gamma_i 2p^{\min\{w_i,m\}}m\frac{p^2-1}{3p}\right)\\
&\quad\times\left(1+\gamma_{d+1}+\gamma_{d+1}2p^{\min\{w_{d+1},m\}}m\frac{p^2-1}{3p}\right)\\
&=\frac{1}{p^m}\prod_{i=1}^{d+1}\left(1+\gamma_i+\gamma_i 2p^{\min\{w_i,m\}}m\frac{p^2-1}{3p}\right).
\end{aligned}
$$

□

The reduced fast CBC construction

So far we have seen how to construct a generating vector $\boldsymbol{g}$ of the point set $\mathcal{P}(\boldsymbol{g}, x^m)$. In fact Algorithm 1.2.1 can be made much faster using results of [23, 84, 85]. In this section we are investigating and improving Algorithm 1.2.1 and additionally analysing the computational cost of the improved algorithm.

Walsh functions are a suitable tool for analysing the computational cost of CBC algorithms for constructing polynomial lattice point sets. Let $\omega = \mathrm{e}^{2\pi \mathrm{i}/p}$, $x \in [0,1)$ and h a nonnegative integer with prime base p representation $x = x_1/p + x_2/p^2 + \cdots$ and $h = h_0 + h_1 p + \cdots + h_r p^r$, respectively. Then we define

$$\mathrm{wal}_h : [0,1) \to \mathbb{C}, \mathrm{wal}_h(x) := \omega^{h_0 x_1 + \cdots + h_r x_{r+1}}.$$

The Walsh function system $\{\mathrm{wal}_h \mid h = 0, 1, \ldots\}$ is a complete orthonormal basis in $L_2([0,1))$ which has been used in the analysis of the discrepancy of digital nets several times before, see for example [25, 46, 70]. For further information on Walsh functions see [26, Appendix A].

Let $d \geq 1$, $N = p^m$. For $P(\boldsymbol{g}, f) = \{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{p^m-1}\}$ with $\boldsymbol{x}_n = (x_n^{(1)}, \ldots, x_n^{(s)})$ we have the formula (see [25, Sect. 4])

$$\frac{1}{p^m} \sum_{n=0}^{p^m-1} \prod_{i=1}^{s} \mathrm{wal}_{h_i}(x_n^{(i)}) = \begin{cases} 1 & \text{if } \boldsymbol{g} \cdot \boldsymbol{h} \equiv 0 \pmod{f}, \\ 0 & \text{otherwise,} \end{cases} \tag{1.49}$$

where h_i are nonnegative integers with base p representation $h_i = h_0^{(i)} + h_1^{(i)} p + \cdots + h_r^{(i)} p^r$. We identify these nonnegative integers h_i with the polynomials $h_i(x) = h_0^{(i)} + h_1^{(i)} x + \cdots + h_r^{(i)} x^r$, which are elements of $G_{p,m}$. The vectors $\boldsymbol{h}$ in (1.49) are then from $G_{p,m}^s$ such that $\boldsymbol{h} = (h_1(x), \ldots, h_s(x))$. Equation (1.49) allows us to rewrite $R_{\boldsymbol{\gamma}}^d(\boldsymbol{g}, x^m)$ in the following way:

$$R_{\boldsymbol{\gamma}}^d(\boldsymbol{g}, x^m) = -\prod_{i=1}^{d}(1+\gamma_i) + \frac{1}{p^m} \sum_{n=0}^{p^m-1} \prod_{i=1}^{d} \sum_{h=0}^{p^m-1} r_p(h, \gamma_i) \mathrm{wal}_h \left(\phi_m \left(\frac{n x^{w_i} g_i}{x^m} \right) \right).$$

Note that $r_p(h, \gamma)$ is defined as in (1.41) and we identify the integer in base p representation $h = h_0 + h_1 p + \cdots + h_r p^r$ with the polynomial $h(x) = h_0 + h_1 x + \cdots + h_r x^r$. If we set $\psi(\frac{n x^{w_i} g_i}{x^m}) := \sum_{h=1}^{p^m-1} r_p(h) \mathrm{wal}_h(\phi_m(\frac{n x^{w_i} g_i}{x^m}))$ we get that

$$R_{\boldsymbol{\gamma}}^d(\boldsymbol{g}, x^m) = -\prod_{i=1}^{d}(1+\gamma_i) + \frac{1}{p^m} \sum_{n=0}^{p^m-1} \prod_{i=1}^{d} \left(1 + \gamma_i + \gamma_i \psi \left(\frac{n x^{w_i} g_i}{x^m} \right) \right)$$

$$= -\prod_{i=1}^{d}(1+\gamma_i) + \frac{1}{p^m}\sum_{n=0}^{p^m-1}\eta_d(n), \tag{1.50}$$

where $\eta_d(n) = \prod_{i=1}^{d}\left(1+\gamma_i+\gamma_i\psi(\frac{nx^{w_i}g_i}{x^m})\right)$.
In [25, Sect. 4] it is proved that we can compute the at most N different values of $\psi(\frac{r}{x^m})$ for $r \in G_{p,m}$ in $\mathcal{O}(N\log N)$ operations.
Let us study one step of the reduced CBC algorithm. Assuming we already have found $(g_1,\dots,g_d) \in G^d_{p,m-\boldsymbol{w}}(x^m)$ we have to minimize $R^{d+1}_{\boldsymbol{\gamma}}((x^{w_1}g_1,\dots,x^{w_{d+1}}g_{d+1}),x^m)$ as a function of $g_{d+1} \in G_{p,m-w_{d+1}}(x^m)$. If $w_{d+1} \geq m$ then $g_{d+1}=1$ and we are done. Let now $w_{d+1} < m$. From (1.50) we have that

$$\begin{aligned}
&R^{d+1}_{\boldsymbol{\gamma}}((x^{w_1}g_1,\dots,x^{w_{d+1}}g_{d+1}),x^m)\\
&\quad= -\prod_{i=1}^{d+1}(1+\gamma_i) + \frac{1}{p^m}\sum_{n=0}^{p^m-1}\left(1+\gamma_{d+1}+\gamma_{d+1}\psi\left(\frac{nx^{w_{d+1}}g_{d+1}}{x^m}\right)\right)\eta_d(n).
\end{aligned}$$

In order to minimize $R^{d+1}_{\boldsymbol{\gamma}}((x^{w_1}g_1,\dots,x^{w_{d+1}}g_{d+1}),x^m)$ it is enough to minimize $T_d(g) := \sum_{n=0}^{p^m-1}\psi(\frac{nx^{w_{d+1}}g}{x^m})\eta_d(n)$. As in [23, Sect. 4] we can represent this quantity using some specific $(p^{m-w_{d+1}-1}(p-1)\times N)$-matrix A and exploiting its additional structure. Let therefore

$$A = \left(\psi\left(\frac{nx^{w_{d+1}}g}{x^m}\right)\right)_{\substack{g\in G_{p,m-w_{d+1}}(x^m),\\ n\in\{0,\dots,N-1\}}} \text{ and } \boldsymbol{\eta}_d = (\eta_d(0),\dots,\eta_d(N-1))^\top.$$

First of all observe that we get $(T_d(g))_{g\in G_{p,m-w_{d+1}}(x^m)} = A\boldsymbol{\eta}_d$. Secondly the matrix A is a block matrix and can be written in the following form

$$A = \left(\Omega^{(m-w_{d+1})}\dots\Omega^{(m-w_{d+1})}\right), \text{ where } \Omega^{(l)} = \left(\psi\left(\frac{nx^{w_{d+1}}g}{x^m}\right)\right)_{\substack{g\in G_{p,m-w_{d+1}}(x^m),\\ n\in\{0,\dots p^l-1\}}}.$$

If $\boldsymbol{x}$ is any vector of size p^m then we compute

$$A\boldsymbol{x} = \Omega^{(m-w_{d+1})}\boldsymbol{x}_1+\dots+\Omega^{(m-w_{d+1})}\boldsymbol{x}_{p^{w_{d+1}}} = \Omega^{(m-w_{d+1})}(\boldsymbol{x}_1+\dots+\boldsymbol{x}_{p^{w_{d+1}}}),$$

where $\boldsymbol{x}_1$ is the vector consisting of the first $p^{m-w_{d+1}}$ components of $\boldsymbol{x}$, $\boldsymbol{x}_2$ is the vector consisting of the next $p^{m-w_{d+1}}$ components of $\boldsymbol{x}$ and so on. Now we apply the machinery of [84, 85] and get that multiplication with $\Omega^{(m-w_{d+1})}$ can be done in $\mathcal{O}((m-w_{d+1})p^{m-w_{d+1}})$ operations. Summarizing we have:

Algorithm 1.2.8.

1. *Compute* $\psi(\frac{r}{x^m})$ *for* $r \in G_{p,m}$.

2. *Set* $\eta_1(n) = \psi(\frac{nx^{w_1}g_1}{x^m})$ *for* $n = 0, \dots, p^m - 1$.

3. *Set* $g_1 = 1,\ d = 2$ *and* $t = \max\{j \in [s] \mid w_j < m\}$. *While* $d \le \min\{s, t\}$,

   (a) *Partition* η_{d-1} *into* p^{w_d} *vectors* $\eta_{d-1}^{(1)}, \dots, \eta_{d-1}^{(p^{w_d})}$ *of length* p^{m-w_d} *and let* $\eta' = \sum_{i=1}^{p^{w_d}} \eta_{d-1}^{(i)}$.

   (b) *Let* $(T_{d-1}(g))_{g \in G_{p,m-w_d}(x^m)} = \Omega^{(m-w_d)}\eta'$.

   (c) *Let* $g_d = \operatorname{argmin}_g T_{d-1}(g)$.

   (d) *Let* $\eta_d(n) = (1 + \gamma_{d-1} + \gamma_{d-1}\psi(\frac{nx^{w_d}g_d}{x^m}))\eta_{d-1}(n)$

   (e) *Increase* d *by 1.*

4. *If* $s \ge t$ *then set* $g_t = g_{t+1} = \ldots = g_s = 1$.

Similar to [23] we obtain from the results in this section the following theorem:

Theorem 1.2.9. *Let* $N = p^m$ *then the cost of Algorithm 1.2.8 is*

$$\mathcal{O}\left(N \log N + \min\{s,t\}N + N \sum_{d=1}^{\min\{s,t\}} (m - w_d)p^{-w_d}\right).$$

1.2.3 Polynomial lattice point sets for irreducible f

Finally we want to consider the special case where f is an irreducible polynomial. So, for this section let f be an irreducible polynomial over $\mathbb{F}_p$ with $\deg(f) = m$. We will prove the subsequent theorem.

Theorem 1.2.10. *Let* $\boldsymbol{\gamma}$ *and* $\boldsymbol{w}$ *as in Theorem 1.2.2 and let* $f \in \mathbb{F}_p[x]$ *be an irreducible polynomial with* $\deg(f) = m$. *Let further* $(g_1, \dots, g_s) \in G^s_{p,m-\boldsymbol{w}}(f)$ *be constructed according to Algorithm 1.2.1. Then we have for every* $d \in [s]$

$$R^d_{\boldsymbol{\gamma}}((x^{w_1}g_1, \dots, x^{w_d}g_d), f) \le \frac{1}{p^m} \prod_{i=1}^{d} \left(1 + \gamma_i + \gamma_i p^{\min\{w_i,m\}} m \frac{p+1}{3}\right).$$

Proof. We will prove this result by induction on d. According to Algorithm 1.2.1 we know that $g_1 = 1$ for $d = 1$. Therefore $R^1_{\boldsymbol{\gamma}}((x^{w_1}g_1), f) = 0$ since for all $h \in G_{p,m}$ we have $\deg(h) < m$ and hence the congruence $hx^{w_1} \equiv 0 \pmod{f}$ has no solutions.
Let $d \in [s-1]$ and assume that we have already found $(g_1, \ldots, g_d) \in G^d_{p,m-\boldsymbol{w}}(f)$. For $\boldsymbol{g} = (x^{w_1}g_1, \ldots, x^{w_d}g_d)$ we have from (1.41) that

$$R^{d+1}_{\boldsymbol{\gamma}}((\boldsymbol{g}, x^{w_{d+1}}g_{d+1}), f) = (1 + \gamma_{d+1})R^d_{\boldsymbol{\gamma}}(\boldsymbol{g}, f) + \theta(g_{d+1}), \tag{1.51}$$

where

$$\theta(g_{d+1}) = \sum_{h_{d+1} \in G_{p,m} \setminus \{0\}} r_p(h_{d+1}, \gamma_{d+1}) \sum_{\substack{\boldsymbol{h} \in G^d_{p,m} \\ \boldsymbol{h} \cdot \boldsymbol{g} \equiv -h_{d+1}x^{w_{d+1}}g_{d+1} \pmod{f}}} \prod_{i=1}^{d} r_p(h_i, \gamma_i).$$

We now proceeded similarly as in the proof of Theorem 1.2.2. Let $g^* \in G_{p,m-w_{d+1}}(f)$ be a minimizer of $R^{d+1}_{\boldsymbol{\gamma}}((\boldsymbol{g}, x^{w_{d+1}}g_{d+1}), f)$ as a function of g_{d+1}. Therefore g^* also minimizes $\theta(g_{d+1})$. Bounding $\theta(g^*)$ by its mean we obtain

$$\begin{aligned}\theta(g^*) \leq & \frac{1}{\#G_{p,m-w_{d+1}}(f)} \sum_{h_{d+1} \in G_{p,m} \setminus \{0\}} r_p(h_{d+1}, \gamma_{d+1}) \\ & \times \sum_{\boldsymbol{h} \in G^d_{p,m}} \left(\prod_{i=1}^{d} r_p(h_i, \gamma_i) \right) \sum_{\substack{g_{d+1} \in G_{p,m-w_{d+1}}(f) \\ \boldsymbol{h} \cdot \boldsymbol{g} \equiv -h_{d+1}x^{w_{d+1}}g_{d+1} \pmod{f}}} 1.\end{aligned}$$

Observe that $\gcd(f, h_{d+1}x^{w_{d+1}}) = 1$. Therefore the congruence $h_{d+1}x^{w_{d+1}}g_{d+1} \equiv -\boldsymbol{h} \cdot \boldsymbol{g} \pmod{f}$ has a unique solution in $G_{p,m}$ but not necessarily in $G_{p,m-w_{d+1}}(f)$. In the case that $-\boldsymbol{h} \cdot \boldsymbol{g} \not\equiv 0 \pmod{f}$ we conclude that the congruence has at most one solution in $G_{p,m-w_{d+1}}(f)$. If $-\boldsymbol{h} \cdot \boldsymbol{g} \equiv 0 \pmod{f}$ the congruence has no solution in $G_{p,m-w_{d+1}}(f)$ since $0 \notin G_{p,m-w_{d+1}}(f)$. Hence we find by an application of [25, Lemma 3.3]

$$\begin{aligned}\theta(g^*) &\leq \frac{1}{\#G_{p,m-w_{d+1}}(f)} \sum_{h_{d+1} \in G_{p,m} \setminus \{0\}} r_p(h_{d+1}, \gamma_{d+1}) \sum_{\boldsymbol{h} \in G^d_{p,m}} \prod_{i=1}^{d} r_p(h_i, \gamma_i) \\ &= \frac{1}{\#G_{p,m-w_{d+1}}(f)} \left[\prod_{i=1}^{d} \left(1 + \gamma_i + \gamma_i m \frac{p^2-1}{3p} \right) \right] \left(\gamma_{d+1} m \frac{p^2-1}{3p} \right).\end{aligned}$$

By (1.51) and the induction hypothesis we have that

$$
\begin{aligned}
R_{\gamma}^{d+1}((\boldsymbol{g}, x^{w_{d+1}} g_{d+1}), f) &= (1+\gamma_{d+1}) R_{\gamma}^{d}(\boldsymbol{g}, f) + \theta(g_{d+1}) \\
&\leq \frac{1}{p^m} \prod_{i=1}^{d} \left(1 + \gamma_i + \gamma_i p^{\min\{w_i, m\}} m \frac{p+1}{3} \right) \\
&\quad \times \left(1 + \gamma_{d+1} + \gamma_{d+1} \frac{p^m}{\# G_{p, m-w_{d+1}}(f)} m \frac{p^2-1}{3p} \right) \\
&\leq \frac{1}{p^m} \prod_{i=1}^{d+1} \left(1 + \gamma_i + \gamma_i p^{\min\{w_i, m\}} m \frac{p+1}{3} \right),
\end{aligned}
$$

where we used in the latter step that $\frac{p^m}{\# G_{p,m-w_{d+1}}(f)} \leq \frac{p}{p-1} p^{\min\{w_{d+1}, m\}}$. This follows from the fact that $\# G_{p,m-w_{d+1}}(f) = p^{m-w_{d+1}} - 1$ if $w_{d+1} < m$ and $\# G_{p,m-w_{d+1}}(f) = 1$ if $w_{d+1} \geq m$. This finishes the proof of Theorem 1.2.10. $\square$

As a consequence of (1.40) and Theorem 1.2.10 we obtain analogous results to Corollary 1.2.3 and Corollary 1.2.4 for an irreducible modulus f.

1.3 Metrical star discrepancy bounds for subsequences of digital Kronecker-sequences

We already pointed out in Subsection 1.1.1 that the star discrepancy is a quantitative measure for the irregularity of distribution of a point set $\mathcal{P}$ and it is also intimately related to the integration error of a QMC algorithm via the celebrated Koksma-Hlawka inequality. Therefore it is natural to study $\mathrm{disc}^*(N,s) = \inf_{\mathcal{P}} D_N^*(\mathcal{P})$, where the infimum is extended over all N-element point sets $\mathcal{P}$ in $[0,1)^s$ and the so-called inverse of the star discrepancy (see Subsection 1.1.6)

$$N^*(\varepsilon, s) = \min\{N \in \mathbb{N} \ : \ \mathrm{disc}^*(N,s) \leq \varepsilon\},$$

where $\varepsilon \in (0,1]$. For fixed dimension $s \geq 2$ it is known that there exist $0 < c_s < C_s$ and $\eta_s \in (0, \frac{1}{2})$ such that

$$c_s \frac{(\log N)^{\frac{s-1}{2}+\eta_s}}{N} \leq \mathrm{disc}^*(N,s) \leq C_s \frac{(\log N)^{s-1}}{N} \quad \text{for all } N \geq 2.$$

In this section we consider a different view point. It was pointed out in several discussions that the excellent asymptotic behaviour of the minimal star discrepancy of N-element point sets is not very useful for practical applications, especially when the dimension s is not small. For example it should be noted that $N \mapsto (\log N)^{s-1}/N$ does not start to decrease until $N = \exp(s-1)$ and this number is already huge for moderately large s. In applications of QMC-algorithms however the dimension s could be in the hundreds (see [24, 73]).
Recall from Subsection 1.1.6 that we know by an outstanding work by Heinrich, Novak, Wasilkowski and Woźniakowski [45] that there exists a constant $C > 0$ (Aistleitner showed in [1, Theorem 1] that one can choose $C = 10$) such that

$$\mathrm{disc}^*(N,s) \leq C\sqrt{\frac{s}{N}} \quad \text{for all } s, N \in \mathbb{N}. \tag{1.52}$$

This bound trades a factor of $N^{-1/2}$ for a gain in the behaviour concerning the dimension s. At this point one should mention that there exists also a slightly weaker bound proven in [45, Theorem 1] which is of the form

$$\mathrm{disc}^*(N,s) \leq C\sqrt{\frac{s}{N}}\,(\log s + \log N)^{1/2}. \tag{1.53}$$

The proof in [1] and also the proof of the slightly weaker bound in [45, Theorem 1] uses the probabilistic method. The main ingredient is the fact that one can obtain extremely small probabilities for the deviation from the mean for sums of independent random variables. This probability can be quantified with the help of Bernstein's (in [1]) or Hoeffding's (in [45]) inequality, respectively. In fact, the point sets in [1, 45] consist of N independently chosen random points from the unit cube $[0,1)^s$.

However, so far no explicit construction of point sets whose star discrepancy satisfies a bound like (1.52) or (1.53) is known. Some authors, initiated in [28], presented algorithmic constructions of point sets with star discrepancy of order (1.53). We refer to the survey [37] for more information and references in this direction. However, all these constructions have the disadvantage that their run times are too large in order to be applied in practical applications with large dimension s. So there is still need for a really explicit construction.

In 2014 Löbbe [75] studied certain lacunary subsequences of Kronecker-sequences $(\{n\boldsymbol{\alpha}\})_{n\geq 0}$, where $\boldsymbol{\alpha} \in \mathbb{R}^s$ and where $\{\cdot\}$ denotes the fractional part applied component-wise to a vector (until now the paper is only available via arXiv.org). Based on the work of Aistleitner [2], Löbbe was able to prove the following remarkable metrical result which can be interpreted as a semi-probabilistic (or semi-constructive) version of (1.53).

For $\boldsymbol{\alpha} \in [0,1)^s$ let $\mathcal{P}_N(\boldsymbol{\alpha}) = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\}$ be the point set consisting of the first N elements of the infinite sequence $(\boldsymbol{x}_n)_{n\geq 1}$ in $[0,1)^s$ with $\boldsymbol{x}_n = \{2^{n-1}\boldsymbol{\alpha}\}$ for $n \in \mathbb{N}$.

Theorem 1.3.1 (Löbbe [75, Theorem 1.1])**.** *Let $N \geq 1$ and $s \geq 2$ be integers. Then for every $\theta \in (0,1)$ there is a quantity $C(\theta) > 0$ such that the star discrepancy of the point set $\mathcal{P}_N(\boldsymbol{\alpha})$ satisfies*

$$D_N^*(\mathcal{P}_N(\boldsymbol{\alpha})) \leq C(\theta)\sqrt{\frac{s \log s}{N}}$$

with probability at least $1-\theta$. The quantity $C(\theta)$ is of order $\mathcal{O}(\log \theta^{-1})$.

The main problem in the proof of this result is to prove independence of certain random variables, which are closely related to the point set $\mathcal{P}(\boldsymbol{\alpha})$, in order to be able to apply Bernstein's inequality. Of course, the elements of the classical Kronecker-sequence are not independent. For this reason the author studied lacunary subsequences of the form $(\{2^{n-1}\boldsymbol{\alpha}\})_{n\geq 1}$ which led then to the desired independence properties.

Theorem 1.3.1 makes an assertion for fixed N, i.e. for finite point sets. In 2007 Dick [20] considered the problem of the dependence of star discrepancy on the dimension s also for infinite sequences and he gave an existence result. Compared to the bound (1.52) for finite point sets the generalisation is penalised with an extra $\sqrt{\log N}$-factor in the discrepancy estimate. Later Aistleitner [2] improved this further and got rid of the $\sqrt{\log N}$-term. In contrast to the probabilistic approaches in, e.g., [45, 20], the proof in [2] is, like in [75], also of a semi-probabilistic nature in the sense that certain coordinates of the points are deterministic others are chosen randomly. This once more shows the relevance of semi-probabilistic constructions in this context.

The following corollary to Theorem 1.3.1 addresses a metrical result for infinite sequences:

Corollary 1.3.2. *Let $s \in \mathbb{N}$ with $s \geq 2$. Then for every $\delta \in (0,1)$ there is a quantity $C(\delta) > 0$ such that the star discrepancy of $\mathcal{P}_N(\boldsymbol{\alpha})$ satisfies*

$$D_N^*(\mathcal{P}_N(\boldsymbol{\alpha})) \leq C(\delta)(\log N)\sqrt{\frac{s \log s}{N}} \quad \text{for all } N \geq 2$$

with probability at least $1-\delta$. We have $C(\delta) = \mathcal{O}\left(\log \delta^{-1}\right)$.

Concerning the proof of Corollary 1.3.2 we will refer to Section 1.3.2.

There is an interesting connection of Corollary 1.3.2 to the theory of normal numbers which is worth to be mentioned: it is well-known that a real number α is normal to base 2, if and only if the sequence $(\{2^{n-1}\alpha\})_{n\geq 1}$ is uniformly distributed modulo one (see [64, Chapter 1, Theorem 8.1]). Hence the $\boldsymbol{\alpha}$'s which satisfy the discrepancy estimate in Corollary 1.3.2 are s-tuples of normal numbers to base 2. (By another well-known result due to Borel [15] almost all numbers $\alpha \in [0,1]$ are normal to every base $b \geq 2$.)

It should also be mentioned, that metrical bounds on the star discrepancy of classical Kronecker-sequences for fixed s have been given by Beck in [10].

In the following subsection we study digital Kronecker-sequences which are a "non-Archimedean analogue" to classical Kronecker-sequences and which fit into the class of digital (t,s)-sequences. This concept was introduced by Niederreiter [80, Section 4] and further investigated by Larcher and Niederreiter [68]. We will give a digital analogue of Theorem 1.3.1. In the next section we provide the necessary definitions and we formulate the metrical discrepancy estimate. The proof of our result will be presented in Section 1.3.2.

1.3.1 Digital Kronecker-sequences and formulation of the main result

We will continue by introducing digital Kroecker-sequences over $\mathbb{F}_q$ where q is a prime. Recall that we have already defined several notions which are related to the field of formal Laurent series $\mathbb{F}_q((t^{-1}))$ in Subsection 1.1.4. Among others we described a way for uniquely associating a ploynomial $n(t)$ to each integer n.

Remark 1.3.3. Note that due to readability we will from now on alter the notation from $\mathbb{F}_p((x^{-1}))$ to $\mathbb{F}_q((t^{-1}))$.

With the help of this notation we are now able to state the subsequent definition:

Definition 1.3.4. For a given s-tuple $\boldsymbol{f} = (f_1, \ldots, f_s)$ of elements of $\mathbb{F}_q((t^{-1}))$ the sequence $\mathcal{S}(\boldsymbol{f}) = (\boldsymbol{y}_n)_{n\geq 0}$ given by

$$\boldsymbol{y}_n = \phi(\{n\boldsymbol{f}\}) = (\phi(\{nf_1\}), \ldots, \phi(\{nf_s\})) \quad \text{for all } n \in \mathbb{N}_0$$

is called a *digital Kronecker-sequence* over $\mathbb{F}_q$. Note that the multiplication of the polynomial n and the Laurent series f_j is carried out in $\mathbb{F}_q((t^{-1}))$. (Obviously it suffices to choose $\boldsymbol{f} \in (\overline{\mathbb{F}}_q((t^{-1})))^s$.) Moreover, the operations $\{\cdot\}$ and ϕ are understood component-wise if they are applied to vectors.

In order to prove a metrical result for digital Kronecker-sequences we need to introduce a suitable probability measure on $(\overline{\mathbb{F}}_q((t^{-1})))^s$.

Definition 1.3.5. By μ we denote the normalized Haar-measure on $\overline{\mathbb{F}}_q((t^{-1}))$ and by μ_s the s-fold product measure on $(\overline{\mathbb{F}}_q((t^{-1})))^s$.

Remark 1.3.6. The measure μ has the following rather simple shape: If we identify the elements $\sum_{k=1}^{\infty} g_k t^{-k}$ of $\overline{\mathbb{F}}_q((t^{-1}))$ where $g_k \neq q-1$ for infinitely many k in the natural way with the real numbers $\sum_{k=1}^{\infty} g_k q^{-k} \in [0,1)$ (see Subsection 1.1.4), then, by neglecting the countably many elements where $g_k \neq q-1$ only for finitely many k, μ corresponds to the Lebesgue measure λ on $[0,1)$. For example, the "cylinder set" $C(c_1, \ldots, c_m)$ consisting of all elements $g = \sum_{k=1}^{\infty} g_k t^{-k}$ from $\overline{\mathbb{F}}_q((t^{-1}))$ with $g_k = c_k$ for $k = 1, \ldots, m$ and arbitrary $g_k \in \mathbb{F}_q$ for $k \geq m+1$ has measure $\mu(C(c_1, \ldots, c_m)) = q^{-m}$.

Metrical results for the star discrepancy of digital Kronecker-sequences for fixed dimension s can be found in [66, 71]. In the following we provide a non-Archimedean version of the result of Löbbe [75].

We pick $\boldsymbol{f} \in (\overline{\mathbb{F}}_q((t^{-1})))^s$ randomly and determine the point set $\mathcal{P}_N(\boldsymbol{f}) = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\}$ consisting of the first N elements of the infinite sequence $(\boldsymbol{x}_n)_{n\geq 1}$ in $[0,1)^s$ with $\boldsymbol{x}_n = \phi(\{t^{n-1}\boldsymbol{f}\})$ for $n \in \mathbb{N}$.

Theorem 1.3.7. *Let q be a prime number and let $N, s \in \mathbb{N}$ with $N, s \geq 2$. Then for every $\varepsilon \in (0,1)$ there is a quantity $C(q,\varepsilon) > 0$ such that the star discrepancy of the point set $\mathcal{P}_N(\boldsymbol{f})$ satisfies*

$$D_N^*(\mathcal{P}_N(\boldsymbol{f})) \leq C(q,\varepsilon)\sqrt{\frac{s \log s}{N}}$$

with μ_s-probability at least $1-\varepsilon$. The quantity $C(q,\varepsilon)$ is of order $\mathcal{O}_q(\log \varepsilon^{-1})$.

The proof of this result will be presented in the next section. It should be mentioned that with some more effort the quantity $C(q,\varepsilon)$ could be given explicitly.

Again Theorem 1.3.7 makes an assertion for fixed N, i.e. for finite point sets. From this we can again deduce a metrical result for infinite sequences:

Corollary 1.3.8. *Let q be a prime number and let $s \in \mathbb{N}$ with $s \geq 2$. Then for every $\delta \in (0,1)$ there is a quantity $C(q,\delta) > 0$ such that the star discrepancy of $\mathcal{P}_N(\boldsymbol{f})$ satisfies*

$$D_N^*(\mathcal{P}_N(\boldsymbol{f})) \leq C(q,\delta)(\log N)\sqrt{\frac{s \log s}{N}} \quad \text{for all } N \geq 2$$

with probability at least $1-\delta$ and $C(q,\delta) = \mathcal{O}_q(\log \delta^{-1})$.

The proof of Corollary 1.3.8 will be presented in Section 1.3.2.

1.3.2 The proof of Theorem 1.3.7

The proof of Theorem 1.3.7 is inspired by the techniques used in [75]. The difficulty here is that we are concerned with polynomial arithmetic over finite fields instead of the usual integer arithmetic.

Throughout the proof we tacitly assume that all components of $\boldsymbol{f}$ belong to the class of Laurent series $\sum_{k=1}^{\infty} g_k t^{-k}$ of $\overline{\mathbb{F}}_q((t^{-1}))$ where $g_k \neq q-1$ for infinitely many k. Recall that with the notation introduced in Subsection 1.1.4 this means that

$$\boldsymbol{f} \in \left(\overline{\mathbb{F}}_q^*((t^{-1}))\right)^s = \left(\overline{\mathbb{F}}_q((t^{-1})) \setminus \mathcal{C}\right)^s \text{ and } \mu(\mathcal{C}) = 0.$$

Some auxiliary results

As in [1, 75] the proof will be based on Bernstein's inequality for sums of independent random variables.

Lemma 1.3.9 ([12], Bernstein inequality). *Let $N \in \mathbb{N}$ and $X_1, \ldots, X_N$ be independent random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with $\mathbb{E}(X_i) = 0$ and $|X_i| \leq C$ for $i \in \{1, \ldots, N\}$ and some $C > 0$. Then we have for any $t > 0$*

$$\mathbb{P}\left(\left|\sum_{i=1}^{N} X_i\right| > t\right) \leq 2\ \exp\left(-\frac{t^2}{2\sum_{i=1}^{N}\mathbb{E}(X_i^2) + \frac{2Ct}{3}}\right).$$

Another very important tool in our analysis are bracketing covers whose definition is recalled below. As usual, for $\boldsymbol{a} = (a_1, \ldots, a_s)$ and $\boldsymbol{b} = (b_1, \ldots, b_s)$ in $[0,1]^s$ we write $\boldsymbol{a} \leq \boldsymbol{b}$ if and only if $a_i \leq b_i$ for all $i \in [s]$.

Definition 1.3.10. Let $\delta > 0$. A subset $\tau \subseteq [0,1]^s \times [0,1]^s$ is called a *δ-bracketing cover* if for every $\boldsymbol{x} \in [0,1]^s$ there exists $(\boldsymbol{v}, \boldsymbol{w}) \in \tau$ such that $\boldsymbol{v} \leq \boldsymbol{x} \leq \boldsymbol{w}$ and $\lambda([\boldsymbol{0}, \boldsymbol{w}) \backslash [\boldsymbol{0}, \boldsymbol{v})) \leq \delta$.

The following result about the number of elements of a δ-bracketing cover is due to Gnewuch:

Lemma 1.3.11 (Gnewuch [36, Theorem 1.15]). *For any $s \in \mathbb{N}$ and any $\delta > 0$ there exists a δ-bracketing cover τ with*

$$|\tau| \leq \tfrac{1}{2}(2\mathrm{e})^s(\delta^{-1} + 1)^s\,.$$

From this result Löbbe deduced the following corollary:

Corollary 1.3.12 (Löbbe [75, Corollary 2.3]). *Let $s, h \in \mathbb{N}$ and $q \geq 2$, then there exists a q^{-h}-bracketing cover τ_h with*

1. *$|\tau_h| \leq \frac{1}{2}(2\mathrm{e})^s(q^{h+2} + 1)^s$, and*

2. *for $(\boldsymbol{v}, \boldsymbol{w}) \in \tau_h$ and every $i \in [s]$ there exist $a_i \in \{0, 1, \ldots, q^{h+1+\lceil \log_q s \rceil}\}$ and $b_i \in \{0, 1, \ldots, q^{h+2+\lceil \log_q s \rceil}\}$ such that*

$$v_i = \frac{a_i}{q^{h+1+\lceil \log_q s \rceil}} \quad \text{and} \quad w_i = \frac{w_i}{q^{h+2+\lceil \log_q s \rceil}}.$$

Preliminaries

Let $N, s \in \mathbb{N}$ and fix some $H \in \mathbb{N}$. For $h \in \{1, \ldots, H\}$ let τ_h be a q^{-h}-bracketing cover of $[0,1)^s$ with elements described as in Corollary 1.3.12. Let $\boldsymbol{y} \in [0,1)^s$. We are going to define inductively a finite sequence of points $\boldsymbol{\beta}_h(\boldsymbol{y}) \in [0,1)^s$ for $h \in \{0, \ldots, H+1\}$ in the following way:

1. Let $\boldsymbol{\beta}_H(\boldsymbol{y}), \boldsymbol{\beta}_{H+1}(\boldsymbol{y}) \in [0,1)^s$ with $\boldsymbol{\beta}_H(\boldsymbol{y}) \leq \boldsymbol{y} \leq \boldsymbol{\beta}_{H+1}(\boldsymbol{y})$ and the tuple $(\boldsymbol{\beta}_H(\boldsymbol{y}), \boldsymbol{\beta}_{H+1}(\boldsymbol{y})) \in \tau_H$.
2. For $h \in \{1, \ldots, H-1\}$ let $\boldsymbol{\beta}_h(\boldsymbol{y}) \in [0,1)^s$ be such that there exists a point $\boldsymbol{w} \in [0,1)^s$ with $\boldsymbol{\beta}_h(\boldsymbol{y}) \leq \boldsymbol{\beta}_{h+1}(\boldsymbol{y}) \leq \boldsymbol{w}$ and $(\boldsymbol{\beta}_h(\boldsymbol{y}), \boldsymbol{w}) \in \tau_h$.
3. Set $\boldsymbol{\beta}_0(\boldsymbol{y}) = \mathbf{0} = (0, \ldots, 0)$, the s-dimensional zero-vector.
4. Additionally we choose the points $\boldsymbol{\beta}_h$ such that the following property is fulfilled. For $\boldsymbol{x}, \boldsymbol{y} \in [0,1)^s$ and $h \in \{0, \ldots, H-1\}$ we have that

$$\boldsymbol{\beta}_{h+1}(\boldsymbol{y}) = \boldsymbol{\beta}_{h+1}(\boldsymbol{x}) \Rightarrow \boldsymbol{\beta}_h(\boldsymbol{y}) = \boldsymbol{\beta}_h(\boldsymbol{x}).$$

An illustration of a possible configuration in two dimensions of the points $\boldsymbol{\beta}_0, \boldsymbol{\beta}_1(\boldsymbol{y}), \ldots, \boldsymbol{\beta}_{H+1}(\boldsymbol{y})$ is given in Figure 1.1. Observe that the hatched areas have measure at most q^{-h} and q^{-H}, respectively due to the construction procedure described above.

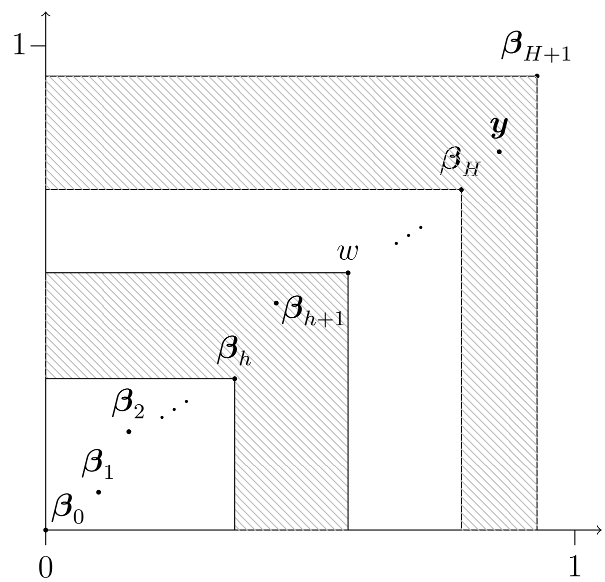

Figure 1.1: Illustration of the points $\boldsymbol{\beta}_i(\boldsymbol{y}) = \boldsymbol{\beta}_i$ for $i \in \{0, \ldots, H+1\}$ and $s = 2$.

