DISSERTATION

Exponential sum estimates and Fourier analytic methods for digitally based dynamical systems

Abschätzungen für Exponentialsummen und Fourier-analytische Methoden für auf Zahlensystemen beruhende dynamische Systeme

Ausgeführt zum Zwecke der Erlangung des akademischen Grades eines Doktors der technischen Wissenschaften unter der Leitung von

Univ.Prof. Michael Drmota
Institut für
Diskrete Mathematik und Geometrie
Technische Universität Wien

und

Prof. Joël Rivat
Institut de
Mathématiques de Luminy
Université d'Aix-Marseille

eingereicht an der Technischen Universität Wien
Fakultät für Mathematik und Geoinformation

von

Clemens MÜLLNER
0925841

Wien, am 17. Januar 2017

Faculté des sciences de Luminy
École Doctorale de Mathématiques et d'Informatique de Marseille
Institut de Mathématiques de Luminy

T H È S E

pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ D'AIX-MARSEILLE

Spécialité: Mathématiques

par

Clemens MÜLLNER

Titre:

Exponential sum estimates and Fourier analytic methods for digitally based dynamical systems

Estimation de sommes d'exponentielles et méthodes d'analyse de Fourier pour les systèmes dynamiques basés sur les développements digitaux

Les directeurs de thèse: Joël RIVAT et Michael DRMOTA

Les rapporteurs: Cécile DARTYGE et Christian ELSHOLTZ

Jury:	Cécile DARTYGE	Rapporteur
	Michael DRMOTA	Directeur de thèse
	Christian ELSHOLTZ	Rapporteur
	Mariusz LEMÁNCZYK	Examinateur
	Joël RIVAT	Directeur de thèse

Diese Arbeit wurde im Rahmen eines gemeinsam betreuten Promotionsverfahrens

basierend auf der Vereinbarung

Convention de co-tutelle de thèse

zwischen der
Technischen Universität Wien
und der
Université d'Aix-Marseille

ausgeführt.

Ce travail est présenté dans le cadre d'une

Convention de co-tutelle de thèse

entre

l'Université d'Aix-Marseille

et

l'Université Technique de Vienne (Technische Universität Wien).

Contents

Abstract in German, French and English **5**

0 Introduction in German and English **9**

0.1 Deutsche Einleitung 9

0.2 Introduction in English 14

1 Automatic Sequences **19**

1.1 Introduction to Automatic Sequences 19

1.1.1 Examples of Automatic Sequences 20

1.1.2 Basic Properties of Automatic Sequences 23

1.2 Equivalent Definitions for Automatic Sequences 24

1.2.1 Fixed Points of Uniform Morphisms 25

1.2.2 The k-Kernel 25

1.3 Frequencies 27

1.4 Recent Results for Automatic Sequences 28

1.4.1 Synchronizing Automata 28

1.4.2 Rudin–Shapiro Sequence 29

1.4.3 The Thue–Morse Sequence along Squares 31

1.4.4 Invertible Automatic Sequence 31

2 Naturally Induced Transducers **33**

2.1 Automaton to Transducer 34

2.2 Connection between an automaton and its induced transducer 38

2.3 Length restrictions for naturally induced transducers 40

2.4 Arithmetic restriction for naturally induced transducers 45

2.5 Reduction to special naturally induced Transducers 50

3 Automatic Sequences fulfill the Sarnak Conjecture 53

3.1 The Sarnak Conjecture 53
3.2 Reduction of Theorem 3.12 55
3.3 A general Möbius Principle 61
3.4 Fourier Estimates 63
3.5 Carry Lemma 68
3.6 Proof of Proposition 3.2.2 69
3.7 Technical and Auxiliary Results 71
3.7.1 Technical Results 71
3.7.2 Van-der-Corput's inequality 74
3.7.3 Vaaler's method 74
3.8 Proof of Theorem 3.3.3 and Theorem 3.3.4 75

4 Subsequences of Automatic sequences 83

4.1 Linear subsequences of automatic sequences 84
4.2 Prime Number Theorem for Automatic Sequences 89
4.3 Technical Results 92
4.3.1 Frequency of T along primes 93

5 Normality of digital sequences along squares 97

5.1 Outline 98
5.2 Digital Functions 99
5.3 Bounds on Fourier Transforms 104
5.3.1 Auxiliary Results for the Bounds of the Fourier Transforms 104
5.3.2 Fourier estimates 105
5.3.3 Proof of Proposition 5.3.7 108
5.3.4 Proof of Proposition 5.3.8 114
5.4 Auxiliary Results 120
5.4.1 Sums of geometric series 120
5.4.2 Gauss sums 121
5.4.3 Carry Lemmas 122

5.5 Proof of the Main Theorem 125
5.5.1 The case $K \in \mathbb{Z}$ 125
5.5.2 The case $K \notin \mathbb{Z}$ 128

Abstract in German, French and English

Deutsche Kurzfassung

Die vorliegende Dissertation wurde durch zwei Vermutungen, einer von Gelfond aus den 60er Jahren des 20. Jahrhunderts und einer von Sarnak aus dem Jahr 2010, die in den letzen 10 Jahren in der Zahlentheorie große Beachtung gefunden haben, stark beeinflusst.

Im Jahr 1968 bewies Gelfond erstmals Gleichverteilungsresultate der Ziffernsumme modulo m in allgemeinen arithmetischen Progressionen und stellte weiters drei Probleme. Die erste handelt von der gemeinsamen Verteilung der Ziffernsummen bezüglich verschiedener Basen und wurde 1999 vollständig von Kim gelöst. Die zweite und dritte Frage beschäftigen sich mit der Ziffernsumme von Primzahlen und polynomialen Teilfolgen. Mauduit und Rivat lösten diese Probleme für Prim- und Quadratzahlen im Jahre 2010 bzw. 2009. Weiters ist zu erwähnen, dass im Jahr 2013 Drmota, Mauduit und Rivat das Resultat betreffend der Folge der Ziffernsumme der Quadratzahlen verallgemeinerten. Sie zeigten, dass jeder Block - d.h. Teilfolge der Länge L - asymptotisch gleich häufig auftritt.

Die Sarnaksche Vermutung betrifft die Möbius Funktion, welche eine wichtige multiplikative Funktion in der Zahlentheorie und der Kombinatorik ist. Sie kodiert die multiplikative Struktur der natürlichen Zahlen und bildet das inverse Element zur Eins-Funktion bezüglich der dirichletschen Faltung. Er vermutet, dass die Möbiusfunktion nicht mit "einfachen" - konkret deterministischen - Funktionen korrelieren. Das heißt insbesondere dass die Möbius Funktion nicht durch solche Funktionen approximiert werden kann.

Die vorliegende Dissertation behandelt die Verteilung von automatischen Folgen entlang spezieller Teilfolgen und andere Eigenschaften von automatischen Folgen. Automatische Folgen sind Folgen $a(n)$ über einem endlichen Alphabet, die die Ausgabe eines endlichen Automaten sind. Diese Art von Folgen erhielten in den letzten 10 oder 15 Jahren große Aufmerksamkeit. Es gibt sehr enge Beziehungen zu dynamischen Systemen, zu Ziffernentwicklungen, zu gleichverteilten Folgen und auch zur Zahlentheorie.

Ein Hauptresultat dieser Arbeit besagt, dass alle automatischen Folgen die Sarnak-Vermutung erfüllen. Durch eine leicht abgewandelte Herangehensweise behandeln wir auch die Verteilung von automatischen Folgen entlang der Teilfolge der Primzahlen, was als Verallgemeinerung des zweiten Gelfondproblems gesehen werden kann.

Im Zuge der Behandlung von allgemeinen automatischen Folgen wird eine neue Struktur für deterministische endliche Automaten entwickelt, welche eine neue Sichtweise für Automaten bzw. automatische Folgen ermöglicht.

Im letzten Teil der Dissertation erweitern wir das Resultat von Drmota, Mauduit und Rivat auf digitale Folgen - die insbesondere automatische Folgen sind. Auch dies stellt eine Verallgemeinerung des dritten Gelfondproblems dar.

Résumé

La présente thèse a été fortement influencée par deux conjectures, l'une de Gelfond datant des années 1960 et l'autre de Sarnak en 2010. Ces deux conjectures ont soulevé un grand intérêt au cours des dix dernières années.

En 1968, Gelfond a prouvé que la somme des chiffres modulo m est asymtotiquement équirépartie dans des progressions arithmétiques, et il a formulé trois problèmes nouveaux. Le premier, qui portait sur la répartition jointe des sommes des chiffres relative à différentes bases, a été intégralement résolu par Kim en 1999. Le deuxième et le troisième problèmes traitent des sommes des chiffres pour les nombres premiers et les suites polynomiales. En ce qui concerne les nombres premiers et les carrés, Mauduit et Rivat ont résolu ces problèmes en 2010 et 2009, respectivement. Il convient en outre de mentionner qu'en 2013, Drmota, Mauduit et Rivat ont réussi généraliser le résultat concernant la suite des sommes des chiffres des carrés. Ils ont démontré que chaque bloc apparaît asymptotiquement avec la même fréquence.

La conjecture de Sarnak concerne la fonction de Möbius, une fonction multiplicative très importante tant en théorie des nombres qu'en combinatoire. Elle encode la structure multiplicative des entiers naturels et constitue l'inverse pour la convolution de Dirichlet de la fonction constante égale à 1. Selon cette conjecture, il n'y a pas de corrélation entre la fonction de Möbius et des fonctions « simples », plus précisément déterministes. Cela implique notamment que ces dernières fonctions ne peuvent servir d'approximation pour la fonction de Möbius.

La présente thèse traite de la répartition de suites automatiques le long de sous-suites particulières ainsi que d'autres propriétés de suites automatiques. On désigne comme suite automatique des suites a(n) sur un alphabet fini qui sont produites par un automate fini. Ce genre de suites a suscité un grand intérêt ces derniers 10 à 15 ans. Elles sont très étroitement liées aux systèmes dynamiques, aux développements digitaux, aux suites équiréparties, et également à la théorie des nombres.

Selon l'un des résultats principaux du présent travail, toutes les suites automatiques vérifient la conjecture de Sarnak. Moyennant une approche légèrement modifiée, nous traitons également la répartition de suites automatiques le long de la suite des nombres premiers, ce qui peut être considéré comme un cas général du deuxième problème de Gelfond.

Dans le cadre du traitement de suites automatiques générales, nous avons mis au point une nouvelle structure destinée aux automates finis déterministes ouvrant une vision nouvelle pour les automates et/ou les suites automatiques.

Dans la dernière partie de la thèse, nous étendons les résultat de Drmota, Mauduit et Rivat concernant les suites digitales qui sont un cas particulier des suites automatiques. Cette approche peut également être considérée comme une généralisation du troisième problème de Gelfond.

Abstract

The present dissertation was inspired by two conjectures, one by Gelfond stated in the 1960's and one of Sarnak in 2010. These two conjectures have received great attention throughout the last ten years.

In 1968 Gelfond proved that the sum of digits modulo m is asymptotically equally distributed along arithmetic progressions. Furthermore, he stated three problems which are nowadays called Gelfond problems. The first problem deals with the common distribution of the sum of digit function with respect to different bases and was completely solved by Kim in 1999. The second and third questions are concerned with the sum of digits of prime numbers and polynomial subsequences. Mauduit and Rivat were able to solve these problems for primes and squares in 2010 and 2009 respectively. It should also be noted that in 2013 Drmota, Mauduit and Rivat generalized the result concerning the sequence of the sum of digits of squares. They showed that each block appears asymptotically equally frequently.

Sarnak's conjecture is related to the Möbius function, which is an important multiplicative function in number theory and combinatorics. It encodes the multiplicative structure of the natural numbers and forms the inverse element to the one-function with respect to the Dirichlet-convolution. He conjectured that the Möbius function does not correlate with " simple " - more precisely deterministic - functions. This means that the Möbius function can not be approximated by such functions.

This dissertation deals with the distribution of automatic sequences along special subsequences and other properties of automatic sequences. Automatic sequences are sequences $a(n)$ over a finite alphabet, which are the output of a finite automaton. This type of sequences has received great attention in the last 10 or 15 years. There are very close relations to dynamical systems, to digital representations, to equidistributed sequences and also to number theory.

A main result of this thesis is that all automatic sequences satisfy the Sarnak conjecture. Through a slightly modified approach, we also deal with the distribution of automatic sequences along the subsequence of primes. This can be seen as a generalization of the second Gelfond problem.

In the course of the treatment of general automatic sequences, a new structure for deterministic finite automata is developed, which allows a new view for automata or automatic sequences.

In the last part of the dissertation, we extend the results of Drmota, Mauduit and Rivat to digital sequences - which are in particular automatic sequences. This is also a generalization of the third Gelfond problem.

Chapter 0

Introduction in German and English

0.1 Deutsche Einleitung

Das zentrale Thema dieser Arbeit sind automatische Folgen. Um diese definieren zu können, benötigen wir zuerst noch zwei weitere Konzepte. Das erste Konzept ist die *Zifferndarstellung* einer ganzen Zahl n, auch *Zahlensystem* genannt. In einem allgemeinen Kontext ist ein solches System eine injektive Abbildung von den natürlichen Zahlen in die Menge von Folgen von *Ziffern*.

Das bekannteste Zahlensystem ist das Dezimalsystem, das von einem großen Teil der Weltbevölkerung verwendet wird. Weniger bekannt sind zum Beispiel das Binärsystem und das Hexadezimalsystem, die jedoch durch die Entwicklung digitaler Rechenmaschinen stark an Bedeutung gewonnen haben. Eine natürliche Verallgemeinerung dieser Zahlensysteme ist die Darstellung von Zahlen in Basis k, wobei $k \geq 2$: Jede natürliche Zahl n kann eindeutig in der folgenden Form geschrieben werden

$$n = \sum_{i \geq 0} \varepsilon_i^{(k)}(n) k^i,$$

wobei $\varepsilon_i^{(k)} \in \{0, \ldots, k-1\}$ und $\varepsilon_i^{(k)} = 0$ für fast alle i. Die Darstellung von Zahlen in Basis k ist von essenzieller Bedeutung in dieser Dissertation und wir bezeichnen mit $(n)_k$ die Zifferndarstellung von n in Basis k ohne führende Nullen: $(n)_k = (\varepsilon_r^{(k)}(n), \ldots, \varepsilon_0^{(k)}(n))$ wobei $r = \lfloor \log_k(n) \rfloor$. Außerdem bezeichnen wir für $\mathbf{w} = (w_0, \ldots, w_r)$ die entsprechende natürliche Zahl $[\mathbf{w}]_k = \sum_{i=0}^{r} k^i w_{r-i}$.

Es ist weiters zu erwähnen, dass es etliche andere Zahlensysteme gibt, z.B. mit negativer Basis $-k$ oder irrationaler Basis $\beta \in \mathbb{R}^+ \setminus \mathbb{Q}$.

Als nächstes definieren wir deterministische endliche Automaten. Ein *deterministischer endlicher Automat* ist ein Quadrupel $A = (Q, \Sigma, \delta, q_0)$, wobei Q eine endliche Menge an Zuständen ist, Σ ist das endliche Eingabealphabet, $\delta : Q \times \Sigma \to Q$ ist die Übergangsfunktion und $q_0 \in Q$ ist der Startzustand - wir beschränken uns meistens auf den Fall $\Sigma = \{0, \ldots, k-1\}$. Die Übergangsfunktion δ ordnet jedem Paar bestehend aus einem Zustand $q \in Q$ und einem

Eingabesymbol $a \in \Sigma$ einen Nachfolgezustand $p \in Q$ zu. Wir erweitern die Funktion δ in natürlicher Weise zu einer Funktion - die wir ebenfalls δ nennen - $\delta : Q \times \Sigma^* \to Q$, wobei Σ^* die Wörter mit Buchstaben in Σ ist:

$$\delta(q, ab) = \delta(\delta(q, a), b)$$

für alle $a, b \in \Sigma^*$. Insbesondere besteht für ein Wort $\mathbf{w}$ der Länge n, $\delta(q, \mathbf{w})$ aus n Anwendungen der ursprünglichen Funktion δ. Wir lesen Wörter immer von links nach rechts, d.h. wenn wir als Eingabe die Zifferndarstellung einer Zahl verwenden, lesen wir zuerst die Ziffer mit dem höchsten Stellenwert.

Ein *deterministischer endlicher Automat mit Output* (DFAO) $A = (Q, \Sigma, \delta, q_0, \Delta, \tau)$ ist ein DFA mit einer zusätzlichen Funktion $\tau : Q \to \Delta$. Wir beschränken uns jedoch meistens auf komplexwertige Funktionen $\tau : Q \to \mathbb{C}$. In diesem Fall schreiben wir $A = (Q, \Sigma, \delta, q_0, \tau)$.

Nun sind wir in der Lage automatische Folgen zu definieren.

Definition 0.1.1. Eine Folge $\mathbf{a} = (a_n)_{n\in\mathbb{N}}$ heißt *k-automatische Folge* genau dann wenn es einen DFAO $A = (Q, \Sigma = \{0, \dots, k-1\}, \delta, q_0, \Delta, \tau)$ gibt, sodass $a_n = \tau(\delta(q_0, (n)_k))$.

Automatische Folgen fanden in den letzten 10 bis 15 Jahren große Beachtung und sind das zentrale Thema dieser Dissertation. Es gibt Verbindungen zu dynamischen Systemen, Zifferndarstellungen, gleichverteilte Folgen und auch Zahlentheorie. In dieser Arbeit werden wir alle diese Themen aufgreifen.

Es gibt zwei Entwicklungen in der Zahlentheorie, die diese Dissertation stark beeinflusst haben. Die erste ist eine Vermutung von Peter Sarnak aus dem Jahr 2010 betreffend die Möbius-Funktion. Die Möbius-Funktion ist wie folgt definiert,

$$\mu(n) = \begin{cases} (-1)^k & \text{wenn } n \text{ quadratfrei ist und} \\ & k \text{ die Anzahl der verschiedenen Primfaktoren von } n \text{ ist} \\ 0 & \text{sonst.} \end{cases}$$

Wir sagen dass eine Folge $\mathbf{a}$ orthogonal zur Möbius-Funktion ist wenn

$$\left| \sum_{n<N} a_n \mu(n) \right| = o\left(\sum_{n<N} |a_n| \right)^{1}.$$

Es gibt ein altes - relativ vages - Prinzip (genannt Möbius Randomness Law) , das besagt, dass jede "vernünftige", beschränkte komplexwertige Folge orthogonal zur Möbiusfunktion ist [28, Seite 388]. Die Idee hinter diesem Prinzip ist, dass die Möbius Funktion so zufällig das Vorzeichen ändert, dass es zu hinreichenden Auslöschungen kommt. Die Sarnak Vermutung besagt, dass eine spezielle Klasse von "einfachen" Folgen orthogonal zur Möbiusfunktion ist.

[1] Wir verwenden hier die klein-o Notation. Seien $f, g : \mathbb{N} \to \mathbb{R}^+$ Funktionen. Wir schreiben $f = o(g)$ genau dann, wenn $\lim_{n\to\infty} f(n)/g(n) = 0$ gilt.