Note that the sequence of points $\boldsymbol{\beta}_h(\boldsymbol{y})$ is well defined for $h \in \{0, \ldots, H+1\}$ since we choose τ_h to be a q^{-h}-bracketing cover. For $\boldsymbol{y} \in [0,1)^s$ we observe the following properties for the finite sequence $\boldsymbol{\beta}_h(\boldsymbol{y})$:
We have

1. $\mathbf{0} = \boldsymbol{\beta}_0(\boldsymbol{y}) \leq \boldsymbol{\beta}_1(\boldsymbol{y}) \leq \cdots \leq \boldsymbol{\beta}_H(\boldsymbol{y}) \leq \boldsymbol{y} \leq \boldsymbol{\beta}_{H+1}(\boldsymbol{y}) \leq \mathbf{1}$;
2. for all $h \in \{0, \ldots, H-1\}$ there exists $\boldsymbol{w} \in [0,1)^s$ such that $\boldsymbol{\beta}_h(\boldsymbol{y}) \leq \boldsymbol{\beta}_{h+1}(\boldsymbol{y}) \leq \boldsymbol{w}$ and $(\boldsymbol{\beta}_h(\boldsymbol{y}), \boldsymbol{w}) \in \tau_h$. Additionally we have that $(\boldsymbol{\beta}_H(\boldsymbol{y}), \boldsymbol{\beta}_{H+1}(\boldsymbol{y})) \in \tau_H$;
3. for all $h \in \{0, \ldots, H\}$ and $i \in [s]$ we have that

$$(\boldsymbol{\beta}_h(\boldsymbol{y}))_i = q^{-(h+1+\lceil \log_q s \rceil)} a_{h,i}$$

and

$$(\boldsymbol{\beta}_{H+1}(\boldsymbol{y}))_i = q^{-(H+2+\lceil \log_q s \rceil)} b_{H+1,i}$$

for $a_{h,i} \in \left\{0, 1, \ldots, q^{h+1+\lceil \log_q s \rceil}\right\}$ and $b_{H+1,i} \in \left\{0, 1, \ldots, q^{H+2+\lceil \log_q s \rceil}\right\}$.

The properties 1. and 2. are an immediate consequence of the definition of the $\boldsymbol{\beta}_h(\boldsymbol{y})$ and property 3. follows directly from Corollary 1.3.12.

Moreover, for $h \in \{0, \ldots, H\}$ we define

$$K_h(\boldsymbol{y}) := [\mathbf{0}, \boldsymbol{\beta}_{h+1}(\boldsymbol{y})) \backslash [\mathbf{0}, \boldsymbol{\beta}_h(\boldsymbol{y})) \tag{1.54}$$

and observe that the $K_h(\boldsymbol{y})$ are pairwise disjoint sets. By the definition respectively property 2. of $\boldsymbol{\beta}_h(\boldsymbol{y})$ we obtain

$$\bigcup_{h=0}^{H-1} K_h(\boldsymbol{y}) \subseteq [\mathbf{0}, \boldsymbol{y}) \subseteq \bigcup_{h=0}^{H} K_h(\boldsymbol{y}) \quad \text{and} \quad \lambda(K_h(\boldsymbol{y})) \leq q^{-h}. \tag{1.55}$$

An illustration of the points $\boldsymbol{\beta}_0(\boldsymbol{y}), \ldots, \boldsymbol{\beta}_{H+1}(\boldsymbol{y})$ together with the corresponding sets $K_h(\boldsymbol{y})$ for $h \in \{0, \ldots, H\}$ in two dimensions can be found in Figure 1.2.

Finally for $\{0, \ldots, H\}$ define $S_h := \{K_h(\boldsymbol{y}) \ : \ \boldsymbol{y} \in [0,1)^s\}$. Note that by definition of the $\boldsymbol{\beta}_h(\boldsymbol{y})$ and Corollary 1.3.12 we have

$$|S_H| = \left|\left\{(\boldsymbol{\beta}_H(\boldsymbol{y}), \boldsymbol{\beta}_{H+1}(\boldsymbol{y})) \ : \ \boldsymbol{y} \in [0,1)^s\right\}\right| \leq |\tau_H| \leq \frac{1}{2}(2\mathrm{e})^s (q^{H+2}+1)^s.$$

With point 4. in the definition of the $\boldsymbol{\beta}_h(\boldsymbol{y})$ we get for $h \in \{0, \ldots, H-1\}$ that

$$|S_h| = \left|\left\{\boldsymbol{\beta}_{h+1}(\boldsymbol{y}) \ : \ \boldsymbol{y} \in [0,1)^s\right\}\right| \leq |\tau_{h+1}| \leq \frac{1}{2}(2e)^s (q^{h+3}+1)^s.$$

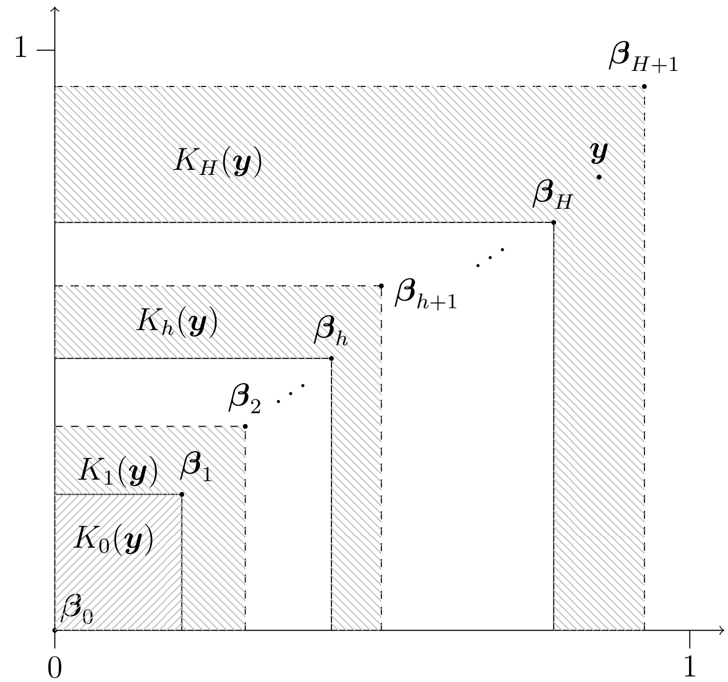

Figure 1.2: Illustration of $K_0(\boldsymbol{y}), \ldots, K_H(\boldsymbol{y})$ for $s = 2$.

Fix $\boldsymbol{y} \in [0, 1)^s$. In order to simplify the notation from now on we will write $\boldsymbol{\beta}_h$ and K_h instead of $\boldsymbol{\beta}_h(\boldsymbol{y})$ and $K_h(\boldsymbol{y})$, respectively . Then by (1.55) we get that

$$\sum_{n=1}^{N} \mathbb{1}_{[\mathbf{0},\boldsymbol{y})}(\boldsymbol{x}_n) \geq \sum_{n=1}^{N} \mathbb{1}_{[\mathbf{0},\boldsymbol{\beta}_H)}(\boldsymbol{x}_n) = \sum_{h=0}^{H-1} \sum_{n=1}^{N} \Big(\mathbb{1}_{K_h}(\boldsymbol{x}_n) - \lambda(K_h) \Big) + N \sum_{h=0}^{H-1} \lambda(K_h) \tag{1.56}$$

and

$$\sum_{n=1}^{N} \mathbb{1}_{[\mathbf{0},\boldsymbol{y})}(\boldsymbol{x}_n) \leq \sum_{n=1}^{N} \mathbb{1}_{[\mathbf{0},\boldsymbol{\beta}_{H+1})}(\boldsymbol{x}_n) = \sum_{h=0}^{H} \sum_{n=1}^{N} \Big(\mathbb{1}_{K_h}(\boldsymbol{x}_n) - \lambda(K_h) \Big) + N \sum_{h=0}^{H} \lambda(K_h). \tag{1.57}$$

Let us define the functions $\Delta_{K_h} : [0, 1)^s \to [-1, 1]$, $\Delta_{K_h}(\boldsymbol{x}) := \mathbb{1}_{K_h}(\boldsymbol{x}) - \lambda(K_h)$ for $h \in \{0, \ldots, H\}$. A crucial step for the proof of the main result will be to use Bernstein's inequality to give a lower bound on the probability that the inequality

$$\left|\sum_{n=1}^{N} \Delta_{K_h}(\boldsymbol{x}_n)\right| \leq t_h$$

holds simultaneously for all $h \in \{0, \ldots, H\}$ and for some $t_h > 0$ to be specified later. First of all observe that $\mathbb{E}(\Delta_{K_h}(\boldsymbol{x}_n)) = 0$, $\mathbb{E}(\Delta_{K_h}(\boldsymbol{x}_n)^2) = \lambda(K_h)(1 - \lambda(K_h))$ and $|\Delta_{K_h}(\boldsymbol{x}_n)| \leq 1$ for all $h \in \{0, \ldots, H\}$ and $n \in \{1, \ldots, N\}$. Unfortunately for $h \in \{0, \ldots, H\}$ the random variables $\Delta_{K_h}(\boldsymbol{x}_1), \Delta_{K_h}(\boldsymbol{x}_2), \ldots, \Delta_{K_h}(\boldsymbol{x}_N)$ are not independent in general. We will see how to overcome this problem in the next section.

Independence of $\Delta_{K_h}(\boldsymbol{x}_n)$

Before we begin we point out the following easy algebraic characterization of Laurent series whose image under ϕ belongs to a certain type of intervals: for $p \in \overline{\mathbb{F}}_q^*((t^{-1}))$ of the form $p = p_1 t^{-1} + p_2 t^{-2} + p_3 t^{-3} + \cdots$, for $r \in \mathbb{N}$ and $k \in \{0, \ldots, q^r - 1\}$ with q-adic expansion $k = k_0 + k_1 q + \cdots + k_{r-1} q^{r-1}$ we have that

$$\phi(p) \in \left[\frac{k}{q^r}, \frac{k+1}{q^r}\right) \Leftrightarrow p_1 = k_{r-1},\ p_2 = k_{r-2}, \ldots, p_r = k_0.$$

Throughout the proof the underlying probability measure is the measure μ_s from Definition 1.3.5. However, out of habit we will in the following denote the probability by $\mathbb{P}$.

Lemma 1.3.13. *Let* $\kappa_h := \log_2(h + 2 + \lceil \log_q s \rceil)$ *and let* $\gamma \in \{0, \ldots, 2^{\kappa_h} - 1\}$. *Moreover, let*

$$Q(N, \kappa_h, \gamma) := \{n \in \{1, \ldots, N\} \ : \ n \equiv \gamma \pmod{2^{\kappa_h}}\}.$$

Then for $n_1, \ldots, n_l \in Q(N, \kappa_h, \gamma)$ *and* $l \in \{1, \ldots, |Q(N, \kappa_h, \gamma)|\}$ *the random variables* $\Delta_{K_h}(\boldsymbol{x}_{n_1}), \Delta_{K_h}(\boldsymbol{x}_{n_2}), \ldots, \Delta_{K_h}(\boldsymbol{x}_{n_l})$ *are independent, i.e.*

$$\mathbb{P}(\Delta_{K_h}(\boldsymbol{x}_{n_1}) = c_1, \ldots, \Delta_{K_h}(\boldsymbol{x}_{n_l}) = c_l) = \prod_{r=1}^{l} \mathbb{P}(\Delta_{K_h}(\boldsymbol{x}_{n_r}) = c_r).$$

Proof. The proof is based on the ideas from [75]. We will show the case $l = 2$. The general case follows by induction. Let $h \in \{0, \ldots, H\}$, $\gamma \in \{0, \ldots, 2^{\kappa_h} - 1\}$

and $n, m \in Q(N, \kappa_h, \gamma)$ with $n > m$. We want to show that $\Delta_{K_h}(\boldsymbol{x}_n), \Delta_{K_h}(\boldsymbol{x}_m)$ are independent. To this end we consider the following decomposition of $[0,1)^s$:

$$\Sigma_{n-1} := \left\{ \prod_{i=1}^{s} \left[\frac{a_i}{q^{n-1}}, \frac{a_i+1}{q^{n-1}} \right) \ : \ a_i \in \{0, \ldots, q^{n-1}-1\} \right\}.$$

Since the underlying structure of the sequence $(\boldsymbol{x}_k)_{k\geq 1}$ is $\overline{\mathbb{F}}_q((t^{-1}))$ we are considering the preimage of Σ_{n-1}.

$$\Lambda_{n-1} := \left\{ \phi^{-1}(S) \cap \left(\overline{\mathbb{F}}_q^*((t^{-1})) \right)^s \ : \ S \in \Sigma_{n-1} \right\},$$

where ϕ is given in (1.26). For $A = (a_{i,j})_{i=1,j=1}^{s,n-1} \in \mathbb{F}_q^{s\times(n-1)}$ let us define

$$B_A := \prod_{i=1}^{s} \left\{ g \in \overline{\mathbb{F}}_q^*((t^{-1})) \ : \ (g_1, \ldots, g_{n-1}) = (a_{i,1}, \ldots, a_{i,n-1}) \right\},$$

where $(a_{i,1}, \ldots, a_{i,n-1})$ is the i-th row of A. One can easily check that

$$\Lambda_{n-1} = \left\{ B_A \ : \ A \in \mathbb{F}_q^{s\times(n-1)} \right\}.$$

For matrices $A_1, A_2 \in \mathbb{F}_q^{s\times(n-1)}$, $A_j = (a_{j,i,k})_{i=1,k=1}^{s,n-1}$ for $j \in \{1,2\}$ we define

$$\alpha_{A_1,A_2} : (\overline{\mathbb{F}}_q((t^{-1})))^s \to (\overline{\mathbb{F}}_q((t^{-1})))^s, \tag{1.58}$$

$$(g^{(1)}, \ldots, g^{(s)}) \mapsto (g^{(1)} + u^{(1)}_{A_1A_2}, \ldots, g^{(s)} + u^{(s)}_{A_1A_2}), \tag{1.59}$$

where for $i \in [s]$, $u^{(i)}_{A_1A_2} = \sum_{k=1}^{\infty} u^{(i)}_{A_1A_2,k} t^{-k} \in \overline{\mathbb{F}}_q((t^{-1}))$ and

$$u^{(i)}_{A_1A_2,k} = \begin{cases} a_{2,i,k} - a_{1,i,k} & \text{if } 1 \leq k \leq n-1, \\ 0 & \text{if } k > n-1. \end{cases}$$

With this definition we have

$$\alpha_{A_1,A_2}(B_{A_1}) = B_{A_2}.$$

Before we can prove the independence of $\Delta_{K_h}(\boldsymbol{x}_n)$ and $\Delta_{K_h}(\boldsymbol{x}_m)$ we need to show four claims:

Claim 1.3.14. Let $c \in \mathbb{R}$, $A_1, A_2 \in \mathbb{F}_q^{s\times(n-1)}$, $\boldsymbol{f} \in B_{A_1}$ and $(\boldsymbol{y}_n)_{n\geq 1}$ in $[0,1)^s$ with $\boldsymbol{y}_n = \phi(\{t^{n-1}\overline{\boldsymbol{f}}\})$ with $\overline{\boldsymbol{f}} = \alpha_{A_1A_2}(\boldsymbol{f})$. Then we have that

$$\mathbb{P}\big(\Delta_{K_h}(\boldsymbol{x}_n) = c \mid \boldsymbol{f} \in B_{A_1}) = \mathbb{P}(\Delta_{K_h}(\boldsymbol{y}_n) = c \mid \overline{\boldsymbol{f}} \in B_{A_2}\big).$$

Proof of Claim 1.3.14: For $i \in [s]$ we have that

$$y_{n,i} = \phi(\{t^{n-1}\overline{f}^{(i)}\}) = \phi(\{t^{n-1}f^{(i)} + t^{n-1}u^{(i)}_{A_1A_2})\}) = \phi(\{t^{n-1}f^{(i)}\}) = x_{n,i}.$$

Note that the second last equality is true because $u^{(i)}_{A_1A_2,k} = 0$ for $k \geq n$. Additionally it holds that $\boldsymbol{f} \in B_{A_1} \Leftrightarrow \overline{\boldsymbol{f}} \in B_{A_2}$. Therefore the claim follows. □

Claim 1.3.15. Let $h \in \{0, \ldots, H\}$ and $\boldsymbol{p} = (p^{(1)}, \ldots, p^{(s)}) \in (\overline{\mathbb{F}}^*_q((t^{-1})))^s$ with $p^{(i)} = \sum_{j=1}^{\infty} p^{(i)}_j t^{-j}$. Then the $s(h+2+\lceil \log_q s \rceil)$ coefficients $p^{(i)}_1, \ldots, p^{(i)}_{h+2+\lceil \log_q s \rceil}$ for $i \in [s]$ determine if $\phi(\boldsymbol{p}) \in K_h$.

Proof of Claim 1.3.15: For $\boldsymbol{p} = (p^{(1)}, \ldots, p^{(s)}) \in (\overline{\mathbb{F}}^*_q((t^{-1})))^s$, we have that

$$\begin{aligned}\phi(\boldsymbol{p}) \in K_h &\Leftrightarrow \phi(\boldsymbol{p}) \in [\boldsymbol{0}, \boldsymbol{\beta}_{h+1}) \backslash [\boldsymbol{0}, \boldsymbol{\beta}_h) \\ &\Leftrightarrow \forall i \in [s] : \phi(p^{(i)}) \in [0, \beta^{(i)}_{h+1}) \text{ and} \\ &\qquad \exists j \in [s] : \phi(p^{(j)}) \in [\beta^{(j)}_h, 1),\end{aligned} \tag{1.60}$$

where $\boldsymbol{\beta}_{h+1} = (\beta^{(1)}_{h+1}, \ldots \beta^{(s)}_{h+1})$ with

$$\beta^{(i)}_{h+1} = \frac{b_i}{q^{h+2+\lceil \log_q s \rceil}} \quad \text{for some} \quad b_i \in \{0, 1, \ldots, q^{h+2+\lceil \log_q s \rceil} - 1\}$$

and similarly $\boldsymbol{\beta}_h = (\beta^{(1)}_h, \ldots, \beta^{(s)}_h)$ with

$$\beta^{(i)}_h = \frac{\bar{b}_i}{q^{h+1+\lceil \log_q s \rceil}} \quad \text{for some} \quad \bar{b}_i \in \{0, 1 \ldots, q^{h+1+\lceil \log_q s \rceil} - 1\}.$$

We can write

$$[0, \beta^{(i)}_{h+1}) = \bigcup_{k=0}^{b_i - 1} \left[\frac{k}{q^{h+2+\lceil \log_q s \rceil}}, \frac{k+1}{q^{h+2+\lceil \log_q s \rceil}}\right).$$

Hence $\phi(p^{(i)}) \in [0, \beta^{(i)}_{h+1})$ if and only if there exists a $k \in \{0, \ldots, b_i - 1\}$ such that

$$\phi(p^{(i)}) \in \left[\frac{k}{q^{h+2+\lceil \log_q s \rceil}}, \frac{k+1}{q^{h+2+\lceil \log_q s \rceil}}\right).$$

Since $\phi(p^{(i)}) = \sum_{j=1}^{\infty} p^{(i)}_j q^{-j}$ the last condition is satisfied if and only if

$$p^{(i)}_1 = k_{h+1+\lceil \log_q s \rceil}, p^{(i)}_2 = k_{h+\lceil \log_q s \rceil}, \ldots, p^{(i)}_{h+2+\lceil \log_q s \rceil} = k_0,$$

whenever k has q-adic expansion $k = k_0 + k_1 q + \cdots + k_{h+1+\lceil \log_q s \rceil} q^{h+1+\lceil \log_q s \rceil}$.

In the same vein we can write

$$[\beta_h^{(j)}, 1) = \bigcup_{\ell=\bar{b}_j}^{q^{h+1+\lceil \log_q s \rceil}-1} \left[\frac{\ell}{q^{h+1+\lceil \log_q s \rceil}}, \frac{\ell+1}{q^{h+1+\lceil \log_q s \rceil}} \right).$$

Hence $\phi(p^{(j)}) \in [\beta_h^{(j)}, 1)$ if and only if there exists a $\ell \in \{\bar{b}_j, \ldots, q^{h+1+\lceil \log_q s \rceil} - 1\}$ such that

$$\phi(p^{(j)}) \in \left[\frac{\ell}{q^{h+1+\lceil \log_q s \rceil}}, \frac{\ell+1}{q^{h+1+\lceil \log_q s \rceil}} \right).$$

Since $\phi(p^{(j)}) = \sum_{k=1}^{\infty} p_k^{(j)} q^{-k}$ the last condition is satisfied if and only if

$$p_1^{(j)} = l_{h+\lceil \log_q s \rceil}, p_2^{(j)} = l_{h+\lceil \log_q s \rceil - 1}, \ldots, p_{h+1+\lceil \log_q s \rceil}^{(j)} = l_0,$$

whenever ℓ has q-adic expansion $\ell = l_0 + l_1 q + \cdots + l_{h+\lceil \log_q s \rceil} q^{h+\lceil \log_q s \rceil}$.

Together with (1.60) it follows that the coefficients $p_1^{(i)}, \ldots, p_{h+2+\lceil \log_q s \rceil}^{(i)}$ for $i \in [s]$ determine whether or not $\phi(\boldsymbol{p})$ belongs to K_h. This proves the second claim. $\square$

Recall that $m \in Q(N, \kappa_h, \gamma)$ and $m < n$. Define

$$\delta_m : \overline{\mathbb{F}}_q((t^{-1})) \to \overline{\mathbb{F}}_q((t^{-1})),\ p \mapsto \{t^{m-1} p\}.$$

Claim 1.3.16. For all $h \in \{0, \ldots, H\}$ and for all $A \in \mathbb{F}_q^{s \times (n-1)}$ we have that Δ_{K_h} is constant on $\phi(\delta_m(B_A))$.

Proof of Claim 1.3.16: Let $\boldsymbol{p} = (p^{(1)}, \ldots, p^{(s)}) \in B_A$ with $p^{(i)} = \sum_{j=1}^{\infty} p_j^{(i)} t^{-j}$. Note that for each $i \in [s]$ the first $n-1$ coefficients $p_1^{(i)}, \ldots, p_{n-1}^{(i)}$ of $p^{(i)}$ are equal to the entries in the i-th row of A. Now we have

$$\delta_m \left(\sum_{j=1}^{\infty} p_j^{(i)} t^{-j} \right) = \left\{ \sum_{j=1}^{\infty} p_j^{(i)} t^{m-1-j} \right\} = \sum_{j=1}^{\infty} p_{m-1+j}^{(i)} t^{-j}.$$

Because of Claim 1.3.15, the coefficients $p_m^{(i)}, \ldots, p_{m+h+1+\lceil \log_q s \rceil}^{(i)}$ for $i \in [s]$ determine if $\phi(\delta_m(\boldsymbol{p})) \in K_h$. Since $n - m \geq h + 2 + \lceil \log_q s \rceil$ these coefficients are fixed by the choice of B_A. Hence it follows that $\phi(\delta_m(B_A)) \cap K_h \in \{\emptyset, \phi(\delta_m(B_A))\}$. Therefore the function $\Delta_{K_h}(\boldsymbol{x}) = \mathbb{1}_{K_h}(\boldsymbol{x}) - \lambda(K_h)$ is constant on $\phi(\delta_m(B_A))$. This proves the claim. $\square$

Define for $c \in \mathbb{R}$,

$$\Lambda_{K_h,c} := \{B_A \in \Lambda_{n-1} \ : \ \Delta_{K_h}(\phi(\delta_m(B_A))) = c\}.$$

Note that $\Lambda_{K_h,c}$ is well-defined according to Claim 1.3.16.

Claim 1.3.17. Let $c \in \mathbb{R}$ and $h \in \{0, \dots, H\}$. Then we have

$$\Delta_{K_h}(\boldsymbol{x}_m) = c \Leftrightarrow \exists B_A \in \Lambda_{K_h,c} \ \text{ such that } \ \boldsymbol{f} \in B_A.$$

Proof of Claim 1.3.17: Let $c \in \mathbb{R}$ and suppose that there exists $B_A \in \Lambda_{K_h,c}$ such that $\boldsymbol{f} \in B_A$. Since $\boldsymbol{x}_m = \phi(\delta_m(\boldsymbol{f}))$ we have

$$\Delta_{K_h}(\boldsymbol{x}_m) = \Delta_{K_h}(\phi(\delta_m(\boldsymbol{f}))).$$

Since $\delta_m(\boldsymbol{f}) \in \delta_m(B_A)$ we get that $\Delta_{K_h}(\boldsymbol{x}_m) = c$. Now assume that $\Delta_{K_h}(\boldsymbol{x}_m) = c$ which is equivalent to $\Delta_{K_h}(\phi(\delta_m(\boldsymbol{f}))) = c$. Now there exists $A' \in \mathbb{F}_q^{s\times(n-1)}$ such that $B_{A'} \in \Lambda_{K_h,c}$ and $\delta_m(\boldsymbol{f}) \in \delta_m(B_{A'})$ and we get that $(f_{m+1}^{(i)}, \dots, f_{n-1}^{(i)}) = (a'_{i,m+1}, \dots, a'_{i,n-1})$ for $i \in [s]$.
On the other hand there exists $A \in \mathbb{F}_q^{s\times(n-1)}$ with $\boldsymbol{f} \in B_A$. We obtain that $(f_{m+1}^{(i)}, \dots, f_{n-1}^{(i)}) = (a_{i,m+1}, \dots, a_{i,n-1})$ for $i \in [s]$. Altogether we have that

$$(a'_{i,m+1}, \dots, a'_{i,n-1}) = (a_{i,m+1}, \dots, a_{i,n-1}) \quad \text{for} \quad i \in [s].$$

Now it follows by Claim 1.3.15 that

$$\Delta_{K_h}(\phi(\delta_m(B_A))) = \Delta_{K_h}(\phi(\delta_m(B_{A'}))) = c.$$

This means that $B_A \in \Lambda_{K_h,c}$ and $\boldsymbol{f} \in B_A$ and this proves the claim. $\square$

Now we can prove the independence of $\Delta_{K_h}(\boldsymbol{x}_n)$ and $\Delta_{K_h}(\boldsymbol{x}_m)$ for $n, m \in Q(N, \kappa_h, \gamma)$ with $n > m$. For $c \in \mathbb{R}$ and $B_{A'} \in \Lambda_{n-1}$ we have

$$\begin{aligned}
\mathbb{P}(\Delta_{K_h}(\boldsymbol{x}_n) = c) &= \sum_{B_A \in \Lambda_{n-1}} \mathbb{P}(\Delta_{K_h}(\boldsymbol{x}_n) = c \mid \boldsymbol{f} \in B_A)\mathbb{P}(\boldsymbol{f} \in B_A) \\
&= \mathbb{P}(\Delta_{K_h}(\boldsymbol{y}_n) = c \mid \overline{\boldsymbol{f}} \in B_{A'}) \sum_{B_A \in \Lambda_{n-1}} \mathbb{P}(\boldsymbol{f} \in B_A) \\
&= \mathbb{P}(\Delta_{K_h}(\boldsymbol{y}_n) = c \mid \overline{\boldsymbol{f}} \in B_{A'}), \qquad (1.61)
\end{aligned}$$

where we used Claim 1.3.14 with $A_1 = A$ and $A_2 = A'$. By Claim 1.3.17 and (1.61) we get for $c_1, c_2 \in \mathbb{R}$ and $A' \in \mathbb{F}_q^{s\times(n-1)}$ that

$$\begin{aligned}
\mathbb{P}(\Delta_{K_h}(\boldsymbol{x}_n) = & c_2 \mid \Delta_{K_h}(\boldsymbol{x}_m) = c_1) = \frac{\mathbb{P}(\Delta_{K_h}(\boldsymbol{x}_n) = c_2, \Delta_{K_h}(\boldsymbol{x}_m) = c_1)}{\mathbb{P}(\Delta_{K_h}(\boldsymbol{x}_m) = c_1)} \\
= & \frac{\sum_{B_A \in \Lambda_{n-1}} \mathbb{P}(\Delta_{K_h}(\boldsymbol{x}_n) = c_2, \Delta_{K_h}(\boldsymbol{x}_m) = c_1 \mid \boldsymbol{f} \in B_A)\mathbb{P}(\boldsymbol{f} \in B_A)}{\mathbb{P}(\Delta_{K_h}(\boldsymbol{x}_m) = c_1)} \\
= & \sum_{B_A \in \Lambda_{K_h, c_1}} \mathbb{P}(\Delta_{K_h}(\boldsymbol{x}_n) = c_2 \mid \boldsymbol{f} \in B_A) \frac{\mathbb{P}(\boldsymbol{f} \in B_A)}{\mathbb{P}(\Delta_{K_h}(\boldsymbol{x}_m) = c_1)} \\
= & \mathbb{P}(\Delta_{K_h}(\boldsymbol{y}_n) = c_2 \mid \overline{\boldsymbol{f}} \in B_{A'}) \frac{\sum_{B_A \in \Lambda_{K_h, c_1}} \mathbb{P}(\boldsymbol{f} \in B_A)}{\mathbb{P}(\Delta_{K_h}(\boldsymbol{x}_m) = c_1)} \\
= & \mathbb{P}(\Delta_{K_h}(\boldsymbol{y}_n) = c_2 \mid \overline{\boldsymbol{f}} \in B_{A'}) \\
= & \mathbb{P}(\Delta_{K_h}(\boldsymbol{x}_n) = c_2).
\end{aligned} \tag{1.62}$$

This implies the desired result. $\square$

Applying Bernstein's inequality and finalizing the proof of Theorem 1.3.7

We may assume that $N \geq s \log_q s$ since otherwise the discrepancy bound is trivial. First of all we set

$$H = \left\lceil \frac{1}{2} \log_q \left(\frac{N}{s \log_q s} \right) \right\rceil \in \mathbb{N}. \tag{1.63}$$

With this choice we obtain

$$\frac{1}{q^H} \leq \sqrt{\frac{s \log_q s}{N}} \quad \text{and} \quad q^{2H} \leq q^2 \frac{N}{s \log_q s}.$$

Recall the definition of $Q(N, \kappa_h, \gamma) = \{n \in \{1, \ldots, N\} : n \equiv \gamma \mod 2^{\kappa_h}\}$ and note that $Q(N, \kappa_h, \gamma)$ for $\gamma \in \{0, \ldots, 2^{\kappa_h} - 1\}$ are a partition of $\{1, \ldots, N\}$ and

$$|Q(N, \kappa_h, \gamma)| \leq \left\lfloor \frac{N}{2^{\kappa_h}} \right\rfloor + \xi \quad \text{for some } \xi \in \{0, 1\}.$$

With the help of Lemma 1.3.13 we are able to apply Bernstein's inequality (see Lemma 1.3.9). For $h \in \{0, \ldots, H\}$ we get that

$$\mathbb{P}\left(\left| \sum_{n=1}^{N} \Delta_{K_h}(\boldsymbol{x}_n) \right| > t_h \right) \leq \sum_{\gamma=0}^{2^{\kappa_h}-1} \mathbb{P} \left(\left| \sum_{n \in Q(N, \kappa_h, \gamma)} \Delta_{K_h}(\boldsymbol{x}_n) \right| > \frac{t_h}{2^{\kappa_h}} \right)$$

$$
\begin{aligned}
&\le 2\sum_{\gamma=0}^{2^{\kappa_h}-1}\exp\left(-\frac{t_h^2/2^{2\kappa_h}}{2|Q(N,\kappa_h,\gamma)|\lambda(K_h)(1-\lambda(K_h))+2t_h/(3\cdot 2^{\kappa_h})}\right)\\
&\le 2^{\kappa_h+1}\exp\left(-\frac{t_h^2/2^{\kappa_h}}{2(1+2^{\kappa_h}/N)Nq^{-h}+2t_h/3}\right).
\end{aligned}
$$

Since

$$
\begin{aligned}
\frac{2^{\kappa_h}}{N} &= \frac{h+2+\lceil\log_q s\rceil}{N}\le\frac{1}{N}\left(\frac{1}{2}\log_q\left(\frac{N}{s\log_q s}\right)+4+\log_q s\right)\\
&\le \frac{1}{2}\frac{\log_q N}{N}+4+\frac{1}{s}\le 5,
\end{aligned}
$$

we obtain

$$
\mathbb{P}\left(\left|\sum_{n=1}^{N}\Delta_{K_h}(\boldsymbol{x}_n)\right|>t_h\right)\le 2^{\kappa_h+1}\exp\left(-\frac{t_h^2/2^{\kappa_h}}{12Nq^{-h}+\frac{2}{3}t_h}\right). \tag{1.64}
$$

For the choice of t_h we will distinguish two cases

$$
t_h:=\begin{cases} C_1\sqrt{Nshq^{-h}2^{\kappa_h}} & \text{if } h\in\{1,\dots,H\}\\ C_2\sqrt{Ns2^{\kappa_0}} & \text{if } h=0\end{cases} \tag{1.65}
$$

for constants $C_1, C_2>0$ to be specified later.

Let us consider first the case $h\in\{1,\dots,H\}$. By $\kappa_h=\log_2(h+2+\lceil\log_q s\rceil)$ we get that

$$
\begin{aligned}
2^{\kappa_h}q^h h \le& 2^{\kappa_H}q^H H\le H^2q^H(4+\log_q s)\le 2q^{2H}(4+\log_q s)\\
\le& 2q^2\frac{N}{s}\left(1+\frac{4}{\log_q s}\right)\le c(q)\frac{N}{s},
\end{aligned}
$$

where $c(q)=2q^2\left(1+\frac{4}{\log_q 2}\right)$. Thus we obtain for $h\in\{1,\dots,H\}$

$$
t_h=C_1\sqrt{Nshq^{-h}2^{\kappa_h}}\le C_1\sqrt{c(q)}q^{-h}N.
$$

Furthermore we get

$$
\frac{t_h^2/2^{\kappa_h}}{12q^{-h}N+\frac{2}{3}t_h}\ge\frac{C_1^2Nshq^{-h}}{12q^{-h}N+\frac{2}{3}C_1\sqrt{c(q)}q^{-h}N}=\frac{C_1^2sh}{12+\frac{2}{3}C_1\sqrt{c(q)}}. \tag{1.66}
$$

Combining (1.64) and (1.66) we get

$$\mathbb{P}\left(\left|\sum_{n=1}^{N}\Delta_{K_h}(\boldsymbol{x}_n)\right| > C_1\sqrt{Nhq^{-h}2^{\kappa_h}s}\right) \leq 2\exp\left(\kappa_h\log 2 - \frac{C_1^2}{12+\frac{2}{3}C_1\sqrt{c(q)}}\ sh\right). \tag{1.67}$$

Consider the case $h=0$, i.e. $t_0 = C_2\sqrt{N2^{\kappa_0}s}$. We have

$$2^{\kappa_0}s \leq (3+\log_q s)s \leq Nc(q).$$

After continuing with the same steps as in the first case we end up with

$$\mathbb{P}\left(\left|\sum_{n=1}^{N}\Delta_{K_0}(\boldsymbol{x}_n)\right| > C_2\sqrt{N2^{\kappa_0}s}\right) \leq 2\exp\left(\kappa_0\log 2 - \frac{C_2^2}{12+\frac{2}{3}C_2\sqrt{c(q)}}\ s\right). \tag{1.68}$$

Recall that $\boldsymbol{\beta}_h$ and K_h are dependent on a point $\boldsymbol{y}\in[0,1)^s$, respectively. Moreover, we defined $S_h = \{K_h(\boldsymbol{y}) \ :\ \boldsymbol{y}\in[0,1)^s\}$ with $|S_h| \leq \frac{1}{2}(2\mathrm{e})^s(q^{h+3}+1)^s$. Additionally we define

$$A_{K_h,N,s} := \left\{\boldsymbol{f}\in(\overline{\mathbb{F}}_q^*((t^{-1})))^s \ :\ \left|\sum_{n=1}^{N}\Delta_{K_h}(\boldsymbol{x}_n)\right| > t_h\right\}$$

with t_h defined as in (1.65) and set

$$C_3 := \frac{C_1^2}{12+\frac{2}{3}C_1\sqrt{c(q)}}, \qquad \text{and} \qquad C_4 := \frac{C_2^2}{12+\frac{2}{3}C_2\sqrt{c(q)}}. \tag{1.69}$$

Then with (1.67) and (1.68) we have

$$\begin{aligned}\mathbb{P}\left(\bigcap_{h=0}^{H}\bigcap_{K_h\in S_h}\left\{\left|\sum_{n=1}^{N}\Delta_{K_h}(\boldsymbol{x}_n)\right|\leq t_h\right\}\right) &= 1-\mathbb{P}\left(\bigcup_{h=0}^{H}\bigcup_{K_h\in S_h}A_{K_h,N,s}\right)\\ &\geq\ 1-\sum_{K_0\in S_0}\mathbb{P}(A_{K_0,N,s}) - \sum_{h=1}^{H}\sum_{K_h\in S_h}\mathbb{P}(A_{K_h,N,s})\\ &\geq\ 1-|S_0|2\mathrm{e}^{\kappa_0\log 2 - C_4 s} - \sum_{h=1}^{H}|S_h|2\mathrm{e}^{\kappa_h\log 2 - C_3 sh}\end{aligned}$$

$$\geq \quad 1-(2q^3+2)^s \mathrm{e}^{s(1-C_4)+\kappa_0 \log 2} - \sum_{h=1}^{H} (2q^{h+3}+2)^s \mathrm{e}^{s(1-C_3 h)+\kappa_h \log 2}. \tag{1.70}$$

We will now choose $C_1 = C_1(\varepsilon)$ and $C_2 = C_2(\varepsilon)$ such that

$$(2q^3+2)^s \mathrm{e}^{s(1-C_4)+\kappa_0 \log 2} \leq \frac{\varepsilon}{2} \tag{1.71}$$

and

$$(2q^{h+3}+2)^s \mathrm{e}^{s(1-C_3 h)+\kappa_h \log 2} \leq \frac{\varepsilon}{2^{h+1}}. \tag{1.72}$$

From (1.70), (1.71) and (1.72) we then obtain that

$$\mathbb{P}\left(\bigcap_{h=0}^{H} \bigcap_{K_h \in S_h} \left\{ \left| \sum_{n=1}^{N} \Delta_{K_h}(\boldsymbol{x}_n) \right| \leq t_h \right\} \right) \geq 1-\varepsilon.$$

Inequality (1.71) is equivalent to

$$C_4 \geq \frac{1}{s} \left(s \log(2q^3+2) + s + \log(2 + \lceil \log_q s \rceil) + \log \frac{2}{\varepsilon} \right).$$

This is certainly satisfied for the choice

$$C_4 = \log(2q^3+2) + 2 + \log \frac{2}{\varepsilon} = \log \left(\frac{4(q^3+1)\mathrm{e}^2}{\varepsilon} \right).$$

With (1.69) it follows that we have to choose

$$C_2 = C_4 \frac{1}{3} \sqrt{c(q)} + \sqrt{C_4^2 \frac{1}{9} c(q) + 12 C_4} = \mathcal{O}_q \left(\log \varepsilon^{-1} \right).$$

Inequality (1.72) is equivalent to

$$C_3 \geq \frac{1}{sh} \left(s \log(2q^{h+3}+2) + s + \log(h + 2 + \lceil \log_q s \rceil) + \log 2^h + \log \frac{2}{\varepsilon} \right).$$

This is certainly satisfied for the choice

$$C_3 = \log(2(q^4+1)) + 2 + \frac{\log 2}{2} + \log \frac{2}{\varepsilon} = \log \left(\frac{4\sqrt{2}(q^4+1)\mathrm{e}^2}{\varepsilon} \right).$$

With (1.69) it follows that we have to choose

$$C_1 = C_3 \frac{1}{3}\sqrt{c(q)} + \sqrt{C_3^2 \frac{1}{9} c(q) + 12 C_3} = \mathcal{O}_q\left(\log \varepsilon^{-1}\right).$$

Finally by (1.68), (1.67), (1.57) we obtain with probability at least $1-\varepsilon$

$$\begin{aligned}
\sum_{n=1}^{N} \mathbb{1}_{[\mathbf{0},\boldsymbol{y})}(\boldsymbol{x}_n) &\leq \sum_{n=1}^{N}\sum_{h=0}^{H} \Delta_{K_h}(\boldsymbol{x}_n) + N\lambda([\mathbf{0},\boldsymbol{\beta}_{H+1}))\\
&\leq \sum_{n=1}^{N} \Delta_{K_0}(\boldsymbol{x}_n) + \sum_{h=1}^{H}\sum_{n=1}^{N} \Delta_{K_h}(\boldsymbol{x}_n) + N\left(\lambda([\mathbf{0},\boldsymbol{y})) + \lambda([\mathbf{0},\boldsymbol{\beta}_{H+1})) - \lambda([\mathbf{0},\boldsymbol{y}))\right)\\
&\leq \sqrt{Ns}\left(C_2\sqrt{2^{\kappa_0}} + \sum_{h=1}^{H} C_1\sqrt{2^{\kappa_h}}\sqrt{hq^{-h}}\right) + N(\lambda([\mathbf{0},\boldsymbol{y})) + q^{-H})\\
&\leq \sqrt{Ns}\left(C_2\sqrt{2^{\kappa_0}} + \sum_{h=1}^{\infty} C_1\sqrt{2^{\kappa_h}}\sqrt{hq^{-h}}\right) + \sqrt{Ns\log_q s} + N\lambda([\mathbf{0},\boldsymbol{y}))\\
&\leq \sqrt{Ns}\left(C_2\sqrt{2^{\kappa_0}} + \sum_{h=1}^{\infty} C_1\sqrt{(h+3)hq^{-h}} + \sqrt{\log_q s}\sum_{h=1}^{\infty} C_1\sqrt{hq^{-h}}\right)\\
&\qquad + \sqrt{Ns\log_q s} + N\lambda([\mathbf{0},\boldsymbol{y}),
\end{aligned} \tag{1.73}$$

where we used that $\lambda\left([\mathbf{0},\boldsymbol{\beta}_{H+1})\right) - \lambda\left([\mathbf{0},\boldsymbol{y})\right) \leq \lambda(K_H(y)) \leq q^{-H}$.