Conjecture 0.1.2 (Sarnak Vermutung). *[40] Sei μ die Möbiusfunktion. Jede komplexwertige beschränkte Folge $\xi(n)$ die von einem deterministischem dynamischen System* [2] *(X, S) realisiert wird - d.h. es gibt eine stetige Funktion $f : X \to \mathbb{C}$ und einen Startwert $x_0 \in X$, sodass $\xi(n) = f(T^n(x_0))$ gilt -, erfüllt*

$$\sum_{n \leq N} \xi(n)\mu(n) = o(N).$$

Zahlreiche Publikationen behandeln die Orthogonalität bestimmter Folgen oder dynamischer Systeme zur Möbiusfunktion.

Der zweite große Einfluss auf diese Dissertation sind die sogenannten Gelfond-Probleme. Diese betreffen sehr spezielle automatische Folgen, nämlich die Ziffernsummenfunktion von n in Basis k modulo m. Das berühmteste Beispiel dafür ist die Thue-Morse Folge die dem Fall $k = m = 2$ entspricht. Wir definieren

$$\mathbf{t}_{k,m} := (s_k(n) \bmod m)_{n \in \mathbb{N}},$$

wobei $s_k(n)$ die Ziffernsumme von n in Basis k ist:

$$s_k(n) := \sum_{i \geq 0} \varepsilon_i^{(k)}(n).$$

Das erste Resultat für die Verteilung von $\mathbf{t}_{k,m}$ stammt von Gelfond. Er zeigt in [22] dass für $ggT(k-1, m) = 1$ und jedes $\ell \in \{0, \ldots, m-1\}$, $\eta > 0$ existiert, sodass

$$|\{n < N : \mathrm{s}_k(an + b) \equiv \ell \bmod m\}| = \frac{N}{m} + \mathcal{O}(N^{1-\eta})$$

gilt. Das bedeutet, dass die linearen Teilfolgen von $\mathbf{t}_{k,m}$ gleichverteilt sind auf den Werten $\{0, \ldots, m-1\}$.

In seiner Arbeit formuliert Gelfond drei weitere Probleme, die später als Gelfond-Probleme bezeichnet wurden. Alle diese Probleme betreffen die Verteilung von $\mathbf{t}_{k,m}$ entlang spezieller Teilfolgen.

1. Seien $k_1, k_2 \geq 2$ teilerfremde Zahlen und $\mathrm{ggT}(k_1 - 1, m_1) = \mathrm{ggT}(k_2 - 1, m_2) = 1$, dann gilt

   $$|\{n < N : \mathrm{s}_{k_1}(n) \equiv \ell_1 \bmod m_1, \mathrm{s}_{k_2}(n) \equiv \ell_2 \bmod m_2\}| = \frac{N}{m_1 m_2} + \mathcal{O}(N^{1-\eta})$$

   für alle ℓ_1, ℓ_2 und ein $\eta > 0$.

2. Sei $k \geq 2$ und $\mathrm{ggT}(k-1, m) = 1$, dann gilt

   $$|\{p < N : p \in \mathbb{P} \wedge \mathrm{s}_k(p) \equiv \ell \bmod m\}| = \frac{\pi(N)}{m} + \mathcal{O}(N^{1-\eta})$$

   für alle ℓ und ein $\eta > 0$. $\pi(x)$ bezeichnet hier die Anzahl der Primzahlen $< x$.

[2] Ein dynamisches System ist ein Paar (X, S) wobei X ein kompakter metrischer Raum ist und $S : X \to X$ eine stetige Abbildung ist. Ein dynamisches System ist deterministisch wenn es topologische Entropie 0 hat.

3. Sei $k \geq 2$ und $\mathrm{ggT}(k-1, m) = 1$ dann gilt für jedes Polynom $P(x)$ mit ganzzahligen Koeffizienten

$$|\{n < N : \mathrm{s}_k(P(n)) \equiv \ell \bmod m\}| = \frac{N}{m} + \mathcal{O}(N^{1-\eta})$$

für alle ℓ und ein $\eta > 0$.

Bereits im Jahr 1972 löste Besineau das erste Problem [4]. Kim verallgemeinerte das Resultat auf k- additive Funktionen, d.h. Funktionen f die $f(ak^\ell + b) = f(a) + f(b)$ erfüllen, wenn $a \geq 1, \ell \geq 0, 0 \leq b < k^\ell$ gilt, und konnte auch einen expliziten Fehlerterm angeben[31].

Es dauerte jedoch fast vierzig Jahre bis das zweite und dritte Problem gelöst bzw. teilweise gelöst wurden. Das zweite Problem wurde von Mauduit und Rivat im Jahr 2010 gelöst [37]. Das dritte Problem wurde im Jahr 2009 von Mauduit und Rivat für quadratische Polynome gelöst [36]. Außerdem gibt es eine Lösung für Basis k wobei k eine Primzahl ist, die groß ist in Anbetracht des Polynomgrades. Diese Resultate beruhen auf der Behandlung von Exponentialsummen mit Fourier-theoretischen Methoden, die von Mauduit und Rivat entwickelt wurden. Dies stellt einen Durchbruch in diesem Gebiet dar und wird in dieser Dissertation verwendet.

Eine natürliche Verallgemeinerung der Ziffernsummenfunktion sind digitale Funktionen. Eine Folge $\mathbf{a}$ heißt digital[3] in Basis k, wenn es $r \geq 1$ und $F : \mathbb{Z}^r \to \mathbb{C}$ gibt mit $F(0, \ldots, 0) = 0$ und

$$a_n = \sum_{i \geq 0} F(\varepsilon_i^{(k)}(n), \ldots, \varepsilon_{i+r-1}^{(k)}(n)).$$

Diese Summe ist wohldefiniert, da $F(0, \ldots, 0) = 0$ und fast alle $\varepsilon_i = 0$. Man kann auch eine Variation dieser Definition betrachten. Dafür definiert man $\varepsilon_{-i}^{(k)}(n) := 0$ für $i \geq 1$ und erweitert die Summe auf $i \in \mathbb{Z}$.

Im ersten Kapitel dieser Dissertation geben wir eine Einleitung zu automatischen Folgen. Wir präsentieren einige Beispiele und klassische Resultate für automatische Folgen. Eine wesentlich detailliertere Behandlung von automatischen Folgen findet sich in [1]. Weiters beschreiben wir zwei wichtige Klassen von Automaten bzw. automatischen Folgen, nämlich invertierbare und synchronisierende Automaten bzw. Folgen. Außerdem besprechen wir aktuelle Resultate für Teilfolgen spezieller automatischer Folgen, die teilweise auf Fourier-theoretische Methoden zurückgehen und in den späteren Kapiteln verwendet werden.

Im zweiten Kapitel präsentieren wir eine neue Struktur für Automaten, genannt "naturally induced transducers". Diese Konstruktion kombiniert Ideen von invertierbaren und synchronisierenden Automaten. Wir werden sehen, dass diese Aspekte beinahe unabhängig voneinander behandelt werden können. Der synchronisierende Teil kann verhältnismäßig leicht behandelt werden, es reichen dafür bereits elementare Methoden. Der invertierbare Teil ist deutlich schwieriger zu behandeln und wir werden Resultate beweisen, die später für Verteilungs-Resultate von speziellen Teilfolgen verwendet werden.

[3] Digitale Folgen werden manchmal auch als stark block-additive Funktionen bezeichnet.

Das dritte Kapitel behandelt die Sarnak Vermutung. Wir nennen einige Beispiele, die die Sarnak Vermutung erfüllen, und formulieren eine Version der Sarnak Vermutung für Folgen. Es gibt eine kanonische Art einer beschränkten komplexwertigen Folge ein dynamisches System zuzuordnen. Wir sagen daher, dass eine Folge die Sarnak Vermutung erfüllt, wenn jede Folge des zugehörigen dynamischen Systems die Sarnak Vermutung erfüllt. Der Rest des Kapitels dient dazu das folgende Resultat zu beweisen.

Theorem 0.1.3. *Sei μ die Möbiusfunktion, $(a_n)_{n\in\mathbb{N}}$ eine automatische Folge und sei (X, S) das dynamische System assoziiert mit $(a_n)_{n\in\mathbb{N}}$. Dann gilt für alle Folgen $\xi(n) := f(S^n(x))$, mit $x \in X$ und $f \in C(X, \mathbb{C})$,*

$$\sum_{n\leq N} \xi(n)\mu(n) = o(N).$$

Für den Beweis verwenden wir die Struktur, die wir in Kapitel 2 beschrieben haben. Hier müssen wir den synchronisierenden und den invertierbaren Teil behandeln. Der synchronisierende Teil kann relativ einfach behandelt werden. Die wesentliche Schwierigkeit besteht in der Behandlung des invertierbaren Teils. Dafür verwenden und adaptieren wir die Ideen von Mauduit und Rivat, die sie verwendet haben um zu zeigen, dass die Rudin-Shapiro Folge orthogonal zur Möbiusfunktion ist.

Kapitel 4 und 5 behandeln Erweiterungen der Gelfond Probleme auf automatische Folgen. In Kapitel 4 beschreiben wir zuerst die Verteilung von automatischen Folgen entlang von arithmetischen Progressionen. Hierbei verwenden wir insbesondere „naturally induced transducer“ und dieses Resultat dient hauptsächlich der Präsentation der verwendeten Methode.
Danach behandeln wir die Verteilung von automatischen Folgen entlang der Primzahlen - jeweils unter einer technischen Bedingung. Dies entspricht dem 2. Gelfond Problem und ist deutlich aufwändiger. Die verwendete technische Bedingung mag anfangs vielleicht willkürlich oder zu stark erscheinen. Es ist jedoch unklar für welche automatischen Folgen man überhaupt die Existenz einer Grenzverteilung für die Teilfolge entlang der Primzahlen bzw arithmetischer Progressionen erwarten kann. Wir beweisen das folgende Resultat.

Theorem 0.1.4. *Sei $A = (Q', \Sigma, \delta', q_0', \tau)$ ein stark zusammenhängender deterministischer endlicher Automat mit Output (DFAO), wobei $\Sigma = \{0, \dots, k-1\}$ und $\delta'(q_0', 0) = q_0'$. Dann existiert eine Grenzverteilung für die Teilfolge $(a_p)_{p\in\mathbb{P}}$ entlang der Primzahlen.*

Der Beweis dieses Satzes erlaubt die Bestimmung der Grenzverteilung, wobei diese von der Verteilung von linearen Teilfolgen der automatischen Folge abhängt. Die Beweisidee ist sehr ähnlich zu den Methoden, die in Kapitel 3 verwendet werden, es ist allerdings etwas mehr Aufwand notwendig um die Grenzverteilung zu bestimmen.

In Kapitel 5 verallgemeinern wir ein Resultat von Drmota, Mauduit und Rivat [14], das besagt, dass $(\mathbf{t}_{2,2}(n^2))_{n\in\mathbb{N}}$ normal ist, auf digitale Folgen, die insbesondere automatische Folgen sind:

Theorem 0.1.5. *Sei b eine digitale Folge in Basis q und $m' \in \mathbb{N}$, wobei $\mathrm{ggT}(q-1, m') = 1$ und $\mathrm{ggT}(m', \mathrm{ggT}(b(n))_{n\in\mathbb{N}}) = 1$. Dann ist $(b(n^2) \mod m')_{n\in\mathbb{N}}$ normal, d.h. jede Teilfolge der Länge k kommt mit asymptotischer Häufigkeit $(m')^{-k}$ vor.*

Hierfür folgen wir großteils der Argumentation von Drmota, Mauduit und Rivat [14]. Sie reduzieren das Problem, unter Zuhilfenahme einer sogenannten "carry-property" für die Ziffernsummenfunktion, zu spezielle Aussagen über Fouriertransformierte. Dieser Teil kann leicht an die veränderte Situation angepasst werden, jedoch stellen die entsprechenden Aussagen über die Fouriertransformierten für digitale Funktionen eine große Herausforderung dar und können nicht leicht von den entsprechenden Resultaten für die Ziffernsummenfunktion abgeleitet werden.

0.2 Introduction in English

A central concept in this thesis are *automatic sequences*. Therefore, we need to define the *digital representation* of an integer n, also known as a *numeration system*. In a general setting, such a system is just an injective map from the non-negative integers to a set of sequences of *digits*.

The most well-known numeration system is the decimal system, used throughout everyday life by a large part of the world population. Slightly less well-known are the binary system and also the hexadecimal system, which have massively gained importance due to the development of digital computing machines. A natural generalization of these numeration systems is to take an arbitrary bases k, where $k \geq 2$ is an integer: every non-negative integer n can be written in a unique way as

$$n = \sum_{i \geq 0} \varepsilon_i^{(k)}(n) k^i,$$

where $\varepsilon_i^{(k)} \in \{0, \dots, k-1\}$ and $\varepsilon_i^{(k)} = 0$ for all but finitely many i. This concept of base-k representation will be important throughout this thesis and we denote by $(n)_k$ the digital expansion of n in base k without leading zeros, i.e. $(n)_k = (\varepsilon_r^{(k)}(n), \dots, \varepsilon_0^{(k)}(n))$ where $r = \lfloor \log_k(n) \rfloor$. This finite sequence can be interpreted as a word over the alphabet $\{0, 1, \dots, k-1\}$. We denote for $\mathbf{w} = (w_0, \dots, w_r)$ the corresponding natural number $[\mathbf{w}]_k := \sum_{i=0}^{r} k^i w_{r-i}$.

There exist many generalizations, for example digital representations in base $(-k)$ or β where $\beta \in \mathbb{R}^+ \setminus \mathbb{Q}$.

A deterministic finite automaton, or DFA, is defined to be a quadruple $A = (Q, \Sigma, \delta, q_0)$, where Q is a finite set of states, Σ is the finite input alphabet - we restrict ourselves mostly to the case $\Sigma = \{0, \dots, k-1\}$ - , $\delta : Q \times \Sigma \to Q$ is the transition function and $q_0 \in Q$ is the initial state. The transition function δ assigns to each state $q \in Q$ and letter of the input alphabet $a \in \Sigma$ a successive state $p \in Q$. We extend the function δ in a natural way to a function which we also call $\delta : Q \times \Sigma^* \to Q$, where Σ^* denotes the sets of words with letters in Σ:

$$\delta(q, ab) = \delta(\delta(q, a), b)$$

for all $a, b \in \Sigma^*$. Thus, for a word $\mathbf{w}$ of length n we find that $\delta(q, \mathbf{w})$ consists of n applications of the original function δ. We always read words from left to right i.e. for digital representations of numbers we start with the most significant digit.

A DFAO $A = (Q, \Sigma, \delta, q_0, \Delta, \tau)$ is a DFA with an additional output function $\tau : Q \to \Delta$. In this thesis we restrict ourselves mostly to complex-valued output functions; i.e. $\tau : Q \to \mathbb{C}$. We omit $\Delta = \mathbb{C}$ in this case.

We have now described the necessary objects to define an *automatic sequence*.

Definition 0.2.1. We say that a sequence $\mathbf{a} = (a_n)_{n\in\mathbb{N}}$ is a k-automatic sequence, if and only if there exists a DFAO $A = (Q, \Sigma = \{0, \dots, k-1\}, \delta, q_0, \Delta, \tau)$ such that $a_n = \tau(\delta(q_0, (n)_k))$.

Automatic sequences have drawn a lot of attention during the last 10 or 15 years and are the main subject of this thesis. There are close relations to dynamical systems, to digital expansions, to uniformly distributed sequences and also to number theory. We will find links to all of these topics throughout this thesis.

There are two two lines of research that have influenced this thesis. The first one is a recent conjecture by Peter Sarnak in 2010, concerning the Möbius function μ. The Möbius function is defined by

$$\mu(n) = \begin{cases} (-1)^k & \text{if } n \text{ is square-free and} \\ & k \text{ is the number of prime factors} \\ 0 & \text{otherwise.} \end{cases}$$

We say that a sequence $\mathbf{a}$ is orthogonal to $\mu(n)$ if

$$\left| \sum_{n<N} a_n \mu(n) \right| = o\left(\sum_{n<N} |a_n| \right).^4$$

There exists an old - relatively vague - principle (the Möbius Randomness Principle, see for example [28, p. 338]), which states that every "reasonable" bounded complex sequence is orthogonal to the Möbius function. The reasoning behind this principle is that the Möbius function changes signs so randomly that it induces sufficient cancellation. Peter Sarnak conjectures that a special class of "simple" sequences is orthogonal to the Möbius function.

Conjecture 0.2.2 (Sarnak Conjecture)**.** *[40] Let μ be the Möbius function. For any bounded complex sequence $\xi(n)$ observed by a deterministic flow* [5] *(X, S), it holds that*

$$\sum_{n\leq N} \xi(n)\mu(n) = o(N).$$

Numerous papers are dedicated to show orthogonality to the Möbius function for certain types of sequences or deterministic flows.

[4] Here we use the little-o notation. Let $f, g : \mathbb{N} \to \mathbb{R}^+$. We say that $f = o(g)$ if and only if $\lim_{n\to\infty} f(n)/g(n) = 0$.

[5] A dynamical system F is a pair (X, S) where X is a compact metric space and $S : X \to X$ is a continuous map. For a dynamical system to be deterministic means that - roughly speaking - there are only few different orbits.

The second major influence are the so called Gelfond-problems, which concern very special automatic sequences, namely the sum of digits in base k modulo m. For $k = m = 2$, one obtains the *Thue-Morse* sequence. We define

$$\mathbf{t}_{k,m} = (\mathrm{s}_k(n) \bmod m)_{n\in\mathbb{N}},$$

where $s_k(n)$ denotes the sum of digits of n in base k:

$$s_k(n) := \sum_{i\geq 0} \varepsilon_i^{(k)}(n).$$

The first distributional property of $\mathbf{t}_{k,m}$ was found by Gelfond [22] who showed that if $\gcd(k-1, m) = 1$, then for every $\ell \in [0, \ldots, m-1]$,

$$|\{n < N : \mathrm{s}_k(an + b) \equiv \ell \bmod m\}| = \frac{N}{m} + \mathcal{O}(N^{1-\eta})$$

holds for some $\eta > 0$. This means exactly that linear sub-sequences of $\mathbf{t}_{k,m}$ are uniformly distributed on the values $\{0, 1, \ldots, m-1\}$.

Additionally, Gelfond formulated three problems which are known as the *Gelfond Problems*. All of these problems cover distributional properties of $\mathbf{t}_{k,m}$.

1. If $k_1, k_2 \geq 2$ are co-prime integers and $\gcd(k_1 - 1, m_1) = \gcd(k_2 - 1, m_2) = 1$ then

$$|\{n < N : \mathrm{s}_{k_1}(n) \equiv \ell_1 \bmod m_1, \mathrm{s}_{k_2}(n) \equiv \ell_2 \bmod m_2\}| = \frac{N}{m_1 m_2} + \mathcal{O}(N^{1-\eta})$$

   holds for all ℓ_1, ℓ_2 and some $\eta > 0$.

2. If $k \geq 2$ and $\gcd(k - 1, m) = 1$ then

$$|\{p < N : p \in \mathbb{P} \wedge \mathrm{s}_k(p) \equiv \ell \bmod m\}| = \frac{\pi(N)}{m} + \mathcal{O}(N^{1-\eta})$$

   for all ℓ and some $\eta > 0$. Here $\pi(x)$ denotes the number of primes $< x$.

3. If $k \geq 2$ and $\gcd(k - 1, m) = 1$ then for each integer polynomial $P(x)$

$$|\{n < N : \mathrm{s}_k(P(n)) \equiv \ell \bmod m\}| = \frac{N}{m} + \mathcal{O}(N^{1-\eta})$$

   for all ℓ and some $\eta > 0$.

In 1972, Besineau was able to solve the first problem [4]. Kim was able to generalize this result to k-additive functions - i.e. functions which fulfill $f(ak^\ell + b) = f(a) + f(b)$ for $a \geq 1$, $\ell \geq 1, 0 \leq b < k^\ell$ - and was also able to formulate an explicit error term [31].

However, it took almost 40 years until the second and third problem were solved or came close to a solution. The second problem was finally solved by Mauduit and Rivat in 2010 [37].

In 2009, the third problem was solved for quadratic polynomials by Mauduit and Rivat [36]. Additionally, there is a solution by Drmota, Mauduit and Rivat [14] for prime numbers k which are sufficiently large with respect to the degree of $P(x)$. The treatment of exponential sums with Fourier-theoretic methods that has been developed by Mauduit and Rivat was a breakthrough in this field and will also be utilized throughout this thesis.

A natural generalization of the sum of digits function are digital functions. We say that a sequence $\mathbf{a}$ is a *digital sequence*[6] in base k if there exist an integer $r \geq 1$ and a function $F : \mathbb{Z}^r \to \mathbb{C}$ such that $F(0, \dots, 0) = 0$ and

$$a_n = \sum_{i \geq 0} F(\varepsilon_i(n), \dots, \varepsilon_{i+r-1}(n)).$$

This sum is well defined since $F(0, \dots, 0) = 0$. One could also consider an alternative definition, where we expand the sum to $i \in \mathbb{Z}$ and define $\varepsilon_{-i}(n) := 0$ for all $i \geq 1$.

In the first chapter, we give an introduction to automatic sequences. We present some examples of automatic sequences and display some classical results for automatic sequences. A more detailed study of automatic sequences is given in [1]. Furthermore, we will discuss two important subclasses of automatic sequences - namely invertible and synchronizing automatic sequences. Finally, we will mention some recent results about subsequences of certain automatic sequences. Most of these results rely on some kind of Fourier-analytic treatment and these methods will also be used in this thesis.

In Chapter 2, we present a new structure for strongly connected DFA, namely naturally induced transducers. This structure combines the ideas of invertible and synchronizing automata. We will show that these aspects behave almost independent of each other. The synchronizing part can be analyzed by elementary methods. The invertible part is much more challenging and we obtain results that allow us to find distributional results later.

Chapter 3 is completely dedicated to the Sarnak Conjecture. We give examples that satisfy the Sarnak Conjecture and state the Sarnak Conjecture for sequences. Therefore, we assign to a sequence a dynamical system in a canonical way. We say that the sequence satisfies the Sarnak Conjecture if every sequence obtained by this dynamical system satisfies the Sarnak Conjecture. Subsequently we will prove the following theorem.

Theorem 0.2.3. *Let μ be the Möbius function, $(a_n)_{n\in\mathbb{N}}$ an automatic sequence and let (X, S) be the symbolic dynamical system associated with $(a_n)_{n\in\mathbb{N}}$. Then for all sequences $\xi(n) := f(S^n(x))$, with $x \in X$ and $f \in C(X, \mathbb{C})$, we have*

$$\sum_{n \leq N} \xi(n)\mu(n) = o(N).$$

To obtain this result, we use the structural result found in Chapter 2. Therefore, we need to analyze the synchronizing and the invertible part of naturally induced transducers. The synchronizing aspect can be treated quite easily. The invertible part takes some effort to deal with. Therefore, we use and adopt the strong framework developed by Mauduit and

[6] Sometimes digital sequences are also referred to as strongly block-additive function

Rivat which was used to show that the Rudin-Shapiro sequence is orthogonal to the Möbius function.

Chapter 4 and 5 are inspired by the Gelfond problems and their natural extension to other sequences - in this case some special automatic sequences. In Chapter 4, we study at first the distribution of automatic sequences along arithmetic progressions. Therefore, we use the concept of naturally induced transducers. This result was included in this chapter to illustrate the used method to find the limiting distribution of subsequences of automatic sequences. Thereafter, we describe the distribution of automatic sequences along prime numbers under some technical conditions that we also need for the first result of this chapter. This corresponds to the second Gelfond problem and is much more difficult to achieve. The technical conditions may seem too restrictive, but actually, it is not even clear for which automatic sequences there exists a limiting distribution if one considers the subsequence along the primes. Nevertheless, we are able to show the following result.

Theorem 0.2.4. *Let $A = (Q', \Sigma, \delta', q_0', \tau)$ be a strongly connected deterministic finite automaton with output (DFAO) with $\Sigma = \{0, \dots, k-1\}$ and $\delta'(q_0', 0) = q_0'$. Then the frequencies of the letters for the prime subsequence $(a_p)_{p\in\mathcal{P}}$ exist.*

The proof of this theorem allows to determine these frequencies which correlate in some sense to the distribution of the automatic sequence along arithmetic subsequences. The methods to show this result are quite similar to the methods used in Chapter 3. However, one needs a bit more effort to determine the limiting distribution.

In Chapter 5, we generalize the result of Drmota, Mauduit and Rivat, who showed that $(\mathbf{t}_{2,2}(n^2))_{n\in\mathbb{N}}$ is normal, to digital sequences, which are a special class of automatic sequences:

Theorem 0.2.5. *Let b be a strongly block-additive function in base q and $m' \in \mathbb{N}$ with $\gcd(q-1, m') = 1$ and $\gcd(m', \gcd(b(n))_{n\in\mathbb{N}}) = 1$. Then $(b(n^2) \mod m')_{n\in\mathbb{N}}$ is normal i.e. every subsequence of length k appears with asymptotic frequency $(m')^{-k}$.*

The proof mainly follows the ideas of Drmota, Mauduit and Rivat in [14]. Exponential sums are manipulated and a so called "carry property" is used to reduce the problem to statements about Fourier-transforms. This part can be adopted easily. However, the statements about the occurring Fourier-transforms pose the main difficulties and can not be adopted easily.

Chapter 1

Automatic Sequences

We use this chapter to highlight classical properties of automatic sequences and also to discuss some new results concerning automatic sequences. Throughout this chapter, we give only few complete proofs and refer mostly to the literature.

In Section 1.1 we define automatic sequences give some basic properties and examples. In Section 1.2 we discuss some other equivalent definitions of automatic sequences that are sometimes useful. Section 1.3 contains some results for frequencies of (automatic) sequences. In the last section we present new results concerning the distribution of subsequences of automatic sequences. These are particularly interesting as they build the base for large parts of this thesis.

The results of Sections 1.1 to 1.3 are taken from [1], where a much more detailed study of automatic sequences is given.

1.1 Introduction to Automatic Sequences

We start this section by a short discussion of base-k representations of natural numbers.
It is a well-known fact that - for every $k \geq 2$ - every non-negative integer n can be written in a unique way as

$$n = \sum_{i \geq 0} \varepsilon_i^{(k)}(n) k^i,$$

where $\varepsilon_i^{(k)} \in \{0, \ldots, k-1\}$ and $\varepsilon_i^{(k)} = 0$ for all but finitely many i. There exists an explicit formula for $\varepsilon_i^{(k)}$:

$$\varepsilon_i^{(k)}(n) = \left\lfloor \frac{n}{k^i} \right\rfloor - k \left\lfloor \frac{n}{k^{i+1}} \right\rfloor.$$

We denote by $(n)_k$ the digital expansion of n in base k without leading zeros, i.e. $(n)_k = (\varepsilon_r^{(k)}(n), \ldots, \varepsilon_0^{(k)}(n))$ where $r = \lfloor \log_k(n) \rfloor$. This finite sequence can be interpreted as a word over the alphabet $\{0, 1, \ldots, k-1\}$. Conversely, we denote for $\mathbf{w} = (w_0, \ldots, w_r)$ the corresponding natural number $[\mathbf{w}]_k := \sum_{i=0}^{r} k^i w_{r-i}$.

Let Σ be a finite set. A word of length n is a map from $\{0,\dots,n-1\}$ to Σ. Similarly we define an infinite word to be a map from $\mathbb{N}\to\Sigma$. If $n=0$, we get the empty word, denoted by ϵ. The set of all finite words made up of letters chosen from Σ is denoted by Σ^*. If $\mathbf{w}$ is a finite word, we denote its length by $|\mathbf{w}|$.

The fundamental operation on words is the *concatenation*. We concatenate two finite words $\mathbf{v},\mathbf{w}$ by juxtaposing their symbols and we denote this by $\mathbf{vw}$.

Example. Let $\Sigma=\{a,b\}, \mathbf{v}=aab$ and $\mathbf{w}=bba$. Then $\mathbf{vw}=aabbba$.

To describe *automatic sequences*, we need to define deterministic finite automata first .

A *deterministic finite automaton*, or *DFA*, is defined to be a quadruple $A=(Q,\Sigma,\delta,q_0)$, where Q is a finite set of states, Σ is the finite input alphabet - we restrict ourselves to the case $\Sigma=\{0,\dots,k-1\}$ - , $\delta: Q\times\Sigma\to Q$ is the transition function and $q_0\in Q$ is the initial state.

We extend δ to a function $\delta: Q\times\Sigma^*\to Q$, in a natural way, by

$$\delta(q,ab)=\delta(\delta(q,a),b)$$

for all $a,b\in\Sigma^*$. Note that we always read words from left to right i.e. for digital representations of numbers we start with the most significant digit. By definition $\delta(q,\mathbf{w})$ consists of $|\mathbf{w}|$ "steps" for every $\mathbf{w}\in\Sigma^*$.

A DFAO $A=(Q,\Sigma,\delta,q_0,\Delta,\tau)$ is a DFA with an additional output function $\tau: Q\to\Delta$. In this thesis we restrict ourselves mainly to complex-valued output functions; i.e. $\tau: Q\to\mathbb{C}$. We omit to denote $\Delta=\mathbb{C}$ in this case.

We have now described the necessary setting to define an *automatic sequence*.

Definition 1.1.1. We say that a sequence $\mathbf{a}=(a_n)_{n\in\mathbb{N}}$ is a *k-automatic sequence*, if and only if there exists a DFAO $A=(Q,\Sigma=\{0,\dots,k-1\},\delta,q_0,\Delta,\tau)$ such that $a_n=\tau(\delta(q_0,(n)_k))$.

1.1.1 Examples of Automatic Sequences

In this part, we give some examples of automatic sequences. We start by a very easy and fundamental example of automatic sequences.

Theorem 1.1.2. *Every periodic sequence is k-automatic for all $k\geq 2$.*

Proof. Let $\mathbf{a}=(a_n)_{n\geq 0}$ be a periodic sequence with period t. We define the k-automaton $A=(Q,\Sigma,\delta,q_0,\Delta,\tau)$, where $\Sigma=\{0,\dots,k-1\}$ as follows:

$$\begin{aligned} Q &= \{0,\dots,t-1\},\\ q_0 &= 0,\\ \delta(q,b) &= (kq+b) \bmod t \qquad \forall q\in Q, b\in\Sigma,\\ \tau(q) &= a_q \qquad \forall q\in Q. \end{aligned}$$

One finds easily by induction on the length of $\mathbf{w}$ that

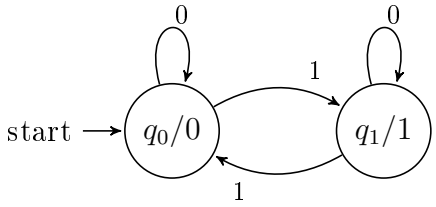
$$\delta(0, \mathbf{w}) = [\mathbf{w}]_k \bmod t$$

and the result follows directly. □

Another prominent examples of automatic sequences is the Thue–Morse sequence $\mathbf{t} = (t_n)_{n\geq 0}$, given by $t_n = s_2(n) \bmod 2$, where s_2 is the base-2 sum-of-digits function[1]. We present the automaton that produces the Thue–Morse sequence using a *transition diagram*.

A transition diagram is a directed graph. The vertices of the transition diagram correspond to the states of the automaton and the labeled edges corresponds to the transition function. More precisely, a labeled edge from q_1 to q_2 with label a corresponds to $\delta(q_1, a) = q_2$. The initial state is indicated by an unlabeled arrow entering the state. The value of the output function τ of a state q is written in the same vertex as the state, separated by /.

Example (Thue–Morse automaton). The transition diagram of the Thue–Morse automaton is given below.



The Thue–Morse sequence is an example of a digital sequence, which are the main subject of the last chapter of this thesis. We use here the definition given in [1, p. 83].

Definition 1.1.3. We call a sequence $(a_n)_{n\geq 0}$ *digital* in base k and of length m if there exists a function $F : \{0, \ldots, k-1\}^m \to \mathbb{C}$, where $F(0, \ldots, 0) = 0$ and

$$a_n = \sum_{i\geq 0} F(\varepsilon_{i+m-1}^{(k)}(n), \ldots, \varepsilon_i^{(k)}(n)).$$

This sum is well defined since $F(0, \ldots, 0) = 0$ and can be rewritten as

$$a_n = \sum_{i=0}^{r-1} F(\varepsilon_{i+m-1}^{(k)}(n), \ldots, \varepsilon_i^{(k)}(n)), \tag{1.1}$$

where $r = \lfloor \log_k(n) \rfloor$.

Theorem 1.1.4. *Let $(a_n)_{n\geq 0}$ be a digital sequence in base k. Then $(a_n \bmod m')_{n\geq 0}$ is a k-automatic sequence for every $m' \in \mathbb{N}$.*

[1] Let $(n)_k = (\varepsilon_r^{(k)}(n), \ldots, \varepsilon_0^{(k)}(n))$ be the representation of n in base k. Then, we define the sum-of-digits function in base k as follows,

$$s_k(n) := \sum_{i=0}^{r} \varepsilon_i^{(k)}(n).$$

Proof. We define the DFAO $A = (Q, \Sigma, \delta, q_0, \tau)$ as follows.
As always we set $\Sigma = \{0, \dots, k-1\}$. We see that a_n is defined by a summation of $r+1 = \lfloor \log_k(n) \rfloor + 1$ terms and $\delta(q_0, (n)_k)$ consists of $r+1$ steps. The idea is that the i-th step in computing $\delta(q_0, (n)_k)$ corresponds to the addition of the $(r-i)$-th summand in (1.1). Thus we need to "remember" the $m-1$ last digits that we have read and the value of the partial summation up to this point.

Let $Q = \{0, \dots, k-1\}^{m-1} \times \mathbb{Z}_{m'}$ and $q_0 = (0^{m-1}, 0)$. The first component corresponds to the last $m-1$ digits when reading $(n)_k$ from the most significant digit to the least significant digit. The second component gives the value of the summation up to this step. Therefore, we define for $\varepsilon \in \Sigma$

$$\delta(((a_0, \dots, a_{m-2}), b), \varepsilon) := ((a_1, \dots, a_{m-2}, \varepsilon), (b + F(a_0, \dots, a_{m-2}, \varepsilon)) \bmod m').$$

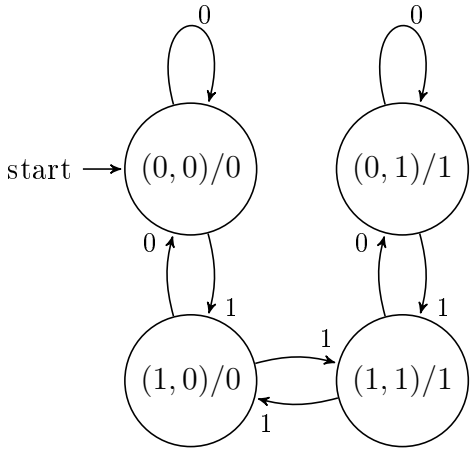
Furthermore, we define $\tau((a, b)) = b$ for all $a \in \{0, \dots, k-1\}^{m-1}$ and $b \in \{0, \dots, m'-1\}$. One finds easily by induction on the length of $\mathbf{w}$ that $\delta(q_0, \mathbf{w}) = (\mathbf{w}', a_{[\mathbf{w}]_k})$ where $\mathbf{w}'$ is the word consisting of the last $m-1$ letters of $0^{m-1}\mathbf{w}$. Thus we have $\tau(\delta(q_0, \mathbf{w})) = a_{[\mathbf{w}]_k}$. □

Another prominent example of a digital sequence is the Rudin–Shapiro sequence $(r_n)_{n \geq 0}$.

Example (Rudin–Shapiro sequence). Let $(n)_2 = (\varepsilon_r^{(2)}(n), \dots, \varepsilon_0^{(2)}(n))$ be the representation of n in base 2. The n-th element of the Rudin–Shapiro sequence is then given by

$$r_n = \left(\sum_{i=0}^{r-1} \varepsilon_i^{(2)}(n) \varepsilon_{i+1}^{(2)}(n) \right) \bmod 2.$$

By following the proof of Theorem 1.1.4 we construct the automaton corresponding to the Rudin–Shapiro sequence. In the first component we remember the last digit and in the second one the partial summation result.



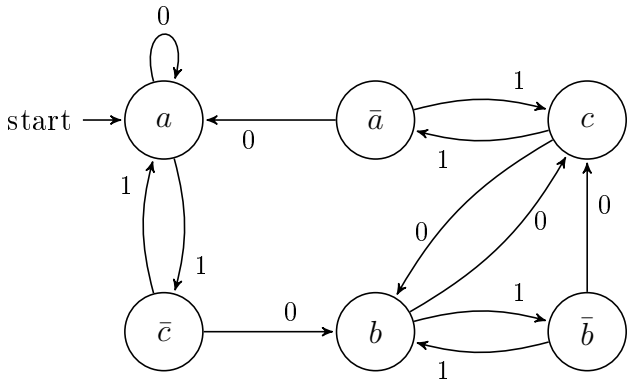
Lastly, we want to mention an example of an automatic sequence, that has a connection to game theory.

Example (The Tower of Hanoi Sequence). We consider the classical Tower of Hanoi problem. There are three pegs and N disks with radii $1, 2, \dots, N$. Initially they are all placed on peg

1 one with decreasing radii from bottom to top. A *legal move* consists of taking the top disk of a peg and moving it to another peg that is empty or the top disk has a larger radius. The goal of the problem is to move all N disks from peg 1 to another peg by only using legal moves.
It is well known that the optimal solution consists of $2^N - 1$ moves. We denote by a, b, c the moves that take the topmost disk of peg 1 (resp. peg 2 and peg 3) and moves it to peg 2 (resp. peg 3 and peg 1). Furthermore we denote by $\bar{a}, \bar{b}, \bar{c}$ the inverse moves, e.g. $\bar{a}$ moves the topmost disk from peg 2 to peg 1. The optimal sequence to move 3 disks from peg 1 to peg 2 is given by

$$a\bar{c}ba\bar{c}\bar{b}a.$$

One can show that there exists an infinite sequence of moves such that the first $2^N - 1$ moves give an optimal solutions to move N disks from peg 1 to peg 2 if N is odd or to peg 3 if N is even. For a more detailed treatment see [1]. Indeed one finds that this sequence is automatic and given by the following automaton.