By the choices for C_1 and C_2 we obtain that

$$\frac{1}{N}\sum_{n=1}^{N} \mathbb{1}_{[\mathbf{0},\boldsymbol{y})}(\boldsymbol{x}_n) - \lambda([\mathbf{0},\boldsymbol{y})) \leq C_5(q,\varepsilon)\sqrt{\frac{s\log s}{N}}, \tag{1.74}$$

where $C_5(q,\varepsilon) = \mathcal{O}_q(\log \varepsilon^{-1})$.

If we use (1.56) instead of (1.57) and the fact that $\lambda\left([\mathbf{0},\boldsymbol{y})\right) - \lambda([\mathbf{0},\boldsymbol{\beta}_H)) \leq \lambda(K_H(y)) \leq q^{-H}$ we get that

$$\frac{1}{N}\sum_{n=1}^{N} \mathbb{1}_{[\mathbf{0},\boldsymbol{y})}(\boldsymbol{x}_n) - \lambda([\mathbf{0},\boldsymbol{y})) \geq -C_5(q,\varepsilon)\sqrt{\frac{s\log s}{N}} \tag{1.75}$$

with $C_5(q,\varepsilon)$ as before.

Finally (1.74) and (1.75) imply

$$\left| \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_{[\mathbf{0},\boldsymbol{y})}(\boldsymbol{x}_n) - \lambda([\mathbf{0},\boldsymbol{y}))\right| \leq C_5(q,\varepsilon)\sqrt{\frac{s\log s}{N}}.$$

Since $\boldsymbol{y} \in [0,1]^s$ was arbitrary we get that

$$D_N^*(\mathcal{P}_N(\boldsymbol{f})) \leq C_5(q,\varepsilon)\sqrt{\frac{s\log s}{N}}.$$

holds with probability at least $1-\varepsilon$. This finishes the proof. $\square$

The proof of Corollaries 1.3.2 and 1.3.8

Since the proofs of the two corollaries are very similar we only present the proof of Corollary 1.3.8.

Let $c(q) > 0$ be such that $C(q,\varepsilon)$ from Theorem 1.3.7 satisfies $C(q,\varepsilon) \leq c(q)\log \varepsilon^{-1}$. For $\delta \in (0,1)$ and $N \geq 2$ let $\varepsilon_N = 6\delta/(\pi N)^2$ and

$$A_N := \left\{ \boldsymbol{f} \in (\overline{\mathbb{F}}_q((t^{-1})))^s \;:\; D_N^*(\mathcal{P}_N(\boldsymbol{f})) \leq c(q)\log \varepsilon_N^{-1}\sqrt{\frac{s\log s}{N}} \right\}.$$

According to Theorem 1.3.7 we have $\mathbb{P}(A_N) \geq 1-\varepsilon_N$.

Set

$$A := \left\{ \boldsymbol{f} \in (\overline{\mathbb{F}}_q((t^{-1})))^s \;:\; D_N^*(\mathcal{P}_N(\boldsymbol{f})) \leq c(q)\log \varepsilon_N^{-1}\sqrt{\frac{s\log s}{N}} \;\text{ for all } N \geq 2 \right\}.$$

Then obviously $A = \bigcap_{N\geq 2} A_N$ and hence

$$\mathbb{P}(A^c) = \mathbb{P}\left(\bigcup_{N\geq 2} A_N^c\right) \leq \sum_{N\geq 2}\mathbb{P}(A_N^c) \leq \sum_{N\geq 2}\varepsilon_N \leq \delta,$$

where A^c is the complement of A in $(\overline{\mathbb{F}}_q((t^{-1})))^s$ and similarly for A_N^c. Hence $\mathbb{P}(A) \geq 1-\delta$ and the result follows. $\square$

1.3.3 Generalisation of Theorem 1.3.7

In the previous subsection we analysed the behaviour of the point set $\mathcal{P}_N(\boldsymbol{f}) = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\}$ with

$$\boldsymbol{x}_n = \phi(\{t^{n-1}\boldsymbol{f}\}),$$

where $\boldsymbol{f} \in \mathbb{F}_q((t^{-1}))^s$ and

$$\phi : \overline{\mathbb{F}}_q((t^{-1})) \to [0,1), \quad \sum_{i=1}^{\infty} g_i t^{-i} \mapsto \sum_{i=1}^{\infty} g_i q^{-i}.$$

We want to prove a slightly generalised version of Theorem 1.3.7 in this subsection. Let therefore $k \in \mathbb{N}$ and $\gamma_k(t) \in \mathbb{F}_q[t]$, $\boldsymbol{f} \in \mathbb{F}_q((t^{-1}))^s$ and consider the points

$$\boldsymbol{x}_n = \phi(\{\gamma_n \boldsymbol{f}\}) \text{ for } n \in \{1, \ldots, N\}. \tag{1.76}$$

It is the aim of this subsection to prove the subsequent theorem which is a generalised version of Theorem 1.3.7.

Theorem 1.3.18. *Let q be a prime number and let $N, s \in \mathbb{N}$ with $N, s \geq 2$. Let $\mathcal{P}_N(\boldsymbol{f}) = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\}$ be defined as in (1.76).*
If there exists a constant $c > 0$ independent of N and s such that

$$\forall k \in \mathbb{N}\ \exists \beta_k \in \mathbb{F}_q[t] : \gamma_k(t) = \beta_k(t) t^{k-1} \textit{ and } \deg(\beta_k) \leq c \log_q s \tag{1.77}$$

then for every $\varepsilon \in (0,1)$ there is a quantity $C(q,\varepsilon) > 0$ such that the star discrepancy of $\mathcal{P}_N(\boldsymbol{f})$ satisfies

$$D_N^*(\mathcal{P}_N(\boldsymbol{f})) \leq C(q,\varepsilon) \sqrt{\frac{s \log s}{N}}$$

with probability at least $1 - \varepsilon$. The quantity $C(q,\varepsilon)$ is of order $\mathcal{O}_q(\log \varepsilon^{-1})$.

Observe that for $\beta_k = 1$ we arrive at $\gamma_k(t) = t^{k-1}$ which results in the case we already studied before. For the proof of Theorem 1.3.18 we are going to use the same machinery as introduced in Subsection 1.3.2. Recall that in order to be able to apply this machinery we need to show a certain independence relation for the family of random variables $\mathbb{1}_{K_h}(\boldsymbol{x}_n)$ for $n \in \{1, \ldots, N\}$ and where K_h is defined as in (1.54). The following lemma is crucial for establishing the above-mentioned independence relation and is therefore also essential for the proof of Theorem 1.3.18. We stick to the notation of the previous subsection.

Lemma 1.3.19. *Let $l, n, m \in \mathbb{N}$, $n > m$ and let $D = \bigcup_{i=1}^{r} B_i$ and $B_i \in \Sigma_l$. If $n - 1 - \deg(\gamma_m) \geq l$ then $\mathbb{1}_D(\boldsymbol{x}_n)$ and $\mathbb{1}_D(\boldsymbol{x}_m)$ are stochastically independent.*

Proof. We follow exactly the lines and notation of the proof of Lemma 1.3.13. This means the proof is split again into 4 claims:

Claim 1.3.20. Let $c \in \mathbb{R}$, $A_1, A_2 \in \mathbb{F}_q^{s\times(n-1)}$, $\boldsymbol{f} \in B_{A_1}$ and $(\boldsymbol{y}_n)_{n\geq 1}$ in $[0,1)^s$ with $\boldsymbol{y}_n = \phi(\{\gamma_n \overline{\boldsymbol{f}}\})$ with $\overline{\boldsymbol{f}} = \alpha_{A_1A_2}(\boldsymbol{f})$. Then we have that

$$\mathbb{P}\big(\Delta_D(\boldsymbol{x}_n) = c \mid \boldsymbol{f} \in B_{A_1}\big) = \mathbb{P}(\Delta_D(\boldsymbol{y}_n) = c \mid \overline{\boldsymbol{f}} \in B_{A_2}\big) \,.$$

Proof. We have for all $i \in [s]$

$$y_n^{(i)} = \phi(\{\gamma_n \tilde{f}^{(i)}\}) = \phi(\{\gamma_n f^{(i)} + \gamma_n u_{A_1A_2}^{(i)}\}) = \phi(\{\gamma_n f^{(i)}) = x_n^{(i)}. \tag{1.78}$$

Note that we used the fact that $\gamma_n u_{A_1A_2}^{(i)} \in \mathbb{F}_q[t]$ which follows by definition of $\alpha_{A_1A_2}$ (see (1.58)) and $\gamma_n = t^{n-1}\beta_n$ for some $\beta_n \in \mathbb{F}_q[t]$. Additionally, we have that $\boldsymbol{f} \in B_{A_1} \Leftrightarrow \tilde{\boldsymbol{f}} \in B_{A_2}$ and the claim follows. □

Claim 1.3.21. Let $\boldsymbol{p} = (p^{(1)}, \ldots, p^{(s)}) \in (\overline{\mathbb{F}}_q^*((t^{-1})))^s$ with $p^{(i)} = \sum_{j=1}^{\infty} p_j^{(i)} t^{-j}$. Then the sl coefficients $p_1^{(i)}, \ldots, p_l^{(i)}$ for $i \in [s]$ determine if $\phi(\boldsymbol{p}) \in D$.

Proof. Let $\boldsymbol{p} = (p^{(1)}, \ldots, p^{(s)}) \in (\overline{\mathbb{Z}}_q^*((t^{-1})))^s$ and recall that $D = \bigcup_{i=1}^{r} B_i$ with $B_i \in \Sigma_l$ and therefore B_i is of the form $\prod_{i=1}^{s} [\frac{a_i}{q^l}, \frac{a_i+1}{q^l})$ for $a_i \in \{0, 1, \ldots, q^l - 1\}$. We already know from the proof of Claim 1.3.15 that for $k \in \{0, 1, \ldots, q^l - 1\}$ with q-adic expansion $k = \sum_{j=0}^{l-1} k_j q^j$ we have

$$\phi(p^{(i)}) \in \left[\frac{k}{q^r}, \frac{k+1}{q^r}\right) \Leftrightarrow p_1^{(i)} = k_{l-1}, p_2^{(i)} = k_{l-2}, \ldots, p_l^{(i)} = k_0.$$

Therefore it follows immediately that $p_1^{(1)}, p_2^{(1)}, \ldots, p_l^{(1)}, \ldots, p_1^{(s)}, p_2^{(s)}, \ldots, p_l^{(s)}$ determine if $\phi(\boldsymbol{p}) \in D$. □

Claim 1.3.22. For all $A \in \mathbb{F}_q^{s\times(n-1)}$ we have that Δ_D is constant on $\phi(\{\gamma_m B_A\})$.

Proof. Let $\boldsymbol{p} = (p^{(1)}, \ldots, p^{(s)}) \in B_A$ with $p^{(i)} = \sum_{j=1}^{\infty} p_j^{(i)} t^{-j}$. Note that for each $i \in \{1, \ldots, s\}$ the first $n-1$ coefficients $p_1^{(i)}, \ldots, p_{n-1}^{(i)}$ of $p^{(i)}$ are equal to the entries in the i-th row of A. By interpreting γ_m as a formal Laurent series with coefficients $\gamma_{m,u}$ we obtain by condition (1.77) that $\gamma_{m,u} = 0$ for $u > -m + 1$. Therefore we have

$$\left\{\gamma_m \sum_{j=1}^{\infty} p_j^{(i)} t^{-j}\right\} = \left\{\sum_{u=-\deg(\gamma_m)}^{-m+1} \gamma_{m,u} t^{-u} \sum_{j=1}^{\infty} p_j^{(i)} t^{-j}\right\} = \left\{\sum_{u=-\deg(\gamma_m)}^{-m+1} \sum_{j=1+u}^{\infty} \gamma_{m,u} p_{j-u}^{(i)} t^{-j}\right\}$$
$$= \left\{\sum_{j=1-\deg(\gamma_m)}^{\infty} \left(\sum_{u=-\deg(\gamma_m)}^{\min\{-m+1,j-1\}} \gamma_{m,u} p_{j-u}^{(i)}\right) t^{-j}\right\}$$
$$= \sum_{j=1}^{\infty} \left(\sum_{u=-\deg(\gamma_m)}^{-m+1} \gamma_{m,u} p_{j-u}^{(i)}\right) t^{-j} = \sum_{j=1}^{\infty} c_j^{(i)} t^{-j},$$

where we set $c_j^{(i)} := \sum_{u=-\deg(\gamma_m)}^{-m+1} \gamma_{m,u} p_{j-u}^{(i)}$. Because of Claim 1.3.21, the coefficients $c_1^{(i)}, \ldots, c_l^{(i)}$ for $i \in [s]$ determine if $\phi(\{\gamma_m \boldsymbol{p}\}) \in D$. But $c_1^{(i)}, \ldots, c_l^{(i)}$ are determined by $p_m^{(i)}, \ldots, p_{l+\deg(\gamma_m)}^{(i)}$ for $i \in [s]$. Since we have $n - 1 - \deg(\gamma_m) \geq l$ by assumption these coefficients are fixed by the choice of B_A. Hence it follows that $\phi(\{\gamma_m B_A\}) \cap D \in \{\emptyset, \phi(\{\gamma_m B_A\})\}$. Therefore the function $\Delta_D(\boldsymbol{x}) = \mathbb{1}_D(\boldsymbol{x}) - \lambda(K_h)$ is constant on $\phi(\{\gamma_m B_A\})$. This proves the claim. □

Define for $c \in \mathbb{R}$,

$$\Lambda_{D,c} := \{B_A \in \Lambda_{n-1} \ : \ \Delta_D(\phi(\{\gamma_m B_A\})) = c\}.$$

Note that $\Lambda_{D,c}$ is well-defined according to Claim 1.3.22.

Claim 1.3.23. Let $c \in \mathbb{R}$. Then we have

$$\Delta_D(\boldsymbol{x}_m) = c \Leftrightarrow \exists B_A \in \Lambda_{D,c} \ \text{ such that } \ \boldsymbol{f} \in B_A.$$

Proof. The implication from right to left follows directly from the definition of $\Lambda_{D,c}$. For the second direction let $c \in \mathbb{R}$ and assume that $\Delta_D(\boldsymbol{x}_m) = c$ which is equivalent to $\Delta_D(\phi(\{\gamma_m \boldsymbol{f}\})) = c$. Since Λ_{n-1} is a partition of $\overline{\mathbb{F}}_q((t^{-1}))$ there exists $A \in \mathbb{F}_q^{s\times(n-1)}$ such that $\boldsymbol{f} \in B_A$. By Claim 1.3.21 we get that $c = \Delta_D(\phi(\{\gamma_m \boldsymbol{f}\})) = \Delta_D(\phi(\{\gamma_m \boldsymbol{g}\}))$ for all $\boldsymbol{g} \in B_A$. Therefore we get that $B_A \in \Lambda_{D,c}$. □

By using the Claims 1.3.20-1.3.23 and following the exact same arguments as in the case where $\beta_k = 1$ and $\gamma_k(t) = t^{k-1}$ we can conclude that (1.61) and (1.62) are also valid for the more general case. This finishes the proof of Lemma 1.3.19. □

Let us now finish the proof of Theorem 1.3.18. First of all note that due to Corrollary 1.3.12 we can write for each $h \in \{0, \dots, H\}$

$$K_h = \bigcup_{i=1}^{v} B_{h,i}, \tag{1.79}$$

for some $B_{h,1}, \dots, B_{h,v} \in \Sigma_w$ and $w = h + 2 + \lceil \log_q s \rceil$. Further let us set $\rho_h := \log_2(h + 2 + \lceil \log_q s \rceil + \max_{k \in \{1,\dots,N\}} \deg(\beta_k))$ and define for $\gamma \in \{0, 1, \dots, 2^{\rho_h} - 1\}$

$$Q(N, \gamma, \rho_h) := \{n \in \{1, \dots, N\} \mid n \equiv \gamma \pmod{2^{\rho_h}}\}. \tag{1.80}$$

Observe that for $n, m \in Q(N, \gamma, \rho_h)$ and $n \geq m$ we have $2^{\rho_h} \leq n - m$ and therefore

$$h + 2 + \lceil \log_q s \rceil \leq n - m - \deg(\beta_m) = n - 1 - (m - 1 + \deg(\beta_m)) = n - 1 - \deg(\gamma_m).$$

Hence we get by Lemma 1.3.19 that for all $n, m \in Q(N, \gamma, \rho_h)$ the random variables $\mathbb{1}_{K_h}(\overline{\boldsymbol{x}}_n)$ and $\mathbb{1}_{K_h}(\overline{\boldsymbol{x}}_m)$ are stochastically independent. In order to apply the machinery of Subsection 1.3.2 it is sufficient if we have additionally to the independence property that there exist constants $C_1, C_2 > 0$ such that for all $N, s \in \mathbb{N}$ and $h \in \{0, \dots, H\}$

$$C_1 2^{\kappa_h} \leq 2^{\rho_h} \leq C_2 2^{\kappa_h}, \tag{1.81}$$

where $\kappa_h = \log_2(h + 2 + \lceil \log_q s \rceil)$. Note that the lower bound for 2^{ρ_h} is satisfied for $C_1 = 1$ by definition of ρ_h and the upper bound is a direct consequence of condition (1.77) since we have

$$2^{\rho_h} = h + 2 + \lceil \log_q s \rceil + \max_{k \in \{1,\dots,N\}} \deg(\beta_k) \leq C_2(h + 2 + \lceil \log_q s \rceil) = C_2 2^{\kappa_h},$$

for some $C_2 > 0$. Therefore we are able to use the exact same steps as in Subsection 1.3.2 and this finishes the proof of Theorem 1.3.18.

1.4 Conclusions and further research

Let us conclude Chapter 1 with a brief summary of the main results of Section 1.2 and Section 1.3 followed by a small discussion of possible extensions and generalisations. In Section 1.2 we studied the efficient construction of polynomial lattice point sets which yield a small weighted star discrepancy. For $p \in \mathbb{P}, m \in \mathbb{N}$ and certain choices of the modulus $f(x) \in \mathbb{F}_p[x], \deg(f) = m$ we were able to provide an algorithm (Algorithm 1.2.8) which constructs a generating vector $\boldsymbol{g} \in (\mathbb{F}_p[x])^s$ such that the corresponding N-element lattice point set $\mathcal{P}_N(\boldsymbol{g}, f)$ where $N = p^m$ satisfies the following weighted star discrepancy bound for all $\delta > 0$

$$D^*_{N,\boldsymbol{\gamma}}(\boldsymbol{g}, f) = \mathcal{O}(N^{-1+\delta}).$$

Additionally the construction cost of Algorithm 1.2.8 is of the order of magnitude

$$\mathcal{O}\left(N \log N + \min\{s,t\}N + N \sum_{d=1}^{\min\{s,t\}} (m - w_d)p^{-w_d}\right),$$

where the quantity t is depending on the weight sequence $\boldsymbol{\gamma}$ and the sequence $\boldsymbol{w}$ helps to control the size of the search sets for the generating vector. Roughly speaking t becomes constant if the weights are decaying fast enough. To put it differently, the computational cost becomes independent of the dimension s eventually if the weights are decreasing sufficiently fast. This speed up is due to the fact that Algorithm 1.2.8 reduces the search sets for the components of $\boldsymbol{g}$. The reduction of the search sets is controlled by the sequence $\boldsymbol{w}$, which has to be chosen in accordance to the weight sequence $\boldsymbol{\gamma}$ such that $\sum_{i=1}^{\infty} \gamma_i p^{w_i} < \infty$.

The first generalisation which comes to ones mind is of course to increase the class for the possible choices of the modulus $f(x)$. So far the results stated in Section 1.2 are valid for the case $f(x) = x^m$ and f irreducible with $\deg(f) = m$, respectively. It should be possible to follow the lines of Section 1.2 in the more general case where $f \in \mathbb{Z}_p[x]$ and $\deg(f) = m$. One of the main tasks will be to handle the technical details which will result from the more evolved structure of the modulus f.

At this point one should mention that there exists a variety of different CBC construction which focus on various aspects and properties of the final point set which are useful in different contexts. For example there exists a construction method by Dick and Kritzer (see [21]) which is called projection corrected CBC construction. The main goal of this method is to construct generating vectors (of lattice point sets) which perform good in the context of QMC methods and additionally get rid of certain undesirable projection properties of the final point set which have been observed by several authors in numerical calculations. (More precise, it was observed that in some situations the standard CBC algorithm produces generating vectors which have repeated components. The reason of this phenomenon could be numerical issues of the CBC algorithm but is currently not known.) A natural question would be to ask if it is possible to combine the projection-corrected CBC construction described in [21] with the construction method in Section 1.2.

In Section 1.3 we studied N-element point sets $\mathcal{P}$ whose star discrepancy shows a sub-exponential behaviour in the dimension s, which is

$$D_N^*(\mathcal{P}) = \mathcal{O}\left(\sqrt{\frac{s}{N}}\right), \tag{1.82}$$

where the implied constant is independent of s and N. So far no (efficient) explicit constructions of such point sets exist. In 2014 Löbbe proved that the point set $\mathcal{P}_N = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\}$ with $\boldsymbol{x}_n = \{2^{n-1}\boldsymbol{\alpha}\}$ satisfies a bound of order $\mathcal{O}(\sqrt{s \log s/N})$ with high probability (see Theorem 1.3.1). We were able to carry over the results of Löbbe to a digital analogue of the point set $\mathcal{P}_N(\boldsymbol{\alpha})$. More precise, we studied the point set $\mathcal{P}_N(\boldsymbol{f}) = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\}$, where $\boldsymbol{x}_n = \phi(\{t^{n-1}\boldsymbol{f}\})$ for some s-tuple $\boldsymbol{f} = (f_1, \ldots, f_s) \in (\mathbb{F}_q((t^{-1})))^s$ of formal Laurent series and ϕ defined as in (1.26). We were capable of proving the following metric result:

Theorem. *Let q be a prime number and let $N, s \in \mathbb{N}$ with $N, s \geq 2$. Then for every $\varepsilon \in (0,1)$ there is a quantity $C(q,\varepsilon) > 0$ such that the star discrepancy of the point set $\mathcal{P}_N(\boldsymbol{f})$ satisfies*

$$D_N^*(\mathcal{P}_N(\boldsymbol{f})) \leq C(q,\varepsilon)\sqrt{\frac{s \log s}{N}} \tag{1.83}$$

μ_s-with probability at least $1-\varepsilon$. The quantity $C(q,\varepsilon)$ is of order $\mathcal{O}_q(\log \varepsilon^{-1})$.

Moreover, we were able to prove a generalisation of the theorem above (see Subsection 1.3.3). We considered points of the form $\boldsymbol{x}_n = \phi(\{\gamma_n \boldsymbol{f}\})$ and $(\gamma_k)_{k\in\mathbb{N}} \in (\mathbb{Z}_q[t])^{\mathbb{N}}$ such that

$$\forall k \in \mathbb{N}\ \exists \beta_k \in \mathbb{F}_q[t] : \gamma_k(t) = \beta_k(t)t^{k-1} \text{ and } \deg(\beta_k) \leq c \log_q s, \tag{1.84}$$

where $c > 0$ is a constant independent of N and s.

In order to ensure the existence of point sets satisfying (1.82) the authors of [45] used a probabilistic method. Roughly speaking this means they considered a point set consisting of N i.i.d. random variables and proved with the help of Hoeffding's inequality, which quantifies the deviation of the mean from a sum of independent random variables, that (1.82) is satisfied with positive probability. Observe that the discrepancy bound in (1.82) and the bound in (1.83) of our main result differ by a factor of $\sqrt{\log s}$. This additional factor seems to be the price one has to pay for the additional structure of the point set $\mathcal{P}(\boldsymbol{f})$. More precise, if one wants to use the same framework as in [45] one has to work around the problem that the additional structure of the point set destroys the independence of the considered random variables. In our case we followed the approach of Löbbe [75] and clustered the random variables into groups where they are pairwise independent and finally apply Bernstein's inequality, which is a more general form of Hoeffding's inequality, for each group of random variables. This workaround lead to the additional factor of $\sqrt{\log s}$. Additionally this approach is somehow responsible for the condition (1.77) in Theorem 1.3.18, where we considered points of the form $\boldsymbol{x}_n = \phi(\{\gamma_n \boldsymbol{f}\})$ and $(\gamma_k)_{k\in\mathbb{N}}$ satisfying (1.84).

Recall that both the Hoeffding and Bernstein inequality are of the same type, i.e. both inequalities measure deviations from the mean for sums of independent random variables. It would be an interesting task to adapt the framework of Section 1.3 in the following way: Choose a suitable inequality of similar type, which can also be applied to sums of dependent random variables, such that one can improve (1.83) or is able to loosen the constraint in Theorem 1.3.18.

Chapter 2

The Sudler product of sines

2.1 Introduction

The main quantity of interest in this chapter will be the following sequence of trigonometric products

$$P_N(\alpha) := \prod_{r=1}^{N} |2\sin(\pi r\alpha)|, \tag{2.1}$$

where $N \in \mathbb{N}$, and $\alpha \in \mathbb{R}$ is fixed. More precise, we will analyse the asymptotic behaviour of $P_N(\alpha)$ for special choices of α.

The study of the sequence $P_N(\alpha)$ goes back to the late 1950s, when questions about its asymptotic behaviour were raised by Erdős and Szekeres [32]. Another early exposition on $P_N(\alpha)$ was given by Sudler [99] in the 1960s, giving rise to the name *Sudler product*. The continued analysis of $P_N(\alpha)$ has been carried out in a number of different fields in both pure and applied mathematics (such as partition theory [99, 107], Padé approximation [29] and continued fractions [76] , as well as KAM theory and the theory of strange non-chaotic attractors [8, 38, 57, 65]). Further the Sudler product seems to play a role in interpolation theory in [52, 53] and the analytic continuation of Dirichlet series [56, 105]. See [17, 30, 76, 87] for a connection to q-series or [11, 16, 34] for more recent studies on $P_N(\alpha)$. Quite recently similar trigonometric products have been analysed in the context of uniform distribution and discrepancy theory [3, 54]. The above given list is by no means complete but this broad interest in the Sudler product has lead to a range of different notations and terminologies, making it challenging to get a full picture of what is actually known. Nevertheless we will try to give a small overview on some key results concerning the Sudler product of sines in Subsection 2.1.2. This overview is mainly based on a more detailed survey of central results on $P_N(\alpha)$, which can be found in [104]. But before we do so let us give a more detailed insight into a couple of items of the long list above of topics which are related to the Sudler product. We will mainly focus on the number theoretic topics that have been mentioned in the list before.

2.1.1 Relations to several fields of number theory

As we have already pointed out in the beginning of this section, the sequence $P_N(\alpha)$ has been studied independently in several disciplines of pure and applied mathematics. In this subsection we want to give some examples of fascinating connections to other mathematical fields and topics:

- We start with an interesting relation to q-series. Before we are able to characterise more precisely what actually a q-series is, it is helpful to understand the property of being a q-analogue. Roughly speaking a q-analogue is a generalisation of some mathematical expression which results in the already known expression in the limit $q \to 1^-$. (Usually one has that $0 < q < 1$.) There are q-analogues for a variety of mathematical objects e.g. binomial coefficients, factorial, Fibbonaci numbers, ... and this generalised objects form the basis for a whole "q-calculus". Consider for example the Pochhammer symbol (shifted factorial) for $a \in \mathbb{C}$ and $N \in \mathbb{N}$ which is given by

  $$(a)_N := a(a+1)\cdots(a+N-1).$$

  The q-analogue of the Pochhammer symbol is called the q-Pochhammer symbol and defined by

  $$(a;q)_N := (1-a)(1-aq)\cdots(1-aq^{N-1}). \tag{2.2}$$

  Note that we have $\lim_{q\to 1^-}(q^a;q)_N/(1-q)^N = (a)_N$. A q-series is now a series where expressions of the form (2.2) appear in its summands. For more detailed information on this topic see for example [5, 33, 35].
  A very important class of q-series are the basic hypergeometric series: For $r, s \in \mathbb{N}$ and $|q| < 1$ we define

  $${}_r\varphi_s\begin{pmatrix} a_1, \dots, a_r \\ b_1, \dots, b_s \end{pmatrix}; q, z\Bigg) = \sum_{n=0}^{\infty} \frac{(a_1;q)_n(a_2;q)_n\cdots(a_r;q)_n}{(b_1;q)_n(b_2;q)_n\cdots(b_s;q)_n(q;q)_n}\left[(-1)^n q^{\binom{n}{2}}\right]^{1+s-r} z^n, \tag{2.3}$$

  where $q \neq 0$ if $r > s+1$ and $a_1, \dots, a_r$ and $b_1, \dots, b_s \in \mathbb{C}$ such that $b_i \neq q^{-m}$ for $m \in \mathbb{N}$ and $i \in [s]$.

These mathematical objects are a powerful tool which have applications in number theory, combinatorics, computer algebra and mathematical physics (e.g. for illustrating the possible number of states on a lattice). Moreover, basic hypergeometric series are a q-analogue of the much better understood generalised hypergeometric series and hypergeometric series, see for example [6, 9, 44, 58, 59, 90, 96, 100]. The topic of q-series has been a fruitful research area in the last decades and unfortunately it is beyond the scope of this thesis to go more into detail concerning this interesting field. If we consider now the special case of (2.2) where $a = q$ and $q = \exp(2\pi\mathrm{i}\alpha)$ we get after some calculations that

$$|(q;q)_N| = \prod_{r=1}^{N} |1 - q^r| = \prod_{r=1}^{N} |1 - \exp(2\pi\mathrm{i}\alpha r)| = P_N(\alpha), \tag{2.4}$$

which connects the product $P_N(\alpha)$ to the world of q-series.

- Consider the partition function $p : \mathbb{N} \to \mathbb{N}$, which counts the number of ways to write an integer n as a sum of positive integers, where the order of the summands is not significant. For example $p(4) = 4$ since $4 = 3+1 = 2+2 = 2+1+1 = 1+1+1+1$. The Euler function defined as $\phi : \mathbb{R} \to \mathbb{R}$,

$$\phi(q) := \prod_{n=1}^{\infty} (1 - q^n) = (q;q)_\infty$$

is closely related to the partition function $p(n)$ and the (generalised) pentagonal numbers. By Euler's pentagonal theorem (cf. [7]) it is well known that

$$\begin{aligned} \phi(q) = \prod_{n=1}^{\infty}(1-q^n) &= 1 + \sum_{n=1}^{\infty} (-1)^n \left(q^{n(3n-1)/2} + q^{n(3n+1)/2} \right) \\ &= 1 - q - q^2 + q^5 + q^7 - q^{12} - q^{15} + \cdots, \end{aligned} \tag{2.5}$$

where the sequence of exponents in (2.5) can be described by $n(3n-1)/2$ for $n = 1, -1, 2, -2, 3, -3, \ldots$. These numbers are called the generalised pentagonal numbers. If we define $p_o(n)$ and $p_e(n)$ as the amount of partitions of n with an odd or even number of summands then the following relation, which was observed by Legendre, is equivalent to (2.5):

$$p_o(n) - p_e(n) = \begin{cases} (-1)^j & n = j(3j \pm 1)/2 \\ 0 & \text{else.} \end{cases}$$

Additionally one can express the generating function of the partition numbers $p(n)$ with the help of the Euler function ϕ:

$$\phi(q)\sum_{n=0}^{\infty}p(n)q^n = 1.$$

Furthermore, if we introduce the notation $\tilde{p}(n)$ for the number of partitions of n into distinct parts it is known that

$$\prod_{n=1}^{\infty}(1+q^n) = \sum_{n=0}^{\infty}\tilde{p}(n)q^n.$$

Now recall that if we set $q = \exp(2\pi \mathrm{i}\alpha)$ we have that

$$|1-q^n| = |2\sin(\pi n\alpha)| \text{ and } |1+q^n| = |2\cos(\pi n\alpha)| = \left|\frac{2\sin(\pi n 2\alpha)}{2\sin(\pi n\alpha)}\right|.$$

Hence, the asymptotic behaviour of the products $P_N(\alpha)$ (and $P_N(2\alpha)/P_N(\alpha)$) has interesting connections to the theory of partitions.

- Moreover, slightly altered versions of the Sudler product appear in connection with discrepancy theory and related topics. For example in [54] the authors studied the discrepancy of a certain hybrid sequence which lead to similar trigonometric products. More precise, in order to get information on the star discrepancy of the two-dimensional sequence $z_k = (\{k\alpha\}, x_k)$, where α is an irrational in $(0,1)$ and x_k a digital Niederreiter sequence (see [78]), the authors had to investigate lacunary trigonometric products of the form

$$\prod_{r=0}^{N-1}\left|\cos(2^r\pi\alpha + \gamma_j\pi/2)\right|,$$

where $\gamma_j \in \{0,1\}$.

- Another interesting approach is to study the Sudler product in the context of uniform distribution by considering the more general product

$$\overline{P}_N((x_r)_{r\in\mathbb{N}}) = \prod_{r=1}^{N} 2\sin(\pi x_r), \tag{2.6}$$

where $(x_r)_{r\in\mathbb{N}}$ is a uniformly distributed sequence in the unit interval. In particular this means that $P_N(\alpha) = \overline{P}_N((x_r)_{r\in\mathbb{N}})$, where $(x_r)_{r\in\mathbb{N}}$ is the famous and well studied one-dimensional Kronecker-sequence $(\{r\alpha\})_{r\in\mathbb{N}}$ and $\alpha \in (0,1)$ and irrational (observe that $P_N(\alpha) = \prod_{r=1}^{N} |2\sin(\pi r\alpha)| = \prod_{r=1}^{N} 2\sin(\pi\{r\alpha\})$). This approach was carried out by Aistleitner et al. in [3]. One of the main result of this article states that if $(x_r)_{r\in\mathbb{N}}$ is a uniformly distributed sequence in the unit interval then

$$\overline{P}_N((x_r)_{r\in\mathbb{N}}) \leq \left(\frac{N}{\Delta_N}\right)^{2\Delta_N}, \tag{2.7}$$

where $\Delta_N := ND_N^*((x_r)_{r\in\mathbb{N}})$ and N is sufficiently large.

- Last but not least we would like to point out that the sequence $(P_N(\alpha))_{N\in\mathbb{N}}$ is also studied in various fields related to physics and applied mathematics (e.g. string theory, KAM theory or the study of strange non-chaotic attractors (SNA)). All of this areas try to understand some aspects of the growth of $P_N(\alpha)$. For more information in this direction we refer the interested reader to [8, 38, 57, 65] and the references therein.