1.1.2 Basic Properties of Automatic Sequences

In the definition of automatic sequences we demanded that the input of the automaton is the canonical representation of n, i.e. without leading zeros. However, there always exists a (possibly different) automaton that has the same output as the original automaton and does not change the output when adding leading zeros.

Theorem 1.1.5. *Let $(a_n)_{n\geq 0}$ be a k-automatic sequence. Then there exists a k-DFAO A such that $a_n = \tau(\delta(q_0, (n)_k))$ for all $n \geq 0$, where also $\delta(q_0, 0) = q_0$. This implies $\tau(\delta(q_0, (n)_k)) = \tau(\delta(q_0, 0^i(n)_k))$ for all $i \geq 0$, i.e. the output does not depend on the representation of n.*

Proof. Let $A' = (Q', \{0, \dots, k-1\}, \delta', q_0', \Delta, \tau')$ be an automaton generating $\mathbf{a}$. We define a new automaton $A = (Q, \{0, \dots, k-1\}, \delta, q_0, \Delta, \tau)$ as follows:

$$\begin{aligned} Q &= Q' \cup \{q_0\} \\ \delta(q, a) &= \delta'(q, a) \qquad \forall q \in Q', a \in \{0, \dots, k-1\} \\ \delta(q_0, 0) &= q_0 \end{aligned}$$

$$\begin{aligned}\delta(q_0, a) &= \delta(q_0', a) \qquad \forall a \in \{1, \ldots, k-1\}\\ \tau(q_0) &= \tau'(q_0')\\ \tau(q) &= \tau'(q) \qquad \forall q \in Q'.\end{aligned}$$

We find that for all $n > 0$ and $i \geq 0$

$$\begin{aligned}\tau(\delta(q_0, 0^i(n)_k)) = \tau(\delta(q_0, (n)_k)) = \tau(\delta'(q_0', (n)_k)) = \tau'(\delta'(q_0', (n)_k))\\ \tau(\delta(q_0, 0^i)) = \tau(q_0) = \tau'(q_0') = \tau'(\delta'(q_0', (0)_k)).\end{aligned}$$

□

Furthermore, we can rename the output of an automatic sequence and obtain again an automatic sequence.

Theorem 1.1.6. *Let $\mathbf{u} = (u_n)_{n\geq 0}$ be a k-automatic sequence with values in Δ and let $\rho : \Delta \to \Delta$. Then the sequence $\rho(\mathbf{u}) = (\rho(u_n))_{n\geq 0}$ is also k-automatic.*

Proof. By the definition of k-automatic, there exists a DFAO $A = (Q, \{0, \ldots, k-1\}, \delta, q_0, \Delta, \tau)$ such that $u_n = \tau(\delta(q_0, (n)_k))$. Now consider the DFAO $A' = (Q, \{0, \ldots, k-1\}, \delta, q_0, \Delta, \rho \circ \tau)$. Clearly, this DFAO generates $\rho(\mathbf{u})$. □

Finally, we show that the product of two automatic sequences is again automatic.

Theorem 1.1.7. *Let $\mathbf{a} = (a_n)_{n\geq 0}$ and $\mathbf{b} = (b_n)_{n\geq 0}$ be two k-automatic sequences with values in Δ_1, Δ_2, respectively. Then $\mathbf{a} \times \mathbf{b} = ([a_n, b_n])_{n\geq 0}$ is k-automatic.*

Proof. Let $A = (Q, \Sigma, \delta, q_0, \Delta, \tau)$ generate $\mathbf{a}$ and $A' = (Q', \Sigma, \delta', q_0', \Delta', \tau')$ generate $\mathbf{b}$. Then $A'' = (Q \times Q', \Sigma, \delta'', [q_0, q_0'], \Delta \times \Delta', \tau'')$ generates $\mathbf{a} \times \mathbf{b}$ where

$$\begin{aligned}\delta''([q, q'], c) &= [\delta(q, c), \delta'(q', c)] \qquad \forall q \in Q, q' \in Q', c \in \Sigma\\ \tau''([q, q']) &= [\tau(q), \tau'(q')] \qquad \forall q \in Q, q' \in Q'.\end{aligned}$$

□

1.2 Equivalent Definitions for Automatic Sequences

Automatic sequences are interesting objects and there are other equivalent approaches to define this class of sequences. We will mention here two alternative ways to define automatic sequences. Even more characterizing properties can be found in [1].

1.2.1 Fixed Points of Uniform Morphisms

We start by defining morphisms.

Definition 1.2.1. Let Σ and Δ be alphabets. A *morphism* is a map h from Σ^* to Δ^* such that $h(xy) = h(x)h(y)$ for every word $x, y \in \Sigma^*$.

Clearly, if h is a morphism, then we have $h(\epsilon) = \epsilon$. Furthermore, once h is defined for all elements of Σ, it can be uniquely extended to a map from Σ^* to Δ^*. Therefore, we give a morphism by specifying its action on Σ.

We are specially interested in the case where $\Sigma = \Delta$. This allows us to iterate a morphism. We define $h^0(\mathbf{w}) = \mathbf{w}$ and $h^{n+1} = h(h^n(\mathbf{w}))$ for all $\mathbf{w} \in \Sigma^*$.

Throughout this work, we work with uniform morphisms. We call a morphism *k-uniform* if $|h(a)| = k$ for all $a \in \Sigma$. A *coding* is a 1-uniform morphism.

Remark. Suppose we have a k-uniform morphism ϕ such that $\phi(a) = a\mathbf{v}$ for some $a \in \Delta, \mathbf{v} \in \Delta^*$. Then, we can construct an infinite sequence/word as a fixed point of ϕ by taking the limit

$$\mathbf{w} := \lim_{n\to\infty} \phi^n(a).$$

Theorem 1.2.2 (Cobham [8])**.** *Let $k \geq 2$. Then a sequence $\mathbf{u} = (u_n)_{n\geq 0}$ is k-automatic if and only if it is the image of a coding of a fixed point of a k-uniform morphism.*

Proof. A detailed proof can be found for example in [1, Theorem 6.3.2]. However, we want to give the basic idea on how to show this result.

Suppose $\mathbf{u}$ is the image of a fixed point of a k-uniform morphism. More precisely, let $\mathbf{u} = \tau(\mathbf{w})$ for a coding $\tau : \Delta \to \Delta'$ and $\mathbf{w} = \phi(\mathbf{w})$ for a k-uniform morphism $\phi : \Sigma^* \to \Sigma^*$. We write $\mathbf{w} = w_0 w_1 w_2 \ldots$ where all $w_i \in \Delta$. We define now a DFAO $A = (\Delta, \{0, \ldots, k-1\}, \delta, q_0, \tau)$ where $q_0 = w_0$ and $\delta(q, a)$ is the a-th letter of $\phi(q)$. One finds easily that $w_n = \delta(q_0, (n)_k)$ for all $n \geq 0$. Thus we find

$$u_n = \tau(w_n) = \tau(\delta(q_0, (n)_k)).$$

Let $\mathbf{u}$ be k-automatic, then it is generated by a DFAO $A = (Q, \{0, \ldots, k-1\}, \delta, q_0, \Delta, \tau)$. By Theorem 1.1.5 we may assume that $\delta(q_0, 0) = q_0$. We define a morphism $\phi : Q^* \to Q^*$ by

$$\phi(q) = \delta(q, 0)\delta(q, 1) \ldots \delta(q, k-1)$$

for all $q \in Q$. One can then easily show that $(\delta(q_0, (n)_k))_{n\geq 0}$ is a fixed point of ϕ and the theorem follows. □

1.2.2 The k-Kernel

The next characterization is given in terms of the so called *k-kernel* and will be useful to treat arithmetic subsequences of automatic sequences.

Definition 1.2.3. Let $\mathbf{u} = (u_n)_{n\geq 0}$ be an infinite sequence. We define the *k-kernel* of $\mathbf{u}$ to be the set of subsequences

$$K_k(\mathbf{u}) = \{(u_{k^i \cdot n+j})_{n\geq 0} : i \geq 0 \text{ and } 0 \leq j < k^i\}$$

Theorem 1.2.4. *Let $k \geq 2$. The sequence $\mathbf{u} = (u_n)_{n\geq 0}$ is k-automatic if and only if $K_k(\mathbf{u})$ is finite.*

Proof. See [1, Theorem 6.6.1] for a detailed proof. □

Example. We calculate the k-kernel of the Thue–Morse sequence $\mathbf{t} = (t_n)_{n\geq 0}$.
Recall that $t_n = s_2(n) \bmod 2$. We find for all $i \geq 0, 0 \leq j < 2^i$ that $s_2(n2^i+j) = s_2(n)+s_2(j)$ and, therefore, $t_{n2^i+j} = (t_n + s_2(j)) \bmod 2$. This shows that the 2-kernel of the Thue–Morse sequence is given by $\{\mathbf{t}, (\mathbf{t}+1 \bmod 2)\}$.

We are now able to show that any arithmetic subsequence of an automatic sequence is again automatic.

Theorem 1.2.5. *Let $\mathbf{u} = (u(n))_{n\geq 0}$ be a k-automatic sequence. Then for all integers $a, b \geq 0$ the subsequence $(u(an+b))_{n\geq 0}$ is also k-automatic.*

For this proof it is particularly useful to use the characterization of automatic sequences by the k-kernel.

Proof. If $a = 0$, then $u(an+b) = u(b)$ is a constant sequence and therefore trivially k-automatic. Let now $a > 0$. Since $\mathbf{u} = (u(n))_{n\geq 0}$ is k-automatic, it has a finite k-kernel which we denote by

$$K_k(\mathbf{u}) = \{(u_1(n))_{n\geq 0}, (u_2(n))_{n\geq 0}, \ldots, (u_r(n))_{n\geq 0}\}.$$

Our goal is to show that the k-kernel of $\mathbf{v} = (v(n))_{n\geq 0} = (u(an+b))_{n\geq 0}$ is finite. Therefore we consider the following set of at most $r \cdot (a+b)$ sequences

$$S = \{(u_i(an+c))_{n\geq 0} : 1 \leq i \leq r, 0 \leq c < a+b\}.$$

We claim that every element of the k-kernel of $\mathbf{v}$ is an element of S, which is sufficient to prove the theorem. Therefore, we need to consider $(v(k^e n+j))_{n\geq 0}$ where $e \geq 0$ and $0 \leq j < k^e$. Using the division algorithm, we can write

$$ja + b = dk^e + f,$$

where $0 \leq f < k^e$ and $0 \leq d < a+b$. This gives

$$v(k^e n+j) = u(a(k^e n+j)+b) = u(k^e(an+d)+f).$$

By the finiteness of the k-kernel of $\mathbf{u}$ there exists i such that

$$u(k^e m+f) = u_i(m)$$

for all $m \geq 0$. Putting $m = an+d$ yields in total that

$$(v(k^e n+j))_{n\geq 0} = u(k^e(an+d)+f)_{n\geq 0} = u_i(an+d)_{n\geq 0} \in S.$$

□

1.3 Frequencies

Some of the main goals of this thesis concern the frequency of letters in subsequences of automatic sequences. Therefore, we give a brief overview about results of *frequencies* of (automatic) sequences. Once again, a more detailed treatment can be found in [1].

Let $\mathbf{x} = (x_n)_{n\geq 0}$ be an infinite sequence with values in Δ. We define the *frequency of the letter* $a \in \Delta$ as

$$\mathrm{Freq}_{\mathbf{x}}(a) = \lim_{N\to\infty} \frac{|\{n < N : x_n = a\}|}{N}$$

if this limit exists. Furthermore, we define the *logarithmic density of the letter* a as

$$\mathrm{LogFreq}_{\mathbf{x}}(a) = \lim_{N\to\infty} \frac{1}{\log N} \sum_{\substack{n<N \\ x_n=a}} \frac{1}{n}$$

if the limit exists.

These two limits are quite closely connected as the following proposition shows.

Proposition 1.3.1. *If the frequency of a in the sequence $\mathbf{x}$ exists, then the logarithmic frequency of a in $\mathbf{x}$ also exists, and these two frequencies are equal.*

Proof. See [1, Proposition 8.4.4]. □

The frequency of a letter need not exist for general automatic sequences. However, Cobham showed in 1972 that the logarithmic density always exists.

Theorem 1.3.2. *Let $\mathbf{x}$ be an automatic sequence. Then all letters occurring in x have a logarithmic frequency.*

Proof. See [1, Corollary 8.4.9]. □

Although the frequencies need not exist for automatic sequences, we can find a restriction for what values the frequencies can take.

Theorem 1.3.3. *Let $\mathbf{x}$ be an automatic sequence. If the frequency of a letter exists, then it is a rational number.*

Proof. See [1, Theorem 8.4.5]. □

Remark. As every periodic sequence is k-automatic for every $k \geq 2$, we see that every $q \in \mathbb{Q} \cap [0, 1]$ appears as the frequency of a letter of an automatic sequence.

Finally, we give a property of automatic sequences that ensures the existence of the frequencies of all letters.

Definition 1.3.4. We say that an automaton is *strongly connected* if and only if for every $q_1, q_2 \in Q$ exists a word $\mathbf{w} \in \Sigma^*$ such that $\delta(q_1, \mathbf{w}) = q_2$.

Furthermore, we say an automaton is *primitive* if and only if there exists $n \geq 0$ such that for all $q_1, q_2 \in Q$ exists a word $\mathbf{w} \in \Sigma^*$ such that $|\mathbf{w}| = n$ and $\delta(q_1, \mathbf{w}) = q_2$.

We say an automatic sequence $\mathbf{a}$ is *primitive* if and only if there exists a primitive DFAO that generates $\mathbf{a}$.

Remark. We see easily that a strongly connected automaton where also $\delta(q_0, 0) = q_0$ holds is already primitive.

Theorem 1.3.5. *Let $\mathbf{x}$ be a primitive automatic sequence. Then the frequencies of all letters exist, and are nonzero.*

Proof. See [1, Theorem 8.4.7]. □

1.4 Recent Results for Automatic Sequences

We present in this section some recent results concerning (special) automatic sequences that are of importance for this thesis.

1.4.1 Synchronizing Automata

There is a recent paper by Deshouillers, Drmota and the author [11] that treats a special class of automatic sequences, namely synchronizing automatic sequences, which are important for this thesis.

We call a DFAO $A = (Q, \Sigma, \delta, q_0, \Delta, \tau)$ *synchronizing* if there exists a *synchronizing word* $\mathbf{w}_0 \in \Sigma^*$ whose action resets A, i.e. $\mathbf{w}_0$ leaves the automaton in one specific state, no matter which state in Q it is applied to: $\delta(q, \mathbf{w}_0) = \delta(q_0, \mathbf{w}_0)$ for all $q \in Q$. Note that the output of a synchronizing automaton for an input word only depends on the last occurrence of the synchronizing word and the part thereafter. Furthermore, we call an automatic sequence *synchronizing* if there exists a synchronizing DFAO that generates this sequence.

Remark. Let A be a DFAO and $\mathbf{a}$ be the corresponding automatic sequence. Suppose there exists $n \geq 1$ such that every word $\mathbf{w} \in \Sigma^n$ is synchronizing. Then, $\mathbf{a}$ is a k^n periodic function.

This class of automata is particularly interesting as Berlinkov showed in [2, 3] that "almost all" automata are synchronizing.

We will mention now the most important results of this paper:

The first result gives bounds on how often synchronizing words occur and will also be used later.

Lemma 1.4.1 (Lemma 2.2 [11])**.** *Let A be a synchronizing DFAO. There exists $\eta > 0$ such that the number of words of length n which are not synchronizing is at most $O(k^{n(1-\eta)})$.*

This lemma states that if you take a word of length n at random, the probability that it is not synchronizing is very small - i.e $O(k^{-n\eta})$.

The main idea to treat synchronizing automatic sequences is to approximate them by periodic sequences. This yields the following results.

Theorem 1.4.2 (Theorem 4.4 [11]). *Suppose the* $\mathbf{a} \in \Delta^{\mathbb{N}}$ *is a synchronizing automatic sequence. Then for every* $\alpha \in \Delta$ *the frequency* $\mathrm{Freq}_{\mathbf{a}}(\alpha)$ *exists.*

Although not explicitly stated in the original proof, one can use the arguments to show that there exists $\eta > 0$ such that

$$|\#\{n < N : a_n = \alpha\} - N \cdot \mathrm{Freq}_{\mathbf{a}}(\alpha)| \ll N^{1-\eta}. \tag{1.2}$$

The authors of [11] also show that the frequencies for certain subsequences of synchronizing automatic sequences exist, namely the subsequences along the set of primes and along positive integer valued polynomials.

Furthermore, they show that - under some technical conditions - the frequencies for the product of two synchronizing automatic sequences exists and is given by the product of the frequencies of the two synchronizing automatic sequences.

Lastly, Deshouillers, Drmota and the author show that synchronizing automatic sequences are orthogonal to the Möbius function and even fulfill the Sarnak conjecture - we will discuss this again in Chapter 3.

1.4.2 Rudin–Shapiro Sequence

Next we discuss a result by Mauduit and Rivat [38]. They study sequences with digit properties and in particular the Rudin–Shapiro sequence $(r_n)_{n\geq 0}$ - which we already discussed - along the subsequence of primes, i.e. $(r_p)_{p\in\mathbb{P}}$. Therefore, they developed a strong framework that allows to find results for subsequences along primes and also the correlation with the Möbius function. From now on, we fix an integer k which we use as a base for digital expansions. This framework requires that the sequences fulfill the following two conditions.

The first condition is called carry property and states in some sense that a disturbance of "low digits" usually does not influence the contribution of "high digits". Therefore they define for $\alpha \in \mathbb{N}$ and $f : \mathbb{N} \to \mathbb{U}$ a truncated version of f by $f_\alpha(n) := f(n \bmod k^\alpha)$.

Definition 1.4.3. A function $f : \mathbb{N} \to \mathbb{U}$ has the carry property if uniformly for $(\lambda, \alpha, \rho) \in \mathbb{N}^3$ with $\rho < \lambda$, the number of integers $0 \leq \ell < k^\lambda$ such that there exists $(n_1, n_2) \in \{0, \dots, k^\alpha - 1\}^2$ with

$$f(\ell k^\alpha + n_1 + n_2)\overline{f(\ell k^\alpha + n_1)} \neq f_{\alpha+\rho}(\ell k^\alpha + n_1 + n_2)\overline{f_{\alpha+\rho}(\ell k^\alpha + n_1)} \tag{1.3}$$

is at most $O(k^{\lambda-\rho})$ where the implied constant may depend only on k and f.

The second property states that the - slightly generalized - Discrete Fourier Transform is uniformly small.

Definition 1.4.4. Given a non decreasing function $\gamma : \mathbb{R} \to \mathbb{R}$ satisfying $\lim_{\lambda\to\infty} \gamma(\lambda) = +\infty$ and $c > 0$ we denote by $F_{\gamma,c}$ the set of functions $f : \mathbb{N} \to \mathbb{U}$ such that for $(\alpha, \lambda) \in \mathbb{N}^2$ with $\alpha \leq c\lambda$ and $t \in \mathbb{R}$:

$$\left| k^{-\lambda} \sum_{u<k^\lambda} f(uk^\alpha)\, \mathrm{e}(-ut) \right| \leq k^{-\gamma(\lambda)}. \tag{1.4}$$

They show that these properties are sufficient to find the following statements.

Theorem 1.4.5. *Let $\gamma : \mathbb{R} \to \mathbb{R}$ be a non decreasing function satisfying $\lim_{\lambda\to\infty} \gamma(\lambda) = +\infty$, and $f : \mathbb{N} \to \mathbb{U}$ be a function satisfying Definition 1.4.3 and $f \in \mathcal{F}_{\gamma,c}$ for some $c \geq 10$ in Definition 1.4.4. Then for any $\theta \in \mathbb{R}$ we have*

$$\left| \sum_{n\leq x} \Lambda(n) f(n)\, \mathrm{e}(\theta n) \right| \ll c_1(k)(\log x)^{c_2(k)} x k^{-\gamma(2\lfloor(\log x)/(80\log k)\rfloor/20}, \tag{1.5}$$

for some explicit constants c_1, c_2.

Λ denotes here the Mangoldt function which is defined by $\Lambda(n) = \log p$ if $n = p^\ell$ for some $p \in \mathbb{P}$ and $\ell \in \mathbb{N}$ and $\Lambda(n) = 0$ otherwise. This is not explicitly a theorem about prime numbers, but one can find a corresponding theorem for primes by partial summation.

Theorem 1.4.6. *Let $\gamma : \mathbb{R} \to \mathbb{R}$ be a non decreasing function satisfying $\lim_{\lambda\to\infty} \gamma(\lambda) = +\infty$, and $f : \mathbb{N} \to \mathbb{U}$ be a function satisfying Definition 1.4.3 and $f \in \mathcal{F}_{\gamma,c}$ for some $c \geq 10$ in Definition 1.4.4. Then for any $\theta \in \mathbb{R}$ we have*

$$\left\| \sum_{n\leq x} \mu(n) f(n)\, \mathrm{e}(\theta n) \right\| \ll c_1(k)(\log x)^{\frac{9}{4}+\frac{1}{4}\max(\omega(k),2)} x k^{-\gamma(2\lfloor(\log x)/(80\log k)\rfloor/20}, \tag{1.6}$$

for the same constants c_1, c_2 as in Theorem 1.4.5.

The proof relies first on Vaughan's identity [29]. Thus it is sufficient to estimate bilinear sums of the form

$$\sum_n \sum_m a_m b_n f(mn).$$

One of the main ideas of the proof is to use the carry property - after applying the Cauchy–Schwarz inequality or variants of it - to reduce the problem to sums over short intervals. Then, one can use a Fourier-analytic treatment as the Fourier transform is uniformly small.

This paper is of particular interest as we will later generalize its results slightly. This generalization will play a substantial role for Chapter 3 and Chapter 4.

1.4.3 The Thue–Morse Sequence along Squares

The next result is due to Drmota, Mauduit and Rivat [14] and concerns the subsequence of the Thue–Morse sequence along squares, i.e. $(t_{n^2})_{n\geq 0}$.

It is a well-known result that the subword complexity of automatic sequences, i.e. the number of different blocks of length n that appear within the sequence, grows at most linearly.

Theorem 1.4.7. *[Corollary 10.3.2 [1]] Let* $\mathbf{a}$ *be an automatic sequence. Then the subword complexity of* $\mathbf{a}$*,*

$$p_{\mathbf{a}}(n) = \#\{(a_i, \ldots, a_{i+n-1}) : i \geq 0\}$$

is sub-linear, i.e.

$$p_{\mathbf{a}}(n) = O(n).$$

For a random sequence $\mathbf{u} \in \{0,1\}^{\mathbb{N}}$ one finds that $p_{\mathbf{u}}(n) = 2^n$ almost surely. Thus, we see that automatic sequences are far from being random.

However, the situation changes completely when one considers the subsequence along squares. They showed that not only $p_{(t_{n^2})_{n\geq 0}}(L) = 2^L$, but were able to show how often such a block appears.

Theorem 1.4.8. *[Theorem 1 [14]] The sequence* $(t_{n^2})_{n\geq 0}$ *is normal, i.e. for any* $L \geq 1$ *and any* $(b_0, \ldots, b_{L-1}) \in \{0,1\}^L$*, we have*

$$\lim_{N\to\infty} \frac{1}{N} \#\{i < N : t_{i^2} = b_0, \ldots, t_{(i+L-1)^2} = b_{L-1}\} = \frac{1}{2^L}.$$

To prove this result, they start by relating the problem to a statement on exponential sums. They apply then Van-der-Corput type inequalities and use results on carry propagation to reduce the problem to sums that only depend on few digits of $(n^2, (n+1)^2, \ldots, (n+L-1)^2)$. Thereafter they have to handle quadratic exponential sums and statements about Fourier terms.

1.4.4 Invertible Automatic Sequence

In order to define the class of invertible automatic sequences, we need another concept concerning automata, namely transition matrices.

Let $A = (Q = \{q_0, \ldots, q_{t-1}, \Sigma = \{0, \ldots, k-1\}, \delta, q_0, \Delta, \tau)$ be a DFAO. We define the *transition matrices* $M_\varepsilon \in \mathbb{Z}^{t\times t}$ for $\varepsilon \in \Sigma$ to be the matrix, such that the entries are given by $m^\varepsilon_{ij} = \mathbf{1}_{[\delta(q_j,\varepsilon)=q_i]}$.

One finds easily that for a transition matrix, in each column there exists exactly one 1 and all other entries are 0. We denote as usual the j-th standard vector in $\mathbb{Z}^t$ (that is, all entries are 0 except the j-th, which is equal to 1) by e_j. We see easily that $\delta(q_j, \varepsilon) = q_i$ for $\varepsilon \in \Sigma, i, j < t$ if and only if $e_i = M_\varepsilon \cdot e_j$. This can easily be extended to $\mathbf{w} = w_0 \cdots w_\ell \in \Sigma^*$: $\delta(q_j, \mathbf{w}) = q_i$ if and only if $e_i = M_{w_\ell} \cdots M_{w_0} \cdot e_j$.

Definition 1.4.9. Let $\mathbf{u}$ be a k-automatic sequence. Then we call $\mathbf{u}$ an *invertible k-automatic sequence* if there exists a primitive DFAO that generates $\mathbf{u}$ and all transition matrices are invertible.

This class of automatic sequences was introduced by Drmota and Morgenbesser in [13], where the subsequences along the squares were studied. Furthermore, Drmota also adopted and applied the framework developed by Mauduit and Rivat to study the subsequence along the primes and the correlation with the Möbius function [13].

One of the main ideas to obtain these results is to utilize the group structure that is given by the transition matrices. Furthermore, these results rely among other things again on carry properties and Fourier estimates concerning the transition matrices.

Invertible automatic sequence are in some sense the opposite of synchronizing automatic sequences:
For an invertible automaton, we find that $\delta(., \mathbf{w})$ is invertible for all $\mathbf{w} \in \Sigma^*$, i.e. for all $q \in Q, \mathbf{w} \in \Sigma^*$ exists exactly one q' such that $\delta(q', \mathbf{w}) = q$. Thus, we can backtrack any element under the transition of any word.
For a synchronizing automaton with synchronizing word $\mathbf{w}_0$ we find that $\delta(., \mathbf{w}_0)$ is as far from invertible as possible, as all states are mapped to exactly one state and all "information is lost".

Chapter 2

Naturally Induced Transducers

We develop in this chapter for any strongly connected automaton a new structure - namely naturally induced transducers - that mimics the behavior of the automaton. This structure combines aspects of invertible and synchronizing automata, and makes an easier treatment of the automatic sequence possible. We will see that the two aspects are almost independent of each other and can be analyzed independently. The synchronizing part can be analyzed by elementary methods. However, the invertible part is much more challenging and we will make an effort to understand it sufficiently well to be able to obtain distributional results. To illustrate the effectiveness of this concept, we will derive the densities of automatic sequences along arithmetic subsequences - under some technical conditions.

At the beginning of this chapter, we give some more necessary definitions. A finite-state transducer $\mathcal{T}$ is a sextuple $(Q, \Sigma, \delta, q_0, \Delta, \lambda)$, where Q is a finite set of states, Σ is the input alphabet, δ is the transition function, q_0 is the initial state, Δ is the output alphabet and $\lambda : Q \times \Sigma \to \Delta^*$ is the output function. We will restrict ourselves to $\lambda : Q \times \Sigma \to \Delta$ and $\Sigma = \{0, \dots, k-1\}$.

A transducer can be viewed as a mean to define functions: on input $\mathbf{w} = w_1 w_2 \dots w_r$ the transducer enters states $q_0 = \delta(q_0, \varepsilon), \delta(q_0, w_1), \dots, \delta(q_0, w_1 w_2 \dots w_r)$ and produces the outputs

$$\lambda(q_0, w_1), \lambda(\delta(q_0, w_1), w_2), \dots, \lambda(\delta(q_0, w_1 w_2 \dots w_{r-1}), w_r).$$

The function $T(\mathbf{w})$ is then defined as

$$T(\mathbf{w}) := \prod_{j=0}^{r-1} \lambda(\delta(q_0, w_1 w_2 \dots w_j), w_{j+1}).$$

We also define the slightly more general form,

$$T(q, \mathbf{w}) := \prod_{j=0}^{r-1} \lambda(\delta(q, w_1 w_2 \dots w_j), w_{j+1}).$$

Throughout this work we assume that $(\Delta, \circ)$ is a group generated by $im(\lambda)$ and the product in the definition of T is a in general non-commutative product according to $\circ$.

We define for a set $M \subseteq \Delta$ the inverse set and the multiplication of two sets $M_1, M_2 \subseteq \Delta$ as usual,

$$M^{-1} := \{g^{-1} | g \in M\}$$
$$M_1 \cdot M_2 := \{g_1 \circ g_2 | g_1 \in M_1, g_2 \in M_2\}.$$

We say a transducer (or analogously a DFA) is **synchronized** if and only if

$$\exists q \in Q, \mathbf{w}_q \in \Sigma^* \quad \forall q_1 \in Q : \delta(q_1, \mathbf{w}_q) = q.$$

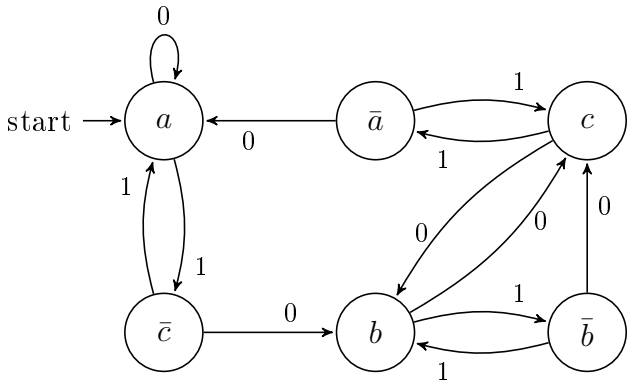
We call $\mathbf{w}_q$ a **synchronizing word**.
We define for $\sigma \in S_n$ and a n-tuple $\mathbf{x} = (x_1, \ldots, x_n)$

$$\sigma \cdot \mathbf{x} := (x_{\sigma^{-1}(1)}, \ldots, x_{\sigma^{-1}(n)}).$$

2.1 Automaton to Transducer

In this section, we develop a special way to construct for an automaton $A = (Q', \Sigma, \delta', q_0')$ a "naturally induced" transducer $\mathcal{T}_A = (Q, \Sigma, \delta, q_0, \Delta, \lambda)$ which contains all the information of A. The main idea is to merge some states of A into a single state of $\mathcal{T}_A$ consistent with δ and δ'. To illustrate this concept, we start with a motivating example:

Example. We treat the sequence of moves, for the Tower of Hanoi problem, which we already mentioned in Chapter 1. The automaton generating this sequence looks as follows.



One finds easily

$$\delta'(\{a,b,c\},0) = \{a,c,b\} \qquad \delta'(\{a,b,c\},1) = \{\bar{c},\bar{b},\bar{a}\}$$
$$\delta'(\{\bar{a},\bar{b},\bar{c}\},0) = \{a,c,b\} \qquad \delta'(\{\bar{a},\bar{b},\bar{c}\},1) = \{c,b,a\}.$$

Consequently, we want to merge the triple (a,b,c) into a single state q_0 (i.e. $q_0 = (a,b,c)$) and $q_1 = (\bar{a},\bar{b},\bar{c})$. We find

$$\delta'(q_0,0) = (23) \cdot q_0 \qquad \delta'(q_0,1) = (13) \cdot q_1$$
$$\delta'(q_1,0) = (23) \cdot q_0 \qquad \delta'(q_1,1) = (13) \cdot q_0.$$

We define ($\Delta := S_3$)

$$\begin{aligned} \delta(q_0,0) &= q_0 & \delta(q_0,1) &= q_1 \\ \delta(q_1,0) &= q_0 & \delta(q_1,1) &= q_0 \\ \lambda(q_0,0) &= (23) & \lambda(q_0,1) &= (13) \\ \lambda(q_1,0) &= (23) & \lambda(q_1,1) &= (13) \end{aligned}$$

and find that $\delta'(q_i,\varepsilon) = \lambda(q_i,\varepsilon)\cdot\delta(q_i,\varepsilon)$ holds. This is already the first example of a naturally induced transducer.

To formalize this idea we first define induced transducers; we denote here by $\pi_1(q)$ the first coordinate of q.

Definition 2.1.1. We call a transducer $\mathcal{T}_A$ **induced** by a DFA $A = (Q', \Sigma, \delta', q_0')$ if and only if
$\mathcal{T}_A = (Q, \Sigma, \delta, q_0, \Delta, \lambda)$ where

1) $\exists n_0 \in \mathbb{N} : Q \subseteq (Q')^{n_0}$

2) $\pi_1(q_0) = q_0'$

3) $\Delta < S_{n_0}$ is the subgroup generated by λ

4) $\forall q \in Q, a \in \Sigma : \delta'(q,a) = \lambda(q,a)\cdot\delta(q,a)$.

We call $\mathcal{T}_A$ a **naturally induced transducer** if furthermore

5) $\forall i \neq j \leq n_0, q \in Q : \pi_i(q) \neq \pi_j(q)$,

6) $\forall q_1 \neq q_2 \in Q \not\exists \sigma \in S_{n_0} : q_1 = \sigma \cdot q_2$,

7) $\mathcal{T}_A$ is strongly connected

8) $\mathcal{T}_A$ is synchronizing

holds.

Remark. Properties 5)-7) only assure that $\mathcal{T}_A$ is chosen minimal to certain aspects.

Example. We call $\mathcal{T}_A$ the **trivial** transducer induced by A when:
$n_0 = 1, Q = Q', \Delta = \{id\}, \forall q \in Q, a \in \Sigma : \delta(q,a) = \delta'(q,a), \lambda(q,a) = id$.
The trivial induced transducer is a naturally induced transducer if and only if A is synchronizing and strongly connected.

Proposition 2.1.2. *For every strongly connected automaton A, there exists a naturally induced transducer $\mathcal{T}_A = (Q, \Sigma, \delta, q_0, \Delta, \lambda)$. All other naturally induced transducers can be obtained by changing the order on the elements of Q and possibly changing the initial state q_0.*

Proof. The given proof is not constructive, but could be modified to be constructive as all appearing objects are finite. We start by defining $n_0(A) := \min\{|\delta'(Q', \mathbf{w})| : \mathbf{w} \in \Sigma^*\}$ and

$$S(A) := \{M \subseteq Q' : |M| = n_0(A), \exists \mathbf{w} \in \Sigma^* \text{ with } \delta'(Q', \mathbf{w}) = M\},$$

which are exactly the sets of size $n_0(A)$ that are reachable from Q' using δ'. One observes easily that an automaton is synchronizing if and only if $n_0 = 1$ and invertible[1] if and only if $n_0 = |Q'|$.

Now we are prepared to construct a naturally induced transducer for A, similar to the previous example.
We choose $n_0 = n_0(A)$. Furthermore, we choose any ordering $\leq_{Q'}$ of Q' such that q'_0 is the minimal element. For every $M \in S(A)$ we define a corresponding n_0-tuple q_M which consists of the elements of M ordered by $\leq_{Q'}$. We define $Q := \{q_M : M \in S(A)\}$. (As A is strongly connected it follows directly by property 4) that) Q' is covered by $S(A)$ and we define q_0 as an arbitrary element $q_M \in Q$ such that $q'_0 \in M$. (Note that $\pi_1(q_M) = q'_0$ since q'_0 is the minimal element).

Take an arbitrary $q_M \in Q$. By minimality of $|M|$ for $a \in \Sigma$ we have $\delta'(M, a) \in S(A)$. Thus we define $\delta(q_M, a) := q_{\delta'(M,a)}$. We define $\lambda' : (Q')^{n_0} \to S_{n_0}$ such that for all q $\lambda'(q) \cdot q$ is ordered with respect to $\leq_{Q'}$ and $\lambda : Q \times \Sigma \to S_{n_0}, \lambda(q_M, a) = \lambda'(\delta'(q_M, a))^{-1}$.

Let Δ be the subgroup of S_{n_0} generated by λ. This defines an induced transducer $\mathcal{T}_A = (Q, \Sigma, \delta, q_0, \Delta, \lambda)$.

It remains to show that this induced transducer is a naturally induced transducer. 5) and 6) follow directly from the construction of the transducer. Every $q_M \in Q$ corresponds to $M \in S(A)$ for which there exists $\mathbf{w}_M$ such that $\delta'(Q', \mathbf{w}_M) = M$. Thus it follows directly that for all $q_{M'} \in Q$ it holds $\delta(q_{M'}, \mathbf{w}_M) = q_M$ which shows that the constructed transducer is synchronizing and also strongly connected.
To prove the second part of this proposition we assume that $\overline{\mathcal{T}_A} = (\overline{Q}, \Sigma, \overline{\delta}, \overline{q_0}, \overline{\Delta}, \overline{\lambda})$ is an arbitrary naturally induced transducer. We start by showing $n_0 = \overline{n_0}$.

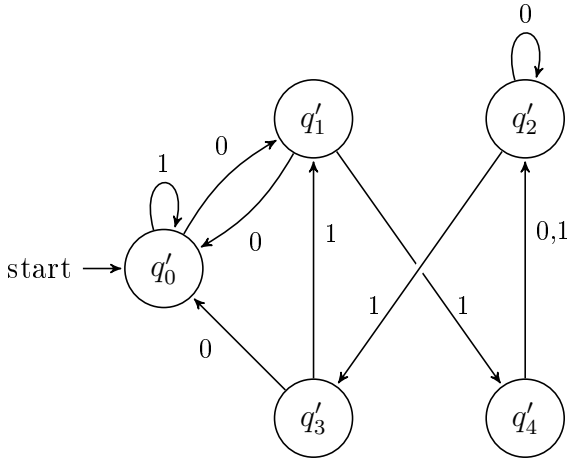
- Assume $\overline{n_0} > n_0$: By the definition of n_0 there exists some $\mathbf{w} \in \Sigma^*$ such that $|\delta'(Q', \mathbf{w})| = n_0$. Consequently $\overline{\delta}(\overline{q}, \mathbf{w}) \in (Q')^{\overline{n_0}}$ violates 5) for every $\overline{q} \in \overline{Q}$.
- Assume $\overline{n_0} < n_0$: As $\overline{\mathcal{T}_A}$ is synchronizing there exists $\mathbf{w} \in \Sigma^*$ and $\overline{q_1}$ such that for all $\overline{q} \in \overline{Q}$ we have $\overline{\delta}(\overline{q}, \mathbf{w}) = \overline{q_1}$. $\overline{q_1}$ corresponds to the set of its components $M \subseteq Q'$. Since $\overline{Q}$ covers Q' this implies that $|\delta'(Q', \mathbf{w})| \leq \overline{n_0} = \left|\delta'(\overline{Q}, \mathbf{w})\right| < n_0(A)$ which contradicts the definition of n_0.