2.1.2 Key results

In this section we provide a very brief overview of some key results related to the sequence $P_N(\alpha)$. For a more detailed overview we refer to the introduction of [104]. Let us start with the well understood case when α is rational. We summarize some basic facts in the subsequent lemma.

Lemma 2.1.1. *Let $N \in \mathbb{N}$, $\alpha \in \mathbb{R}$ and $P_N(\alpha)$ defined as in* (2.1)*. Then we have*

1. *$P_N(\alpha) = P_N(\{\alpha\})$, where $\{\cdot\}$ denotes the fractional part.*
2. *If $\alpha \in \mathbb{Q}$ with $\alpha = p/q$ and $\gcd(p,q) = 1$ then*

$$P_N(\alpha) = \begin{cases} 0 & \text{if } N \geq q \\ q & \text{if } N = q-1. \end{cases}$$

3. *For $\alpha \in \mathbb{R}$ we have that $P_N(\alpha) \leq 2^N$.*

Proof. Except for the case $N = q-1$ in the second statement of Lemma 2.1.1 all the properties are a direct consequence of the definition of $P_N(\alpha)$ given in (2.1). Now assume that $\alpha = p/q$ with $\gcd(p,q) = 1$. We will show a slightly stronger result which is that $\prod_{r=1}^{q-1} 2\sin(\pi r\alpha) = q$. Observe that $rp \pmod q$ runs through the set

$\{1, \dots, q-1\}$ if r runs through $\{1, \dots, q-1\}$ since $\gcd(p,q) = 1$. Hence we need to show that

$$\prod_{r=1}^{q-1} 2\sin\left(\pi \frac{r}{q}\right) = q.$$

Note that we have $2\sin(\pi r/q) = |1 - \mathrm{e}^{2\pi \mathrm{i} r/q}|$. Moreover, it is a fact that $(x-1)\prod_{r=1}^{q-1}(x - \mathrm{e}^{2\pi \mathrm{i} r/q}) = x^q - 1 = (x-1)\sum_{r=0}^{q-1} x^r$. Using this observations we obtain

$$\prod_{r=1}^{q-1} 2\sin\left(\pi \frac{r}{q}\right) = \left|\prod_{r=1}^{q-1}(1 - \mathrm{e}^{2\pi \mathrm{i} r/q})\right| = \sum_{r=0}^{q-1} 1 = q \tag{2.8}$$

and the result follows. Of course (2.8) implies that $P_{q-1}(p/q) = q$. $\square$

Due to the first and the second statement of Lemma 2.1.1 we can restrict our attention to the case where $\alpha \in (0,1)$ and irrational. For this case the growth of $P_N(\alpha)$ has been studied in different contexts. One of them was to analyse $\sup_{\alpha \in (0,1)} P_N(\alpha)^{1/N}$, which was first carried out by Sudler in [99], where he could prove that

$$\lim_{N\to\infty} \left(\sup_{\alpha\in(0,1)} P_N(\alpha)\right)^{1/N} = \alpha_0^{-1} \int_0^{\alpha_0} \log|2\sin(\pi\alpha)| \,\mathrm{d}\alpha,$$

where α_0 is the unique solution in $[1/2, 1]$ of the equation $\int_0^{\alpha_0} \alpha \cot(\pi\alpha)\,\mathrm{d}\alpha = 0$. (From numerical calculations we obtain $\alpha_0 = 0.7912\dots$.) Moreover, Sudler could show that $\sup_{\alpha\in(0,1)} P_N(\alpha)$ is achieved at α_N with asymptotic behaviour of the form $\alpha_N \sim \alpha_0/N$ as N grows. Recently Bell [11] adopted the method of Wright [107] and was capable of giving a more precise version of the result of Sudler:

$$\sup_{\alpha\in(0,1)} P_N(\alpha) \sim C_1\sqrt{N}E^N,$$

where $C_1 > 0$ is independent of N and $E = \alpha_0^{-1}\int_0^{\alpha_0} \log|2\sin(\pi\alpha)|$. Furthermore Bell could apply his methods also for L_p-norms and obtained for $p \in [1,\infty)$ that

$$\left(\int_0^1 P_N(\alpha)^p \,\mathrm{d}\alpha\right)^{1/p} \sim C_1(C_2 n^{-3/2})^{1/p}\sqrt{n}E^N,$$

where $C_1, C_2 > 0$ are independent of N.

Another natural question is to ask about the growth rate of $P_N(\alpha)$ for fixed α. Surprisingly Lubinsky [76] was able to prove that for almost all α one has $\lim_{N\to\infty} P_N^{1/N}(\alpha) = 1$ in contrast to the exponential growth of $\sup_\alpha P_N(\alpha)$. At first sight this result seems to be of a counter-intuitive nature. A rather heuristic explanation for this phenomenon is that the exponential behaviour of $\sup_{\alpha\in(0,1)} P_N(\alpha)$ is due to peaks, which do not reflect the typical behaviour of the function. But with increasing N this peaks narrow more and more. This means that for growing N the set of α, which are responsible for the exponential behaviour of $\sup_{\alpha\in(0,1)} P_N(\alpha)$ becomes negligible. Furthermore Lubinsky was able to characterize the sub-exponential growth of $P_N(\alpha)$ more precisely (see [76]). One of his results states that for almost all α and almost all $\varepsilon > 0$ we have

$$P_N(\alpha) \le N^{C(\log\log N)^{1+\varepsilon}}$$

if N is sufficiently large and where $C > 0$ is depending on α and ε. In particular, if α has bounded continued fraction coefficients then he could even prove a polynomial growth rate, i.e. $P_N(\alpha) \le N^C$.

The special case where α is the golden ratio φ gained a lot of attention since the golden ratio is for example in terms of the continued fraction expansion the simplest irrational. Knill and Tangerman investigated the behaviour of the sum $S_N(\varphi) = \sum_{r=1}^{N} \log(2 - 2\cos(2\pi r\varphi))$ in [57] (note that $S_N(\varphi) = \log(P_N(\varphi))$). Quite recently Verschueren and Mestel were capable of proving that the subsequence $(P_{F_n}(\varphi))_{n\ge1}$ of the Sudler product is convergent for $n \to \infty$ and $(F_n)_{n\ge1}$ being the Fibonacci sequence. It is exactly this work of Verschueren and Mestel that will play a central role in Section 2.2 and which we will generalise in Section 2.3. In the next section we will stick to the special case where α equals the golden ratio φ and shift our attention to $\liminf_{N\to\infty} P_N(\varphi)$.

2.2 A positive lower bound for $\liminf_{N\to\infty} P_N(\varphi)$

A long-standing open question raised by Erdős and Szekeres in 1959 is: what can we say about $\liminf_{N\to\infty} P_N(\alpha)$? This question occupied Lubinsky, who studied the product $P_N(\alpha)$ in the context of q-series in [76]. In his paper, Lubinsky shows the following theorem:

Theorem 2.2.1 ([76, Theorem 1.3])**.** *Let $\alpha \in (0,1)$ and irrational with continued fraction expansion $\alpha = [0; a_1, a_2, \dots]$ and let $P_N(\alpha)$ be defined as in* (2.1)*. If $\sup_{j\in\mathbb{N}} a_j = \infty$ then*

$$\liminf_{N\to\infty} P_N(\alpha) = 0. \tag{2.9}$$

Note that (2.9) holds in the presence of unbounded continued fraction coefficients. Moreover, Lubinsky expresses that he "feels certain that it (equation (2.9)) is true in general", i.e. also for α with bounded continued fraction coefficients. The first main goal of this chapter is to show that, in fact, this is not the case. This section is dedicated to prove the following theorem

Theorem 2.2.2. *If $\varphi = (\sqrt{5}-1)/2$, then*

$$\liminf_{N\to\infty} P_N(\varphi) = \liminf_{N\to\infty} \prod_{r=1}^{N} |2\sin(\pi r\varphi)| > 0. \tag{2.10}$$

The number $\varphi = (\sqrt{5}-1)/2$, known as the fractional part of the golden ratio, has the simplest possible continued fraction expansion

$$\varphi = \cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{1+\dots}}} = [0;\overline{1}].$$

This observation is key in establishing Theorem 2.2.2. Nevertheless, we suspect that $\liminf_{N\to\infty} P_N(\alpha) > 0$ also for other quadratic irrationals α (see Section 2.4.1 for a discussion on this topic).

Remark 2.2.3. Observe that in the literature the golden ratio is defined as

$$\varphi = \frac{\sqrt{5}+1}{2}. \tag{2.11}$$

Due to the fact that $P_N(\alpha) = P_N(\{\alpha\})$ we will not distinguish between φ and $\{\varphi\}$ when it is not necessary.

In the following section, we present our strategy for proving Theorem 2.2.2. The proof relies heavily on a paper by Verschueren and Mestel [104], where the asymptotic behaviour of the subsequence $(P_{F_n}(\varphi))_{n\geq 1}$ is investigated for the Fibonacci sequence $(F_n)_{n\geq 1}$. Let us therefore briefly review the connection between the golden ratio φ and the Fibonacci sequence before we present our proof strategy.

2.2.1 The Fibonacci sequence

Throughout this chapter, we denote by φ the (fractional part of the) golden ratio

$$\varphi := \frac{\sqrt{5}-1}{2},$$

and by $(F_n)_{n\geq 1} = (1, 1, 2, 3, 5, 8, 13, \ldots)$ the sequence of Fibonacci numbers. There is an intimate relationship between φ and the Fibonacci sequence; $(F_n)_{n\geq 1}$ is precisely the sequence of best approximation denominators of φ. Moreover, we have the property

$$F_n\varphi = F_{n-1} - (-\varphi)^n, \tag{2.12}$$

for $F_0 := 0$ and $n \in \mathbb{N}$.

Finally, recall that any positive integer N has a unique expansion in terms of the Fibonacci sequence, known as its *Zeckendorf representation* [108].

Definition 2.2.4. *Any $N \in \mathbb{N}$ has a unique Zeckendorf representation*

$$N = \sum_{j=1}^{m} F_{n_j},$$

where $(F_n)_{n\in\mathbb{N}_0}$ is the Fibonacci sequence, and:

(i) $n_1 \geq 2$;

(ii) $n_{j+1} > n_j + 1$ for all $j \in \{1, \ldots, m-1\}$.

Moreover, it is well known that $n_m = \mathcal{O}(\log N)$ (see e.g. [64, p. 126]).

In other words, we can associate to any $N \in \mathbb{N}$ a unique integer sequence $(n_1, \ldots, n_m)$. Note that since $m < n_m$, the length of this sequence is $m = \mathcal{O}(\log N)$.

2.2.2 Strategy

The proof of Theorem 2.2.2 relies on central results in a recent paper by Verschueren and Mestel [104]. In this paper, the authors analyse the asymptotic behaviour of the product sequence $(P_N(\varphi))_{N\geq 1}$ for the golden ratio φ, and show in particular that:

Theorem 2.2.5 ([104, Theorem 3.1]). *The subsequence $(P_{F_n}(\varphi))_{n\geq 1}$ is convergent, and*

$$\lim_{n\to\infty} P_{F_n}(\varphi) = \lim_{n\to\infty} \prod_{r=1}^{F_n} |2\sin(\pi r\varphi)| \simeq 2.407\dots. \tag{2.13}$$

A consequence of Theorem 2.2.5 is that the general product $P_N(\varphi)$ must necessarily obey polynomial bounds[1]

$$N^{C_1} \leq P_N(\varphi) \leq N^{C_2}, \tag{2.14}$$

where $C_1 \leq 0 < 1 \leq C_2$. These bounds are established as follows: Expressing the integer N by its Zeckendorf representation $N = \sum_{j=1}^m F_{n_j}$, we can rewrite $P_N(\varphi)$ as

$$\begin{aligned}
P_N(\varphi) &= \prod_{r=1}^{\sum_{j=1}^m F_{n_j}} |2\sin(\pi r\varphi)| \\
&= \left(\prod_{r=1}^{F_{n_m}} |2\sin(\pi r\varphi)|\right) \left(\prod_{r=F_{n_m}+1}^{F_{n_m}+F_{n_{m-1}}} |2\sin(\pi r\varphi)|\right) \cdots \left(\prod_{r=\sum_{j=2}^m F_{n_j}+1}^{\sum_{j=1}^m F_{n_j}} |2\sin(\pi r\varphi)|\right) \\
&= \prod_{j=1}^{m} \prod_{r=1}^{F_{n_j}} |2\sin(\pi(r\varphi + k_j\varphi))|\,,
\end{aligned} \tag{2.15}$$

where $k_j = \sum_{s=j+1}^m F_{n_s}$ for $1 \leq j \leq m-1$ and $k_m = 0$ (see [104, p. 220] for further details). Verschueren and Mestel then show that:

Lemma 2.2.6 (see [104, p. 220-221]). *There exist real constants $0 < K_1 \leq 1 \leq K_2$ (independent of N) bounding all terms in* (2.15), *i.e. so that*

$$K_1 \leq \prod_{r=1}^{F_{n_j}} |2\sin(\pi(r\varphi + k_j\varphi))| \leq K_2, \tag{2.16}$$

for all $1 \leq j \leq m$.

---

[1] This was first established by Lubinsky in [76] using a different approach.

It immediately follows from Lemma 2.2.6 and (2.15) that

$$K_1^m \leq P_N(\varphi) \leq K_2^m.$$

Finally, since the Zeckendorf representation of N has length $m = \mathcal{O}(\log N)$, we get (2.14) for some constants $C_1 < C_2$. It follows immediately from Theorem 2.2.5 that $C_1 \leq 0$ (and an argument of why $C_2 \geq 1$ is given in [104, p. 219]).

Our strategy for concluding that $\liminf_{N\to\infty} P_N(\varphi) > 0$ is to evaluate the subproducts in (2.15) more carefully for large values of j.

Lemma 2.2.7. *There exists a threshold value $J \in \mathbb{N}$ (independent of N) such that for all terms in* (2.15) *where $j > J$, we have*

$$\prod_{r=1}^{F_{n_j}} |2\sin(\pi(r\varphi + k_j\varphi))| \geq 1.$$

Combining Lemmas 2.2.6 and 2.2.7, we find that

$$P_N(\varphi) = \prod_{j=1}^{m} \prod_{r=1}^{F_{n_j}} |2\sin(\pi(r\varphi + k_j\varphi))| \geq K_1^J > 0,$$

confirming Theorem 2.2.2.

The proof of Lemma 2.2.7 is given in Section 2.2.4. It requires a certain decomposition of the product $\prod_{r=1}^{F_{n_j}} |2\sin\pi(r\varphi + k_j\varphi)|$ into three more manageable subproducts. This decomposition is inspired by the work of Verschueren and Mestel, and is thoroughly described in the following section.

2.2.3 Decomposition

It is shown in [104, Lemma 5.1] that the product $P_{F_n}(\varphi)$ can be split into three subproducts

$$P_{F_n}(\varphi) = \prod_{r=1}^{F_n} |2\sin(\pi r\varphi)| = A_n B_n C_n, \tag{2.17}$$

where

$$A_n = |2F_n \sin(\pi\varphi^n)|, \tag{2.18}$$

$$B_n = \prod_{t=1}^{F_n - 1} \left| \frac{s_{nt}}{2\sin(\pi t/F_n)} \right|, \tag{2.19}$$

$$C_n = \prod_{t=1}^{F_n-1} \left(1 - \frac{s_{n0}^2}{s_{nt}^2}\right)^{1/2}, \tag{2.20}$$

and where

$$s_{nt} := 2\sin\left(\pi\left(\frac{t}{F_n} - \varphi^n\left(\left\{\frac{tF_{n-1}}{F_n}\right\} - \frac{1}{2}\right)\right)\right). \tag{2.21}$$

Observe that $s_{nt} = s_{n(F_n-t)}$ for $t \in \{1, \dots, F_n - 1\}$ (see [104, Lemma 4.1]). Now it follows immediately for odd F_n that $C_n = \prod_{t=1}^{(F_n-1)/2}(1 - s_{n0}^2/s_{nt}^2)$ and for even F_n we get that

$$\prod_{t=1}^{F_n-1} \left(1 - \frac{s_{n0}^2}{s_{nt}^2}\right)^{1/2} = \left(1 - \frac{s_{n0}^2}{s_{n(F_n/2)}^2}\right)^{1/2} \prod_{t=1}^{(F_n-2)/2} \left(1 - \frac{s_{n0}^2}{s_{nt}^2}\right). \tag{2.22}$$

Further, by definition of s_{nt} we obtain $\lim_{n\to\infty}(1 - s_{n0}^2/s_{n(F_n/2)}^2) = 1$. In other words it does not make a difference for the asymptotic behaviour of C_n if we consider the case where F_n is even or where F_n is odd. In order to avoid such a case distinction and improve readability of what follows we stick to the lines of [104] and introduce the following generalised sum and product notation:
Given a summable sequence $(b_r)_{r\in\mathbb{N}}$, we define the step function $f(t) = b_r$ for $t \in [r, r+1)$. Then for any $x, y \in \mathbb{R}$ where $x \leq y$, we let

$$\sum_{r=x}^{y} b_r := \int_x^y f(t)\, dt. \tag{2.23}$$

Moreover, if $f(t) > 0$ on $[x, y]$, we let

$$\prod_{r=x}^{y} b_r := \exp\left(\sum_{r=x}^{y} \log b_r\right) = \exp\left(\int_x^y \log f(t)\, dt\right). \tag{2.24}$$

This allows us to define sums and products with real, rather than just integer, upper and lower bounds. Note in particular that this definition coincides with normal summation and product notation whenever $x, y \in \mathbb{Z}$.

With this notation in mind we can rewrite (2.20) as

$$C_n = \prod_{t=1}^{(F_n-1)/2} \left(1 - \frac{s_{n0}^2}{s_{nt}^2}\right).$$

A similar decomposition as given in (2.17) can be established for a perturbed version of $P_{F_n}(\varphi)$. Let us introduce the notation

$$P_{F_n}(\varphi, \varepsilon) = \prod_{r=1}^{F_n} |2\sin(\pi(r\varphi + \varepsilon))| \,, \tag{2.25}$$

where ε is some fixed, real number. We claim the following:

Lemma 2.2.8. *We have*

$$P_{F_n}(\varphi, \varepsilon) = \overline{A}_n(\varepsilon) B_n \overline{C}_n(\varepsilon), \tag{2.26}$$

where

$$\overline{A}_n(\varepsilon) = 2F_n|\sin\left(\pi((-\varphi)^n - \varepsilon)\right)|, \tag{2.27}$$

$$\overline{C}_n(\varepsilon) = \prod_{t=1}^{(F_n-1)/2} \left(1 - \frac{v_n^2(\varepsilon)}{s_{nt}^2}\right), \tag{2.28}$$

B_n *and* s_{nt} *are given in* (2.19) *and* (2.21) *respectively, and*

$$v_n(\varepsilon) := v_n := 2\sin\left(\pi\left(\frac{(-\varphi)^n}{2} - \varepsilon\right)\right). \tag{2.29}$$

Proof. By definition we have that

$$\begin{aligned}
(P_{F_n}(\varphi, \varepsilon))^2 =& \big(2\sin(\pi(F_n\varphi + \varepsilon))\big)^2 \prod_{r=1}^{F_n-1} \big(2\sin(\pi(r\varphi + \varepsilon))\big)^2 \\
=& \big(2\sin(\pi(F_n\varphi + \varepsilon))\big)^2 \prod_{r=1}^{F_n-1} \big(2\sin(\pi(r\varphi + \varepsilon))\big)\big(2\sin(\pi((F_n - r)\varphi + \varepsilon))\big) \\
=& \big(2\sin(\pi(F_n\varphi + \varepsilon))\big)^2 \prod_{r=1}^{F_n-1} 2\big(\cos(\pi(2r\varphi - F_n\varphi)) - \cos(\pi(F_n\varphi + 2\varepsilon))\big),
\end{aligned}$$

where we have used the identity $\sin x \sin y = (\cos(x-y) - \cos(x+y))/2$ for the final step. Recall from (2.12) that $F_n\varphi = F_{n-1} - (-\varphi)^n$ for all $n \in \mathbb{N}$. Thus, we get

$$\begin{aligned}
(P_{F_n}(\varphi, \varepsilon))^2 =& \big(2\sin(\pi(F_n\varphi + \varepsilon))\big)^2 \\
&\times \prod_{r=1}^{F_n-1} 2\big(\cos(\pi(2r\varphi - F_{n-1} + (-\varphi)^n)) - \cos(\pi(F_{n-1} - (-\varphi)^n + 2\varepsilon))\big)
\end{aligned}$$

$$=\big(2\sin(\pi(F_n\varphi+\varepsilon))\big)^2(-1)^{(F_{n-1}+1)(F_n-1)}$$
$$\times\prod_{r=1}^{F_n-1}2\big(-\cos(\pi(2r\varphi+(-\varphi)^n))+\cos(\pi((-\varphi)^n-2\varepsilon))\big).$$

Note that $\gcd(F_{n-1},F_n)=1$, and this implies $(-1)^{(F_{n-1}+1)(F_n-1)}=1$. We now use the identity $\cos(x)=1-2\sin^2(x)/2$ to obtain

$$(P_{F_n}(\varphi,\varepsilon))^2=\big(2\sin(\pi(F_n\varphi+\varepsilon))\big)^2$$
$$\times\prod_{r=1}^{F_n-1}4\left(\sin^2\left(\pi\left(r\varphi+\frac{(-\varphi)^n}{2}\right)\right)-\sin^2\left(\pi\left(\frac{(-\varphi)^n}{2}-\varepsilon\right)\right)\right).$$

Applying again the identity (2.12) we get

$$2\sin\left(\pi\left(r\varphi+\frac{(-\varphi)^n}{2}\right)\right)=2\sin\left(\pi\left(r\frac{F_{n-1}}{F_n}-(-\varphi)^n\left(\frac{r}{F_n}-\frac{1}{2}\right)\right)\right).$$

Note that if r runs through $\{1,\ldots,F_n-1\}$ then so does $t=F_{n-1}r \pmod{F_n}$. Furthermore, recall the well known identity $F_{n-1}^2\equiv(-1)^n \pmod{F_n}$. Using the substitution $t=F_{n-1}r \pmod{F_n}$ we obtain that

$$\begin{aligned}
&\left|2\sin\left(\pi\left(r\frac{F_{n-1}}{F_n}-(-\varphi)^n\left(\frac{r}{F_n}-\frac{1}{2}\right)\right)\right)\right|\\
&\quad=\left|2\sin\left(\pi\left(\frac{t}{F_n}-(-\varphi)^n\left(\frac{(-1)^n tF_{n-1}\pmod{F_n}}{F_n}-\frac{1}{2}\right)\right)\right)\right|\\
&\quad=\left|2\sin\left(\pi\left(\frac{t}{F_n}-(-\varphi)^n\left(\left\{\frac{(-1)^n tF_{n-1}}{F_n}\right\}-\frac{1}{2}\right)\right)\right)\right|\\
&\quad=\left|2\sin\left(\pi\left(\frac{t}{F_n}-\varphi^n\left(\left\{\frac{tF_{n-1}}{F_n}\right\}-\frac{1}{2}\right)\right)\right)\right|=|s_{nt}|,
\end{aligned}\tag{2.30}$$

where we have used in the second last step that $f(x)=\{x\}-1/2$ is an odd function and s_{nt} is defined as in (2.21). Using (2.29) and (2.30) we finally have

$$\begin{aligned}
(P_{F_n}(\varphi,\varepsilon))^2&=\big(2\sin(\pi(F_n\varphi+\varepsilon))\big)^2\prod_{t=1}^{F_n-1}\left(s_{nt}^2-v_n^2\right)\\
&=\big(2\sin(\pi(F_n\varphi+\varepsilon))\big)^2\prod_{t=1}^{F_n-1}s_{nt}^2\prod_{t=1}^{F_n-1}\left(1-\frac{v_n^2}{s_{nt}^2}\right)
\end{aligned}$$

$$
\begin{aligned}
&= \big(2\sin(\pi(F_n\varphi+\varepsilon))\big)^2 \prod_{t=1}^{F_n-1} s_{nt}^2 \prod_{t=1}^{F_n-1}\left(1-\frac{v_n^2}{s_{nt}^2}\right) F_n^2 \left(\prod_{t=1}^{F_n-1} 2\sin\left(\pi\frac{t}{F_n}\right)\right)^{-2} \\
&= \left(\overline{A}_n(\varepsilon) B_n \overline{C}_n(\varepsilon)\right)^2,
\end{aligned}
$$

where we have used the notation introduced in (2.24) and the well known product formula (see Lemma 2.1.1)

$$
\prod_{r=1}^{q-1} 2\sin\left(\pi r \frac{p}{q}\right) = q,
$$

for positive integers $p, q \geq 1$ with $\gcd(p,q) = 1$. □

2.2.4 Proof of Lemma 2.2.7

Let us now turn to Lemma 2.2.7. Fix some $N \in \mathbb{N}$, and let

$$
N = \sum_{j=1}^{m} F_{n_j}
$$

be its unique Zeckendorf representation. The product $P_N(\varphi)$ may be decomposed as

$$
P_N(\varphi) = \prod_{j=1}^{m} \prod_{r=1}^{F_{n_j}} \left|2\sin(\pi(r\varphi + k_j\varphi))\right|,
$$

where $k_j = \sum_{s=j+1}^{m} F_{n_s}$ for $1 \leq j \leq m-1$ and $k_m = 0$ (see (2.16)). Using the notation introduced in (2.25), we get

$$
P_N(\varphi) = \prod_{j=1}^{m} P_{F_{n_j}}(\varphi, k_j\varphi).
$$

By applying again the identity $F_n\varphi = F_{n-1} - (-\varphi)^n$ from (2.12), we have

$$
k_j\varphi = \sum_{s=j+1}^{m} \left(F_{n_s-1} - (-\varphi)^{n_s}\right),
$$

and thus

$$
P_N(\varphi) = \prod_{j=1}^{m} P_{F_{n_j}}(\varphi, \varepsilon_j), \quad \varepsilon_j = -\sum_{s=j+1}^{m} (-\varphi)^{n_s}. \tag{2.31}
$$

Recall that by Lemma 2.2.8 we have

$$P_{F_{n_j}}(\varphi, \varepsilon_j) = \overline{A}_{n_j}(\varepsilon_j) B_{n_j} \overline{C}_{n_j}(\varepsilon_j), \tag{2.32}$$

where the terms $\overline{A}_{n_j}(\varepsilon_j)$, B_{n_j} and $\overline{C}_{n_j}(\varepsilon_j)$ are given in (2.27), (2.19) and (2.28), respectively. The claim in Lemma 2.2.7 is that $P_{F_{n_j}}(\varphi, \varepsilon_j) \geq 1$ whenever j exceeds some threshold value (independent of N). We will continue by analysing and bounding each of the three terms $\overline{A}_{n_j}(\varepsilon_j)$, B_{n_j} and $\overline{C}_{n_j}(\varepsilon_j)$ from below.

The term $\overline{A}_{n_j}(\varepsilon_j)$

The following lemma gives a suitable lower bound for the term $\overline{A}_{n_j}(\varepsilon_j)$.

Lemma 2.2.9. *Let $\overline{A}_{n_j}(\varepsilon_j)$ be given in* (2.27)*. We have*

$$\overline{A}_{n_j}(\varepsilon_j) = \frac{2\pi}{\sqrt{5}}(1 + p_j)\left(1 + \mathcal{O}(\varphi^{2n_j})\right), \tag{2.33}$$

where the implied constant is independent of n_j,

$$p_j := -\varepsilon_j(-\varphi)^{-n_j} = \sum_{s=j+1}^{m} (-\varphi)^{n_s - n_j}, \tag{2.34}$$

and $p_j \in [-\varphi^2, \varphi]$.

Proof. Since $n_1 \geq 2$ and any two consecutive elements n_s and n_{s+1} must necessarily satisfy $n_{s+1} - n_s \geq 2$ (recall Definition 2.2.4), we have that

$$p_j = \sum_{s=j+1}^{m} (-\varphi)^{n_s - n_j} \leq \sum_{s=j+1}^{\infty} \varphi^{2(s-j)} = \varphi$$

and

$$p_j = \sum_{s=j+1}^{m} (-\varphi)^{n_s - n_j} \geq -\sum_{s=j+1}^{\infty} \varphi^{2(s-j)+1} = -\varphi^2.$$

Thus we get that $p_j \in [-\varphi^2, \varphi]$.

Moreover, it follows that $|(-\varphi)^{n_j} - \varepsilon_j| = \varphi^{n_j}(1 + p_j) = \mathcal{O}(\varphi^{n_j})$. Hence, by the definition of $\overline{A}_{n_j}(\varepsilon_j)$ and applying $\sin x = x(1 + \mathcal{O}(x^2))$ we get

$$\overline{A}_{n_j}(\varepsilon_j) = 2F_{n_j}\left|\sin(\pi((-\varphi)^{n_j} - \varepsilon_j))\right| = 2\pi F_{n_j}\left|((-\varphi)^{n_j} - \varepsilon_j)\left(1 + \mathcal{O}(\varphi^{2n_j})\right)\right|$$

$$= 2\pi(1+p_j)F_{n_j}\varphi^{n_j}\left(1+\mathcal{O}(\varphi^{2n_j})\right) = \frac{2\pi}{\sqrt{5}}(1+p_j)\left(1+\mathcal{O}(\varphi^{2n_j})\right),$$

where we have used that $F_n = (\varphi^{-n} - (-\varphi)^n)/\sqrt{5}$ for $n \in \mathbb{N}$.

□

The term B_{n_j}

We will exploit the fact that B_{n_j} is not depending on ε_j in contrast to the terms $\overline{A}_{n_j}$ and $\overline{C}_{n_j}$. We want to prove the following general lower bound for B_n:

Lemma 2.2.10. *Let $n \in \mathbb{N}$ and B_n be given in (2.19). Then we have*

$$B_n > \frac{\sqrt{5}}{2\pi\varphi^2}\frac{P_{F_n+2}(\varphi)}{P_{F_n+1}(\varphi)}\left(1-\mathcal{O}(\varphi^{2n})\right). \tag{2.35}$$

Proof. Note that by definition we have

$$\begin{aligned}\frac{P_{F_{n+2}}(\varphi)}{P_{F_{n+1}}(\varphi)} &= \prod_{r=F_{n+1}+1}^{F_{n+2}} |2\sin(\pi r\varphi)| = \prod_{r=1}^{F_n} |2\sin(\pi(r\varphi+\varphi F_{n+1}))| \\ &= \prod_{r=1}^{F_n} |2\sin\left(\pi(r\varphi-(-\varphi)^{n+1})\right)| = P_{F_n}(\varphi,-(-\varphi)^{n+1}) \\ &= \overline{A}_n(-(-\varphi)^{n+1})B_n\overline{C}_n(-(-\varphi)^{n+1}),\end{aligned}$$

where we applied Lemma 2.2.8 for the special case $\varepsilon = -(-\varphi)^n$. Therefore we obtain

$$B_n = \frac{P_{F_{n+2}}(\varphi)}{P_{F_{n+1}}(\varphi)}\left(\overline{A}_n(-(-\varphi)^{n+1})\overline{C}_n(-(-\varphi)^{n+1})\right)^{-1}. \tag{2.36}$$

First of all observe that by following the same lines as in the proof of Lemma 2.2.9 one can easily derive that $\overline{A}_n(-(-\varphi)^{n+1}) = 2\pi\varphi^2/\sqrt{5}(1+\mathcal{O}(\varphi^{2n}))$. Moreover, we have that $\overline{C}_n(-(-\varphi)^{n+1}) < 1$, which can be seen in the following way. Verschueren and Mestel proved in [104, Lemma 4.1] that $s_{nt} \geq s_{n0}$ for $t \in \{0,\ldots,F_n-1\}$. This leads us to

$$\left|\frac{v_n(-(-\varphi)^n)}{s_{nt}}\right| \leq \left|\frac{v_n(-(-\varphi^n))}{s_{n0}}\right| = \left|\frac{\sin(\frac{\pi}{2}\varphi^{n+3})}{\sin(\frac{\pi}{2}\varphi^n)}\right| < 1.$$

Hence, by (2.28) the product $\overline{C}_n(-(-\varphi)^{n+1}) < 1$ and we finally arrive at

$$B_n > \frac{\sqrt{5}}{2\pi\varphi^2}\frac{P_{F_{n+2}}(\varphi)}{P_{F_{n+1}}(\varphi)}\left(1-\mathcal{O}(\varphi^{2n})\right).$$

□

The term $\overline{C}_{n_j}(\varepsilon_j)$

We now shift our attention to the term $\overline{C}_{n_j}(\varepsilon_j)$. Our goal is to prove:

Lemma 2.2.11. *Let $\overline{C}_{n_j}(\varepsilon_j)$ be given in* (2.28)*. We have*

$$\overline{C}_{n_j}(\varepsilon_j) \geq 1 - \frac{1}{7}(1+2p_j)^2 - \mathcal{O}(\varphi^{n_j/5}), \tag{2.37}$$

with p_j as in (2.34) *and where the implied constant is independent of n_j.*

The proof of Lemma 2.2.11 is more elaborate than the proofs of Lemma 2.2.9 and Lemma 2.2.10, and we start by stating two preliminary results.

Lemma 2.2.12 ([104, Lemma 4.3])**.** *For $n \geq 2$ and real numbers a_t, $t = 1, 2, \ldots, n$, satisfying $A := \sum_{t=1}^{n} |a_t| < 1$, we have*

$$1 - A < \prod_{t=1}^{n}(1 - |a_t|) < \frac{1}{1-A}.$$

Lemma 2.2.12 is used in [104] to show that the product C_n in (2.20) can be expressed as

$$C_n = \prod_{t=1}^{\infty}\left(1 - \frac{1}{u_t^2}\right) - \mathcal{O}(\varphi^{n/5}),$$

where

$$u_t := 2\left(\sqrt{5}t - \{t\varphi\} + \frac{1}{2}\right). \tag{2.38}$$

We use it here to verify that a similar expression can be given for the perturbed product $\overline{C}_n(\varepsilon)$ whenever the perturbation ε is sufficiently small.