Assume now that there exists some $q \in Q$ such that no element of $\overline{Q}$ is a permutation of q. The elements of q form a set $M_q \in S(A)$ and thus there exists $\mathbf{w}_q$ such that $\delta'(Q', \mathbf{w}_q) = M_q$. $\overline{\delta}(\overline{q}, \mathbf{w}_q)$ contains exactly the elements of M_q which means that it is a permutation of q (by property 5)).
The elements of $\overline{Q}$ that are permutations of elements of Q form a strongly connected component. Thus, the proposition follows by property 7). □

[1] For the definition of invertible automata see Definition 1.4.9.

Remark. Let $\mathcal{T} = (Q, \Sigma, \delta, q_0, \Delta, \lambda)$ be a transducer fulfilling 1)-5) and 7)-8). We define an equivalence relation $\sim$ on Q by: $q_1 \sim q_2$ if and only if $\exists \sigma$: $q_1 = \sigma \cdot q_2$. Then $\mathcal{T}/\sim$ is a naturally induced transducer.

Example. We consider the following more complex automaton with $\Sigma = \{0, 1\}$ and find a naturally induced transducer.



One finds that $n_0(A) = 3$ and $S(A) = \{M_1, M_2\}$ with

$$M_1 := \delta'(Q', 0) = \{q_0', q_1', q_2'\}$$
$$M_2 := \delta'(Q', 01) = \{q_0', q_3', q_4'\}.$$

We construct now a naturally induced transducer $\mathcal{T}_A$ and start by defining $Q = \{q_0, q_1\}$ where

$$q_0 = (q_0', q_1', q_2')$$
$$q_1 = (q_0', q_3', q_4').$$

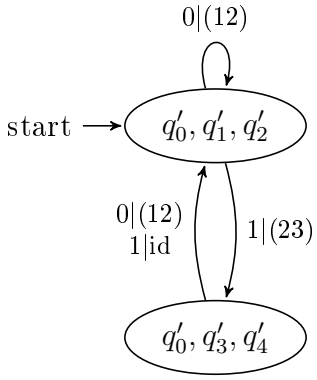
We find

$$\delta'(q_0, 0) = (q_1', q_0', q_2'), \quad \delta'(q_0, 1) = (q_0', q_4', q_3')$$
$$\delta'(q_1, 0) = (q_1', q_0', q_2'), \quad \delta'(q_1, 1) = (q_0', q_1', q_2')$$

and therefore

$$\delta(q_0, 0) = q_0, \quad \delta(q_0, 1) = q_1$$
$$\delta(q_1, 0) = q_0, \quad \delta(q_1, 1) = q_0$$
$$\lambda(q_0, 0) = (12), \quad \lambda(q_0, 1) = (23)$$
$$\lambda(q_1, 0) = (12), \quad \lambda(q_1, 1) = id.$$

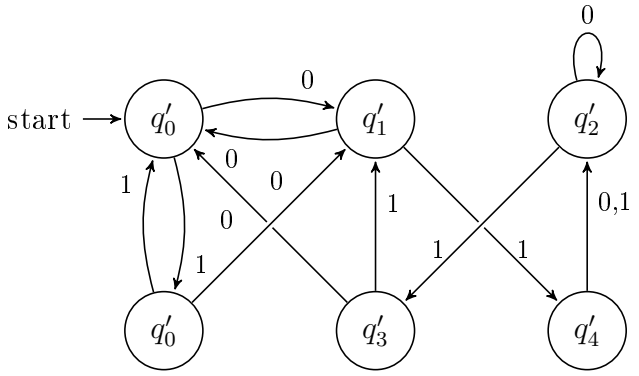
Thus we constructed a transducer $\mathcal{T}_A$ (as described in the proof of Proposition 2.1.2):



One observes easily that this is a naturally induced transducer.

Remark. It is sometimes useful to consider the automaton where each state occurs as often as it occurs in the subsets of $S(A)$ and "group" them according to $S(A)$.

Example. For the example above this gives the following automaton.



We end this section with a technical lemma used later.

Lemma 2.1.3. *Let $\mathcal{T}_A$ be a naturally induced transducer. There exists $N_0 \in \mathbb{N}$ such that for all $q_1, q_2 \in Q$ and $n \geq N_0$ there exists $\mathbf{w} \in \Sigma^n$ such that $\delta(q_1, \mathbf{w}) = q_2$.*

Proof. Let $\mathbf{w}_q$ be a synchronizing word. We denote by $dist(x, y) := \min\{n | \exists \mathbf{w} \in \Sigma^n : \delta(x, \mathbf{w}) = y\}$ and define $N_0 := |\mathbf{w}_q| + \max_{q' \in Q} dist(q, q')$.
Take $q_1, q_2 \in Q$ and $n \geq N_0$. There exists $\mathbf{w}_1$ from q to q_2 of length $dist(q, q_2)$. We consider $\mathbf{w} = 0^{n-|\mathbf{w}_q|-dist(q_0,q_2)}\mathbf{w}_q\mathbf{w}_1$ and see that this is a path from q_1 to q_2 of length n (0^x denotes the word consisting of x consecutive zeros). □

2.2 Connection between an automaton and its induced transducer

Property 4) of an induced transducer relates one step of A to one step of $\mathcal{T}_A$ and λ/T. In this section we show how arbitrary many steps of A relate to $\mathcal{T}_A$. Therefore we start by a technical lemma that describes how the permutation of one state $q \in Q$ influences the behavior of the induced transducer.

Lemma 2.2.1. *Let A be a DFA and $\mathcal{T}_A$ an induced transducer. Furthermore, let $\sigma \in S_{n_0}, a \in \Sigma$ and $q = (x_1, \ldots, x_{n_0}) \in Q$. It holds*

$$\delta'(\sigma \cdot q, a) = (\sigma \circ \lambda(q, a)) \cdot \delta(q, a).$$

Proof. We find for $1 \leq k \leq n_0$

$$\begin{aligned}\pi_k(\delta'(\sigma \cdot (x_1, \ldots, x_{n_0}), a)) &= \delta'(x_{\sigma^{-1}(k)}, a) = \pi_{\sigma^{-1}(k)}(\delta'(q, a)) \\ &= \pi_{\sigma^{-1}(k)}(\lambda(q, a) \cdot \delta(q, a)) = \pi_k(\sigma \cdot (\lambda(q, a) \cdot \delta(q, a))).\end{aligned}$$

Thus, it holds

$$\delta'(\sigma \cdot q, a) = \sigma \cdot (\lambda(q, a) \cdot \delta(q, a)) = (\sigma \circ \lambda(q, a)) \cdot \delta(q, a).$$

□

Now we are ready to show a very important connection between an automaton and its induced transducer.

Proposition 2.2.2. *Let A be a strongly connected automaton and $\mathcal{T}_A$ an induced transducer. For every $\mathbf{w} \in \Sigma^*$ we have*

$$\delta'(q_0', \mathbf{w}) = \pi_1(T(q_0, \mathbf{w}) \cdot \delta(q_0, \mathbf{w})).$$

Proof. We actually prove that for all $q \in Q$ and $\mathbf{w} \in \Sigma^*$ it holds that

$$\delta'(q, \mathbf{w}) = T(q, \mathbf{w}) \cdot \delta(q, \mathbf{w}),$$

which obviously implies the statement. We use induction on the length of $\mathbf{w}$.

- $|\mathbf{w}| = 0$: Obviously $\delta'(q, \varepsilon) = q = (id(\delta(q, \varepsilon)))$ holds.
- $\mathbf{w} = \mathbf{w}'x$: We define $q_1 := \delta(q, \mathbf{w}')$ and find

  $$\begin{aligned}\delta'(q, \mathbf{w}'x) &= \delta'(\delta'(q, \mathbf{w}'), x) = \delta'(T(q, \mathbf{w}') \cdot q_1, x) = (T(q, \mathbf{w}') \circ \lambda(q_1, x)) \cdot \delta(q_1, x) \\ &= T(q, \mathbf{w}'x) \cdot \delta(q, \mathbf{w}'x)).\end{aligned}$$

  The third equality holds by Lemma 2.2.1.

□

This Proposition allows us to reconstruct the output of A by knowing the output of $\mathcal{T}_A$ and T.

Example. We continue our previous example and find

$$\begin{aligned}\delta'(q_0', 0110) = q_2' &= \pi_1(T(q_0, 0110) \cdot \delta(q_0, 0110)) = \pi_1(((12) \circ (23) \circ id \circ (12)) \cdot q_0) \\ &= \pi_1((13) \cdot q_0) = \pi_1(q_2', q_1', q_0') = q_2'.\end{aligned}$$

As one has some freedom on how to choose the order of the elements of Q it is natural to ask how this choice influences the induced transducer.

Proposition 2.2.3. *Let A be a DFA and $\mathcal{T}_A = (Q, \Sigma, \delta, q_0, \Delta, \lambda)$ an induced transducer. Let $\overline{\mathcal{T}}_A$ be another induced transducer obtained by changing the order on every tuple $q \in Q$ by a permutation σ_q, i.e. $\overline{q} = \sigma_q \cdot q$ (where still $\pi_1(\sigma_{q_0} \cdot q_0) = q_0'$ holds). For $\mathbf{w} \in \Sigma^*$ it holds that*

$$\overline{T}(\overline{q}, \mathbf{w}) = \sigma_q \circ T(q, a) \circ \sigma^{-1}_{\delta(q,\mathbf{w})}.$$

Furthermore, if $\mathcal{T}_A$ is a naturally induced transducer, then so is $\overline{\mathcal{T}}_A$.

Proof. Let $q \in Q, a \in \Sigma$. We define for all $q \in Q, a \in \Sigma$

$$\begin{aligned} \overline{q} &:= \sigma_q \cdot q \\ \overline{\delta}(\overline{q}, a) &:= \overline{\delta(q, a)} \\ \overline{\lambda}(\overline{q}, a) &:= \sigma_q \circ \lambda(q, a) \circ \sigma^{-1}_{\delta(q,a)}, \end{aligned}$$

$\overline{Q} := \cup_{q \in Q} \overline{q}$ and $\overline{\Delta}$ is again the group generated by $\overline{\lambda}$.
We claim that $\overline{\mathcal{T}}_A = (\overline{Q}, \Sigma, \overline{\delta}, \overline{q_0}, \overline{\Delta}, \overline{\lambda})$ is again an induced transducer. 1)-3) follow immediately so it remains to show 4). Therefore, we compute by Lemma 2.2.1

$$\begin{aligned} \delta'(\sigma_q \cdot q, a) &= (\sigma_q \circ \lambda(q, a)) \cdot \delta(q, a) \\ &= (\sigma_q \circ \lambda(q, a) \circ \sigma^{-1}_{\delta(q,a)}) \cdot (\sigma_{\delta(q,a)} \cdot \delta(q, a)) \\ &= \overline{\lambda}(\overline{q}, a) \cdot \delta(q, a). \end{aligned}$$

Thus $\overline{\mathcal{T}}_A$ is indeed an induced transducer and the stated equation follows easily by induction on the length of $\mathbf{w}$.

The last statement is very easy to verify. □

2.3 Length restrictions for naturally induced transducers

As Δ is generated by λ, one might assume that all elements of Δ can be obtained by T. We show in this section that this is not true in general but for certain naturally induced transducers. Furthermore, we will show that restrictions on the length of $\mathbf{w}$ gives restriction on what elements of Δ can be obtained by $T(q, \mathbf{w})$.
The main result of this section – which contains all the facts mentioned above – is the following theorem.

Theorem 2.3.1. *Let A be a strongly connected automaton. There exists a minimal $d \in \mathbb{N}$, $m_0 \in \mathbb{N}$, a naturally induced transducer $\mathcal{T}_A$ and a subgroup G of Δ such that the following two conditions hold.*

- *For all $q \in Q, \mathbf{w} \in (\Sigma^d)^*$ we have $T(q, \mathbf{w}) \in G$.*

- *For all $g \in G, q, \overline{q} \in Q$ and $n \geq m_0$ it holds that*

$$\{T(q, \mathbf{w}) : \mathbf{w} \in \Sigma^{nd}, \delta(q, \mathbf{w}) = \overline{q}\} = G.$$

d and m_0 only depend on A, but not on its initial state q_0'.

The other results of this section are rather technical and only used in this and the following section.
According to Proposition 2.1.2 we assume in this section that $\mathcal{T}_A$ is a naturally induced transducer of a strongly connected automaton. Thus, Δ is finite.
Paths $\mathbf{w}$ with the property

$$\delta(q, \mathbf{w}) = q, T(q, \mathbf{w}) = id \tag{2.1}$$

will play an important role in this section and we start by exploring some of its properties.

Remark. Obviously concatenating two paths $\mathbf{w}_1, \mathbf{w}_2$ fulfilling (2.1) gives again a path fulfilling (2.1).

We start this section by showing that every path can be extended to a path with this property.

Lemma 2.3.2. *For every $q \in Q, \mathbf{w} \in \Sigma^*$ there exists an **inverse path** $\overline{\mathbf{w}}^q \in \Sigma^*$ such that (2.1) holds for $\mathbf{w}\overline{\mathbf{w}}^q$.*

Proof. Since $\mathcal{T}_A$ is strongly connected, there exists $\mathbf{w}' \in \Sigma^*$ such that $\delta(q, \mathbf{w}\mathbf{w}') = \delta(\delta(q, \mathbf{w}), \mathbf{w}') = q$.
We define $g := T(q, \mathbf{w}\mathbf{w}')$. Since G is finite we know that $g^n = id$ for some $n \in \mathbb{N}$. Since $\delta(q, (\mathbf{w}\mathbf{w}')^n) = q$ and $T(q, (\mathbf{w}\mathbf{w}')^n) = id$ we know that the desired property holds for $\overline{\mathbf{w}}^q := \mathbf{w}'(\mathbf{w}\mathbf{w}')^{n-1}$. □

We define $M_q := \{n \in \mathbb{N} | \exists \mathbf{w} \in \Sigma^n, \delta(q, \mathbf{w}) = q, T(q, \mathbf{w}) = id\}$ and it follows directly by Lemma 2.3.2 that M_q is non-empty for all q.

Lemma 2.3.3. *Let A be a strongly connected automaton and $\mathcal{T}_A$ an induced transducer. There exists $d = d(\mathcal{T}_A) \in \mathbb{N}$ such that for all $q \in Q : M_q \subset d \cdot \mathbb{N}$ and $d \cdot \mathbb{N} \setminus M_q$ is finite.*

Proof. The first remark in this section shows directly that M_q is closed under addition for all $q \in Q$. Thus, it is easy to show that for every $q \in Q$ there exists d_q such that $M_q \subset d_q \cdot \mathbb{N}$ and $d_q \cdot \mathbb{N} \setminus M_q$ is finite.
It remains to show that all d_q coincide. Let $q_1, q_2 \in Q$. There exist $m \in \mathbb{N}$, $\mathbf{w}_1 \in \Sigma^m$ and $\mathbf{w}_2 \in \Sigma^{m+d_{q_1}}$ such that (2.1) holds for $\mathbf{w}_1, \mathbf{w}_2$ (for $q = q_1$). Furthermore, by Lemma 2.1.3 and Lemma 2.3.2, there exist $\mathbf{w}, \overline{\mathbf{w}}^{q_2}$ such that $\delta(q_2, \mathbf{w}) = q_1, \delta(q_1, \overline{\mathbf{w}}^{q_2}) = q_2$ and $T(q_2, \mathbf{w}\overline{\mathbf{w}}^{q_2}) = id$. Therefore, $\mathbf{w}\mathbf{w}_i\overline{\mathbf{w}}^{q_2}$ fulfill (2.1) for $i = 1, 2$ (and $q = q_2$). These two paths belonging to M_{q_2} have length $|\mathbf{w}| + |\overline{\mathbf{w}}^{q_2}| + m$ and $|\mathbf{w}| + |\overline{\mathbf{w}}^{q_2}| + m + d_{q_1}$ respectively. As $M_{q_2} \subseteq d_{q_2}\mathbb{N}$ we see that $d_{q_2} \leq d_{q_1}$ holds. Since q_1, q_2 are arbitrary elements of Q the statement follows. □

Remark. Changing the order of the tuples $q \in Q$, d does not change $d(\mathcal{T}_A)$ (since $1 \mapsto \sigma_q \circ id \circ \sigma_q^{-1} = id$). Thus we find that $d(\mathcal{T}_A)$ only depends on A and we write $d = d(A) = d(\mathcal{T}_A)$.

As mentioned before, we are interested in what elements of Δ occur for paths with certain length. One might assume that some periodic behavior (with period d) might occur when restricting to paths of certain length and, furthermore, that for long enough paths some closeness properties hold.

To find some precise statements, we define

$$M_{q\overline{q},\ell} := \{T(q, \mathbf{w}) | \mathbf{w} \in \Sigma^\ell, \delta(q, \mathbf{w}) = \overline{q}\}.$$

$M_{q\overline{q},\ell}$ are all the group elements that occur when reading a word of length ℓ that corresponds to a path from q to $\overline{q}$. Furthermore, we define $G_{q\overline{q}}(\ell) := \bigcup_{\ell' \equiv \ell(d)} M_{q\overline{q},\ell'}$. The following lemma shows that this union is actually a limit.

Lemma 2.3.4. *For all $q, \overline{q} \in Q$ and $0 \leq \ell \leq d-1$ it holds that, $G_{q\overline{q}}(\ell) = M_{q\overline{q},\ell+kd}$ for almost all $k \in \mathbb{N}$.*

Proof. It is sufficient to show that for each $g \in G_{q\overline{q}}(\ell)$ we have $g \in M_{q\overline{q},\ell+kd}$ for almost all $k \in \mathbb{N}$.
For $g \in G_{q\overline{q}}(\ell)$, there exists $\mathbf{w} \in \Sigma^{\ell+k_0 d}$ such that $\delta(q, \mathbf{w}) = \overline{q}$ and $T(q, \mathbf{w}) = g$. Lemma 2.3.3 implies that we have for almost all $k' \in \mathbb{N}$: There exists $\mathbf{w}_{k'} \in \Sigma^{k'd}$ which fulfills (2.1). Since $\mathbf{w}_{k'}\mathbf{w}$ is a path from q to $\overline{q}$ with weight g, we have for almost all $k \geq k_0$ (and therefore for almost all $k \in \mathbb{N}$), $g \in M_{q\overline{q},\ell+kd}$. □

We are interested in how the different $G_{q\overline{q}}(\ell)$ are related to each other. We see easily that for $q_1, q_2, q_3 \in Q$, $G_{q_1 q_2}(\ell_1) \cdot G_{q_2 q_3}(\ell_2) \subseteq G_{q_1 q_3}(\ell_1 + \ell_2)$ holds, but we actually even find equality:

Lemma 2.3.5. *For all $\ell_1, \ell_2 \in \mathbb{N}$ and $q_1, q_2, q_3 \in Q$ it holds that*

$$G_{q_1 q_3}(\ell_1 + \ell_2) = G_{q_1 q_2}(\ell_1) \cdot G_{q_2 q_3}(\ell_2).$$

Furthermore, $|G_{q_0 q_0}(0)| = |G_{q\overline{q}}(\ell)|$ holds for all $q, \overline{q} \in Q$ and $0 \leq \ell \leq d-1$.

Proof. We start by showing $G_{q\overline{q}}(\ell)^{-1} = G_{\overline{q}q}(-\ell)$ for all $q, \overline{q} \in Q, \ell \in \mathbb{N}$.
We know by Lemma 2.3.2 that for each $\mathbf{w}$ with $\delta(q, \mathbf{w}) = \overline{q}$, there exists $\overline{\mathbf{w}}^q$. Thus we conclude that $|\mathbf{w}| + |\overline{\mathbf{w}}^q| \equiv 0 \bmod d$ and, therefore, $G_{q\overline{q}}(\ell)^{-1} \subseteq G_{\overline{q}q}(-\ell)$. By inverting both sides, we also find $G_{q\overline{q}}(\ell) \subseteq G_{\overline{q}q}(-\ell)^{-1}$ and since $q, \overline{q}$ and ℓ are chosen arbitrarily, $G_{q\overline{q}}(\ell)^{-1} = G_{\overline{q}q}(-\ell)$ holds for all $q, \overline{q} \in Q, \ell \in \mathbb{N}$.

The inclusion $\supseteq$ is trivial. Let $g \in G_{q_1 q_3}(\ell_1 + \ell_2)$. By definition there exists $\mathbf{w}$ such that $\delta(q_1, \mathbf{w}) = q_3, T(q_1, \mathbf{w}) = g$ and $|\mathbf{w}| \equiv \ell_1 + \ell_2 \bmod d$. Lemma 2.1.3 shows the existence of $\mathbf{w}_1$ with $\delta(q_1, \mathbf{w}_1) = q_2$ and $|\mathbf{w}_1| \equiv \ell_1 \bmod d$. We find, by the first part of this proof, an inverse path $\overline{\mathbf{w}_1}^{q_1}$ from q_2 to q_1 such that $|\overline{\mathbf{w}_1}^{q_1}| \equiv -\ell_1 \bmod d$ holds. We see that $\mathbf{w}_2 := \overline{\mathbf{w}_1}^{q_1}\mathbf{w}$ is a path from q_2 to q_3 with $|\mathbf{w}_2| \equiv \ell_2 \bmod d$ and

$$T(q_1, \mathbf{w}_1\mathbf{w}_2) = T(q_1, \mathbf{w}_1) \cdot T(q_2, \mathbf{w}_2) = T(q_1, \mathbf{w}_1) \cdot T(q_2, \overline{\mathbf{w}_1}^{q_1}) \cdot T(q_1, \mathbf{w})$$

$$= T(q_1, \mathbf{w}_1\overline{\mathbf{w}_1}^{q_1}) \cdot g = g.$$

This finishes the proof of the first statement. We see by the first part of this lemma, that $G_{q_1q_3}(\ell) \supseteq G_{q_1q_2}(\ell') \cdot g_2$ and $G_{q_1q_3}(\ell) \supseteq g_1 \cdot G_{q_2q_3}(\ell')$ for $g_1 \in G_{q_1q_2}(\ell - \ell'), g_2 \in G_{q_2q_3}(\ell - \ell')$. Thus we find

$$|G_{q_1q_3}(\ell)| \geq |G_{q_1q_2}(\ell')|$$
$$|G_{q_1q_3}(\ell)| \geq |G_{q_2q_3}(\ell')|$$

for all $q_1, q_2, q_3 \in Q$ and $\ell, \ell' \in \mathbb{N}$. We can use this fact to show that

$$|G_{q_0q_0}(0)| \geq |G_{q_0q}(0)| \geq |G_{qq}(\ell)| \geq |G_{qq_0}(0)| \geq |G_{q_0q_0}(0)|.$$

Therefore we see that $|G_{qq}(\ell)| = |G_{q_0q_0}(0)|$ for all $q \in Q$ and $\ell \in \mathbb{N}$. To complete the proof we see that

$$|G_{q_0q_0}(0)| = |G_{qq}(0)| \geq |G_{q\overline{q}}(\ell)| \geq |G_{\overline{q}\overline{q}}(0)| = |G_{q_0q_0}(0)|.$$

□

Remark. Lemma 2.3.5 gives a strong structural results. Take, for example, any element $g_2 \in G_{q_2q_3}(\ell_2)$. As $|G_{q_1q_3}(\ell_1 + \ell_2)| = |G_{q_1q_2}(\ell_1)|$, it follows that $G_{q_1q_3}(\ell_1 + \ell_2) = G_{q_1q_2}(\ell_1) \cdot g_2$. A similar result holds for $g_1 \in \mathrm{G}_{q_1q_2}(\ell_1)$.

Corollary 2.3.6. *$G_{qq}(0)$ is a subgroup of G for all $q \in Q$.*

$G_{qq}(0)$ being a subgroup is the first example for the appearance of some closeness property occurring in this setting.
In this section, we also prove that all $G_{q\overline{q}}(\ell)$ are cosets of $G_{q_0q_0}(0)$, but first we mention a slightly stronger result than Lemma 2.3.4 which we need in the next section.

Lemma 2.3.7. *There exists $m_0 = m_0(A)$ such that for all $q, \overline{q} \in Q, 0 \leq \ell \leq d-1, g \in G_{q\overline{q}}(\ell)$ and $k \geq m_0$ exist $\mathbf{w}_1, \mathbf{w}_2 \in \Sigma^{\ell+kd}$ such that $\mathbf{w}_1 \neq \mathbf{w}_2$ and $\delta(q, \mathbf{w}_i) = \overline{q}, T(q, \mathbf{w}_i) = g$.*

Proof. Lemma 2.3.4 guarantees the existence of some m_1 such that for all $q, \overline{q} \in Q, 0 \leq \ell \leq d-1, g \in G_{q\overline{q}}(\ell)$ and $k \geq m_1$ exists $\mathbf{w} \in \Sigma^{\ell+kd}$ such that $\delta(q, \mathbf{w}) = \overline{q}, T(q, \mathbf{w}) = g$. We show that the desired property holds for $2m_1$ which obviously implies the statement.

- Case $|Q| \geq 2$:
  We fix $q_1 \neq q_2 \in Q^2$. A combination of Lemma 2.3.4 and Lemma 2.3.5 yields the existence of some paths $\mathbf{w}_1', \mathbf{w}_2'$ of length m_1d such that $\delta(q, \mathbf{w}_i') = q_i$ as well as paths $\mathbf{w}_1'', \mathbf{w}_2''$ of length $\ell + m_1d$ such that $\delta(q_i, \mathbf{w}_i'') = \overline{q}$ and $T(q, \mathbf{w}_i'\mathbf{w}_i'') = g$.
- Case $|Q| = 1, |G_{q_0q_0}(0)| \geq 2$:
  Thus we have $Q = \{q_0\}$. We fix $g_1 \neq g_2 \in G_{q_0q_0}(0)$. A combination of Lemma 2.3.4 and Lemma 2.3.5 yields the existence of some paths $\mathbf{w}_1', \mathbf{w}_2'$ of length m_1d and paths $\mathbf{w}_1'', \mathbf{w}_2''$ of length $\ell + m_1d$ such that $T(q_0, \mathbf{w}_i') = g_i$ and $T(q_0, \mathbf{w}_i'\mathbf{w}_i'') = g$.

- Case $|Q| = |G_{q_0q_0}(0)| = 1$:
  Obviously taking any two different $\mathbf{w}'_i \in \Sigma^{m_1 d}$ and $\mathbf{w} \in \Sigma^{\ell+m_1 d}$ such that $T(q_0, \mathbf{w}'_i) = id, T(q_0, \mathbf{w}) = g$, gives $T(q_0, \mathbf{w}'_i\mathbf{w}) = g$.

Thus we constructed always two different paths with the desired properties. □

We have seen in Proposition 2.2.3 that permuting the elements of Q again gives an induced transducer. We now show that by choosing the right permutations, we are able to make the actual structure of the induced transducer more apparent.

Lemma 2.3.8. *Let $\mathcal{T}_A$ be a naturally induced transducer. There exists a naturally induced transducer $\overline{\mathcal{T}}_A$ such that $id \in \overline{G}_{\overline{q}_0\overline{q}}(0)$ holds for each $\overline{q} \in \overline{Q}$.*

Proof. The idea is to find permutations σ_q such that the induced transducer defined in Proposition 2.2.3 has the desired properties. We take $\sigma_q \in G_{q_0q}(0)$ for all $q \neq q_0$ and $\sigma_{q_0} = id$. By Proposition 2.2.3, we find $\overline{G}_{\overline{q}_0\overline{q}}(0) = G_{q_0q} \cdot \sigma_q^{-1}$, which directly finishes the proof. □

We consider, from now on, only naturally induced transducers for which $id \in G_{q_0q}(0)$ holds for all $q \in Q$.

We define $G := G_{q_0q_0}(0)$ and find that the condition above is already sufficient to make $G_{q\overline{q}}(\ell)$ independent of $q, \overline{q}$, which makes the actual structure more apparent.

Proposition 2.3.9. *There exists $g_0 \in \Delta$ such that we have for all $q, \overline{q} \in Q, \ell \in \mathbb{N}$*

$$G_{q\overline{q}}(\ell) = G \cdot g_0^\ell = g_0^\ell \cdot G.$$

Remark. It follows easily that $d|ord(g_0)$ and for $d > 1, g_0 \notin G$.
We also find that $\Delta = G \cdot \{g_0^\ell : \ell \in \mathbb{N}\} = \{g_0^\ell : \ell \in \mathbb{N}\} \cdot G$, i.e. $G \lhd \Delta$.
This proposition shows together with Lemma 2.3.4 that $M_{q\overline{q},\ell}$ depends only on (ℓ mod d) for large ℓ.

Proof. As a first step, we prove that for all $q, \overline{q} \in Q$ we have $G_{q\overline{q}}(0) = G$:
We know by Lemma 2.3.5 that $G_{q_0q}(0) = G_{q_0q_0}(0) \cdot G_{q_0q}(0)$ and without loss of generality by Lemma 2.3.8, $id \in G_{q_0q_0}(0)$. Thus we have $G_{q_0q_0}(0) \cdot G_{q_0q}(0) \supseteq G \cdot id$ and, by comparing the cardinality, we see that $G_{q_0q}(0) = G$. Analogously, one finds that $G_{q_0\overline{q}}(0) = G$. Lemma 2.3.5 gives again $G_{q_0\overline{q}}(0) = G_{q_0q}(0) \cdot G_{q\overline{q}}(0)$ (i.e. $G = G \cdot G_{q\overline{q}}(0)$) and one concludes by a reasoning similar to above $G_{q\overline{q}}(0) = G$.

Now we consider $G_{q_0q_0}(1)$. We find by Lemma 2.3.5

$$G_{q_0q_0}(1) = G \cdot G_{q_0q_0}(1) = G_{q_0q_0}(1) \cdot G.$$

We now select an arbitrary element g_0 of $G_{q_0q_0}(1)$ and find

$$G_{q_0q_0}(1) = G \cdot g_0 = g_0 \cdot G.$$

One shows easily by induction that $G_{q_0q_0}(\ell) = G \cdot g_0^\ell = g_0^\ell \cdot G$ holds for all $\ell \in \mathbb{N}$.

It just remains to note that

$$G_{q\overline{q}}(\ell) = G_{qq_0}(0) \cdot G_{q_0q_0}(\ell) \cdot G_{q_0\overline{q}}(0) = G \cdot (g_0^\ell \cdot G) \cdot G = g_0^\ell \cdot G.$$

□

Thus we can prove now Theorem 2.3.1.

Proof of Theorem 2.3.1: We choose all the variables as defined throughout this section. The statement of this theorem follows easily by a combination of Lemma 2.3.4 and Proposition 2.3.9. □

Remark. For the period p of A, one finds easily that $p|d$. However, $p = d$ need not hold as the subsequent example shows.

Example. We continue our example for the tower of Hanoi problem. The corresponding automaton is of period 1 and we have already found the naturally induced transducer and we recall the functions δ, λ.

$$\begin{aligned} \delta(q_0,0) &= q_0 & \delta(q_0,1) &= q_1 \\ \delta(q_1,0) &= q_0 & \delta(q_1,1) &= q_0 \\ \lambda(q_0,0) &= (23) & \lambda(q_0,1) &= (13) \\ \lambda(q_1,0) &= (23) & \lambda(q_1,1) &= (13), \end{aligned}$$

where $q_0 = (a,b,c)$ and $q_1 = (\bar{a},\bar{b},\bar{c})$. One finds easily that $d = 2$ together with $G = \{id, (123), (132)\}$ and $g_0 \cdot G = \{(12), (23), (13)\}$.

2.4 Arithmetic restriction for naturally induced transducers

Theorem 2.3.1 shows that all elements of G occur when we restrict ourselves to paths whose length is divisible by d. We show in this section that – similar to Section 2.3 – restrictions on $[\mathbf{w}]_k \bmod (k^d - 1)$ lead to restrictions on what elements of G occur.
The main result of this section is the following theorem.

Theorem 2.4.1. *Under the same conditions as in Theorem 2.3.1, there exist natural numbers k_0, m_0', ℓ_0 and d' (where $\gcd(k,d') = 1$) together with a naturally induced transducer $\mathcal{T}_A$ fulfilling the properties of Theorem 2.3.1, a subgroup G_0 of G and $g_0' \in G$ fulfilling the following two conditions.*

- *For all $q, \overline{q} \in Q$ it holds that*

$$\{T(q,\mathbf{w}) : q \in Q, \mathbf{w} \in (\Sigma^{dk_0})^*, \delta(q,\mathbf{w}) = \overline{q}, [\mathbf{w}]_k \equiv \ell \bmod d'\} = G_0 \cdot g_0'^\ell = g_0'^\ell \cdot G_0.$$

- *For all $q, \overline{q} \in Q, m \geq m_0', g \in G_0$ exists $d''(q,\overline{q})|k^{\ell_0}$ fulfilling*

$$\gcd\{[\mathbf{w}]_k : \mathbf{w} \in \Sigma^{dk_0 m}, \delta(q,\mathbf{w}) = \overline{q}, T(q,\mathbf{w}) = g\} = d' \cdot d''(q,\overline{q}).$$

All of the variables only depend on A, but not on its initial state q_0'.

The rest of this section is again rather technical and only used to show this theorem.

We only consider G (i.e. weights corresponding to paths whose length is divisible by d) as just the same phenomena occur in the general situation. To assure that we are only working on these elements, we only consider paths whose length is divisible by d.

Lemma 2.3.7 assures, that we have for $\ell \geq m_0$

$$\forall q, \overline{q} \in Q \quad \forall g \in G \quad \exists \mathbf{w}_1, \mathbf{w}_2 \in \Sigma^{d\ell} : \mathbf{w}_1 \neq \mathbf{w}_2, \delta(q,\mathbf{w}_i) = \overline{q}, T(q,\mathbf{w}_i) = g.$$

Therefore, we define for $\ell \geq m_0, g \in G$

$$d_{g,\ell}^{q,\overline{q}} := \max\{m \in \mathbb{N} | \exists r : \forall \mathbf{w} \in \Sigma^{d\ell} \text{ such that } \delta(q,\mathbf{w}) = \overline{q} \text{ and } T(q,\mathbf{w}) = g \Rightarrow [\mathbf{w}]_k \equiv r \bmod m\}.$$

An equivalent definition for $d_{g,\ell}^{q\overline{q}}$ is the greatest common divisor of all differences of numbers corresponding to paths of length $d\ell$ from q to q' with weight g. We will use one of the two definitions depending on the situation. Our next goal is to show that $d_{g,\ell}^{q\overline{q}}$ converges to some $d(q,\overline{q})$ for all $g \in G$.

Lemma 2.4.2. *Let $q, \overline{q} \in Q$. There exists $d(q,\overline{q})$ such that for all $\ell \geq m_0, g \in \Delta$ we have $d(q,\overline{q})|d_{g,\ell}^{q\overline{q}}$ and there exists m_0' (not depending on $q,\overline{q}$) such that for all $\ell \geq m_0', g \in \Delta$ we have $d_{q,\ell}^{q\overline{q}} = d(q,\overline{q})$.*

Proof. We start by showing that for all $\ell, \ell' \geq m_0, q, \overline{q} \in Q$ and $g_1, g_2 \in G$ it holds that $d_{g_1,\ell+\ell'}^{q\overline{q}} | d_{g_2,\ell}^{q\overline{q}}$. Theorem 2.3.1 shows that there exists $\mathbf{w} \in \Sigma^{\ell'}$ such that $\delta(q,\mathbf{w}) = q$ and $T(q,\mathbf{w}) = g_1 \cdot g_2^{-1}$. We find

$$\begin{aligned} d_{g_1,\ell+\ell'}^{q\overline{q}} &= gcd(\{[\mathbf{w}_1]_k - [\mathbf{w}_2]_k : \mathbf{w}_i \in \Sigma^{d(\ell+\ell')} \text{ with } \delta(q,\mathbf{w}_i) = \overline{q}, T(q,\mathbf{w}_i) = g_1\}) \\ &\quad | \quad gcd(\{[\mathbf{w}\mathbf{w}_1]_k - [\mathbf{w}\mathbf{w}_2]_k : \mathbf{w}_i \in \Sigma^{d\ell} \text{ with } \delta(q,\mathbf{w}\mathbf{w}_i) = \overline{q}, T(q,\mathbf{w}\mathbf{w}_i) = g_1\}) \\ &= gcd(\{[\mathbf{w}_1]_k - [\mathbf{w}_2]_k : \mathbf{w}_i \in \Sigma^{d\ell} \text{ with } \delta(q,\mathbf{w}_i) = \overline{q}, T(q,\mathbf{w}_i) = g_2\}) \\ &= d_{g_2,\ell}^{q\overline{q}} \end{aligned}$$

The minimal value of $d_{g,\ell}^{q\overline{q}}$ for all ℓ is denoted by $d_g^{q\overline{q}}$ (we choose $\ell_0 \geq m_0$ such that $d_{q,\ell_0}^{q\overline{q}} = d_q^{q\overline{q}}$) and since

$$d_g^{q\overline{q}} = d_{g,\ell+\ell_0}^{q\overline{q}} \leq d_{q,\ell_0}^{q\overline{q}} = d_q^{q\overline{q}}$$

there exists m_0' such that for all $\ell \geq m_0'$ it holds that $d_{g,\ell}^{q\overline{q}} = d_g^{q\overline{q}}$. It also follows directly that $d_{g_1}^{q\overline{q}} = d_{g_2}^{q\overline{q}}$ for all $g_1, g_2 \in \Delta$. For $d(q\overline{q}) := d_{id}^{q\overline{q}}$ the result follows directly. □

Let $\mathbf{w} \in \Sigma^\ell$ and suppose $\delta(q, \mathbf{w}) = \overline{q}$, $T(q, \mathbf{w}) = g$, we then denote $r_\ell^{q\overline{q}}(g) := [\mathbf{w}]_k \bmod d(q, \overline{q})$ as we see directly that this definition does not depend on the choice of $\mathbf{w}$, but only on $q, \overline{q}$ and g. More generally, it follows that for $\mathbf{w} \in \Sigma^\ell, \delta(q, \mathbf{w}) = \overline{q}$ we have

$$r_\ell^{q\overline{q}}(T(q, \mathbf{w})) \equiv [\mathbf{w}]_k \bmod d(q, \overline{q}).$$

Remark. Since any word containing a synchronizing word is again synchronizing, and by assumption $\mathcal{T}_A$ is synchronizing and strongly connected, there exists a minimal $\ell_0 = \ell_0(A)$ such that for all $q \in Q$ exists $\mathbf{w}_q \in \Sigma^{\ell_0}$ with

$$\forall \overline{q} \in Q : \delta(\overline{q}, \mathbf{w}_q) = q.$$

We find an important restriction on $d(q, \overline{q})$.