Lemma 2.2.13. *Let $\overline{C}_n(\varepsilon)$ be given in* (2.28)*, and assume that $|\varepsilon| \leq \varphi^{n+1}$. Then*

$$\overline{C}_n(\varepsilon) \geq \prod_{t=1}^{\infty}\left(1 - \frac{(1-2\varepsilon(-\varphi)^{-n})^2}{u_t^2}\right) - \mathcal{O}(\varphi^{n/5}),$$

with u_t given in (2.38) *and where the implied constant is independent of n.*

Proof. Recall that

$$\overline{C}_n(\varepsilon) = \prod_{t=1}^{(F_n-1)/2}\left(1 - \frac{v_n(\varepsilon)^2}{s_{nt}^2}\right), \tag{2.39}$$

where $v_n(\varepsilon)$ is given in (2.29) and s_{nt} is given in (2.21). The assumption on ε implies that $|\varepsilon(-\varphi)^{-n}| \leq \varphi$ and

$$|1 - 2\varepsilon(-\varphi)^{-n}| \leq 1 + 2\varphi = \sqrt{5}. \tag{2.40}$$

It thus follows from $\sin x = x(1 + \mathcal{O}(x^2))$ that

$$\begin{aligned} |v_n(\varepsilon)| &= 2\left|\sin\left(\frac{\pi}{2}(-\varphi)^n(1 - 2\varepsilon(-\varphi)^{-n})\right)\right| \\ &= \pi\varphi^n\left|1 - 2\varepsilon(-\varphi)^{-n}\right|(1 + \mathcal{O}(\varphi^{2n})) \\ &\leq \pi\varphi^n\sqrt{5}\left(1 + \mathcal{O}(\varphi^{2n})\right). \end{aligned} \tag{2.41}$$

We now split the product (2.39) at $\eta_n := \lceil \varphi^{-3n/5} \rceil$, and treat first the terms where $t \geq \eta_n$. As a next step note that from (2.12) and the fact that $\sqrt{5}F_n = \varphi^{-n} - (-\varphi)^n$ it follows for $n \geq 1$ that

$$\frac{F_{n-1}}{F_n} = \varphi + \mathcal{O}(\varphi^{2n}) \text{ and } F_n\varphi^n = \frac{1}{\sqrt{5}}(1 + \mathcal{O}(\varphi^{2n})) \tag{2.42}$$

Now it follows by (2.42) for sufficiently large n that

$$\begin{aligned} s_{nt} &= 2\sin\left(\pi\left(t\sqrt{5}\varphi^n\frac{1}{1 + \mathcal{O}(\varphi^{2n})} - \varphi^n\left(\{t\varphi\} + t\mathcal{O}(\varphi^{2n}) - \frac{1}{2}\right)\right)\right) \\ &= 2\sin\left(\pi t\varphi^n\left(\sqrt{5} - \frac{1}{t}\left(\{t\varphi\} - \frac{1}{2}\right) + \mathcal{O}(\varphi^{2n})\right)\right). \end{aligned} \tag{2.43}$$

Combining (2.41) with (2.43) and $(\pi/2)\sin(x) \geq x$ for $x \in (0, \pi/2)$ we obtain

$$\begin{aligned} \left|\frac{v_n(\varepsilon)}{s_{nt}}\right| &\leq \frac{\pi\varphi^n\sqrt{5}(1 + \mathcal{O}(\varphi^{2n}))}{4\varphi^n t\left|\sqrt{5} - t^{-1}\left(\{t\varphi\} - \frac{1}{2}\right) + \mathcal{O}(\varphi^{2n})\right|} \\ &\leq \frac{\sqrt{5}\pi(1 + \mathcal{O}(\varphi^{2n}))}{4\eta_n\left|\sqrt{5} - \eta_n^{-1}\left(\{t\varphi\} - \frac{1}{2}\right) + \mathcal{O}(\varphi^{2n})\right|} \\ &= \frac{\pi(1 + \mathcal{O}(\varphi^{2n}))}{4\eta_n(1 + \mathcal{O}(\eta_n^{-1}))} = \mathcal{O}(\eta_n^{-1}). \end{aligned} \tag{2.44}$$

Observe that if n is sufficiently large (2.44) implies that $\sum_{t=\eta_n}^{(F_n-1)/2} v_n^2/s_{nt}^2 < 1$ and therefore we are able to apply Lemma 2.2.12 and obtain

$$1 \geq \prod_{t=\eta_n}^{(F_n-1)/2}\left(1 - \frac{v_n^2}{s_{nt}^2}\right) \geq 1 - \sum_{t=\eta_n}^{(F_n-1)/2}\frac{v_n^2}{s_{nt}^2} = 1 - \mathcal{O}(\varphi^{n/5}). \tag{2.45}$$

Now consider the terms in (2.39) where $t < \eta_n$. Using again (2.42) and $\sin(x) = x + \mathcal{O}(x^3)$ we get for sufficiently large n that

$$\begin{aligned} s_{nt} &= 2\pi \left(\frac{t}{F_n} - \varphi^n \left(\left\{ \frac{tF_{n-1}}{F_n} \right\} - \frac{1}{2} \right) \right) + \mathcal{O}(t^3\varphi^{3n}) \\ &= 2\pi\varphi^n \left(\frac{\sqrt{5}t}{1+\mathcal{O}(\varphi^{2n})} - \left\{ t\varphi + t\mathcal{O}(\varphi^{2n}) \right\} + \frac{1}{2} \right) + \mathcal{O}(\varphi^{n/5}) \\ &= 2\pi\varphi^n \left(\sqrt{5}t(1+\mathcal{O}(\varphi^{2n})) - \{t\varphi\} + \frac{1}{2} + t\mathcal{O}(\varphi^{2n}) \right) + \mathcal{O}(\varphi^{n/5}) \\ &= \pi\varphi^n(u_t + \mathcal{O}(\varphi^{n/5})), \end{aligned}$$

with u_t given in (2.38), and combined with (2.41) this implies

$$\begin{aligned} \left| \frac{v_n(\varepsilon)}{s_{nt}} \right| &= \frac{\pi\varphi^n \left|1 - 2\varepsilon(-\varphi)^{-n}\right| (1+\mathcal{O}(\varphi^{2n}))}{\pi\varphi^n(u_t + \mathcal{O}(\varphi^{n/5}))} \\ &= \frac{\left|1 - 2\varepsilon(-\varphi)^{-n}\right|}{u_t} \, \frac{1+\mathcal{O}(\varphi^{2n})}{1+\mathcal{O}(\varphi^{n/5})} \\ &= \frac{\left|1 - 2\varepsilon(-\varphi)^{-n}\right|}{u_t}(1+\mathcal{O}(\varphi^{n/5})). \end{aligned}$$

It follows that

$$\begin{aligned} \prod_{t=1}^{\eta_n} \left(1 - \frac{v_n(\varepsilon)^2}{s_{nt}^2} \right) &= \prod_{t=1}^{\eta_n} \left(1 - \frac{(1-2\varepsilon(-\varphi)^{-n})^2}{u_t^2} - \frac{\mathcal{O}(\varphi^{n/5})}{u_t^2} \right) \\ &= \prod_{t=1}^{\eta_n} \left(1 - \frac{(1-2\varepsilon(-\varphi)^{-n})^2}{u_t^2} \right) \qquad (2.46) \\ &\quad \times \prod_{t=1}^{\eta_n} \left(1 - \frac{\mathcal{O}(\varphi^{n/5})}{u_t^2 - (1-2\varepsilon(-\varphi)^{-n})^2} \right). \qquad (2.47) \end{aligned}$$

Before we now evaluate the two subproducts (2.46) and (2.47) separately note that

$$\begin{aligned} (1-2\varepsilon(-\varphi)^{-n})^2 \sum_{t=\eta_n+1}^{\infty} \frac{1}{u_t^2} &\leq 5 \sum_{t=1}^{\infty} \frac{1}{u_t^2} \leq 5 \left(\frac{1}{u_1^2} + \sum_{t=2}^{\infty} \frac{1}{20(t-1)^2} \right) \\ &= 5 \left(\frac{1}{(\sqrt{5}+1)^2} + \frac{\pi^2}{120} \right) < \frac{5}{7}. \qquad (2.48) \end{aligned}$$

Consider the subproduct (2.47) and observe that by (2.48) we are capable of applying Lemma 2.2.12 and we then have

$$
\begin{aligned}
0 &\le \prod_{t=\eta_n+1}^{\infty}\left(1-\frac{(1-2\varepsilon(-\varphi)^{-n})^2}{u_t^2}\right) \\
&\le \left(1-(1-2\varepsilon(-\varphi)^{-n})^2\sum_{t=1}^{\infty}\frac{1}{u_{t+\eta_n}^2}\right)^{-1} \\
&\le \left(1-\frac{1}{4}\sum_{t=1}^{\infty}\frac{1}{(t+\eta_n-1)^2}\right)^{-1} = \left(1-\mathcal{O}(\eta_n^{-1})\right)^{-1},
\end{aligned}
$$

and thus

$$
\prod_{t=1}^{\eta_n}\left(1-\frac{(1-2\varepsilon(-\varphi)^{-n})^2}{u_t^2}\right) \ge \prod_{t=1}^{\infty}\left(1-\frac{(1-2\varepsilon(-\varphi)^{-n})^2}{u_t^2}\right)\left(1-\mathcal{O}(\eta_n^{-1})\right). \quad (2.49)
$$

Now consider the second subproduct (2.47). Using the bound (2.40), it is easily checked that

$$
\sum_{t=1}^{\infty}\frac{1}{u_t^2-(1-2\varepsilon(-\varphi)^{-n})^2}<\infty.
$$

Thus, for sufficiently large n, we can use Lemma 2.2.12 to conclude that

$$
\begin{aligned}
\prod_{t=1}^{\eta_n}\left(1-\frac{\mathcal{O}(\varphi^{n/5})}{u_t^2-(1-2\varepsilon(-\varphi)^{-n})^2}\right) &> 1-\mathcal{O}(\varphi^{n/5})\sum_{t=1}^{\infty}\frac{1}{u_t^2-(1-2\varepsilon(-\varphi)^{-n})^2} \\
&= 1-\mathcal{O}(\varphi^{n/5}). \quad (2.50)
\end{aligned}
$$

Inserting the bounds (2.49) and (2.50) for the subproducts (2.46) and (2.47), respectively, we get

$$
\prod_{t=1}^{\eta_n}\left(1-\frac{v_n(\varepsilon)^2}{s_{nt}^2}\right) \ge \prod_{t=1}^{\infty}\left(1-\frac{(1-2\varepsilon(-\varphi)^{-n})^2}{u_t^2}\right)\left(1-\mathcal{O}(\eta_n^{-1})\right)\left(1-\mathcal{O}(\varphi^{n/5})\right). \quad (2.51)
$$

Finally, inserting (2.45) and (2.51) in (2.39), and recalling that $\eta_n=\lceil\varphi^{-3n/5}\rceil$, we get

$$
\overline{C}_n(\varepsilon) = \prod_{t=1}^{\eta_n}\left(1-\frac{v_n(\varepsilon)^2}{s_{nt^2}}\right)\prod_{t=\eta_n+1}^{(F_n-1)/2}\left(1-\frac{v_n(\varepsilon)^2}{s_{nt^2}}\right)
$$

$$\geq \prod_{t=1}^{\infty}\left(1-\frac{(1-2\varepsilon(-\varphi)^{-n})^2}{u_t^2}\right)-\mathcal{O}(\varphi^{n/5}).$$

□

We are now equipped to bound $\overline{C}_{n_j}(\varepsilon_j)$ from below.

Proof of Lemma 2.2.11. For $n = n_j$ and $\varepsilon = \varepsilon_j = -\sum_{s=j+1}^{m}(-\varphi)^{n_s}$, we have $|\varepsilon| \leq \varphi^{n_j+1}$, and thus by Lemma 2.2.13 we get

$$\overline{C}_{n_j}(\varepsilon_j) \geq \prod_{t=1}^{\infty}\left(1-\frac{(1+2p_j)^2}{u_t^2}\right)-\mathcal{O}(\varphi^{n_j/5}),$$

with p_j given in (2.34). Recall that $p_j \in [-\varphi^2, \varphi]$, and thus $(1+2p_j)^2 \leq 5$. Moreover, from (2.48) it also follows that $(1+2p_j)^2\sum_{t=1}^{\infty} u_t^{-2} < 1$. Thus, we may apply Lemma 2.2.12 to obtain

$$\overline{C}_{n_j}(\varepsilon_j) \geq 1-\sum_{t=1}^{\infty}\frac{(1+2p_j)^2}{u_t^2}-\mathcal{O}(\varphi^{n_j/5}) > 1-\frac{1}{7}(1+2p_j)^2-\mathcal{O}(\varphi^{n_j/5}).$$

□

Main proof

Let us now confirm that Lemma 2.2.7 indeed follows from Lemmas 2.2.9, 2.2.10 and 2.2.11.

Proof of Lemma 2.2.7. We recall that our goal is to show that

$$P_{F_{n_j}}(\varphi, \varepsilon_j) \geq 1 \tag{2.52}$$

whenever n_j is sufficiently large. We have seen that

$$P_{F_{n_j}}(\varphi, \varepsilon_j) = \overline{A}_{n_j}(\varepsilon_j)B_{n_j}\overline{C}_{n_j}(\varepsilon_j), \tag{2.53}$$

where $\overline{A}_{n_j}(\varepsilon_j)$, B_{n_j} and $\overline{C}_{n_j}(\varepsilon_j)$ are defined in (2.27), (2.19) and (2.28). By Lemmas 2.2.9, 2.2.10 and 2.2.11 we get

$$P_{F_{n_j}}(\varphi, \varepsilon_j) > \varphi^{-2}(1+p_j)\left(1-\frac{1}{7}(1+2p_j)^2\right)\frac{P_{F_{n_j+2}}(\varphi)}{P_{F_{n_j+1}}(\varphi)}\left(1-\mathcal{O}(\varphi^{n_j/5})\right)$$

with p_j given in (2.34). Consider the function

$$g(x) = (1+x)\left(1 - \frac{1}{7}(1+2x)^2\right).$$

It is easily checked that for $x \in [-\varphi^2, \varphi]$, this function satisfies $g(x) > 5/11$. Additionally by the main result of [104] we know that $\lim_{n\to\infty} P_{F_{n+2}}(\varphi)/P_{F_{n+1}}(\varphi) = 1$. So there exists $S \in \mathbb{N}$ such that

$$g(p_j)\frac{P_{F_{n_j+2}}(\varphi)}{P_{F_{n_j+1}}(\varphi)}\left(1 - \mathcal{O}(\varphi^{n_j/5})\right) > \frac{5}{12}$$

whenever $n_j \geq S$. Hence, we obtain that

$$P_{F_{n_j}}(\varphi, \varepsilon_j) > \varphi^{-2}\frac{5}{12} > 1 \text{ for all } n_j \geq S.$$

and in particular, this means that (2.52) holds for all $j \geq J = S/2$ $\square$

Finally, we recall that Theorem 2.2.2 is a consequence of Lemma 2.2.7:

Proof of Theorem 2.2.2. Let N be any natural number, and let $N = \sum_{j=1}^{m} F_{n_j}$ be its unique Zeckendorf representation. We rewrite $P_N(\varphi)$ as

$$P_N(\varphi) = \prod_{r=1}^{N} |2\sin(\pi r\varphi)| = \prod_{j=1}^{m} P_{F_{n_j}}(\varphi, \varepsilon_j),$$

with ε_j given in (2.31).

Assume first that the length of the Zeckendorf representation of N is smaller than the bound J in Lemma 2.2.7, i.e. $m \leq J$. In this case it follows from Lemma 2.2.6 that

$$P_N(\varphi) \geq K_1^m \geq K_1^J. \tag{2.54}$$

for some $0 < K_1 \leq 1$.

Suppose now that $m > J$. Then by Lemmas 2.2.6 and 2.2.7, we have

$$P_N(\varphi) = \left(\prod_{j=1}^{J} P_{F_{n_j}}(\varphi, \varepsilon_j)\right)\left(\prod_{j=J+1}^{m} P_{F_{n_j}}(\varphi, \varepsilon_j)\right) \geq K_1^J \cdot 1^{m-J} \geq K_1^J. \tag{2.55}$$

Combining (2.54) and (2.55) we have $P_N(\varphi) \geq K_1^J$ for all N, where $J \in \mathbb{N}$ and $K_1 > 0$ are absolute constants. It follows that

$$\liminf_{N\to\infty} P_N(\varphi) \geq K_1^J > 0.$$

This finishes the proof of Theorem 2.2.2. $\square$

2.3 Asymptotic behaviour of the Sudler product of sines for quadratic irrationals

We will continue the analysis of the asymptotic behaviour of the Sudler product $P_N(\alpha)$ in this section. This time we put the focus at certain subsequences of the form

$$Q_n(\alpha) := \prod_{r=1}^{q_n} |2\sin(\pi r\alpha)|, \tag{2.56}$$

where $(q_n)_{n\geq 0}$ are the best approximation denominators of α. In a recent paper, Verschueren and Mestel [104] study $Q_n(\alpha)$ in the special case where $\alpha = \varphi = (\sqrt{5}-1)/2$ is the fractional part of the golden mean. For this case, it was suggested by Knill and Tangerman in [57] that the limit value $\lim_{n\to\infty} Q_n(\varphi)$ might exist, and this is confirmed by Verschueren and Mestel.

Theorem 2.3.1 ([104, Theorem 2.2])**.** *If φ denotes the golden mean and $(F_n)_{n\geq 1} = (1,1,2,3,5,\ldots)$ the Fibonacci sequence, then there exists a constant $c > 0$ such that*

$$\lim_{n\to\infty} Q_n(\varphi) = \lim_{n\to\infty} \prod_{r=1}^{F_n} |2\sin(\pi r\varphi)| = c.$$

Verschueren and Mestel [104] conjecture that Theorem 2.3.1 can be extended to all quadratic irrationals. More precisely, they suggest that if the continued fraction expansion of α has period ℓ, then the subsequence $Q_n(\alpha)$ will converge to a periodic sequence whose period length divides ℓ. Our main goal is to verify this claim.

Theorem 2.3.2. *Suppose α has a purely periodic continued fraction expansion $\alpha = [0;\overline{a_1,\ldots,a_\ell}]$ with $a_1,\ldots,a_\ell \in \mathbb{N}$ and period ℓ. Let $(q_n)_{n\geq 1}$ be the sequence of best approximation denominators of α. Then there exist positive constants c_0, $c_1,\ldots,c_{\ell-1}$ such that*

$$\lim_{m\to\infty} Q_{\ell m+k}(\alpha) = \lim_{m\to\infty} \prod_{r=1}^{q_{\ell m+k}} |2\sin(\pi r\alpha)| = c_k$$

for each $k = 0,1,2,\ldots,\ell-1$.

Corollary 2.3.3. *Suppose β has continued fraction expansion of the form $\beta = [a_0;a_1,\ldots,a_h,\overline{a_{h+1},\ldots,a_{h+\ell}}]$ and let $\alpha = [0;\overline{a_{h+1},\ldots,a_{h+\ell}}]$. We then have*

$$\lim_{m\to\infty} Q_{h+\ell m+k}(\beta) = \lim_{m\to\infty} Q_{\ell m+k}(\alpha).$$

The proof of Theorem 2.3.2 (and Corollary 2.3.3) largely follows that given by Verschueren and Mestel for the special case of the golden mean. Nevertheless, we include the proof in full detail for the sake of completeness. We emphasise that the challenge in generalising Theorem 2.3.1 to all quadratic irrationals lies in finding appropriate analogues for $(q_n)_{n\geq 1}$ of certain special properties of the Fibonacci sequence $(F_n)_{n\geq 1} = (1, 1, 2, 3, 5, 8, 13, \ldots)$. Throughout their proof for the golden mean case, Verschueren and Mestel make heavy use of the identities

$$F_n \varphi^n = \frac{1}{\sqrt{5}} + \mathcal{O}(\varphi^{2n})$$

and

$$\frac{F_{n-1}}{F_n} = \varphi + \mathcal{O}(\varphi^{2n}),$$

which do not have obvious analogues for the more general case of a quadratic irrational α. However, we will see that similar identities can indeed be formulated for the sequence $(q_n)_{n\geq 1}$ of best approximation denominators of α, and with these established the proof of Verschueren and Mestel easily carries over.

The existence of $\lim_{m\to\infty} Q_{\ell m+k}$ claimed by Theorem 2.3.2 is verified by splitting the product $Q_{\ell m+k}$ into three more manageable products

$$\begin{aligned} Q_{\ell m+k} &= A_m B_m C_m \\ &= |2q_n \sin(\pi e_k b^m)| \cdot \left| \prod_{t=1}^{q_n-1} \frac{s_{mt}}{2\sin(\pi t/q_n)} \right| \cdot \prod_{t=1}^{q_n-1} \left(1 - \frac{s_{m0}^2}{s_{mt}^2}\right)^{1/2}, \end{aligned} \tag{2.57}$$

where $n = \ell m + k$, e_k is a k-dependent constant, and s_{mt} is the generalised version of the sequence given in (2.21), which will be introduced later on in Section 2.3.3.

Remark 2.3.4. If we consider the special case where $\alpha = \varphi = [0; \overline{1}]$, i.e. $\ell = 1$, $k = 0$, $q_n(\alpha) = F_n$, $b = -\varphi$ and $e_0 = -1$, the decomposition of $Q_{\ell m+k}$ given above reduces to the decomposition of $P_{F_n}(\varphi)$ as stated in (2.17). We will also see in Section 2.3.3 that s_{mt} and the sequence given in (2.21) are connected in the same vein.

The decomposition (2.57) is explained in detail in Section 2.3.3, where we also show the straightforward convergence of A_m as $m \to \infty$. The convergence of B_m and C_m is more involved, and is therefore treated in subsequent Sections 2.3.4 and 2.3.5. Prior to this, in Section 2.3.2, we establish analogues for $(q_n)_{n\geq 1}$ of the above-mentioned Fibonacci identities. In particular, we point out a connection to so-called Lehmer sequences, which we consider to be of independent interest (see Theorem 2.3.13). Finally, we summarize the proofs of Theorem 2.3.2 and Corollary 2.3.3 in Section 2.3.6. First, however, we introduce necessary notation and some general theory on continued fraction expansions in the following section.

2.3.1 Preliminaries

Following Verschueren and Mestel [104], we stick to the generalised sum and product notation which we have already introduced in Section 2.2 (see p.78).

Permutation operators

Whenever we have an ℓ-dimensional, integer-valued vector, e.g. $\boldsymbol{d} = (d_1, \ldots, d_\ell) \in \mathbb{N}^\ell$, we use the corresponding greek letter (in this case δ) to denote the real number with continued fraction expansion $\delta = [0; \overline{d_1, \ldots, d_\ell}]$.

We introduce two families of permutation operators acting on $\mathbb{N}^\ell$: Let $\tau_u : \mathbb{N}^\ell \to \mathbb{N}^\ell$ be defined by

$$\tau_u(\boldsymbol{d}) := (d_{u+1}, \ldots, d_\ell, d_1, \ldots, d_u), \quad u \in \{0, 1, \ldots, \ell - 1\}, \tag{2.58}$$

and similarly $\sigma_u : \mathbb{N}^\ell \to \mathbb{N}^\ell$ be defined by

$$\sigma_u(\boldsymbol{d}) := (d_{u-1}, \ldots, d_1, d_\ell, \ldots, d_u), \quad u \in \{2, 3, \ldots, \ell - 1\}, \tag{2.59}$$

with $\sigma_0(\boldsymbol{d}) = (d_{\ell-1}, \ldots, d_1, d_\ell)$ and $\sigma_1(\boldsymbol{d}) = (d_\ell, \ldots, d_1)$. Moreover, we use δ_{τ_u} and δ_{σ_u} to denote the real numbers with periodic continued fraction expansions given by $\tau_u(\boldsymbol{d})$ and $\sigma_u(\boldsymbol{d})$, respectively. That is, we write

$$\delta_{\tau_u} = [0; \overline{d_{u+1}, \ldots, d_\ell, d_1, \ldots, d_u}] \tag{2.60}$$

and

$$\delta_{\sigma_u} = [0; \overline{d_{u-1}, \ldots, d_1, d_\ell, \ldots, d_u}]. \tag{2.61}$$

Our motivation for introducing the operator τ_u is explained by Lemma 2.3.7 in the following subsection. The need to introduce σ_u is less evident, but will be clear from Lemma 2.3.18 in Section 2.3.2, where we describe the asymptotic behaviour of the sequence of denominator quotients $(q_{n-1}/q_n)_{n\geq 1}$ for a quadratic irrational number.

Continued fraction expansions

We briefly review some facts about continued fraction expansions of real numbers. In general, for any irrational, real $\alpha \in (0, 1)$ whose continued fraction expansion is given by

$$[0; a_1, a_2, \ldots],$$

we denote its nth convergent by p_n/q_n. The numerators p_n and denominators q_n are given recursively by

$$\begin{aligned} q_0 &= 0, \quad q_1 = 1 \qquad q_{n+1} = a_n q_n + q_{n-1}; \\ p_0 &= 1, \quad p_1 = 0 \qquad p_{n+1} = a_n p_n + p_{n-1}. \end{aligned}$$

Note that the indexing of p_n and q_n is offset by one compared to what is normally seen in literature. As a consequence, the nth convergent p_n/q_n is smaller than α for every odd value of n, and greater than α for every even value of n. It follows readily from the recurrences above that

$$p_n q_{n+1} - p_{n+1} q_n = (-1)^n, \tag{2.62}$$

and as a consequence of this identity we have the error bound

$$\left| \alpha - \frac{p_n}{q_n} \right| < \frac{1}{q_{n+1} q_n} \tag{2.63}$$

for the nth convergent of α.

Remark. Whenever it is *not* clear from context, we write $p_n(\alpha)$ and $q_n(\alpha)$ to indicate that these are the best approximation numerators and denominators corresponding to the real number α.

Theorem 2.3.5 (Ostrowski representation)**.** *Let $\alpha \in (0,1)$ be an irrational number with continued fraction expansion $[0; a_1, a_2, \ldots]$ and best approximation denominators $(q_n)_{n \geq 1}$. Then every non-negative integer N has a unique expansion*

$$N = \sum_{n=1}^{z} v_n q_n, \tag{2.64}$$

where:

i) $0 \leq v_1 \leq a_1 - 1$ and $0 \leq v_n \leq a_n$ for $n > 1$.

ii) If $v_n = a_n$ for some n, then $v_{n-1} = 0$.

iii) $z = z(N) = \mathcal{O}(\log(N))$

We refer to (2.64) *as the Ostrowski representation of N in base α.*

The proof of Theorem 2.3.5 can for example be found in [64, p. 126].

Remark 2.3.6. Observe that for the special case where $\alpha = \varphi = [0; \overline{1}]$ and $q_n(\varphi) = F_n$ the Ostrowski representation of N in base φ is exactly the Zeckendorf representation of N (see Definition 2.2.4).

Periodic continued fraction expansions

Suppose now that α is an irrational with ℓ-periodic continued fraction expansion $\alpha = [0; \overline{a_1, \dots, a_\ell}]$. For this special case, further properties of the convergents p_n/q_n of α can be established. The following lemma summarizes useful relations for $(p_n)_{n\geq 0}$ and $(q_n)_{n\geq 0}$ established by Perron in [86, p. 14–17].

Lemma 2.3.7. *Let* $\alpha = [0; \overline{a_1, \dots, a_\ell}]$ *and for every* $k \in \{0, \dots, \ell-1\}$ *let* τ_k *be defined as in* (2.58)*.*

(a) For all $n, m \in \mathbb{N}$*, we have*

$$p_{n+m}(\alpha)q_m(\alpha) - p_m(\alpha)q_{n+m}(\alpha) = (-1)^{m-1}q_n(\alpha_{\tau_{m \bmod \ell}}),$$

where α_{τ_k} *is defined in* (2.60)*.*

(b) For all $r \in \mathbb{N}_0$*, we have*

$$q_{\ell+r}(\alpha) = q_{\ell+1}(\alpha)q_r(\alpha) + q_\ell(\alpha)p_r(\alpha)$$

and

$$p_{\ell+r}(\alpha) = p_\ell(\alpha)p_r(\alpha) + p_{\ell+1}(\alpha)q_r(\alpha).$$

(c) For every $k \in \{1, 2, \dots, \ell-1\}$*, we have*

$$q_{\ell-1}(\alpha_{\tau_k}) = p_\ell(\alpha_{\tau_{k-1}}).$$

Let us now associate to $\alpha = [0; \overline{a_1, \dots a_\ell}]$ the constant

$$c(\alpha) := q_{\ell+1}(\alpha) + p_\ell(\alpha). \tag{2.65}$$

This constant will play an important role as we go forward. As a first application, it appears in the following recursion formula for $(q_n)_{n\geq 0}$.

Lemma 2.3.8. *Let* $\alpha = [0; \overline{a_1, \dots, a_\ell}]$ *and let* $(q_n)_{n\geq 0}$ *be the sequence of best approximation denominators of* α*. For all* $n \geq 2\ell$ *we have*

$$q_n = c(\alpha)q_{n-\ell} + (-1)^{\ell-1}q_{n-2\ell}, \tag{2.66}$$

with $c(\alpha)$ *given in* (2.65)*.*

Proof by induction. For $n = 2\ell$, the right hand side in (2.66) reads

$$c(\alpha)q_\ell + (-1)^{\ell-1}q_0 = q_\ell q_{\ell+1} + q_\ell p_\ell = q_{2l}$$

according to Lemma 2.3.7(b) with $r = \ell$. So (2.66) holds for $n = 2\ell$.

Now let $n = 2\ell + 1$. The right hand side in (2.66) then reads

$$c(\alpha)q_{\ell+1} + (-1)^{\ell-1}q_1 = q_{\ell+1}^2 + p_\ell q_{\ell+1} + (-1)^{\ell-1}q_1 = q_{\ell+1}^2 + p_{\ell+1}q_\ell,$$

where the last equality follows from Lemma 2.3.7(a) with $m = \ell$ and $n = 1$. Again, using Lemma 2.3.7(b) with $r = \ell + 1$, we have

$$q_{2\ell+1} = q_{\ell+1}^2 + p_{\ell+1}q_\ell,$$

so (2.66) holds for $n = 2\ell + 1$.

For general $n > 2\ell + 1$, we have

$$q_{n+1} = a_n q_n + q_{n-1} = a_{n \bmod \ell} q_n + q_{n-1},$$

where we understand a_0 as a_ℓ. Using the induction hypothesis for q_n and q_{n-1}, and the fact that $n \bmod \ell = (n - \ell) \bmod \ell = (n - 2\ell) \bmod \ell$, we get

$$\begin{aligned} q_{n+1} &= c(\alpha)\left(a_{n \bmod \ell} q_{n-\ell} + q_{n-\ell-1}\right) + (-1)^{\ell-1}\left(a_{n \bmod \ell} q_{n-2\ell} + q_{n-2\ell-1}\right) \\ &= c(\alpha)q_{n+1-\ell} + (-1)^{\ell-1}q_{n+1-2\ell}. \end{aligned}$$

This completes the proof of Lemma 2.3.8. □

We complete this section by observing that the constant $c(\alpha)$ is, in a sense, independent of the permutation operators τ_u and σ_u introduced in (2.58) and (2.59), respectively.

Lemma 2.3.9. *Let $\alpha = [0; \overline{a_1, \ldots, a_\ell}]$, and let $c(\alpha)$ be given in* (2.65)*. Moreover, let α_{τ_u} and α_{σ_u} be defined as in* (2.60) *and* (2.61)*. We have*

$$c(\alpha) = c(\alpha_{\tau_u}) = c(\alpha_{\sigma_u}), \tag{2.67}$$

for every $u \in \{0, 1, \ldots, \ell - 1\}$.

Proof. Given an integer vector $\boldsymbol{d} = (d_1, \ldots, d_\ell) \in \mathbb{N}^\ell$ and the corresponding irrational number $\delta = [0; \overline{d_1, \ldots, d_\ell}]$, we define two sequences of matrices $A_n(\boldsymbol{d}), B_n(\boldsymbol{d}) \in \mathbb{N}^{n\times n}$, where $A_0(\boldsymbol{d}) := \begin{pmatrix}1\end{pmatrix}$, $A_1(\boldsymbol{d}) := \begin{pmatrix}d_1\end{pmatrix}$, $B_0(\boldsymbol{d}) := \begin{pmatrix}1\end{pmatrix}$, $B_1(\boldsymbol{d}) := \begin{pmatrix}0\end{pmatrix}$ and

$$A_n(\boldsymbol{d}) = \begin{pmatrix} d_1^{(\ell)} & -1 & 0 & \dots & 0 \\ 1 & d_2^{(\ell)} & -1 & \ddots & \vdots \\ 0 & 1 & d_3^{(\ell)} & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -1 \\ 0 & \dots & 0 & 1 & d_n^{(\ell)} \end{pmatrix}, \; B_n(\boldsymbol{d}) = \begin{pmatrix} 0 & -1 & 0 & \dots & 0 \\ 1 & d_1^{(\ell)} & -1 & \ddots & \vdots \\ 0 & 1 & d_2^{(\ell)} & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -1 \\ 0 & \dots & 0 & 1 & d_{n-1}^{(\ell)} \end{pmatrix},$$

with $d_j^{(\ell)} := d_{j \bmod \ell}$, and where we understand d_0 as d_ℓ. It can be verified that

$$q_{n+1}(\delta) = \det A_n(\boldsymbol{d}) \quad \text{and} \quad p_n(\delta) = \det B_n(\boldsymbol{d}),$$

where $p_n(\delta)/q_n(\delta)$ is the nth convergent of δ (see e.g. [86, p. 10–11]). In particular, for every $u \in \{0, 1, \dots, \ell - 1\}$, we have

$$c(\alpha_{\tau_u}) = \det A_\ell\left(\tau_u(\boldsymbol{a})\right) + \det B_\ell\left(\tau_u(\boldsymbol{a})\right)$$

and

$$c(\alpha_{\sigma_u}) = \det A_\ell\left(\sigma_u(\boldsymbol{a})\right) + \det B_\ell\left(\sigma_u(\boldsymbol{a})\right),$$

where $\boldsymbol{a} = (a_1, \dots, a_\ell)$.

Let us first show that

$$c(\alpha_{\tau_u}) = c(\alpha) \quad \text{for} \quad u = 0, 1, \dots, \ell - 1. \tag{2.68}$$

This is clearly the case when $u = 0$, as τ_0 is the identity operator on $\mathbb{N}^\ell$ and $\alpha_{\tau_0} = \alpha$. For $u > 0$, one can show that

$$\det A_\ell\left(\tau_u(\boldsymbol{a})\right) - \det B_\ell\left(\tau_{u+1}(\boldsymbol{a})\right) = \det A_\ell\left(\tau_{u+1}(\boldsymbol{a})\right) - \det B_\ell\left(\tau_u(\boldsymbol{a})\right).$$

(This is attained by taking the Laplace expansion along appropriate rows and columns of the matrices above.) It follows that

$$\begin{aligned} c(\alpha_{\tau_{u+1}}) &= \det A_\ell\left(\tau_{u+1}(\boldsymbol{a})\right) + \det B_\ell\left(\tau_{u+1}(\boldsymbol{a})\right) \\ &= \det A_\ell\left(\tau_u(\boldsymbol{a})\right) + \det B_\ell\left(\tau_u(\boldsymbol{a})\right) = c(\alpha_{\tau_u}), \end{aligned}$$

and thus (2.68) holds.

Now let us verify that

$$c(\alpha_{\sigma_u}) = c(\alpha) \quad \text{for} \quad u = 0, 1, \dots, \ell - 1. \tag{2.69}$$

We note first that the operator σ_u (for $u \neq 0$) can be expressed as a composition of τ_u and σ_0; namely

$$\sigma_u(\boldsymbol{a}) = \sigma_0(\tau_u(\boldsymbol{a})), \quad u \in \{1, 2, \ldots, \ell - 1\}.$$

Thus, if we can verify that

$$c(\alpha_{\sigma_0}) = c(\alpha), \tag{2.70}$$

then the general case (2.69) will follow from (2.68). In fact, for σ_0 it is easily seen that

$$\det A_\ell(\sigma_0(\boldsymbol{a})) = \det A_\ell(\tau_{\ell-1}(\boldsymbol{a})),$$

and by Laplace expansion one can verify that also

$$\det B_\ell(\sigma_0(\boldsymbol{a})) = \det B_\ell(\tau_{\ell-1}(\boldsymbol{a})).$$

It follows that $c(\alpha_{\sigma_0}) = c(\alpha_{\tau_{\ell-1}})$, which by (2.68) implies (2.70). This completes the proof of Lemma 2.3.9. □

2.3.2 Properties of the sequence $(q_n)_{n\geq 0}$

The main focus in this section is to attain a closed form for the sequence of best approximation denominators $(q_n)_{n\geq 0}$ for the irrational $\alpha = [0; \overline{a_1, \ldots, a_\ell}]$. We will see that having such a closed form enables us to formulate analogues of known properties for the Fibonacci sequence $(F_n)_{n\geq 0} = (0, 1, 1, 2, 3, 5, 8, \ldots)$, most notably of

$$F_n \varphi^n = \frac{1}{\sqrt{5}} + \mathcal{O}(\varphi^{2n}) \text{ for } n \geq 0; \tag{2.71}$$

$$\frac{F_{n-1}}{F_n} = \varphi + \mathcal{O}(\varphi^{2n}) \text{ for } n > 0, \tag{2.72}$$

where $\varphi = (\sqrt{5} - 1)/2$ is the fractional part of the golden mean. These two identities play a crucial role in the proof of Theorem 2.3.1 by Verschueren and Mestel. Likewise, the analogous identities for the sequence $(q_n)_{n\geq 0}$, formulated in Lemmas 2.3.14 and 2.3.18 below, will be important for the proof of Theorem 2.3.2.

A connection to Lehmer sequences

We begin by establishing a closed form for the sequence $(q_n)_{n\geq 0}$ of best approximation denominators of α. It turns out that this closed form can be expressed in terms of a *Lehmer sequence*. Lehmer sequences were first introduced in [72], and are defined as follows.

Definition 2.3.10. *Let $R, Q \in \mathbb{Z}$ with $R > 0$ and $R - 4Q > 0$. We define the Lehmer sequence $(L_n(R,Q))_{n\geq 0}$ with parameters R and Q by*

$$\begin{aligned} L_{2n}(R,Q) &:= L_{2n-1}(R,Q) - QL_{2n-2}(R,Q) \\ L_{2n+1}(R,Q) &:= RL_{2n}(R,Q) - QL_{2n-1}(R,Q) \\ L_0(R,Q) &= 0 \\ L_1(R,Q) &= 1. \end{aligned}$$

The closed form of the recurrence in Definition 2.3.10 is

$$L_n(R,Q) = \begin{cases} \dfrac{u^n - v^n}{u - v} & \text{if } n \text{ is odd,} \\ \dfrac{u^n - v^n}{u^2 - v^2} & \text{if } n \text{ is even,} \end{cases} \tag{2.73}$$

where u and v are the unique solutions of the equation $x^2 - \sqrt{R}x + Q = 0$.

We will consider only the Lehmer sequence with parameters $R = c(\alpha)^2$ and $Q = (-1)^\ell$. Accordingly, we write

$$L_n := L_n\left(c(\alpha)^2, (-1)^\ell\right)$$

from now on. If we introduce the constants

$$a = a(\alpha) := \frac{c(\alpha) + \sqrt{c(\alpha)^2 + 4(-1)^{\ell-1}}}{2}; \tag{2.74}$$

$$b = b(\alpha) := \frac{c(\alpha) - \sqrt{c(\alpha)^2 + 4(-1)^{\ell-1}}}{2}, \tag{2.75}$$

for the two distinct solutions of $x^2 - c(\alpha)x + (-1)^\ell = 0$, then it follows from (2.73) that

$$L_n = \begin{cases} \dfrac{a^n - b^n}{a - b} & \text{if } n \text{ is odd,} \\ \dfrac{a^n - b^n}{a^2 - b^2} & \text{if } n \text{ is even.} \end{cases} \tag{2.76}$$

By straightforward calculations one can verify that

$$ab = (-1)^\ell \quad \text{and} \quad a + b = c(\alpha).$$

Moreover, we have $a > 1$, and $b \in (-1, 0)$ if ℓ is odd and $b \in (0, 1)$ if ℓ is even. Finally, as a consequence of Lemma 2.3.9, we have

$$a(\alpha) = a(\alpha_{\tau_u}) = a(\alpha_{\sigma_u});$$

$$b(\alpha) = b(\alpha_{\tau_u}) = b(\alpha_{\sigma_u}),$$

for every $u \in \{0, 1, \ldots, \ell - 1\}$, where we recall the definition of α_{τ_u} and α_{σ_u} from (2.60) and (2.61).

Let us now formulate a closed form of the sequence of best approximation denominators $(q_n)_{n\geq 0}$ for α. In fact, similar closed forms can be established for both $(q_n)_{n\geq 0}$ and $(p_n)_{n\geq 0}$.

Lemma 2.3.11. *For every $n = \ell m + k \geq 2\ell$, where $m \in \mathbb{N}$ and $k \in \{0, 1, \ldots, \ell - 1\}$, the approximation denominator q_n for $\alpha = [0; \overline{a_1, \ldots, a_\ell}]$ is given by*

$$q_n = q_{\ell m + k} = \frac{1}{a - b} \left((a^m - b^m) q_{\ell + k} + (-1)^{\ell - 1} (a^{m-1} - b^{m-1}) q_k \right),$$

where a and b are defined in (2.74) *and* (2.75).

Lemma 2.3.12. *For every $n = \ell m + k \geq 2\ell$, where $m \in \mathbb{N}$ and $k \in \{0, 1, \ldots, \ell - 1\}$, the approximation numerator p_n for $\alpha = [0; \overline{a_1, \ldots, a_\ell}]$ is given by*

$$p_n = p_{\ell m + k} = \frac{1}{a - b} \left((a^m - b^m) p_{\ell + k} + (-1)^{\ell - 1} (a^{m-1} - b^{m-1}) p_k \right),$$

where a and b are defined in (2.74) *and* (2.75).

It is a simple observation that Lemmas 2.3.11 and 2.3.12 may alternatively be formulated in terms of the Lehmer sequence (2.76). As we find this to be of independent interest, we formulate it as a theorem.

Theorem 2.3.13. *For every $n = \ell m + k \geq 2\ell$, where $m \in \mathbb{N}$ and $k \in \{0, 1, \ldots, \ell - 1\}$, the convergents p_n/q_n of $\alpha = [0; \overline{a_1, \ldots, a_\ell}]$ are given by*

$$q_n = q_{\ell m + k} = \gamma_1^{(m)} L_m q_{\ell + k} + (-1)^{\ell - 1} \gamma_2^{(m)} L_{m-1} q_k,$$

and

$$p_n = p_{\ell m + k} = \gamma_1^{(m)} L_m p_{\ell + k} + (-1)^{\ell - 1} \gamma_2^{(m)} L_{m-1} p_k,$$

where a and b are defined in (2.74) *and* (2.75), *L_m is the Lehmer sequence* (2.76), *and*

$$\gamma_1^{(m)} := \begin{cases} c(\alpha) & \text{if } m \text{ is even,} \\ 1 & \text{if } m \text{ is odd,} \end{cases} \qquad \text{and} \qquad \gamma_2^{(m)} = \begin{cases} 1 & \text{if } m \text{ is even,} \\ c(\alpha) & \text{if } m \text{ is odd.} \end{cases}$$

As the proofs of Lemmas 2.3.11 and 2.3.12 are nearly identical, we include only the former.

Proof of Lemma 2.3.11. By Lemma 2.3.8, we have the recursion formula

$$q_n = c(\alpha)q_{n-\ell} + (-1)^{\ell-1}q_{n-2\ell},$$

whenever $n \geq 2\ell$. The corresponding polynomial characteristic equation is

$$x^{2\ell} - c(\alpha)x^\ell + (-1)^\ell = 0. \tag{2.77}$$

Substituting $y = x^\ell$, we get the equation $y^2 - c(\alpha)y + (-1)^\ell = 0$, whose two solutions a and b are given in (2.74) and (2.75), respectively. Now let $e_\ell = e^{2\pi i/\ell}$. The 2ℓ unique solutions of (2.77) are

$$x_v = a^{1/\ell}e_\ell^v \quad \text{and} \quad x_{v+\ell} = b^{1/\ell}e_\ell^v \tag{2.78}$$

for $v = 1, 2, \ldots, \ell$. Accordingly, for an arbitrary $n \geq 2\ell$, the closed form of q_n is

$$q_n = \sum_{v=1}^{2\ell} c_v x_v^n, \tag{2.79}$$

where the constants $c_1, \ldots, c_{2\ell}$ are determined by the 2ℓ first terms $q_0, \ldots, q_{2\ell-1}$. For a given $j \in \{1, 2, \ldots, \ell\}$, inserting q_{j-1} in (2.79) yields

$$q_{j-1} = \sum_{v=1}^{2\ell} c_v x_v^{j-1} = a^{(j-1)/\ell}C_j^{(1)} + b^{(j-1)/\ell}C_j^{(2)},$$

where

$$C_j^{(1)} = \sum_{v=1}^{\ell} c_v e_\ell^{v(j-1)} \quad \text{and} \quad C_j^{(2)} = \sum_{v=\ell+1}^{2\ell} c_v e_\ell^{v(j-1)}.$$

Similarly, we have

$$q_{\ell+j-1} = a^{1+(j-1)/\ell}C_j^{(1)} + b^{1+(j-1)/\ell}C_j^{(2)}.$$

Thus, the system of 2ℓ equations determining the constants $c_1, \ldots, c_{2\ell}$ decouples into ℓ systems of 2 equations in the variables $(C_j^{(1)}, C_j^{(2)})$, with $j = 1, 2, \ldots, \ell$. Solving these ℓ systems, we get

$$\begin{aligned} C_j^{(1)} &= \frac{1}{a^{(j-1)/\ell}} \frac{bq_{j-1} - q_{\ell+j-1}}{b-a}; \\ C_j^{(2)} &= \frac{1}{b^{(j-1)/\ell}} \frac{q_{\ell+j-1} - aq_{j-1}}{b-a}. \end{aligned}$$

Finally, for $n = \ell m + k \geq 2\ell$, we thus have

$$\begin{aligned} q_n = q_{\ell m+k} &= \sum_{v=1}^{2\ell} c_v x_v^{\ell m+k} = a^{m+k/\ell} C_{k+1}^{(1)} + b^{m+k/\ell} C_{k+1}^{(2)} \\ &= \frac{1}{a-b} \left((a^m - b^m) q_{\ell+k} + (-1)^{\ell-1} (a^{m-1} - b^{m-1}) q_k \right), \end{aligned}$$

where in the final equality we have used that $ab = (-1)^\ell$. This completes the proof of Lemma 2.3.11. □

Asymptotic behaviour of q_n and q_{n-1}/q_n

In this section, we show that the closed form of $(q_n)_{n\geq 0}$ that was established in Lemma 2.3.11 can be used to formulate analogues of the Fibonacci identities (2.71) and (2.72) for the more general case of irrationals with a periodic continued fraction expansion. We begin with the simpler task of formulating an analogue of (2.71). Informally speaking, we will see that the constant b in (2.75) plays the role of the fractional part of the golden mean φ.

Lemma 2.3.14. *Let $\ell \in \mathbb{N}$ and $k \in \{0, 1, \ldots, \ell - 1\}$ be fixed integers, and let $(p_n/q_n)_{n\geq 1}$ be the convergents for $\alpha = [0; \overline{a_1, \ldots, a_\ell}]$. For all integers $m \geq 2$, we have*

$$q_{\ell m+k} |b|^m = c_k + \mathcal{O}(|b|^{2m}), \tag{2.80}$$

where

$$c_k := \frac{q_{\ell+k} - b q_k}{a-b}, \tag{2.81}$$

and a and b are given in (2.74) and (2.75). Thus, we have $q_{\ell m+k} = \Theta(|b|^{-m})$ for $m \to \infty$ (note that $|b| \in (0, 1)$).

Remark 2.3.15. Note that in the special case when $\alpha = \varphi = [0; \overline{1}]$ is the fractional part of the golden mean and $q_n = F_n$ is the Fibonacci sequence, we have $b = -\varphi$, and $c_k = c_0 = 1/\sqrt{5}$. Accordingly, Lemma 2.3.14 reduces to the Fibonacci identity (2.71) in this case.

Proof of Lemma 2.3.14. Recall again the closed form

$$q_{\ell m+k} = \frac{1}{a-b} \left((a^m - b^m) q_{\ell+k} + (-1)^{\ell-1} (a^{m-1} - b^{m-1}) q_k \right)$$

from Lemma 2.3.11. Multiplying both sides by $|b|^m$ and using that $ab = (-1)^\ell$, we get

$$q_{\ell m+k}|b|^m = \frac{1}{a-b}\left(q_{\ell+k} - bq_k + (-1)^{\ell m} b^{2m}(aq_k - q_{\ell+k})\right) = c_k + \mathcal{O}(|b|^{2m}),$$

with $c_k > 0$ as in (2.81). $\square$

We now aim to establish an analogue, or extension, of the Fibonacci identity (2.72). This identity, which plays a crucial role in the work of Verschueren and Mestel [104], is a consequence of the fact that F_{n-1}/F_n is the nth convergent of the golden mean φ. Naturally, we cannot expect the same identity to hold for the general case of irrationals with a periodic continued fraction expansion. However, we will see in Lemma 2.3.18 below that a similar identity can indeed be formulated. Needed for this lemma is the following estimation error of the nth convergent p_n/q_n of $\alpha = [0; \overline{a_1, \ldots, a_\ell}]$ in terms of the constant b in (2.75).

Lemma 2.3.16. *Let $n = \ell m + k \geq 2\ell$, where $m \in \mathbb{N}$ and $k \in \{0, 1, \ldots, \ell - 1\}$. Moreover, let $(p_n/q_n)_{n\geq 1}$ be the sequence of convergents for $\alpha = [0; \overline{a_1, \ldots, a_\ell}]$. We have that*

$$q_n\alpha = q_{\ell m+k}\alpha = p_n + e_k b^m, \tag{2.82}$$

where

$$e_k = \frac{(-1)^{k-1}}{q_\ell}\left|aq_k - q_{\ell+k}\right|, \tag{2.83}$$

and a and b are given in (2.74) *and* (2.75), *respectively.*

Remark 2.3.17. Since $q_n = q_{\ell m+k} = \Theta(|b|^{-m})$ for $m \to \infty$ holds by Lemma 2.3.14 it follows in particular from Lemma 2.3.16 that

$$\alpha = \frac{p_n}{q_n} + \mathcal{O}(|b|^{2m}) \text{ for } m \to \infty. \tag{2.84}$$

Proof of Lemma 2.3.16. As a preliminary step (note that this is not part of the statement), we show that (2.82) holds when $m = 1$ and $k = 0$, that is

$$q_\ell\alpha = p_\ell - b. \tag{2.85}$$

It is well known that α is a root of the polynomial $q_\ell x^2 + (q_{\ell+1} - p_\ell)x - p_{\ell+1}$ (see e.g. [86, p. 69]), and since $\alpha > 0$ we must have

$$\alpha = \frac{-q_{\ell+1} + p_\ell + \sqrt{(q_{\ell+1} - p_\ell)^2 + 4p_{\ell+1}q_\ell}}{2q_\ell}$$

$$= \frac{p_\ell}{q_\ell} + \frac{-(q_{\ell+1}+p_\ell) + \sqrt{(q_{\ell+1}+p_\ell)^2 - 4q_{\ell+1}p_\ell + 4p_{\ell+1}q_\ell}}{2q_\ell}.$$

Using that $c(\alpha) = q_{\ell+1} + p_\ell$ and $p_{\ell+1}q_\ell - q_{\ell+1}p_\ell = (-1)^{\ell-1}$, we thus get

$$q_\ell\alpha = p_\ell - \frac{c(\alpha) - \sqrt{c(\alpha)^2 + 4(-1)^{\ell-1}}}{2} = p_\ell - b.$$

Now let us see that (2.82) holds for all $n = \ell m + k \geq 2\ell$. We recall the closed forms

$$p_n = p_{\ell m+k} = \frac{1}{a-b}\left((a^m - b^m)p_{\ell+k} + (-1)^{\ell-1}(a^{m-1} - b^{m-1})p_k\right),$$

and

$$q_n = q_{\ell m+k} = \frac{1}{a-b}\left((a^m - b^m)q_{\ell+k} + (-1)^{\ell-1}(a^{m-1} - b^{m-1})q_k\right),$$

from Lemmas 2.3.12 and 2.3.11, respectively. Multiplying the latter with α, and using (2.85), we get

$$q_n\alpha = \frac{(p_\ell - b)}{q_\ell(a-b)}\left((a^m - b^m)q_{\ell+k} + (-1)^{\ell-1}(a^{m-1} - b^{m-1})q_k\right).$$

For ease of notation, let us write $A = a^m - b^m$ and $B = (-1)^{\ell-1}(a^{m-1} - b^{m-1})$. We then have

$$\begin{aligned} q_n\alpha - p_n &= \frac{1}{q_\ell(a-b)}\left((p_\ell - b)\left(Aq_{\ell+k} + Bq_k\right) - \left(Aq_\ell p_{\ell+k} + Bq_\ell p_k\right)\right) \\ &= \frac{1}{q_\ell(a-b)}\left(A(q_{\ell+k}p_\ell - bq_{\ell+k} - q_\ell p_{\ell+k}) + B(p_\ell q_k - bq_k - q_\ell p_k)\right). \end{aligned}$$

Now recall from Lemma 2.3.7 that

$$q_{\ell+k}p_\ell - q_\ell p_{\ell+k} = (-1)^\ell q_k$$

(part (a) with $m = \ell$ and $n = k$) and

$$p_k q_\ell = q_{\ell+k} - q_{\ell+1}q_k$$

(part (b) with $m = \ell$ and $n = k$). Inserting this above, we get

$$q_n\alpha - p_n = \frac{1}{q_\ell(a-b)}\left(A((-1)^\ell q_k - bq_{\ell+k}) + B(-q_{\ell+k} + q_k(q_{\ell+1} + p_\ell - b))\right)$$

$$= \frac{1}{q_\ell}\left(-\left(\frac{Ab+B}{a-b}\right)q_{\ell+k} + \left(\frac{(-1)^\ell A + Ba}{a-b}\right)q_k\right),$$

where we have used that $q_{\ell+1} + p_\ell - b = c(\alpha) - b = a$. From $ab = (-1)^\ell$, it follows that

$$\frac{Ab+B}{a-b} = b^m \quad \text{and} \quad \frac{(-1)^\ell A + Ba}{a-b} = ab^m.$$

Accordingly, we have

$$q_n\alpha - p_n = \frac{b^m}{q_\ell}(aq_k - q_{\ell+k}).$$

Finally, we know from the general theory of continued fractions that p_n/q_n is greater than α if n is even and smaller than α if n is odd. We thus get

$$\begin{aligned} q_n\alpha - p_n &= (-1)^{n-1}\left|\frac{b^m}{q_\ell}(aq_k - q_{\ell+k})\right| \\ &= (-1)^{\ell m+k-1}\left((-1)^\ell b\right)^m \frac{1}{q_\ell}\left|aq_k - q_{\ell+k}\right| = e_k b^m, \end{aligned}$$

with e_k given in (2.83). This completes the proof of Lemma 2.3.16. $\square$

With Lemma 2.3.16 established, we may now formulate the following analogue of the Fibonacci sequence (2.72) for the general case of irrationals with a periodic continued fraction expansion.

Lemma 2.3.18. *Let $\ell \in \mathbb{N}$ and $k \in \{0, 1, \ldots, \ell-1\}$ be fixed integers, and let $(p_n/q_n)_{n\geq 1}$ be the convergents for $\alpha = [0; \overline{a_1, \ldots, a_\ell}]$. We have that*

$$\frac{q_{\ell m+k-1}}{q_{\ell m+k}} = \frac{p_{\ell m+k}(\alpha_{\sigma_k})}{q_{\ell m+k}(\alpha_{\sigma_k})} = \alpha_{\sigma_k} + \mathcal{O}(|b|^{2m}), \tag{2.86}$$

where b is given in (2.75).

Remark 2.3.19. In the special case when $\alpha = \varphi = [0; \overline{1}]$ is the fractional part of the golden mean and $q_n = F_n$ is the Fibonacci sequence, we have $k = 0$, $\alpha_{\sigma_0} = \alpha = \varphi$ and $b = -\varphi$. Accordingly, Lemma 2.3.18 reduces to the Fibonacci identity (2.72) in this case.

Proof of Lemma 2.3.18. For ease of notation, we write $n = \ell m + k$. Let us first see that $q_{n-1}/q_n = p_n(\alpha_{\sigma_k})/q_n(\alpha_{\sigma_k})$. We treat only the case $k \geq 1$ (the case $k = 0$ is similar). On the one hand, we have

$$\frac{p_n(\alpha_{\sigma_k})}{q_n(\alpha_{\sigma_k})} = [0; \underbrace{a_{k-1}, \ldots, a_1, a_\ell, \ldots, a_k, \ldots, a_{k-1}, \ldots, a_1, a_\ell, \ldots, a_k}_{m \text{ times}}, a_{k-1}, \ldots, a_1].$$

On the other hand, using the recursion formula for q_n, we get

$$\frac{q_{n-1}}{q_n} = \frac{q_{\ell m+k-1}(\alpha)}{q_{\ell m+k}(\alpha)} = \cfrac{1}{a_{k-1} + \cfrac{q_{\ell m+k-2}}{q_{\ell m+k-1}}} = \cfrac{1}{a_{k-1} + \cfrac{1}{a_{k-2} + \cfrac{q_{\ell m+k-3}}{q_{\ell m+k-2}}}}$$
$$= \ldots = [0; a_{k-1}, \ldots, a_1, \underbrace{a_\ell, \ldots, a_1, \ldots, a_\ell, \ldots, a_1}_{m \text{ times}}].$$

Thus, these quotients are equal. Finally, it follows from (2.84) and the fact that $b(\alpha_{\sigma_k}) = b(\alpha)$ for every $k \in \{0, 1, \ldots, \ell - 1\}$ that

$$\frac{p_n(\alpha_{\sigma_k})}{q_n(\alpha_{\sigma_k})} = \alpha_{\sigma_k} + \mathcal{O}(|b(\alpha_{\sigma_k})|^{2m}) = \alpha_{\sigma_k} + \mathcal{O}(|b|^{2m}).$$

□

We conclude this section by a more thorough investigation of the constants c_k and e_k in (2.81) and (2.83). More specifically, we consider the absolute value of their product $|c_k e_k|$. This quantity will repeatedly appear in the proof of Theorem 2.3.2, and the following lemma on $|c_k e_k|$ will then be useful.

Lemma 2.3.20. *For c_k and e_k given in* (2.81) *and* (2.83)*, we have that*

$$|c_k e_k| = \frac{q_\ell(\alpha_{\tau_k})}{a - b} = \frac{q_\ell(\alpha_{\tau_k})}{c(\alpha_{\tau_k}) - 2b}\,, \tag{2.87}$$

with a and b given in (2.74) *and* (2.75)*. It follows that $|c_k e_k| < 1$ for each $k = 0, 1, \ldots, \ell - 1$.*

Proof. Recalling the definition of c_k and e_k, we have

$$|c_k e_k| = \frac{|(q_{\ell+k} - bq_k)(aq_k - q_{\ell+k})|}{q_\ell(a-b)} = \frac{|q_k\left(c(\alpha)q_{\ell+k} + (-1)^{\ell-1}q_k\right) - q_{\ell+k}^2|}{q_\ell(a-b)},$$

where we have used that $a + b = c(\alpha)$ and $ab = (-1)^\ell$. We now look at the numerator in this expression. Using the recursion formula in Lemma 2.3.8, and Lemma 2.3.7(b) (with $r = \ell + k$), we have

$$c(\alpha)q_{\ell+k} + (-1)^{\ell-1}q_k = q_{2\ell+k} = q_{\ell+1}q_{\ell+k} + q_\ell p_{\ell+k}.$$

Inserting this in the numerator, we get

$$|c_k e_k| = \frac{|q_\ell q_k p_{\ell+k} + q_{\ell+k}(q_k q_{\ell+1} - q_{\ell+k})|}{q_\ell(a-b)} = \frac{|q_k p_{\ell+k} - q_{\ell+k} p_k|}{(a-b)},$$

where for the last equality we have used that $q_k q_{\ell+1} - q_{\ell+k} = -q_\ell p_k$ (Lemma 2.3.7(b) with $r = k$). It now follows from Lemma 2.3.7(a) (with $m = k$ and $n = \ell$) and $c(\alpha) = c(\alpha_{\tau_k})$ that

$$|c_k e_k| = \frac{q_\ell(\alpha_{\tau_k})}{a-b} = \frac{q_\ell(\alpha_{\tau_k})}{c(\alpha) - 2b} = \frac{q_\ell(\alpha_{\tau_k})}{c(\alpha_{\tau_k}) - 2b},$$

which confirms (2.87).

From (2.87) it easily follows that $|c_k e_k| < 1$. For $\ell = 1$, we get

$$|c_0 e_0| = \frac{1}{\sqrt{a_1^2 + 4}} < 1.$$

For $\ell \geq 2$, we have $p_\ell(\alpha_{\tau_k}) \geq 1$, and accordingly

$$|c_k e_k| = \frac{q_\ell(\alpha_{\tau_k})}{q_{\ell+1}(\alpha_{\tau_k}) + p_\ell(\alpha_{\tau_k}) - 2b} < \frac{q_\ell(\alpha_{\tau_k})}{q_{\ell+1}(\alpha_{\tau_k}) - 1} \leq 1$$

for each $k = 0, 1, \ldots, \ell - 1$. This completes the proof of Lemma 2.3.20. □

2.3.3 Decomposing $Q_{\ell m+k}(\alpha)$

The aim of this section is to decompose the product of sines $Q_{\ell m+k}(\alpha)$ in (2.56) into three subproducts

$$\begin{aligned} Q_{\ell m+k}(\alpha) &= A_m B_m C_m \\ &= |2q_n \sin(\pi e_k b^m)| \cdot \left| \prod_{t=1}^{q_n - 1} \frac{s_{mt}}{2\sin(\pi t/q_n)} \right| \cdot \prod_{t=1}^{q_n-1} \left(1 - \frac{s_{m0}^2}{s_{mt}^2}\right)^{1/2}, \end{aligned} \tag{2.88}$$

where $n = \ell m + k$ and s_{mt} is a perturbed rational sine function defined in (2.89) below. This decomposition is achieved by substituting the identity $\alpha = p_n/q_n + e_k b^m/q_n$ from Lemma 2.3.16 into the definition of $Q_{\ell m+k}(\alpha)$, which in turn allows us to view $Q_{\ell m+k}(\alpha)$ as a perturbation of the rational sine product $\prod_{r=1}^{q_n-1} |2\sin(\pi r(p_n/q_n))|$. Accordingly, proving Theorem 2.3.2 is a matter of showing that these perturbations have a suitable behaviour. For the latter task, it is a disadvantage that the perturbations $re_k b^m/q_n$ do not sum up to zero. However, by a rebasing of the argument one can attain a set of shifted perturbations $e_k b^m(r/q_n - 1/2)$ which *do* sum up to zero, and this approach eventually leads to the decomposition above.

Important families of sequences

Before we decompose $Q_{\ell m+k}(\alpha)$ into subproducts A_m, B_m and C_m in Section 2.3.3, let us introduce certain families of sequences which enter into the decomposition. For integers $m \geq 1$ and $t \in \{0, 1, \ldots, q_{\ell m+k} - 1\}$, we define:

$$s_{mt} := 2\sin\left(\pi\left(\frac{t}{q_{\ell m+k}} - |e_k b^m|\left(\left\{\frac{tq_{\ell m+k-1}}{q_{\ell m+k}}\right\} - \frac{1}{2}\right)\right)\right) \tag{2.89}$$

$$\xi_{mt} := \left\{t\frac{q_{\ell m+k-1}}{q_{\ell m+k}}\right\} - \frac{1}{2} \tag{2.90}$$

$$\xi_{\infty t} := \{t\alpha_{\sigma_k}\} - \frac{1}{2} \tag{2.91}$$

$$h_{mt} := \cot\left(\frac{\pi t}{q_{\ell m+k}}\right)\sin(\pi|e_k b^m|\xi_{mt}) \tag{2.92}$$

$$h_{\infty t} := \frac{|c_k e_k|\xi_{\infty t}}{t}, \ (t \neq 0) \tag{2.93}$$

Combining (2.89) and (2.90), we get

$$s_{mt} = 2\sin\left(\pi\left(\frac{t}{q_{\ell m+k}} - \xi_{mt}|e_k b^m|\right)\right). \tag{2.94}$$

Remark 2.3.21. Note that if $\alpha = \varphi = [0; \overline{1}]$ we have that $\ell = 1$, $k = 0$, $q_n(\alpha) = F_n$, $b = -\varphi$ and $e_0 = -1$. In this special case s_{mt} reduces to the sequence given in (2.21).

It is clear from the definition of ξ_{mt} that $|\xi_{mt}| \leq 1/2$, and since $|b| < 1$, we recognize s_{mt} as the perturbation of a rational sine, where the perturbation tends to zero as m increases. As we have already seen, the sequence s_{mt} plays a crucial role in the decomposition of $Q_{\ell m+k}(\alpha)$.

The sequences h_{mt} and $h_{\infty t}$ will not enter the story until the convergence of the subproduct B_m is considered in Section 2.3.5. Nevertheless, we introduce them at this early stage.

Lemma 2.3.22. *Let s_{mt}, ξ_{mt}, $\xi_{\infty t}$, h_{mt} and $h_{\infty t}$ be the sequences given in* (2.89) *–*(2.93)*. We have that:*

(a) $s_{mt} = s_{m(q_{\ell m+k}-t)}$, $h_{mt} = h_{m(q_{\ell m+k}-t)}$ *and* $\xi_{mt} = -\xi_{m(q_{\ell m+k}-t)}$ *for all* $t \in \{0, 1, \ldots, q_{\ell m+k} - 1\}$.

(b) $s_{mt} > s_{m0}$ *for all* $t \in \{1, 2, \ldots, q_{\ell m+k} - 1\}$ *if* m *is sufficiently large.*

(c) *For $m \to \infty$ it follows that $\xi_{mt} - \xi_{\infty t} = t\mathcal{O}(b^{2m})$, and thus for any fixed $t \in \mathbb{N}$, we have*

$$\lim_{m\to\infty} \xi_{mt} = \xi_{\infty t}.$$

(d) *For $m \to \infty$ it follows that $h_{mt} - h_{\infty t} = t\mathcal{O}(b^{2m})$, and thus for any fixed $t \in \mathbb{N}$, we have*

$$\lim_{m\to\infty} h_{mt} = h_{\infty t}.$$

Proof. For ease of notation, let us again write $n = \ell m + k$.

We first verify (a). The fact that $\{-x\} = 1 - \{x\}$ for $x \in \mathbb{R} \setminus \mathbb{Z}$ immediately implies $\xi_{mt} = -\xi_{m(q_n-t)}$. Combining this with $\sin(\pi - x) = \sin x$, we get

$$\begin{aligned} s_{m(q_n-t)} &= 2\sin\left(\pi\left(\frac{q_n - t}{q_n} - \xi_{m(q_n-t)}|e_k b^m|\right)\right) \\ &= 2\sin\left(\pi\left(1 - \left(\frac{t}{q_n} - \xi_{mt}|e_k b^m|\right)\right)\right) = s_{mt}, \end{aligned}$$

and likewise since $\cot(\pi - x) = -\cot x$ and $\sin x$ is an odd function, we have

$$h_{m(q_n-t)} = \cot\left(\pi\left(1 - \frac{t}{q_n}\right)\right)\sin\left(-\pi|e_k b^m|\xi_{mt}\right) = h_{mt}$$

for every $t \in \{0, 1, \ldots, q_n - 1\}$.

Now let us verify (b). In light of (a), it is enough to verify $s_{mt} > s_{m0}$ for $t \in \{1, 2, \ldots, \lfloor q_n/2 \rfloor\}$. Writing s_{mt} as in (2.94), and recalling that $|\xi_{mt}| < 1/2$ for these values of t, it is clear that $s_{mt} > s_{m(t-1)}$, and in particular

$$s_{m1} > 2\sin\left(\pi\left(\frac{1}{q_n} - \frac{1}{2}|e_k b^m|\right)\right) > 2\sin\left(\frac{\pi|e_k b^m|}{2}\right) = s_{m0},$$

if only $|e_k b^m| < 1/q_n$. This in turn follows from Lemmas 2.3.14 and 2.3.20, as

$$q_n|e_k b^m| = |c_k e_k| + \mathcal{O}(|b|^{2m}) < 1$$

for sufficiently large values of m.

Finally, we verify (c) and (d). It follows directly from Lemma 2.3.18 that

$$\xi_{mt} = \{t\alpha_{\sigma_k}\} - \frac{1}{2} + \mathcal{O}(tb^{2m}) = \xi_{\infty t} + \mathcal{O}(tb^{2m}),$$

which confirms (c). For property (d), we use $\cot x = (1/x)(1+\mathcal{O}(x^2))$ and $\sin x = x(1+\mathcal{O}(x^2))$ to rewrite h_{mt} as

$$h_{mt} = \frac{q_n|e_k b^m|\xi_{mt}}{t}\left(1+\mathcal{O}(t^2b^{2m})\right),$$

where we have also exploited that $1/q_n = \Theta(|b|^m)$. Moreover, since $q_n|b|^m = c_k + \mathcal{O}(b^{2m})$ by Lemma 2.3.14, we get

$$h_{mt} = \frac{|c_k e_k|\xi_{mt}}{t}\left(1+\mathcal{O}(t^2b^{2m})\right) = \frac{|c_k e_k|\xi_{mt}}{t} + \mathcal{O}(tb^{2m}),$$

and finally by recalling property (c) it follows that $h_{mt} = h_{\infty t} + \mathcal{O}(tb^{2m})$. This confirms (d), and completes the proof of Lemma 2.3.22. □

Decomposition of $Q_{\ell m+k}(\alpha)$

We are now equipped to decompose the sine product $Q_{\ell m+k}(\alpha)$.

Lemma 2.3.23. *Fix $k \in \{0,1,\ldots,\ell-1\}$, and for all integers $m \geq 1$ and $t \in \{0,1,\ldots,q_{\ell m+k}-1\}$ let s_{mt} be given in (2.89). The product of sines $Q_{\ell m+k}(\alpha)$ can be written as*

$$Q_{\ell m+k}(\alpha) = \prod_{r=1}^{q_{\ell m+k}} |2\sin(\pi r\alpha)| = A_m B_m C_m,$$

where:

$$A_m = \left|2q_{\ell m+k}\sin(\pi e_k b^m)\right|, \tag{2.95}$$

$$B_m = \left|\prod_{t=1}^{q_{\ell m+k}-1} \frac{s_{mt}}{2\sin(\pi t/q_{\ell m+k})}\right|, \tag{2.96}$$

$$C_m = \prod_{t=1}^{(q_{\ell m+k}-1)/2}\left(1-\frac{s_{m0}^2}{s_{mt}^2}\right). \tag{2.97}$$

Proof. Again we introduce $n = \ell m+k$ for ease of notation. We then have $Q_{\ell m+k}(\alpha) = Q_n(\alpha) = \prod_{r=1}^{q_n}|2\sin(\pi r\alpha)|$, and

$$Q_n^2(\alpha) = \left(2\sin(\pi q_n\alpha)\right)^2\left(\prod_{r=1}^{q_n-1} 2\sin(\pi r\alpha)\right)^2$$

$$
\begin{aligned}
&= \big(2\sin(\pi q_n\alpha)\big)^2 \prod_{r=1}^{q_n-1} \big(2\sin(\pi r\alpha) 2\sin(\pi(q_n-r)\alpha)\big) \\
&= \big(2\sin(\pi q_n\alpha)\big)^2 \prod_{r=1}^{q_n-1} 2\big(\cos(2\pi r\alpha - \pi q_n\alpha) - \cos(\pi q_n\alpha)\big).
\end{aligned}
$$

For the last equality we have used the identity $\sin(x)\sin(y) = (\cos(x-y) - \cos(x+y))/2$. Inserting $q_n\alpha = p_n + e_k b^m$ from Lemma 2.3.16 in the expression above, we get

$$
\begin{aligned}
Q_n^2(\alpha) &= \big(2\sin(\pi e_k b^m)\big)^2 \prod_{r=1}^{q_n-1} 2(-1)^{p_n}\big(\cos\left(2\pi r\alpha - \pi e_k b^m\right) - \cos(\pi e_k b^m)\big) \\
&= \big(2\sin(\pi e_k b^m)\big)^2 \prod_{r=1}^{q_n-1} 4\left(\sin^2\left(\pi r\alpha - \frac{\pi}{2}e_k b^m\right) - \sin^2\left(\frac{\pi}{2}e_k b^m\right)\right).
\end{aligned}
$$

Observe that we have used the identity $\cos(x) = 1 - 2\sin^2(x/2)$ and the fact that $(-1)^{(p_n+1)(q_n-1)} = 1$. The latter follows from the fact that $\gcd(p_n, q_n) = 1$, and accordingly either (p_n+1) or (q_n-1) is an even number. This concludes the rebasing of the argument described in the introduction to this section.

We now aim to express $Q_n^2(\alpha)$ as a product of perturbed rational sines. Again we use the identity $\alpha = p_n/q_n + e_k b^m/q_n$ from Lemma 2.3.16 to get

$$
\sin^2\left(\pi r\alpha - \frac{\pi}{2}e_k b^m\right) = \sin^2\left(\pi\left(\frac{rp_n}{q_n} + e_k b^m\left(\frac{r}{q_n} - \frac{1}{2}\right)\right)\right).
$$

By the substitution $t = rp_n \bmod q_n$, and recalling from (2.62) that $p_n q_{n-1} = (-1)^n \bmod q_n$, we have

$$
\begin{aligned}
\sin^2\left(\pi r\alpha - \frac{\pi}{2}e_k b^m\right) &= \sin^2\left(\pi\left(\frac{rp_n \bmod q_n}{q_n} + e_k b^m\left(\frac{r}{q_n} - \frac{1}{2}\right)\right)\right) \\
&= \sin^2\left(\pi\left(\frac{t}{q_n} + e_k b^m\left(\frac{(-1)^n t q_{n-1} \bmod q_n}{q_n} - \frac{1}{2}\right)\right)\right) \\
&= \frac{1}{4}s_{mt}^2,
\end{aligned}
$$

with s_{mt} given in (2.89), and where we have used $e_k b^m = (-1)^{n-1}|e_k b^m|$ and

$$
\frac{(-1)^n t q_{n-1} \bmod q_n}{q_n} - \frac{1}{2} = \left\{\frac{(-1)^n t q_{n-1}}{q_n}\right\} - \frac{1}{2} = (-1)^n\left(\left\{\frac{t q_{n-1}}{q_n}\right\} - \frac{1}{2}\right).
$$

As r runs through the values $1, 2, \ldots, q_n - 1$, so does $t = rp_n \bmod q_n$. Accordingly, we get

$$\begin{aligned}
Q_n^2(\alpha) &= (2\sin(\pi e_k b^m))^2 \prod_{t=1}^{q_n-1} \left(s_{mt}^2 - s_{m0}^2\right) \\
&= (2\sin(\pi e_k b^m))^2 \prod_{t=1}^{q_n-1} s_{mt}^2 \prod_{t=1}^{q_n-1} \left(1 - \frac{s_{m0}^2}{s_{mt}^2}\right) \\
&= (2q_n \sin(\pi e_k b^m))^2 \prod_{t=1}^{q_n-1} \frac{s_{mt}^2}{4\sin^2(\pi t/q_n)} \prod_{t=1}^{q_n-1} \left(1 - \frac{s_{m0}^2}{s_{mt}^2}\right).
\end{aligned}$$

For the last equality above we have used the well-known identity

$$\prod_{r=1}^{q-1} 2\sin\left(\frac{\pi r p}{q}\right) = q$$

whenever $p, q \in \mathbb{Z}$ satisfy $\gcd(p,q) = 1$ (see Lemma 2.1.1). Finally, we recall from Lemma 2.3.22 (a) that $s_{mt} = s_{m(q_n-t)}$ and hence $s_{mt}^2 = s_{m(q_n-t)}^2$ for every $t \in \{0, 1, \ldots, q_n - 1\}$. With our generalised notion of products (see p.78), we thus get

$$\prod_{t=1}^{q_n-1} \left(1 - \frac{s_{m0}^2}{s_{mt}^2}\right) = \prod_{t=1}^{(q_n-1)/2} \left(1 - \frac{s_{m0}^2}{s_{mt}^2}\right)^2.$$

Inserting this in the expression for $Q_n^2(\alpha)$ above and taking the square root of both sides, we arrive at

$$Q_n(\alpha) = Q_{\ell m+k}(\alpha) = A_m B_m C_m,$$

where A_m, B_m and C_m are given in (2.95), (2.96) and (2.97), respectively. □

Convergence of A_m

Let us now see that A_m in (2.95) converges as $m \to \infty$. Since $\sin x = x + \mathcal{O}(x^3)$, we have

$$A_m = \left|2q_{\ell m+k}\left(\pi e_k b^m + \mathcal{O}(b^{3m})\right)\right|.$$

By Lemma 2.3.14 and $|b| < 1$ it thus follows that

$$\lim_{m\to\infty} A_m = 2\pi |e_k c_k|, \tag{2.98}$$

where c_k and e_k are the constants given in (2.81) and (2.83), respectively. Alternatively, using the expression for $|e_k c_k|$ given in Lemma 2.3.20, we have

$$\lim_{m\to\infty} A_m = \frac{2\pi q_\ell(\alpha_{\tau_k})}{(a-b)}.$$

2.3.4 Convergence of C_m

In this section we show that the product

$$C_m = \prod_{t=1}^{(q_{\ell m+k}-1)/2} \left(1 - \frac{s_{m0}^2}{s_{mt}^2}\right)$$

is convergent. This is not quite straightforward, as it is not obvious that the sequence $(C_m)_{m\geq 1}$ is monotonically decreasing. However, we will see that $(C_m)_{m\geq 1}$ is comparable to a monotonically decreasing sequence of products bounded below by a positive number.

Theorem 2.3.24. *The sequence C_m converges to the strictly positive limit*

$$\lim_{m\to\infty} C_m = \prod_{t=1}^{\infty} \left(1 - \frac{1}{4\left(t/|c_k e_k| - \xi_{\infty t}\right)^2}\right), \tag{2.99}$$

where $|c_k e_k|$ is given in (2.87) *and $\xi_{\infty t} = \{t\alpha_{\sigma_k}\} - 1/2$.*

We will need Lemma 2.2.12 from Section 2.2.4 for proving Theorem 2.3.24. Recall that it states the following: For $n \geq 2$ and real numbers a_t, $t = 1, 2, \ldots, n$, satisfying $A := \sum_{t=1}^{n} |a_t| < 1$, we have that

$$1 - A < \prod_{t=1}^{n}(1 - |a_t|) < \frac{1}{1-A}.$$

Proof of Theorem 2.3.24. For ease of notation we again write $n = \ell m + k$, and begin by developing estimates for the quotients s_{m0}/s_{mt}. We have

$$s_{m0} = 2\sin\left(\pi|e_k b^m|/2\right) = \pi|e_k b^m|\left(1 + \mathcal{O}(b^{2m})\right),$$

and for $t \geq 1$ it follows from Lemmas 2.3.14 and 2.3.22 (c) that

$$s_{mt} = 2\sin\pi\left(\frac{t}{q_n} - |e_k b^m|\xi_{mt}\right) = 2\sin\pi t|b|^m\left(\frac{1}{c_k} - \frac{|e_k|\xi_{\infty t}}{t} + \mathcal{O}(b^{2m})\right). \tag{2.100}$$

We now split the values of t at $\eta_m = \lceil |b|^{-3m/5} \rceil$, and treat $t \leq \eta_m$ and $t > \eta_m$ separately in order to find appropriate bounds on s_{mt} in (2.100). For $t > \eta_m$, we use $\sin x \geq 2x/\pi$ for $x \in [0, \pi/2]$ to obtain

$$s_{mt} \geq 4t|b|^m\left(\frac{1}{c_k} - \frac{|e_k|\xi_{\infty t}}{t} + \mathcal{O}(b^{2m})\right).$$

Recall that $c_k > 0$ and $|\xi_{\infty t}| \leq 1/2$. Thus, for sufficiently large m (and thereby sufficiently large t), we have $s_{mt} > 2\eta_m |b|^m / c_k$ and

$$\frac{s_{m0}}{s_{mt}} \leq \frac{\pi |e_k b^m| \left(1 + \mathcal{O}(b^{2m})\right)}{2\eta_m |b|^m / c_k} = \frac{\pi |c_k e_k|}{2\eta_m} \left(1 + \mathcal{O}(b^{2m})\right) = \mathcal{O}(\eta_m^{-1}).$$

It follows that

$$\sum_{t=\eta_m+1}^{(q_n-1)/2} \frac{s_{m0}^2}{s_{mt}^2} \leq q_n \cdot \mathcal{O}(\eta_m^{-2}) = \mathcal{O}(|b|^{m/5}),$$

and accordingly this sum is convergent and smaller than one for sufficiently large m. Thus, by Lemma 2.2.12 we get

$$1 \geq \prod_{t=\eta_m+1}^{(q_n-1)/2} \left(1 - \frac{s_{m0}^2}{s_{mt}^2}\right) > 1 - \sum_{t=\eta_m+1}^{(q_n-1)/2} \frac{s_{m0}^2}{s_{mt}^2} \geq 1 - \mathcal{O}(|b|^{m/5}). \tag{2.101}$$

Now consider $t \leq \eta_m$. It is clear from (2.100) that by choosing m sufficiently large, the argument in the sine function s_{mt} can be made arbitrarily small in this case. Applying $\sin x = x + \mathcal{O}(x^3)$, we get

$$\begin{aligned} s_{mt} &= 2\pi |b|^m \left(\frac{t}{c_k} - |e_k| \xi_{\infty t} + \mathcal{O}(t b^{2m})\right) + \mathcal{O}(|b|^{6m/5}) \\ &= \pi |e_k b^m| \left(u_t + \mathcal{O}(|b|^{m/5})\right), \end{aligned}$$

where we have introduced the notation

$$u_t = 2\left(\frac{t}{|c_k e_k|} - \xi_{\infty t}\right) = 2\left(\frac{t}{|c_k e_k|} - \{t\alpha_{\sigma_k}\} + \frac{1}{2}\right). \tag{2.102}$$

We thus have

$$\frac{s_{m0}}{s_{mt}} = \frac{\pi |e_k b^m| \left(1 + \mathcal{O}(b^{2m})\right)}{\pi |e_k b^m| \left(u_t + \mathcal{O}(|b|^{m/5})\right)} = \frac{1 + \mathcal{O}(|b|^{m/5})}{u_t},$$

and moreover

$$\begin{aligned} \prod_{t=1}^{\eta_m} \left(1 - \frac{s_{m0}^2}{s_{mt}^2}\right) &= \prod_{t=1}^{\eta_m} \left(1 - \frac{1}{u_t^2} - \frac{\mathcal{O}(|b|^{m/5})}{u_t^2}\right) \\ &= \prod_{t=1}^{\eta_m} \left(1 - \frac{1}{u_t^2}\right) \prod_{t=1}^{\eta_m} \left(1 - \frac{\mathcal{O}(|b|^{m/5})}{u_t^2 - 1}\right). \end{aligned}$$

We look closer at the two products on the final line above. Since $|\xi_{\infty t}| < 1/2$ and $|c_k e_k| < 1$, we see from (2.102) that $u_t > 1$ for all $1 \le t \le \eta_m$. This guarantees that both products are well-defined. Moreover, we see that u_t behaves as $2t/|c_k e_k|$ for large t. Hence by comparison with $\sum_{t=1}^{\infty} 1/t^2 = \pi^2/6$, the sum $\sum_{t=1}^{\infty} 1/(u_t^2 - 1)$ converges, and it follows that $\sum_{t=1}^{\infty} \mathcal{O}(|b|^{m/5})/(u_t^2 - 1) = \mathcal{O}(|b|^{m/5})$. The latter sum is thus smaller than one, provided m is sufficiently large, and again it follows from Lemma 2.2.12 that

$$1 > \prod_{t=1}^{\eta_m} \left(1 - \frac{\mathcal{O}(|b|^{m/5})}{u_t^2 - 1}\right) \ge 1 - \sum_{t=1}^{\eta_m} \frac{\mathcal{O}(|b|^{m/5})}{u_t^2 - 1} = 1 - \mathcal{O}(|b|^{m/5}). \tag{2.103}$$

For the second product we introduce the notation

$$U_j := \prod_{t=1}^{j} \left(1 - \frac{1}{u_t^2}\right).$$

Since $u_t > 1$ for all t, the sequence $(U_j)_{j \ge 1}$ is monotonically decreasing and bounded below by zero. Thus, the limit $\lim_{j\to\infty} U_j$ exists.

By combining the estimates for $t > \eta_m$ and $t \le \eta_m$, we now have

$$C_m = U_{\eta_m} \prod_{t=1}^{\eta_m} \left(1 - \frac{\mathcal{O}(|b|^{m/5})}{u_t^2 - 1}\right) \prod_{t=\eta_m+1}^{(q_n-1)/2} \left(1 - \frac{s_{m0}^2}{s_{mt}^2}\right).$$

Taking the limit of both sides as $m \to \infty$, and recalling (2.101) and (2.103), we arrive at

$$\lim_{m\to\infty} C_m = \lim_{m\to\infty} U_{\eta_m} = \prod_{t=1}^{\infty} \left(1 - \frac{1}{u_t^2}\right). \tag{2.104}$$

This nearly completes the proof of Theorem 2.3.24. Our claim, however, is that $\lim_{m\to\infty} C_m$ is strictly positive. This will follow from (2.104) and Lemma 2.2.12 if we can verify that

$$\sum_{t=1}^{\infty} \frac{1}{u_t^2} < 1. \tag{2.105}$$

Let us first verify (2.105) for $\ell = 1$. In this case, we have $k = 0$ and $|c_0 e_0| \le 1/\sqrt{5}$ by Lemma 2.3.20. It follows that

$$\sum_{t=1}^{\infty} \frac{1}{u_t^2} \le \frac{1}{u_1^2} + \sum_{t=2}^{\infty} \frac{1}{20(t-1)^2} < \frac{1}{4(\sqrt{5} - 1/2)^2} + \frac{\pi^2}{120} < 1.$$

The case $\ell > 1$ is more involved. However, observe that in this case $q_{\ell+1}(\alpha_{\tau_k}) \geq 2$ and $p_\ell(\alpha_{\tau_k}) \geq 1$ and by (2.75) and (2.65) we get that

$$p_\ell(\alpha_{\tau_k}) - 2b = -q_{\ell+1}(\alpha_{\tau_k}) + \sqrt{(q_{\ell+1}(\alpha_{\tau_k}) + p_\ell(\alpha_{\tau_k}))^2 + 4(-1)^{\ell-1}} > 0.$$

Therefore it follows from Lemma 2.3.20 that

$$\frac{1}{|c_k e_k|} = \frac{q_{\ell+1}(\alpha_{\tau_k}) + p_\ell(\alpha_{\tau_k}) - 2b}{q_\ell(\alpha_{\tau_k})} \geq a_k + \frac{q_{\ell-1}(\alpha_{\tau_k})}{q_\ell(\alpha_{\tau_k})},$$

where we have also used the classical recursion formula for q_ℓ. By Lemma 2.3.18 we get

$$\frac{q_{\ell-1}(\alpha_{\tau_k})}{q_\ell(\alpha_{\tau_k})} = \frac{p_\ell(\alpha_{\sigma_0\tau_k})}{q_\ell(\alpha_{\sigma_0\tau_k})} = \frac{p_\ell(\alpha_{\sigma_k})}{q_\ell(\alpha_{\sigma_k})},$$

(recall from the proof of Lemma 2.3.9 that $\sigma_k = \sigma_0\tau_k$) and thus we have

$$u_1 = 2\left(\frac{1}{|c_k e_k|} - \alpha_{\sigma_k} + \frac{1}{2}\right) \geq 2\left(a_k + \frac{p_\ell(\alpha_{\sigma_k})}{q_\ell(\alpha_{\sigma_k})} - \alpha_{\sigma_k} + \frac{1}{2}\right). \tag{2.106}$$

By the standard error estimate (2.63) for continued fractions, we know that

$$\left|\frac{p_\ell(\alpha_{\sigma_k})}{q_\ell(\alpha_{\sigma_k})} - \alpha_{\sigma_k}\right| < \frac{1}{q_{\ell+1}(\alpha_{\sigma_k}) q_\ell(\alpha_{\sigma_k})} < \frac{1}{2}$$

when $\ell > 1$, and inserting this in (2.106) we find that $u_1 \geq 2a_k \geq 2$. For all other terms in the sum $\sum_{t=1}^{\infty} 1/u_t^2$, the estimate $|c_k e_k| < 1$ from Lemma 2.3.20 suffices. We get

$$\sum_{t=1}^{\infty} \frac{1}{u_t^2} \leq \frac{1}{4} + \sum_{t=2}^{\infty} \frac{1}{4(t-1/2)^2} = \frac{1}{4} + \frac{\pi^2}{8} - 1 < 0.49.$$

This verifies (2.105) for the case $\ell > 1$. Thus, we conclude that $\lim_{m\to\infty} C_m > 0$, and this completes the proof of Theorem 2.3.24. $\square$

2.3.5 Convergence of B_m

The aim of this section is to verify the convergence of

$$B_m = \left|\prod_{t=1}^{q_{\ell m+k}-1} \frac{s_{mt}}{2\sin(\pi t/q_{\ell m+k})}\right|$$

as $m \to \infty$. We will see that this requires greater efforts than verifying the convergence of C_m. In fact, what we will show is that $\log B_m$ converges to a finite limit, and accordingly $\lim_{m\to\infty} B_m$ exists and is strictly positive.

For the remainder of this section, let us again ease notation by writing $n = \ell m + k$. We begin by examining each term of the product B_m. Recalling the definition of s_{mt} from (2.89), we have

$$\begin{aligned}\frac{s_{mt}}{2\sin(\pi t/q_n)} &= \cos(\pi|e_k b^m|\xi_{mt}) - \cot(\pi t/q_n)\sin(\pi|e_k b^m|\xi_{mt}) \\ &= 1 - 2\sin^2(\pi|e_k b^m|\xi_{mt}/2) - h_{mt},\end{aligned}$$

with h_{mt} given in (2.92). Taking $\beta_{mt} := 2\sin^2(\pi|e_k b^m|\xi_{mt}/2)$, it is easily verified that $\beta_{m(q_n-t)} = \beta_{mt}$ for $t \in \{1,\dots,q_n-1\}$. Likewise, we recall from Lemma 2.3.22 (a) that $h_{m(q_n-t)} = h_{mt}$, and thus

$$B_m = \prod_{t=1}^{q_n-1}(1-\beta_{mt}-h_{mt}) = \prod_{t=1}^{(q_n-1)/2}(1-\beta_{mt}-h_{mt})^2.$$

This shows that we need only consider $t \in \{1,\dots,(q_n-1)/2\}$.

Let us now show that rather than analysing B_m, we may choose to analyse the simpler product

$$B_m^* := \prod_{t=1}^{q_n-1}(1-h_{mt}) = \prod_{t=1}^{(q_n-1)/2}(1-h_{mt})^2. \tag{2.107}$$

Taking logarithms, we get

$$\log(1-\beta_{mt}-h_{mt}) = \log(1-h_{mt}) + \log\left(1-\frac{\beta_{mt}}{1-h_{mt}}\right). \tag{2.108}$$

Our claim is that the latter term on the right hand side in (2.108) will not contribute significantly to the sum $\log B_m = \sum_{t=1}^{q_n-1}\log(1-\beta_{mt}-h_{mt})$. To see this, let us first estimate the size of h_{mt} and β_{mt}. Considering only $t \in \{1,\dots,(q_n-1)/2\}$, we use $\cot x < 1/x$ and $\sin x < x$ to obtain

$$|h_{mt}| = \cot(\pi t/q_n)\sin(\pi|e_k b^m|\xi_{mt}) \le \frac{q_n|e_k b^m|\xi_{mt}}{t}.$$

We recall from Lemmas 2.3.14 and 2.3.20 that $q_n|e_k b^m| = |c_k e_k| + \mathcal{O}(b^{2m}) < 1$ for sufficiently large m. As $|\xi_{mt}| < 1/2$, we thus get

$$|h_{mt}| < \frac{1}{2t} < \frac{1}{2}, \tag{2.109}$$

and it follows that $1 - h_{mt} > 1/2$. For β_{mt}, we have

$$\beta_{mt} < 2\left(\frac{\pi|e_k b^m|\xi_{mt}}{2}\right)^2 < \frac{\pi^2(e_k b^m)^2}{8},$$

and thus for sufficiently large values of m we get $|\beta_{mt}/(1 - h_{mt})| < 1$ and

$$\log\left(1 - \frac{\beta_{mt}}{1 - h_{mt}}\right) = -\sum_{j=1}^{\infty} \frac{1}{j}\left(\frac{\beta_{mt}}{1 - h_{mt}}\right)^j = \mathcal{O}(b^{2m}). \tag{2.110}$$

Recalling that $q_n = \Theta(|b|^{-m})$, it now follows from (2.108) and (2.110) that

$$|\log B_m - \log B_m^*| = \left|2\sum_{t=1}^{(q_n-1)/2} \log\left(1 - \frac{\beta_{mt}}{1 - h_{mt}}\right)\right| = \mathcal{O}(|b|^m),$$

and thus $\lim_{m\to\infty} \log B_m = \lim_{m\to\infty} \log B_m^*$. This confirms that we may choose to analyse B_m^* in (2.107) rather than the original product B_m.

Finally, we rewrite $\log B_m^*$ using its Taylor expansion as

$$\begin{aligned}\log B_m^* &= 2\sum_{t=1}^{(q_n-1)/2} \log(1 - h_{mt}) = -2\sum_{t=1}^{(q_n-1)/2}\sum_{j=1}^{\infty} \frac{1}{j} h_{mt}^j \\ &= -2\left(\sum_{t=1}^{(q_n-1)/2} h_{mt} + \sum_{t=1}^{(q_n-1)/2}\sum_{j=2}^{\infty} \frac{1}{j} h_{mt}^j\right) =: -2(H_m^{(1)} + H_m^{(2)}).\end{aligned} \tag{2.111}$$

We go on to study the behaviour of the two sums $H_m^{(1)}$ and $H_m^{(2)}$ separately in the following subsections.

Convergence of $H_m^{(2)}$

Let us first treat the sum

$$H_m^{(2)} = \sum_{t=1}^{(q_n-1)/2}\sum_{j=2}^{\infty} \frac{1}{j} h_{mt}^j.$$

It is an easy task to show that $H_m^{(2)}$ is bounded, but showing convergence requires greater efforts.

We begin by showing that terms where t or j is greater than $|b|^{-m/2}$ will not contribute significantly to $H_m^{(2)}$. Recall from (2.109) that $|h_{mt}| < 1/(2t)$ for sufficiently large m, and thus for $u \geq 2$ we get

$$\left|\sum_{j=u}^{\infty} \frac{1}{j} h_{mt}^j\right| < \sum_{j=u}^{\infty} |h_{mt}^j| = \frac{|h_{mt}^u|}{1-|h_{mt}|} < 2\left(\frac{1}{2t}\right)^u .$$

Now let $u_m = \lfloor |b|^{-m/2} \rfloor$, and choose m so that $2 \leq u_m \leq (q_n - 1)/2$ and (2.109) holds. Note that this is always possible since $u_m = \Theta(|b|^{-m/2})$ and $q_n = q_{\ell m+k} = \Theta(|b|^{-m})$. We then have

$$\left|\sum_{t=u_m+1}^{(q_n-1)/2} \sum_{j=2}^{\infty} \frac{1}{j} h_{mt}^j\right| < \sum_{t=u_m+1}^{(q_n-1)/2} 2\left(\frac{1}{2t}\right)^2 < \frac{1}{2} \sum_{t=u_m+1}^{\infty} \frac{1}{t^2} < \frac{1}{2u_m}$$

and

$$\left|\sum_{t=1}^{u_m} \sum_{j=u_m+1}^{\infty} \frac{1}{j} h_{mt}^j\right| < \sum_{t=1}^{u_m} 2\left(\frac{1}{2t}\right)^{u_m+1} < \left(\frac{1}{2}\right)^{u_m} \sum_{t=1}^{\infty} \frac{1}{t^2} = \frac{\pi^2}{6 \cdot 2^{u_m}}.$$

Both of these sums are $\mathcal{O}(|b|^{m/2})$, and it follows that

$$\begin{aligned} H_m^{(2)} &= \sum_{t=1}^{u_m} \sum_{j=2}^{u_m} \frac{1}{j} h_{mt}^j + \sum_{t=1}^{u_m} \sum_{j=u_m+1}^{\infty} \frac{1}{j} h_{mt}^j + \sum_{t=u_m+1}^{(q_n-1)/2} \sum_{j=2}^{\infty} \frac{1}{j} h_{mt}^j \\ &= \sum_{t=1}^{u_m} \sum_{j=2}^{u_m} \frac{1}{j} h_{mt}^j + \mathcal{O}(|b|^{m/2}). \end{aligned}$$

Thus, we have for $m \to \infty$

$$H_m^{(2)} \sim \sum_{t=1}^{u_m} \sum_{j=2}^{u_m} \frac{1}{j} h_{mt}^j, \tag{2.112}$$

where we recall that this notation means that the limit of $H_m^{(2)}$ equals that of its truncation $\sum_{t=1}^{u_m} \sum_{j=2}^{u_m} h_{mt}^j / j$ as $m \to \infty$.

Now let us see that

$$\sum_{t=1}^{u_m} \sum_{j=2}^{u_m} \frac{1}{j} h_{mt}^j \sim \sum_{t=1}^{u_m} \sum_{j=2}^{u_m} \frac{1}{j} h_{\infty t}^j, \tag{2.113}$$

where $h_{\infty t}$ is given in (2.93). As we are considering only $j, t \le u_m$, we have $jt \le u_m^2 \le |b|^{-m}$, and hence $jtb^{2m} \to 0$ as $m \to \infty$. From Lemma 2.3.22 (d), we therefore get for some constant $c_1 > 0$ that

$$h_{mt}^j \le \left(h_{\infty t} + c_1 tb^{2m}\right)^j = h_{\infty t}^j + \sum_{k=1}^{j} \binom{j}{k} h_{\infty t}^{j-k} \left(c_1 tb^{2m}\right)^k.$$

Recall that by (2.109) we have $|h_{\infty t}| \le 1/2$. Thus by rewriting the above inequality it follows

$$\begin{aligned}
|h_{mt}^j - h_{\infty t}^j| &\le c_1 2^{-j+1} jtb^{2m} + \sum_{k=2}^{j} \binom{j}{k} |h_{\infty t}|^{j-k} \left(c_1 tb^{2m}\right)^k \\
&\le c_1 jtb^{2m} + 2^{-j} \sum_{k=2}^{j} \binom{j}{k} \left(2c_1 tb^{2m}\right)^k \\
&\le c_1 jtb^{2m} + c_2 b^{3m} 2^{-j} \sum_{k=2}^{j} \binom{j}{k} \\
&\le c_3 jtb^{2m}, \qquad (2.114)
\end{aligned}$$

where we have used that $t \le u_m \le b^{-m/2}$ and that $2c_1 b^{3m/2} < 1$ for sufficiently large m. Hence by (2.114) we obtain $h_{mt}^j - h_{\infty t}^j = \mathcal{O}(jtb^{2m})$ and we get that

$$\sum_{t=1}^{u_m} \sum_{j=2}^{u_m} \frac{1}{j}(h_{mt}^j - h_{\infty t}^j) = \sum_{t=1}^{u_m} \sum_{j=2}^{u_m} \mathcal{O}(tb^{2m}) = \mathcal{O}(u_m^3 b^{2m}) = \mathcal{O}(|b|^{m/2}).$$

This confirms (2.113).

Finally, by reusing the arguments that led us to the conclusion that $H_m^{(2)} \sim \sum_{t=1}^{u_m} \sum_{j=2}^{u_m} h_{mt}^j/j$, we find that

$$\sum_{t=1}^{u_m} \sum_{j=2}^{u_m} \frac{1}{j} h_{\infty t}^j \sim \sum_{t=1}^{\infty} \sum_{j=2}^{\infty} \frac{1}{j} h_{\infty t}^j, \qquad (2.115)$$

and recalling that $|h_{\infty t}| = |c_k e_k \xi_{\infty t}/t| < 1/(2t) < 1/2$, we get

$$\sum_{t=1}^{\infty} \sum_{j=2}^{\infty} \frac{1}{j} |h_{\infty t}|^j < \sum_{t=1}^{\infty} \frac{h_{\infty t}^2}{1 - |h_{\infty t}|} < \infty.$$

Thus, the sum on the right hand side in (2.115) is absolutely convergent. We denote its limit by $\Gamma_{\ell,k}^{(2)}$, and from (2.112)–(2.115) it follows that

$$\lim_{m \to \infty} H_m^{(2)} = \Gamma_{\ell,k}^{(2)}. \qquad (2.116)$$

Convergence of $H_m^{(1)}$

We are left with verifying the convergence of

$$H_m^{(1)} = \sum_{t=1}^{(q_n-1)/2} h_{mt}.$$

This rather tedious task is performed in several steps. Eventually we will see that if $\lim_{m\to\infty} H_m^{(1)}$ exists, then it equals the limit of $\sum_{t=1}^{(q_n-1)/2} C_{mt}S_{mt}$, where

$$S_{mt} := \sum_{s=1}^{t} \sin(\pi|e_k b^m|\xi_{\infty s}) \text{ and } C_{mt} := \cot(\pi t/q_n) - \cot(\pi(t+1)/q_n).$$

Careful estimates of S_{mt} and C_{mt} will reveal that the sum $\sum_{t=1}^{(q_n-1)/2} C_{mt}S_{mt}$ indeed converges.

Note first that we may return to standard summation notation at this point, as $h_{m(q_n/2)} = 0$ if q_n is even. Thus, we let $M_n := \lfloor (q_n-1)/2 \rfloor$ and have

$$H_m^{(1)} = \sum_{t=1}^{M_n} h_{mt} = \sum_{t=1}^{M_n} \cot\left(\frac{\pi t}{q_n}\right) \sin(\pi|e_k b^m|\xi_{mt}),$$

regardless of whether q_n is even or odd.

Now let us see that

$$H_m^{(1)} \sim H_m^* := \sum_{t=1}^{M_n} \cot\left(\frac{\pi t}{q_n}\right) \sin(\pi|e_k b^m|\xi_{\infty t}). \tag{2.117}$$

Using that $\sin x = x(1+\mathcal{O}(x^2))$ and Lemma 2.3.22 (c), we get

$$\begin{aligned} H_m^{(1)} - H_m^* &= \sum_{t=1}^{M_n} \cot\left(\frac{\pi t}{q_n}\right) \pi|e_k b^m|(\xi_{mt} - \xi_{\infty t})(1+\mathcal{O}(b^{2m})) \\ &= \sum_{t=1}^{M_n} \cot\left(\frac{\pi t}{q_n}\right) \pi|e_k b^m|\mathcal{O}(tb^{2m}). \end{aligned}$$

From the inequality $\cot x < 1/x$ it thus follows that

$$|H_m^{(1)} - H_m^*| < \mathcal{O}(b^{2m}) \sum_{t=1}^{M_n} q_n|e_k b^m|$$

$$= \mathcal{O}(b^{2m}) \cdot M_n \left(|e_k c_k| + \mathcal{O}(b^{2m})\right) = \mathcal{O}(|b|^m),$$

where we have used Lemma 2.3.14 and the fact that $M_n < q_n = \Theta(|b|^{-m})$. This confirms (2.117).

Finally, let us see that if $\lim_{m\to\infty} H_m^*$ exists, then it equals that of $\sum_{t=1}^{(q_n-1)/2} S_{mt}C_{mt}$ for a certain sum of sines S_{mt} and cotangent difference C_{mt}. As a first step we will make use of the following formula which is also called summation by parts:

$$\sum_{k=1}^{n} a_k b_k = \sum_{k=1}^{n-1} \left((b_k - b_{k+1}) \sum_{i=1}^{k} a_i \right) + b_n \sum_{i=1}^{n} a_i, \tag{2.118}$$

where $n \in \mathbb{N}$ and $a_k, b_k \in \mathbb{R}$ for $k \in [n]$. Using (2.118), we may rewrite H_m^* as

$$\begin{aligned} H_m^* = &\sum_{t=1}^{M_n-1} \left(\cot\left(\frac{\pi t}{q_n}\right) - \cot\left(\frac{\pi(t+1)}{q_n}\right) \right) \sum_{s=1}^{t} \sin(\pi |e_k b^m| \xi_{\infty s}) \\ &+ \cot\left(\frac{\pi M_n}{q_n}\right) \sum_{s=1}^{M_n} \sin(\pi |e_k b^m| \xi_{\infty s}). \end{aligned} \tag{2.119}$$

Consider the second term on the right hand side in this equation. As $|\xi_{\infty s}| < 1/2$ and $\sin x = x(1 + \mathcal{O}(x^2))$, we have

$$\left| \sum_{s=1}^{M_n} \sin(\pi |e_k b^m| \xi_{\infty s}) \right| < \frac{\pi}{2} q_n |e_k b^m| (1 + \mathcal{O}(b^{2m})) = \frac{\pi}{2} |c_k e_k| (1 + \mathcal{O}(b^{2m})),$$

where we have again used that $M_n < q_n = \mathcal{O}(b^{2m})$ and Lemma 2.3.14. It follows that

$$\left| \cot\left(\frac{\pi M_n}{q_n}\right) \sum_{s=1}^{M_n} \sin(\pi |e_k b^m| \xi_{\infty s}) \right| < \frac{\pi}{2} \left| c_k e_k \cot\left(\frac{\pi M_n}{q_n}\right) (1 + \mathcal{O}(b^{2m})) \right|,$$

and recalling that $M_n = \lfloor (q_n - 1)/2 \rfloor$, it is clear that the cotangent term tends to zero as $m \to \infty$. It thus follows from (2.117) and (2.119) that

$$H_m^{(1)} \sim H_m^* \sim \sum_{t=1}^{M_n-1} C_{mt} S_{mt} \tag{2.120}$$

where $C_{mt} = \cot(\pi t/q_n) - \cot(\pi(t+1)/q_n)$ and $S_{mt} = \sum_{s=1}^{t} \sin(\pi |e_k b^m| \xi_{\infty s})$.

The cotangent difference C_{mt}

We establish two estimates for C_{mt}; one rather coarse bound and one more precise estimate. For ease of notation, let us write $\phi = \pi/q_n$. We then have

$$\begin{aligned} 0 < C_{mt} &= \frac{\sin((t+1)\phi)\cos(t\phi) - \cos((t+1)\phi)\sin(t\phi)}{\sin(t\phi)\sin((t+1)\phi)} \\ &= \frac{\sin(\phi)}{\sin(t\phi)\sin((t+1)\phi)}. \end{aligned}$$

Note that when $t < M_n$, we have $(t+1)\phi < \pi/2$, and thus by $2x/\pi < \sin x < x$ for $0 < x < \pi/2$, we obtain

$$0 < C_{mt} < \frac{\pi q_n}{4t(t+1)} < \frac{\pi q_n}{4t^2}. \tag{2.121}$$

This is our coarse bound for C_{mt}.

For $t\phi < 1$, or equivalently $t < q_n/\pi$, we have the finer estimate

$$\begin{aligned} C_{mt} &= \frac{\phi(1 + \mathcal{O}(\phi^2))}{t\phi(1+\mathcal{O}(t^2\phi^2))(t+1)\phi(1+\mathcal{O}(t^2\phi^2))} \\ &= \frac{q_n}{\pi t(t+1)}\left(1 + \mathcal{O}\left(\frac{t^2}{q_n^2}\right)\right) = \frac{q_n}{\pi t(t+1)}\left(1 + \mathcal{O}\left(t^2 b^{2m}\right)\right), \end{aligned}$$

where we have used that $1/q_n = \mathcal{O}(|b|^m)$. Combining this estimate with

$$S_{mt} = \pi |e_k b^m| (1 + \mathcal{O}(b^{2m})) \sum_{s=1}^{t} \xi_{\infty s},$$

and using Lemma 2.3.14, we get

$$C_{mt} S_{mt} = \frac{|c_k e_k|}{t(t+1)}\left(1 + \mathcal{O}(t^2 b^{2m})\right) \sum_{s=1}^{t} \xi_{\infty s}. \tag{2.122}$$

The sum of sines S_{mt}

Now let us find an appropriate bound on S_{mt} in terms of m and t. As illustrated by Verschueren and Mestel in [104, Figure 7.1], this sum appears to grow slowly with increasing values of t, at least for the specal case of $\alpha = \varphi$ the golden mean. As demonstrated by the next lemma, this is also true for the general case where $\ell \geq 1$.

Lemma 2.3.25. *For $t \in \{1, 2, \dots, q_n - 1\}$ and sufficiently large m, the sum*

$$S_{mt} = \sum_{s=1}^{t} \sin(\pi |e_k b^m| \xi_{\infty s})$$

satisfies $|S_{mt}| \le c|b|^m \log t$ for some constant $c > 0$ independent of m.

For proving Lemma 2.3.25, we will need the following result.

Lemma 2.3.26. *Let p/q be a convergent of any real α. Then for any $\theta \in \mathbb{R}$ and $v \in \mathbb{N}$, we have*

$$\left| \sum_{i=1}^{vq} \{\theta + i\alpha\} - \frac{1}{2} \right| < \frac{3v}{2}.$$

Proof. The proof for $v = 1$ is given in [104, Lemma 7.2]. For $v \ge 2$ it follows that

$$\begin{aligned} \left| \sum_{i=1}^{vq} \{\theta + i\alpha\} - \frac{1}{2} \right| &= \left| \sum_{j=0}^{v-1} \sum_{u=1}^{q} \{\theta + (jq + u)\alpha\} - \frac{1}{2} \right| \\ &\le \sum_{j=0}^{v-1} \left| \sum_{u=1}^{q} \{(\theta + jq\alpha) + u\alpha\} - \frac{1}{2} \right| < \sum_{j=0}^{v-1} \frac{3}{2} = \frac{3v}{2}. \end{aligned}$$

□

Proof of Lemma 2.3.25. Recall from Lemma 2.3.5 that there exist unique integers $z, v_1, \dots, v_z \in \mathbb{N}$ such that

$$t = \sum_{s=1}^{z} v_s q_s(\alpha_{\sigma_k}).$$

In order to simplify notation we will drop the argument of $q_s(\alpha_{\sigma_k})$ for the rest of the proof. We will use this representation of t to split the sum S_{mt} into chunks of length $v_s q_s$ as follows. Let us introduce the notation $t_z = 0$ and $t_s = \sum_{u=s+1}^{z} v_u q_u$. Moreover, we define

$$\xi_{\infty r}(\theta) := \{\theta + r\alpha_{\sigma_k}\} - \frac{1}{2}.$$

Note that our $\xi_{\infty r}$ defined in (2.91) is then precisely $\xi_{\infty r}(0)$. With this generalised $\xi_{\infty r}(\theta)$ introduced, we may rewrite S_{mt} as

$$S_{mt} = \sum_{r=1}^{v_z q_z} \sin(\pi |e_k b^m| \xi_{\infty r}(0)) + \sum_{r=1}^{v_{z-1} q_{z-1}} \sin(\pi |e_k b^m| \xi_{\infty r}(v_z q_z \alpha_{\sigma_k}))$$

$$+\sum_{r=1}^{v_{z-2}q_{z-2}} \sin(\pi|e_k b^m|\xi_{\infty r}((v_z q_z + v_{z-1}q_{z-1})\alpha_{\sigma_k})) + \cdots$$

$$= \sum_{s=1}^{z}\sum_{r=1}^{v_s q_s} \sin(\pi|e_k b^m|\xi_{\infty r}(t_s\alpha_{\sigma_k})).$$

Thus, if we also introduce the generalised notation

$$S_{mt}(\theta) = \sum_{r=1}^{t} \sin(\pi|e_k b^m|\xi_{\infty r}(\theta)),$$

then we can express S_{mt} as

$$S_{mt} = S_{mt}(0) = \sum_{s=1}^{z} S_{m(v_s q_s)}(t_s\alpha_{\sigma_k}). \tag{2.123}$$

Finally we use Lemma 2.3.26 to bound the terms in the sum (2.123). Using the estimate $\sin x = x(1+\mathcal{O}(x^2))$, we get

$$\begin{aligned}
\left|S_{m(v_s q_s)}(t_s\alpha_{\sigma_k})\right| &= \left|\pi e_k b^m(1+\mathcal{O}(b^{2m}))\sum_{r=1}^{v_s q_s}\xi_{\infty r}(t_s\alpha_{\sigma_k})\right| \\
&= \pi|e_k b^m|(1+\mathcal{O}(b^{2m}))\left|\sum_{r=1}^{v_s q_s}\{t_s\alpha_{\sigma_k} + r\alpha_{\sigma_k}\} - \frac{1}{2}\right| \\
&\le \frac{3}{2}\pi v_s|e_k b^m|(1+\mathcal{O}(b^{2m})).
\end{aligned}$$

Thus, we have

$$|S_{mt}| \le \sum_{s=1}^{z}|S_{m(v_s q_s)}(t_s\alpha_{\sigma_k})| \le \frac{3}{2}\pi|e_k b^m|(1+\mathcal{O}(b^{2m}))\sum_{s=1}^{z} v_s. \tag{2.124}$$

Recalling from Lemma 2.3.5 that $v_s \le \max\{a_1,\dots,a_\ell\}$ for all s, we have

$$\sum_{s=1}^{z} v_s \le z\cdot\max_{1\le j\le \ell} a_j = \mathcal{O}(z) = \mathcal{O}(\log t),$$

and combined with (2.124) this yields $|S_{mt}| = \mathcal{O}(|b|^m\log t)$. □

We are now equipped to prove the convergence of $H_m^{(1)}$, or equivalently the convergence of $\sum_{t=1}^{M_n-1} C_{mt}S_{mt}$ in (2.120). From (2.121) and Lemma 2.3.25 it follows that

$$|C_{mt}S_{mt}| \leq \frac{\pi q_n}{4t^2} \cdot \mathcal{O}(|b|^m \log t) = \mathcal{O}\left(\frac{\log t}{t^2}\right), \tag{2.125}$$

where we have also used $q_n = \Theta(b^{-m})$ from Lemma 2.3.14. This implies that there exists a constant $K > 0$ such that

$$\sum_{t=1}^{M_n-1} |C_{mt}S_{mt}| \leq K. \tag{2.126}$$

It is not clear that the sequence $(\sum_{t=1}^{M_n-1} C_{mt}S_{mt})_{m\geq 1}$ is monotone, so the bound (2.126) alone does not prove convergence. But let us now compare this sequence to a closely related, absolutely convergent sum.

Let $u_m = \lfloor |b|^{-m/2} \rfloor$, and choose m sufficiently large such that $u_m < q_n/\pi < M_n - 1$. We can then write

$$\sum_{t=1}^{M_n-1} C_{mt}S_{mt} = \sum_{t=1}^{u_m} C_{mt}S_{mt} + \sum_{t=u_m+1}^{M_n-1} C_{mt}S_{mt}. \tag{2.127}$$

It follows from (2.125) that for some constant $c > 0$ we have

$$\left|\sum_{t=u_m+1}^{M_n-1} C_{mt}S_{mt}\right| \leq c\int_{u_m}^{\infty} \frac{\log t}{t^2}\,\mathrm{d}t = c\frac{\log(u_m)+1}{u_m} \to 0 \text{ for } m \to \infty. \tag{2.128}$$

For the first sum on the right hand side in (2.127), we use the finer estimate (2.122) from Section 2.3.5 to obtain

$$\sum_{t=1}^{u_m} C_{mt}S_{mt} = (1 + \mathcal{O}(|b|^m)) \sum_{t=1}^{u_m} \frac{|c_k e_k|}{t(t+1)} \sum_{s=1}^{t} \xi_{\infty s}. \tag{2.129}$$

It follows from (2.126) that both sides in (2.129) are bounded by K in absolute value. Thus, the series $\sum_{t=1}^{u_m} |c_k e_k|/(t(t+1)) \sum_{s=1}^{t} \xi_{\infty s}$ is absolutely convergent, and converges to some real number $\Gamma_{\ell,k}^{(1)}$ as $m \to \infty$. Finally, by combining (2.127)–(2.129), it follows that

$$\lim_{m\to\infty} H_m^{(1)} = \lim_{m\to\infty} \sum_{t=1}^{M_n-1} C_{mt}S_{mt} = \Gamma_{\ell,k}^{(1)}. \tag{2.130}$$

Conclusion

Combining (2.111), (2.116) and (2.130), we finally arrive at

$$\lim_{m\to\infty} \log B_m^* = -2\left(\Gamma_{\ell,k}^{(1)} + \Gamma_{\ell,k}^{(2)}\right).$$

Recalling that $\log B_m^* \sim \log B_m$, it follows that $\log B_m$ converges to a finite limit, and accordingly the product B_m in (2.96) converges to a strictly positive number.

2.3.6 Proof of Theorem 2.3.2

The proof of Theorem 2.3.2 is essentially completed. Nevertheless, we include a brief summary. Theorem 2.3.2 states that if $\alpha = [0; \overline{a_1, \ldots, a_\ell}]$ is an irrational with a periodic continued fraction expansion, then there are positive constants $c_0, \ldots, c_{\ell-1}$ such that

$$\lim_{m\to\infty} Q_{\ell m+k}(\alpha) = \prod_{r=1}^{q_{\ell m+k}} |2\sin \pi r\alpha| = c_k \tag{2.131}$$

for each $k = 0, 1, 2, \ldots, \ell-1$. By Lemma 2.3.23, the product $Q_{\ell m+k}(\alpha)$ for fixed k can be decomposed as

$$Q_{\ell m+k}(\alpha) = A_m B_m C_m, \tag{2.132}$$

where A_m, B_m and C_m are defined in (2.95)–(2.97). We have seen in Section 2.3.3 that

$$\lim_{m\to\infty} A_m = 2\pi|c_k e_k| > 0.$$

Moreover, by Theorem 2.3.24 we have $\lim_{m\to\infty} C_m > 0$, and finally we have seen in Section 2.3.5 that also $\lim_{m\to\infty} B_m > 0$. It thus follows from (2.132) that (2.131) holds for some $C_k > 0$.

Proof of Corollary 2.3.3

We only sketch the proof of Corollary 2.3.3, as it largely follows that of Theorem 2.3.2. Let $\beta = [a_0; a_1, \ldots, a_h, \overline{a_{h+1}, \ldots, a_{h+\ell}}]$ and $\alpha = [0; \overline{a_{h+1}, \ldots, a_{h+\ell}}]$. It is an easy exercise to verify the identity

$$q_{h+u}(\beta) = q_{h+1}(\beta) q_u(\alpha) + q_h(\beta) p_u(\alpha) \tag{2.133}$$

for all $u \geq 0$. By combining (2.133) with Theorem 2.3.13 and Lemmas 2.3.14 and 2.3.16 for the purely periodic case, one can establish the closed form

$$q_{h+\ell m+k}(\beta) = \gamma_1^{(m)} q_{h+\ell+k}(\beta) L_m + (-1)^{\ell-1} \gamma_2^{(m)} q_{h+k}(\beta) L_{m-1}, \tag{2.134}$$

where $L_m = L_m(c(\alpha)^2, (-1)^{l-1})$ is the Lehmer sequence and $\gamma_1^{(m)}$ and $\gamma_2^{(m)}$ are defined as in Theorem 2.3.13. Moreover, one can find constants $c_{h,k}$ and $e_{h,k}$ (independent of m) such that

$$q_{h+\ell m+k}(\beta)|b|^m = c_{h,k} + \mathcal{O}(b^{2m}) \tag{2.135}$$

and

$$q_{h+\ell m+k}(\beta)\beta = p_{h+\ell m+k}(\beta) + e_{h,k}b^m, \tag{2.136}$$

with $b = b(\alpha)$ defined in (2.75). Note that (2.135) is essentially Lemma 2.3.14 for the irrational β, and similarly (2.136) corresponds to Lemma 2.3.16. Further calculations verify that

$$\frac{q_{h+\ell m+k-1}(\beta)}{q_{h+\ell m+k}(\beta)} = \alpha_{\sigma_k} + \mathcal{O}(b^{2m}), \tag{2.137}$$

which is basically Lemma 2.3.18 for β. Thus, we have all tools needed to prove that the limit

$$\lim_{m\to\infty} Q_{h+\ell m+k}(\beta)$$

indeed exists for each $k \in \{0, 1, \ldots \ell - 1\}$. Finally, it turns out that the product $|c_{h,k}e_{h,k}|$ is independent of h, that is

$$|c_{h,k}e_{h,k}| = |c_k e_k|, \tag{2.138}$$

with c_k and e_k given in (2.81) and (2.83). By carefully examining the proof of Theorem 2.3.2, it is clear that (2.138) guarantees that

$$\lim_{m\to\infty} Q_{h+\ell m+k}(\beta) = \lim_{m\to\infty} Q_{\ell m+k}(\alpha),$$

and this completes the proof of Corollary 2.3.3.

2.4 Conclusions and further research

Let us briefly summarize the main results of the second part of this thesis. The central object of Chapter 2 was the following sequence of trigonometric products (referred to as Sudler product)

$$P_N(\alpha) = \prod_{r=1}^{N} |2\sin(\pi r\alpha)|,$$

where $N \in \mathbb{N}$ and $\alpha \in (0,1)$ and irrational. Amongst many others this sequence was also analysed by Lubinsky and one of his results in [76] states that

$$\liminf_{N\to\infty} P_N(\alpha) = 0 \tag{2.139}$$

if α has unbounded continued fraction coefficients. Moreover he suggested that the same result is valid for α with bounded partial quotients. However, the main result of Section 2.2 states that this can not be true. For α being the golden ratio $\varphi = [0;\overline{1}]$ we were capable to prove that

$$\liminf_{N\to\infty} P_N(\varphi) > 0, \tag{2.140}$$

which provides a counter example to the suggestion of Lubinsky.

In the second section of Chapter 2 we shifted our attention to certain subsequences of $(P_N(\alpha))_{N\in\mathbb{N}}$. Verschueren and Mestel showed in [104] that there exists a constant $c > 0$ such that

$$\lim_{n\to\infty} P_{F_n}(\varphi) = c, \tag{2.141}$$

where φ is again the golden ratio and $(F_n)_{n\in\mathbb{N}}$ is the Fibonacci sequence. Additionally they conjectured in [104] that the above result is valid for the whole class of quadratic irrationals. The main goal of the remaining section was to prove exactly this conjecture. More precise, we were capable of proving for a (purely periodic) quadratic irrational α with continued fraction expansion $\alpha = [0;\overline{a_1,\dots,a_\ell}]$ that for each $k \in \{0,\dots,\ell-1\}$

$$\lim_{m\to\infty} P_{q_{\ell m+k}}(\alpha) = \prod_{r=1}^{q_{\ell m+k}(\alpha)} |2\sin(\pi r\alpha)| = c_k,$$

where $q_n(\alpha)$ is the nth best approximation denominator of α and $c_k > 0$ some constant independent of m. There is also a version of this result for quadratic irrationals of the form $\alpha = [a_0; a_1,\dots,a_h,\overline{a_{h+1},\dots,a_{h+\ell}}]$ (see Corollary 2.3.3).

2.4.1 Further research

In this final subsection we want to discuss further research ideas and pose some interesting problems related to the topics we have seen in Chapter 2.

First of all the work of Verschueren and Mestel [104] was essential to proof that $\liminf_{N\to\infty} P_N(\varphi) > 0$. It was exactly this work that has been generalised in Section 2.3 to quadratic irrationals and therefore it would be a natural question to ask whether it is possible to join the ideas of Section 2.2 and Section 2.3 and prove that

$$\liminf_{N\to\infty} P_N(\alpha) > 0 \tag{2.142}$$

for other quadratic irrationals α than the golden ratio. Unfortunately an extension to all quadratic irrationals is too much to hope for. This is reflected by the following theorem which is a direct consequence of the results in [76]. For more information in this direction and a nicely illustrated proof of the subsequent theorem we refer to [39].

Theorem 2.4.1. *Let $\alpha = [0; a_1, a_2, \ldots]$ have bounded continued fraction coefficients, and let $M = \max_{j\in\mathbb{N}} a_j$. Provided M is sufficiently large, there exists some threshold value $K = K(M)$ such that if $a_j \geq K$ infinitely often, then*

$$\liminf_{n\to\infty} P_n(\alpha) = 0.$$

Remark. We want to point out that Lubinsky mentioned in [76] that (2.139) is already true if the continued fraction coefficients of α exceed some absolute constant K infinitely often. Unfortunately the constant K is not described in more detail and it is not clear for us how to derive this absolute constant from his proves. We were only able to derive a weaker version of his statement (see Theorem 2.4.1), where the constant $K(M)$ is not absolute.

In other words the actual size of the continued fraction coefficients of α plays a crucial role if one wants to determine if either (2.139) or (2.142) is true. Additionally note that pursuing the same strategy as it was outlined in Section 2.2.2 for some other quadratic irrational α forces a switch from the Zeckendorf representation of an integer N to a more general representation (e.g. Ostrowski representation (see [4, 92]) or β-expansion (see [91])) which unfortunately results in several technical problems. In order to keep these technicalities at a minimal extent and still gain some new information, a possible next step could be to try to follow the ideas of Section 2.2 and Section 2.3 in the case where $\alpha = [0; \overline{a}]$ and $a \in \mathbb{N}$ with $a \geq 2$ (i.e. $M = K = a$). Numerical experiments indicate that in this case the critical value of a seems to be 5. This phenomenon is illustrated in more detail in Table 2.1.

α	Evolution of minima $(P_n(\alpha), n)$
$[0;\overline{1}]$	$(1.865,\ 1)$
$[0;\overline{2}]$	$(1.928,\ 1)$
$[0;\overline{3}]$	$(1.333,\ 1)$
$[0;\overline{4}]$	$(1.351,\ 1)$
$[0;\overline{5}]$	$(1.138,\ 1)$
$[0;\overline{6}]$	$(0.977,\ 1), (0.907,\ 7), (0.849,\ 44), (0.794,\ 272), (0.742,\ 1\,677), (0.693,\ 10\,335)$
$[0;\overline{7}]$	$(0.852,\ 1), (0.708,\ 8), (0.589,\ 58), (0.491,\ 415), (0.408,\ 2\,964), (0.340,\ 21\,164)$
$[0;\overline{8}]$	$(0.755,\ 1), (0.564,\ 9), (0.422,\ 74), (0.316,\ 602), (0.236,\ 4\,891), (0.177,\ 39\,731)$

Table 2.1: Evolution of minima of $P_n(\alpha)$ for $n = 1, \ldots, 50\,000$.

Second, recall that the asymptotic behaviour of $P_{q_n(\alpha)}(\alpha)$ has been studied for the cases where α is the golden ratio [104] and where α is a quadratic irrational [40]. The underlying structure of the continued fraction expansion of α is essential in both proofs. It is therefore the subject of current research to investigate if it is possible to widen the class of irrationals for which $(P_{q_n(\alpha)}(\alpha))_{\geq 1}$ converges (in some sense). One should point out that we already know by the work of Lubinsky [76] that $\lim_{i\to\infty} P_{q_{n_i}(\alpha)}(\alpha) = 0$ if $\alpha = [a_0; a_1, a_2, \ldots]$ and $(a_{n_i})_{i\in\mathbb{N}}$ is strictly increasing. Using the following notation to represent the continued fraction expansion of Eulers number in the subsequent way

$$\mathrm{e} = [2; 1, 2, 1, 1, 4, 1, 1, 6, 1, 1, 8, \ldots] = [2; 1, 2, \overline{1, 1, 2n+2}]_{n=1}^{\infty},$$

gives rise to the question if $P_{q_n(\mathrm{e})}(\mathrm{e})$ shows a similar behaviour as a quadratic irrational with period $\ell = 3$. This is indeed what numerical experiments suggest, i.e. numerics indicate that $P_{q_n(\mathrm{e})}(\mathrm{e})$ has again three subsequences (see Figure 2.1). However, the limiting behaviour of the corresponding subsequences is still subject to further research.

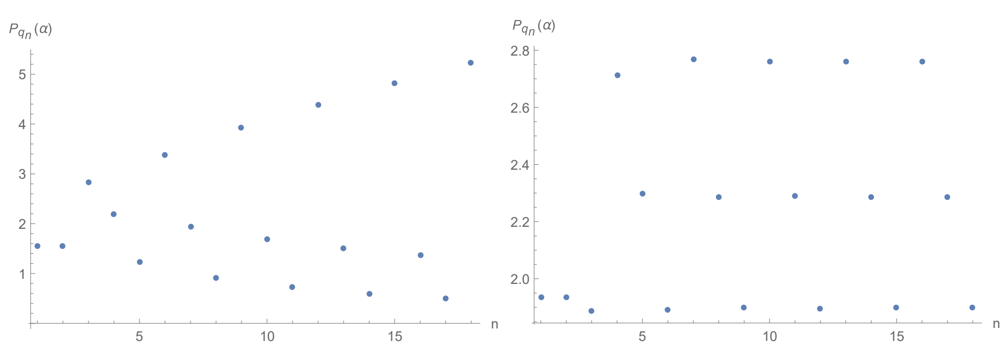

Figure 2.1: Illustration of $P_{q_n(\alpha)}(\alpha)$ for $n \in \{1, \ldots, 18\}$, where for $\alpha = \mathrm{e}$ (left) and $\alpha = [0; \overline{1, 1, 2}]$ (right).

Moreover, one could even try to let go of the periodic structure of the underlying irrational α and allow patterns of variable length in the continued fraction expansion of α. Maybe the simplest example would be

$$\alpha = [0; 1, 2, 1, 1, 2, 1, 1, 1, 2, 1, 1, 1, 1, 2, \ldots]. \tag{2.143}$$

It would be an interesting task to analyse the asymptotic behaviour of $P_{q_n(\alpha)}(\alpha)$ with α given in (2.143).

Last but not least there are several open problems related to the case $\alpha = \varphi$. It would be a nice additional result to determine if the minimum of the sequence $P_N(\varphi)$ exists. Indeed numerical experiments (see Table 2.1) indicate that

$$\min_{N\in\mathbb{N}} P_N(\varphi) = P_1(\varphi) = 1.864\ldots, \tag{2.144}$$

but unfortunately this is not evident from the proof we have seen in Section 2.2. Similar numerical computations give rise to the more general conjecture that for $n \geq 3$ and all $N \in \{F_{n-1}, \ldots, F_n - 1\}$ we have

$$P_{F_{n-1}}(\varphi) \leq P_N(\varphi) \leq P_{F_n-1}(\varphi). \tag{2.145}$$

Exactly this behaviour is illustrated in Figure 2.2.

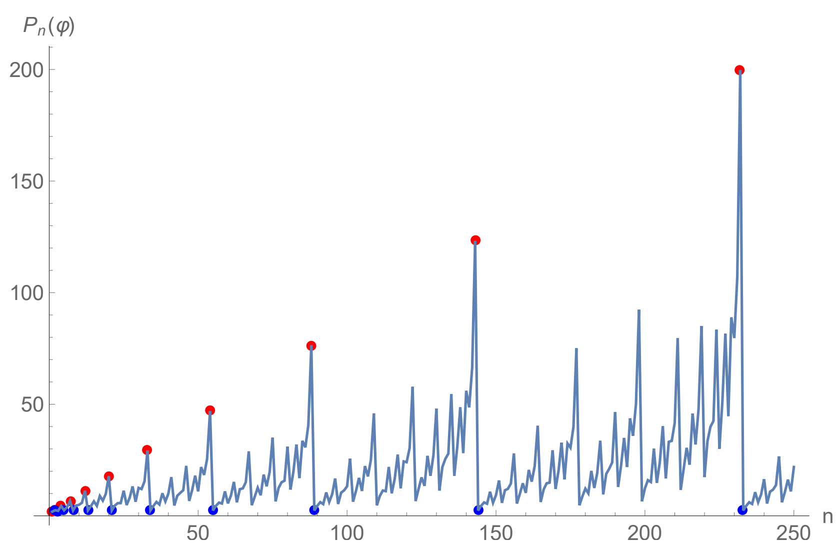

Figure 2.2: Value of $P_n(\varphi)$, with the two subsequences $P_{F_n}(\varphi)$ and $P_{F_n-1}(\varphi)$ indicated by blue and red marks, respectively.

For similar conjectures (where α is not necessarily the golden ratio) see also the introduction of [3] and the references therein. Observe that the upper bound of (2.145) would imply for $n \geq 3$ and all $N \in \{F_{n-1}, \dots, F_n - 1\}$ that

$$P_N(\varphi) \leq P_{F_n - 1}(\varphi) \leq cF_n \leq 2cN, \tag{2.146}$$

where we have additionally used that $P_{F_n-1}(\varphi) \leq cF_n$ (see [104]) for the constant $c = \sqrt{5}/4 \lim_{n\to\infty} P_{F_n}(\varphi) \simeq 1.35$. The linear bound stated in (2.146) is also supported by numerical calculations (with an even better constant) shown in Figure 2.3.

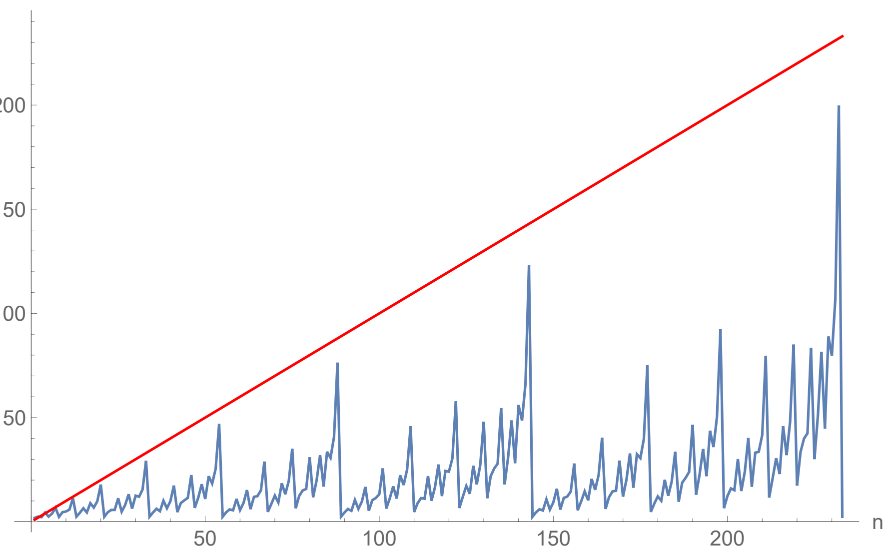

Figure 2.3: Value of $P_n(\alpha)$ for $\alpha = (\sqrt{5}-1)/2$ (blue line) plotted against $f(n) = n$ (red line).

This linear bound in N would be a significant improvement to all known bounds of $P_N(\varphi)$ and is an interesting problem on its own.

Bibliography

[1] Aistleitner, C. Covering numbers, dyadic chaining and discrepancy. *J. Complexity 27*, 6 (2011), 531–540.

[2] Aistleitner, C. On the inverse of the discrepancy for infinite dimensional infinite sequences. *J. Complexity 29*, 2 (2013), 182–194.

[3] Aistleitner, C., Larcher, G., Pillichshammer, F., Eddin, S. S., and Tichy, R. F. On Weyl products and uniform distribution modulo one. *Monatsh. Math. 185*, 3 (2018), 365–395.

[4] Allouche, J.-P., and Shallit, J. *Automatic sequences: theory, applications, generalizations.* Cambridge University Press, Cambridge, 2003.

[5] Andrews, G. E. *q-series: their development and application in analysis, number theory, combinatorics, physics, and computer algebra*, vol. 66 of *CBMS Regional Conference Series in Mathematics.* Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 1986.

[6] Andrews, G. E., Askey, R., and Roy, R. *Special functions*, vol. 71 of *Encyclopedia of Mathematics and its Applications.* Cambridge University Press, Cambridge, 1999.

[7] Andrews, G. E., and Bell, J. Euler's pentagonal number theorem and the Rogers-Fine identity. *Ann. Comb. 16*, 3 (2012), 411–420.

[8] Awata, H., Hirano, S., and Shigemori, M. The partition function of ABJ theory. *Prog. Theoret. Exp. Phys. 5* (2013).

[9] Bailey, W. N. *Generalized hypergeometric series.* Cambridge Tracts in Mathematics and Mathematical Physics, No. 32. Stechert-Hafner, Inc., New York, 1964.

[10] Beck, J. Probabilistic diophantine approximation, I. Kronecker-sequences. *Ann. Math. 140* (1994), 451–502.

[11] Bell, J. Estimates for the norms of products of sines and cosines. *J. Math. Anal. Appl. 405*, 2 (2013), 530–545.

[12] Bernstein, S. *The theory of probabilities.* Gastehizdat Publishing House, Moscow, 1946.

[13] Bilyk, D., Lacey, M. T., and Vagharshakyan, A. On the small ball inequality in all dimensions. *J. Funct. Anal. 254*, 9 (2008), 2470–2502.

[14] Bilyk, D., Temlyakov, V. N., and Yu, R. The L_2 discrepancy of two-dimensional lattices. In *Recent advances in harmonic analysis and applications*, vol. 25 of *Springer Proc. Math. Stat.* Springer, New York, 2013, pp. 63–77.

[15] Borel, E. Les probabilités denombrables et leurs applications arithmétiques. *Rend. Circ. Mar. Palermo 27* (1909), 247–271 (french).

[16] Bourgain, J., and Chang, M.-C. On a paper of Erdős and Szekeres. *J. Anal. Math. 136*, 1 (2018), 253–271.

[17] Buslaev, V. I. On the convergence of the Rogers-Ramanujan continued fraction. *Mat. Sb. 194*, 6 (2003), 43–66.

[18] Bykovskiĭ, V. A. Deviation of Korobov grids. *Izv. Ross. Akad. Nauk Ser. Mat. 76*, 3 (2012), 19–38.

[19] Chen, W. W. L., and Skriganov, M. M. Explicit constructions in the classical mean squares problem in irregularities of point distribution. *J. Reine Angew. Math. 545* (2002), 67–95.

[20] Dick, J. A note on the existence of sequences with small star discrepancy. *J. Complexity 23*, 4-6 (2007), 649–652.

[21] Dick, J., and Kritzer, P. On a projection-corrected component-by-component construction. *J. Complexity 32*, 1 (2016), 74–80.

[22] Dick, J., Kritzer, P., Leobacher, G., and Pillichshammer, F. Constructions of general polynomial lattice rules based on the weighted star discrepancy. *Finite Fields Appl. 13*, 4 (2007), 1045–1070.

[23] Dick, J., Kritzer, P., Leobacher, G., and Pillichshammer, F. A reduced fast component-by-component construction of lattice points for integration in weighted spaces with fast decreasing weights. *J. Comput. Appl. Math. 276* (2015), 1–15.

[24] Dick, J., Kuo, F. Y., and Sloan, I. H. High-dimensional integration: the quasi-Monte Carlo way. *Acta Numer. 22* (2013), 133–288.

[25] Dick, J., Leobacher, G., and Pillichshammer, F. Construction algorithms for digital nets with low weighted star discrepancy. *SIAM J. Numer. Anal. 43*, 1 (2005), 76–95.

[26] Dick, J., and Pillichshammer, F. *Digital nets and sequences. Discrepancy theory and quasi-Monte Carlo integration*. Cambridge University Press, Cambridge, 2010.

[27] Dick, J., and Pillichshammer, F. Optimal L_2 discrepancy bounds for higher order digital sequences over the finite field $\mathbb{F}_2$. *Acta Arith. 162*, 1 (2014), 65–99.

[28] Doerr, B., Gnewuch, M., and Srivastav, A. Bounds and constructions for the star discrepancy via δ-covers. *J. Complexity 21* (2005), 691–709.

[29] Driver, K. A., and Lubinsky, D. S. Summary: convergence of Padé approximants for a q-hypergeometric series (Wynn's Power Series III). *Aequationes Math. 44*, 2-3 (1992), 328.

[30] Driver, K. A., Lubinsky, D. S., Petruska, G., and Sarnak, P. Irregular distribution of $\{n\beta\},\ n = 1, 2, 3\cdots$, quadrature of singular integrands, and curious basic hypergeometric series. *Indag. Math. (N.S.) 2*, 4 (1991), 469–481.

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

[32] Erdős, P., and Szekeres, G. On the product $\prod_{k=1}^{n}(1 - z^{a_k})$. *Acad. Serbe Sci. Publ. Inst. Math. 13* (1959), 29–34.

[33] Exton, H. *q-hypergeometric functions and applications*. Ellis Horwood Series: Mathematics and its Applications. Ellis Horwood Ltd., Chichester; Halsted Press [John Wiley & Sons, Inc.], New York, 1983.

[34] Freimann, G., and Halberstam, H. On a product of sines. *Acta Arith. 49* (1988), 377–385.

[35] Gasper, G., and Rahman, M. *Basic hypergeometric series*, second ed., vol. 96 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2004.

[36] Gnewuch, M. Bracketing numbers for axis-parallel boxes and applications to geometric discrepancy. *J. Complexity 24*, 2 (2008), 154–172.

[37] Gnewuch, M. Entropy, randomization, derandomization, and discrepancy. In *Monte Carlo and quasi-Monte Carlo methods 2010*, vol. 23 of *Springer Proc. Math. Stat.* Springer, Heidelberg, 2012, pp. 43–78.

[38] Grebogi, C., Ott, E., Pelikan, S., and Yorke, J. A. Strange attractors that are not chaotic. *Phys. D 13*, 1-2 (1984), 261–268.

[39] Grepstad, S., Kaltenböck, L., and Neumüller, M. On the asymptotic behaviour of the sine product $\prod_{r=1}^{n} |2\sin(\pi r\alpha)|$. In *Discrepancy theory*, D. Bilyk, J. Dick, and F. Pillichshammer, Eds. Radon Series on Computational and Applied Mathematics, to appear.

[40] Grepstad, S., and Neumüller, M. Asymptotic behaviour of the Sudler product of sines for quadratic irrationals. *J. Math. Anal. Appl. 465*, 2 (2018), 928–960.

[41] Halász, G. On Roth's method in the theory of irregularities of point distributions. In *Recent progress in analytic number theory, Vol. 2 (Durham, 1979)*. Academic Press, London-New York, 1981, pp. 79–94.

[42] Halton, J. H. On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. *Numer. Math. 2* (1960), 84–90.

[43] Hammersley, J. M. Monte Carlo methods for solving multivariable problems. *Ann. New York Acad. Sci. 86* (1960), 844–874 (1960).

[44] Hardy, G. H. A chapter from Ramanujan's notebook. *Proc. Cambridge Philos. Soc. 21* (1923), 492–503.

[45] Heinrich, S., Novak, E., Wasilkowski, G. W., and Woźniakowski, H. The inverse of the star discrepancy depends linearly on the dimension. *Acta Arith. 96*, 3 (2001), 279–302.

[46] Hellekalek, P. General discrepancy estimates: the Walsh function system. *Acta Arith. 67*, 3 (1994), 209–218.

[47] Hickernell, F. J., and Niederreiter, H. The existence of good extensible rank-1 lattices. *J. Complexity 19*, 3 (2003), 286–300. Numerical integration and its complexity (Oberwolfach, 2001).

[48] Hinrichs, A. Covering numbers, Vapnik-Červonenkis classes and bounds for the star discrepancy. *J. Complexity 20*, 4 (2004), 477–483.

[49] Hinrichs, A., Pillichshammer, F., and Schmid, W. C. Tractability properties of the weighted star discrepancy. *J. Complexity 24*, 2 (2008), 134–143.

[50] Hlawka, E. Funktionen von beschränkter Variation in der Theorie der Gleichverteilung. *Ann. Mat. Pura Appl. (4) 54* (1961), 325–333 (german).

[51] Hlawka, E. Zur angenäherten Berechnung mehrfacher Integrale. *Monatsh. Math. 66* (1962), 140–151 (german).

[52] Hlawka, E. Interpolation analytischer Funktionen auf dem Einheitskreis. In *Number Theory and Analysis (Papers in Honor of Edmund Landau)*. Plenum, New York, 1969, pp. 97–118 (german).

[53] Hlawka, E. Über ein Produkt, das in der Interpolation analytischer Funktionen im Einheitskreis auftritt. In *Zahlentheoretische Analysis*, vol. 1114 of *Lecture Notes in Math.* Springer, Berlin, 1985, pp. 19–25 (german).

[54] Hofer, R., and Puchhammer, F. On the discrepancy of two-dimensional perturbed Halton-Kronecker sequences and lacunary trigonometric products. *Acta Arith. 180*, 4 (2017), 365–392.

[55] Joe, S. Construction of good rank-1 lattice rules based on the weighted star discrepancy. In *Monte Carlo and Quasi-Monte Carlo Methods 2004*. Springer, Berlin, 2006, pp. 181–196.

[56] Knill, O., and Lesieutre, J. Analytic continuation of Dirichlet series with almost periodic coefficients. *Complex Anal. Oper. Theory 6*, 1 (2012), 237–255.

[57] Knill, O., and Tangerman, F. Self-similarity and growth in Birkhoff sums for the golden rotation. *Nonlinearity 24*, 11 (2011), 3115–3127.

[58] Koekoek, R., Lesky, P. A., and Swarttouw, R. F. *Hypergeometric orthogonal polynomials and their q-analogues.* Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2010.

[59] Koepf, W. *Hypergeometric summation. An algorithmic approach to summation and special function identities*, second ed. Universitext. Springer, London, 2014.

[60] Koksma, J. F. A general theorem from the theory of uniform distribution modulo 1. *Mathematica, Zutphen. B. 11* (1942), 7–11.

[61] Korobov, N. M. Approximate evaluation of repeated integrals. *Dokl. Akad. Nauk SSSR 124* (1959), 1207–1210.

[62] Korobov, N. M. On number-theoretic methods in approximate analysis. In *Probl. Numer. Math. Comp. Techn. (Russian)*. Gosudarstv. Naučno-Tehn. Izdat. Mašinostr. Lit., Moscow, 1963, pp. 36–44.

[63] Kritzinger, R., and Laimer, H. A reduced fast component-by-component construction of lattice point sets with small weighted star discrepancy. *Unif. Distrib. Theory 10*, 2 (2015), 21–47.

[64] Kuipers, L., and Niederreiter, H. *Uniform distribution of sequences.* John Wiley, New York, 1974. Reprint, Dover Publications, Mineola, NY, 2006.

[65] Kuznetsov, S. P., Pikovsky, A. S., and Feudel, U. Birth of a strange nonchaotic attractor: a renormalization group analysis. *Phys. Rev. E (3) 51*, 3, part A (1995), R1629–R1632.

[66] Larcher, G. On the distribution of an analog to classical Kronecker-sequences. *J. Number Theory 52*, 2 (1995), 198–215.

[67] Larcher, G. Digital point sets: analysis and application. In *Random and quasi-random point sets*, vol. 138 of *Lect. Notes Stat.* Springer, New York, 1998, pp. 167–222.

[68] Larcher, G., and Niederreiter, H. Kronecker-type sequences and nonarchimedean diophantine approximation. *Acta Arith. 63* (1993), 380–396.

[69] Larcher, G., Niederreiter, H., and Schmid, W. C. Digital nets and sequences constructed over finite rings and their application to quasi-Monte Carlo integration. *Monatsh. Math. 121*, 3 (1996), 231–253.

[70] Larcher, G., and Pillichshammer, F. Sums of distances to the nearest integer and the discrepancy of digital nets. *Acta Arith. 106*, 4 (2003), 379–408.

[71] Larcher, G., and Pillichshammer, F. Metrical lower bounds on the discrepancy of digital Kronecker-sequences. *J. Number Theory 135* (2014), 262–283.

[72] Lehmer, D. H. An extended theory of Lucas' functions. *Ann. of Math. (2) 31*, 3 (1930), 419–448.

[73] Leobacher, G., and Pillichshammer, F. *Introduction to quasi-Monte Carlo integration and applications*. Compact Textbook in Mathematics. Birkhäuser/Springer, Cham, 2014.

[74] Lidl, R., and Niederreiter, H. *Introduction to finite fields and their applications*, first ed. Cambridge University Press, Cambridge, 1994.

[75] Löbbe, T. Probabilistic star discrepancy bounds for lacunary point sets. *arXiv:1408.2220* (2014).

[76] Lubinsky, D. S. The size of $(q;q)_n$ for q on the unit circle. *J. Number Theory 76*, 2 (1999), 217–247.

[77] Niederreiter, H. Point sets and sequences with small discrepancy. *Monatsh. Math. 104*, 4 (1987), 273–337.

[78] Niederreiter, H. Low-discrepancy and low-dispersion sequences. *J. Number Theory 30*, 1 (1988), 51–70.

[79] Niederreiter, H. Low-discrepancy point sets obtained by digital constructions over finite fields. *Czechoslovak Math. J. 42(117)*, 1 (1992), 143–166.

[80] Niederreiter, H. *Random number generation and quasi-Monte Carlo methods*, vol. 63 of *CBMS-NSF Series in Applied Mathematics*. SIAM, Philadelphia, PA, 1992.

[81] Novak, E., and Woźniakowski, H. *Tractability of multivariate problems. Volume I: linear information*, vol. 6 of *EMS Tracts in Mathematics*. EMS, Zürich, 2008.

[82] Novak, E., and Woźniakowski, H. *Tractability of multivariate problems. Volume II: standard information for functionals*, vol. 12 of *EMS Tracts in Mathematics*. EMS, Zürich, 2010.

[83] Novak, E., and Woźniakowski, H. *Tractability of multivariate problems. Volume III: standard information for operators*, vol. 18 of *EMS Tracts in Mathematics*. EMS, Zürich, 2012.

[84] Nuyens, D., and Cools, R. Fast algorithms for component-by-component construction of rank-1 lattice rules in shift-invariant reproducing kernel Hilbert spaces. *Math. Comp. 75*, 254 (2006), 903–920.

[85] Nuyens, D., and Cools, R. Fast component-by-component construction of rank-1 lattice rules with a non-prime number of points. *J. Complexity 22*, 1 (2006), 4–28.

[86] Perron, O. *Die Lehre von den Kettenbrüchen (2nd ed.)*. Chelsea Publishing Co., New York, N. Y., 1950 (german).

[87] Petruska, G. On the radius of convergence of q-series. *Indag. Math. (N.S.) 3*, 3 (1992), 353–364.

[88] Pillichshammer, F. Polynomial lattice point sets. In *Monte Carlo and Quasi-Monte Carlo Methods 2010*, vol. 23 of *Springer Proc. Math. Stat.* Springer, Heidelberg, 2012, pp. 189–210.

[89] Proinov, P. D. On irregularities of distribution. *C. R. Acad. Bulgare Sci. 39*, 9 (1986), 31–34.

[90] Rainville, E. D. *Special functions*, first ed. Chelsea Publishing Co., Bronx, N.Y., 1971.

[91] Rényi, A. Representation for real numbers and their ergodic properties. *Acta. Math. Acad. Sci. Hungar. 8* (1957), 477–493.

[92] Rockett, A. M., and Szüsz, P. *Continued fractions.* World Scientific Publishing Co., Inc., River Edge, NJ, 1992.

[93] Roth, K. F. On irregularities of distribution. *Mathematika 1* (1954), 73–79.

[94] Schmidt, W. M. Irregularities of distribution. VII. *Acta Arith. 21* (1972), 45–50.

[95] Schmidt, W. M. Irregularities of distribution. X. In *Number theory and algebra (H. Zassenhaus ed.)*. Academic Press, New York, 1977, pp. 311–329.

[96] Slater, L. J. *Generalized hypergeometric functions.* Cambridge University Press, Cambridge, 1966.

[97] Sloan, I. H., and Reztsov, A. V. Component-by-component construction of good lattice rules. *Math. Comp. 71*, 237 (2002), 263–273.

[98] Sloan, I. H., and Woźniakowski, H. When are quasi-Monte Carlo algorithms efficient for high-dimensional integrals? *J. Complexity 14*, 1 (1998), 1–33.

[99] Sudler, Jr., C. An estimate for a restricted partition function. *Quart. J. Math. Oxford Ser. (2) 15* (1964), 1–10.

[100] Temme, N. M. *Special functions.* Wiley, 1996.

[101] Triebel, H. *Bases in function spaces, sampling, discrepancy, numerical integration*, vol. 11 of *EMS Tracts in Mathematics.* EMS, Zürich, 2010.

[102] van der Corput, J. G. Verteilungsfunktionen I. *Proc. Akad. Amsterdam 38* (1935), 813–821.

[103] van der Corput, J. G. Verteilungsfunktionen II. *Proc. Akad. Amsterdam 38* (1935), 1058–1066.

[104] Verschueren, P., and Mestel, B. Growth of the Sudler product of sines at the golden rotation number. *J. Math. Anal. Appl. 433*, 1 (2016), 200–226.

[105] Wagner, G. On a problem of Erdős in diophantine approximation. *Bull. Lond. Math. Soc. 12*, 2 (1980), 81–88.

[106] Weyl, H. Über die Gleichverteilung von Zahlen mod. Eins. *Math. Ann. 77*, 3 (1916), 313–352 (german).

[107] Wright, E. M. Proof of a conjecture of Sudler's. *Quart. J. Math. Oxford Ser. (2) 15* (1964), 11–15.

[108] Zeckendorf, E. Représentation des nombres naturels par une somme de nombres de Fibonacci ou de nombres de Lucas. *Bull. Soc. Roy. Sci. Liège 41* (1972), 179–182 (french).

Curriculum Vitae

MARIO NEUMÜLLER

Personal Information

Date of Birth	18.08.1991
Place of birth	Linz
Nationality	Austrian

Education

1997-2005	Elementary school
2005-2009	High school in Linz
2010-2013	Bachelor's programme in technical mathematics at JKU Linz
2013-2015	Master's programme in computer mathematics at JKU Linz
2015-2019	Doctoral programme in technical sciences at JKU Linz

Publications

- F. Pillichshammer, M. Neumüller: *Metrical star discrepancy bounds for lacunary subsequences of digital Kronecker-sequences and polynomial tractability*, Unif. Distrib. Theory 13 1 (2018), 65–86.
- R. Kritzinger, H. Laimer, M. Neumüller: *A reduced fast construction of polynomial lattice point sets with low weighted star discrepancy.* In Monte Carlo and quasi-Monte Carlo methods 2016. Springer, Cham, 2018, pp. 377–394.
- S. Grepstad, M. Neumüller: *Asymptotic behaviour of the Sudler product of sines for quadratic irrationals*, J. Math. Anal. Appl. 465 2 (2018), 928–960.
- S. Grepstad, L. Kaltenböck, M. Neumüller: *A positive lower bound for* $\liminf_{N\to\infty} P_N(\varphi)$, Proc. Am. Math. Soc. (to appear).
- S. Grepstad, L. Kaltenböck, M. Neumüller: *On the asymptotic behaviour of the sine product* $\prod_{r=1}^{n} |2\sin(\pi r\alpha)|$. In: D. Bilyk, J. Dick, F. Pillichshammer (Eds.) *Discrepancy theory*, Radon Series on Computational and Applied Mathematics (to appear).

Talks

8/2016, Stanford	*Probabilistic star discrepancy bounds for digital Kronecker-sequences.*
11/2016, Vienna	*Metric star discrepancy bounds for subsequences of digital Kronecker-sequences.*
7/2017, Montreal	*Reduced fast polynomial lattice point sets with small weighted star discrepancy.*
7/2018, Rennes	*Asymptotic behaviour of the Sudler product of sines for quadratic irrationals.*
11/2018, Linz	*Asymptotic behaviour of the Sudler product of sines for quadratic irrationals.*