Lemma 2.4.3. *We find for every $q, \overline{q} \in Q$*

$$k^{\ell_0}(k^d - 1) \equiv 0 \bmod d(q, \overline{q}).$$

Proof. We fix $q, \overline{q}$ and start by considering the case where we concatenate a word $\mathbf{w}$ with some words $\mathbf{w}_1, \mathbf{w}_2$ from left and right, respectively, such that the weight does not change: Let $\mathbf{w} \in \Sigma^{dm_2}, \mathbf{w}_1, \mathbf{w}_2 \in \Sigma^{dm_1}$ $(m_1 \geq m_0', m_2 \geq dm_0')$ such that $\delta(q, \mathbf{w}_1) = q, T(q, \mathbf{w}_1) = id$ and $\delta(\overline{q}, \mathbf{w}_2) = \overline{q}, T(\overline{q}, \mathbf{w}_2) = id$. We see that

$$\begin{aligned} r_{m_2+dm_1}^{q\overline{q}}(T(q, \mathbf{w}_1\mathbf{w})) &= r_{dm_1+m_2}^{q\overline{q}}(T(q, \mathbf{w}\mathbf{w}_2)) \\ [\mathbf{w}_1]_k k^{m_2} + [\mathbf{w}]_k &\equiv [\mathbf{w}]_k k^{dm_1} + [\mathbf{w}_2]_k \bmod d(q, \overline{q}) \\ [\mathbf{w}_1]_k k^{m_2} + [\mathbf{w}_2]_k &\equiv [\mathbf{w}]_k (k^{dm_1} - 1) \bmod d(q, \overline{q}). \end{aligned}$$

The left hand side of the last equation only depends on $\mathbf{w}_1, \mathbf{w}_2$ and m_2. By choosing $dm_2 \geq \ell_0 + 1$, $\mathbf{w} = 0\mathbf{w}_{\overline{q}}$ or $\mathbf{w} = 1\mathbf{w}_{\overline{q}}$ and – if $\mathbf{w}$ is not of sufficient length – adding zeros from the left we find

$$\begin{aligned} [\mathbf{w}_{\overline{q}}]_k (k^{dm_1} - 1) &\equiv (k^{\ell_0} + [\mathbf{w}_{\overline{q}}]_k)(k^{dm_1} - 1) \bmod d(q, \overline{q}) \\ k^{\ell_0}(k^{dm_1} - 1) &\equiv 0 \bmod d(q, \overline{q}). \end{aligned}$$

Comparing the results for m_1 and $m_1 + 1$ completes the proof. □

This allows us to decompose $d(q, \overline{q})$ into two co-prime factors $d'(q, \overline{q}), d''(q, \overline{q})$ such that $d'(q, \overline{q}) | (k^d - 1)$ and $d''(q, \overline{q}) | k^{\ell_0}$. On the one hand $d''(q, \overline{q})$ corresponds to restrictions on how you can reach the state $\overline{q}$, on the other hand $d'(q, \overline{q})$ corresponds to some restrictions $\bmod k^d - 1$, which is of greater interest for us.

Lemma 2.4.4. *For every $q, \overline{q} \in Q$ it holds that $d'(A) := d'(q_0, q_0) = d'(q, \overline{q})$.*

Proof. Since we have already gained some knowledge about $d(q, \overline{q})$ we can adopt Lemma 2.4.2 to our needs.

Let $\ell, \ell_1, \ell_2 \geq m_0'$, $q_1, q_2, \overline{q_1}, \overline{q_2} \in Q$ and paths $\mathbf{w} \in \Sigma^{d\ell_1}$ from q_1 to $\overline{q_1}$ with weight id, $\mathbf{w}' \in \Sigma^{d\ell_2}$ from $\overline{q_2}$ to q_2 with weight id. It follows, as in Lemma 2.4.2,

$$\begin{aligned} d(q_1,q_2) = d_{id,\ell_1+\ell+\ell_2}^{q_1,q_2} &= gcd(\{[\mathbf{w}_1]_k - [\mathbf{w}_2]_k : \mathbf{w}_i \in \Sigma^{d(\ell_1+\ell+\ell_2)} \text{ with } \delta(q_1,\mathbf{w}_i) = q_2, T(q_1,\mathbf{w}_i) = id\}) \\ &\mid \quad gcd(\{[\mathbf{w}\mathbf{w}_1\mathbf{w}']_k - [\mathbf{w}\mathbf{w}_2\mathbf{w}']_k : \\ &\qquad \mathbf{w}_i \in \Sigma^{d\ell} \text{ with } \delta(q_1,\mathbf{w}\mathbf{w}_i\mathbf{w}') = q_2, T(q_1,\mathbf{w}\mathbf{w}_i\mathbf{w}') = id\}) \\ &= gcd(\{k^{d\ell_2}([\mathbf{w}_1]_k - [\mathbf{w}_2]_k) : \mathbf{w}_i \in \Sigma^{d\ell} \text{ with } \delta(\overline{q_1},\mathbf{w}_i) = \overline{q_2}, T(\overline{q_1},\mathbf{w}_i) = id\}) \\ &= k^{d\ell_2} d_{1,\ell}^{\overline{q_1},\overline{q_2}} = k^{d\ell_2} d(\overline{q_1},\overline{q_2}) \end{aligned}$$

for arbitrary $q_1, q_2, \overline{q_1}, \overline{q_2}$. Thus we find $d'(q_1,q_2) | d'(\overline{q_1},\overline{q_2})$. As changing the order of some state $q \in Q$ does not change $d(\mathcal{T}_A)$ ($id \mapsto \sigma_q \circ id \circ \sigma_q^{-1}$), we find that d' only depends on A. □

As mentioned before, we are more interested in d' and define $s_\ell^{q\overline{q}}(g) := r_\ell^{q\overline{q}}(g) \bmod d'$. We find the following important properties:

Lemma 2.4.5. *For all $q_1, q_2, q_3 \in Q, m_1, m_2 \geq m_0'$, and $g_1, g_2 \in G$ follows*

$$s_{m_1+m_2}^{q_1,q_3}(g_1 \cdot g_2) = s_{m_1}^{q_1,q_2}(g_1) + s_{m_2}^{q_2,q_3}(g_2).$$

Proof. This follows directly from the fact that $k^{dm_2} \equiv 1 \bmod k^d - 1$. □

Lemma 2.4.6. *There exists $k_0 \in \mathbb{N}$ such that for every $\ell \geq m_0'$ we have $s_{d\ell}^{q_0q_0}(id) = 0 \Leftrightarrow \ell \equiv 0 \bmod k_0$.*

Proof. We find by Lemma 2.4.5 $s_{d(\ell_1+\ell_2)}^{q_0q_0}(id) = s_{d\ell_1}^{q_0q_0}(id) + s_{d\ell_2}^{q_0q_0}(id) \bmod d'$. The statement follows by the same arguments as we used in the proof of Lemma 2.3.3. □

Remark. One can actually prove that the Lemma above holds for all $\ell \in \mathbb{N}$ for which $s_{d\ell}^{q_0q_0}(id)$ is properly defined. Furthermore, $k_0 | d'$ holds.

We further follow the ideas of Section 2.3 and find the following result.

Lemma 2.4.7. *There exists a naturally induced transducer $\overline{\mathcal{T}}_A$ such that $\overline{s}_{dk_0m_0'}^{\overline{q_0}\overline{q}}(id) = 0$ holds for all $\overline{q} \in \overline{Q}$.*

Proof. We want to find some permutations σ_q such that applying them gives an induced transducer with the desired properties. Let $q \neq q_0$, take $\mathbf{w} \in \Sigma^{dk_0m_0}$ such that $\mathbf{w} \equiv 0 \bmod d', \delta(q_0,\mathbf{w}) = q$ and define $\sigma_q = T(q_0,\mathbf{w})$. W.l.o.g. we restrict ourselves to $\mathbf{w} = \mathbf{w}'\mathbf{w}_q$, where $\mathbf{w}_q$ is again a synchronizing word – we can easily choose a suitable $\mathbf{w}'$.)
By choosing $\sigma_{q_0} = id$ we find by using Proposition 2.2.3 that $\overline{T}(\overline{q_0},\mathbf{w}) = T(q_0,\mathbf{w}) \cdot T(q_0,\mathbf{w})^{-1} = id$.
Note that applying these permutations does not change the property $id \in G_{q_0q}(0)$ used in Section 2.3 □

We consider from now on only naturally induced transducers such that $s^{q_0q}_{k_0m_0'}(id) = 0$ holds for all $q \in Q$.
We find a similar result to Section 2.3 concerning $M_{q\overline{q},\ell}$, which shows that $s^{q\overline{q}}_{d\ell}(g)$ only depends on $\ell \bmod k_0$ and g.

Proposition 2.4.8. *For all $q, \overline{q} \in Q, g \in G$ and $\ell \in \mathbb{N}, \ell \geq m_0'$ we have*

$$s^{q\overline{q}}_{\ell}(g) = s_{\ell \bmod k_0}(g),$$

where $s_\ell(g) := s_{\ell \bmod k_0}(g) := s^{q_0q_0}_{k_0m_0'+(\ell \bmod k_0)}(g)$.

Proof. We find by Lemma 2.4.5 that

$$0 = s^{q_0q_0}_{2k_0m_0'}(id) = s^{q_0q}_{k_0m_0'}(id) + s^{qq_0}_{k_0m_0'}(id) = s^{qq_0}_{k_0m_0'}(id).$$

Let us now assume $\ell \geq 5k_0m_0'$ and we compute

$$\begin{aligned} s^{q\overline{q}}_{\ell}(g) &= s^{qq_0}_{k_0m_0}k^{\ell-k_0m_0}(id) + s^{q_0q_0}_{\ell-2k_0m_0}(g)k^{k_0m_0} + s^{q_0\overline{q}}_{k_0m_0}(id) \\ &= s^{q_0q_0}_{\ell-2k_0m_0}(g) = s^{q_0q_0}_{k_0m_0+(\ell \bmod k_0)}(g) + s^{q_0q_0}_{\ell-(\ell \bmod k_0)-3k_0m_0}(id) = s_\ell(g). \end{aligned}$$

Let now $\ell \geq m_0'$. We find

$$s^{q\overline{q}}_{\ell}(g) = s^{q\overline{q}}_{\ell}(g) + 5s^{\overline{q}\overline{q}}_{\ell}(id) = s^{q\overline{q}}_{\ell+5k_0m_0'}(g) = s_{\ell \bmod k_0}(g),$$

which finishes the proof. □

We define $G_\ell := \{g \in G : s_0(g) = \ell\}$.

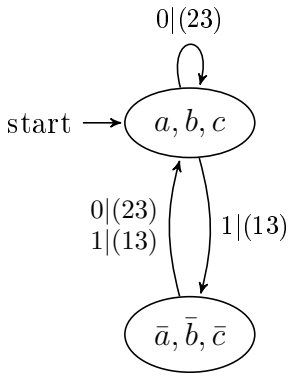
Corollary 2.4.9. *G_0 is a subgroup of G.*

Proof. The statement follows directly by Lemma 2.4.5. □

Thus we can prove now Theorem 2.4.1.

Proof of Theorem 2.4.1: Of course we choose k_0, m_0', ℓ_0, d and G_0 as we defined them throughout this section. To prove the first part we choose g_0' as an arbitrary element of G_1 and the proof follows by the same arguments as in the proof of Theorem 2.3.1. The second part is just the result of Lemma 2.4.4. □

Example. We continue our example for the tower of Hanoi problem. The naturally induced transducer we considered was



We have already seen that $d = 2$ holds. Now we derive d', G_0, g_0'. We find easily that $\mathbf{w}_{q_0} = 0$ is a synchronizing word. We see, that $T(q, \mathbf{w})$ does not depend on q, which simplifies the treatment. Furthermore, it is easy to prove that for all $\mathbf{w} \in (\Sigma^2)^* \cong \{0,1,2,3\}^*$

$$T(q, \mathbf{w}) = (123)^{[\mathbf{w}]_2}.$$

We compute

$$\delta(q_0, 00) = \delta(q_0, 11) = q_0$$
$$T(q_0, 00) = T(q_0, 11) = id.$$

Consequently, we find $d' = 3, G_0 = \{id\}, g_0' = (123), k_0 = 1$ and $d''(q_0, q_0) = 1$. We also mention that $d''(., q_0) = 1$ and $d''(., q_1) = 2$ without proof.

2.5 Reduction to special naturally induced Transducers

We want to show how to reduce the general setting of an automatic sequence to automata with special properties.

Proposition 2.5.1. *Let $\mathbf{a} = (a_n)_{n\in\mathbb{N}}$ be a k-automatic sequence. There exists $p \in \mathbb{N}$ and $A = (Q', \{0, \dots, k^p - 1\}, \delta', q_0')$ such that $\mathbf{a}$ is generated by A and for every Automaton A_i, that is the restriction to a final component [2] of A, it holds that $d(A_i) = k_0(A_i) = 1$.*

Proof. Let $\overline{A} = (\overline{Q}, \{0, \dots, k-1\}, \overline{\delta}, \overline{q_0})$ be an automaton that generates $\mathbf{a}$. We can assume without loss of generality that $\overline{\delta}(\overline{q_0}, 0) = \overline{q_0}$ and every state $\overline{q} \in \overline{Q}$ is reachable from $\overline{q_0}$. Let $\overline{Q}_1, \dots, \overline{Q}_\ell$ be the final components of $\overline{A}$. We denote by $\overline{A}_i$ the automaton that corresponds to the restriction of $\overline{A}$ to $\overline{Q}_i$, with arbitrary initial state.
We define $p = \operatorname{lcm}(d(\overline{A}_1)k_0(\overline{A}_1), \dots, d(\overline{A}_\ell)k_0(\overline{A}_\ell))$ and note that $d(\overline{A}_i)$ and $k_0(\overline{A}_i)$ do not depend on the initial state of A_i – compare Theorem 2.3.1 and Theorem 2.4.1. The idea is now to take the "p-th power" of $\overline{A}$ which gives the desired property.

Therefore we define $A = (\overline{Q}, \{0, \dots, k^p - 1\}, \delta', \overline{q_0})$ where δ' is the extension of $\overline{\delta}$ to $\{0, \dots, k-1\}^p \cong \{0, \dots, k^p - 1\}$ from letters to words of length p. We see that A generates $\mathbf{a}$ as $\overline{\delta}(\overline{q_0}, 0) = \overline{q_0}$ ensures that adding leading zeros does not change the output (of $\overline{A}$).

One observes easily that every final component of A is contained in a final component of $\overline{A}$: The strongly connected components of a directed graph form a new directed acyclic graph where every node corresponds to a strongly connected component. There exists an edge between two such nodes if and only if there is an edge between two states of these strongly connected components. Every final component of A is part of a final component of $\overline{A}$, since there exists a path (whose length is divisible by p) from every strongly connected component, which is not closed under $\overline{\delta}$ to a final component. Therefore, we can restrict ourselves to consider only a final component Q_i'.

It remains to show $d(A_i) = k_0(A_i) = 1$. For this purpose, we want to describe how a naturally induced transducer of A_i looks like. Let $\mathcal{T}_{\overline{A}_j} = (Q, \{0, \dots, k-1\}, \delta, q_0, \Delta, \lambda)$ be a naturally

[2] A final component is a strongly connected component that is closed under δ'.

induced transducer of $\overline{A}_j$. As $d(\overline{A}_j)k_0(\overline{A}_j)|p$ we know by Theorem 2.3.1 that for all $q_1, q_2 \in Q$ and $g \in G$ there exists $\mathbf{w} \in \{0, \ldots, k^p - 1\}^*$ such that $\delta(q_1, \mathbf{w}) = q_2$ and $T(q_1, \mathbf{w}) = g$. This means, in particular, that there exists $\mathbf{w} \in \{0, \ldots, k^p - 1\}^*$ such that $\delta(q_1, \mathbf{w}) = q_2$ and $T(q_1, \mathbf{w}) = id$. Thus Proposition 2.2.2 shows that we can describe the states of A_i as the elements corresponding to some coordinates of $\mathcal{T}_{\overline{A}_j}$: $\pi_i(q_1)$ and $\pi_i(q_2)$ belong to the same strongly connected component for all $q_1, q_2 \in Q$ as $\mathbf{w}$ satisfies $\delta'(\pi_i(q_1), \mathbf{w}) = \pi_i(q_2)$. This implies that

$$Q'_i = \{\pi_\ell(q) | q \in Q, \ell \in I\},$$

holds for some index set $I \subseteq \{1, \ldots, n_0\}$. We claim that $\mathcal{T} := (Q_I, \Sigma^p, \widetilde{\delta}_I, (q_0)_I, G_I, \widetilde{\lambda}_I)$ provides a naturally induced transducer of A_i. Here we again denote by $\widetilde{\delta}$ the extension of δ from letters to words, $\widetilde{\lambda}$ coincides with T for words of length p (for $\mathcal{T}_A$) and G is defined as in Theorem 2.3.1. Furthermore, we use $(.)_I$ to denote the projection to the coordinates of I. Theorem 2.3.1 assures that $\mathcal{T}_{\overline{A}}$ is an induced transducer. To see that $\mathcal{T}_{\overline{A}}$ is synchronizing we just have to take a synchronizing word $\mathbf{w}_{q_0}$ of $\mathcal{T}_A$ and add leading zeros such that it is a word whose length is divisible by p.
However, it is possible that property 6) may not hold, but we already discussed in Section 2.1 that this can be fixed easily.
By the construction of $\mathcal{T}$ we see easily that $d(A_i) = k_0(A_i) = 1$ holds. □

This Proposition can be simplified under suitable conditions.

Corollary 2.5.2. *Let $\overline{A} = (Q', \Sigma, \delta', q'_0)$ be a strongly connected automaton, such that $\delta'(q'_0, 0) = q'_0$. There exists a strongly connected automaton A such that $d(A) = k_0(A) = 1$ and the automatic sequences generated by A and $\overline{A}$ coincide.*

Proof. We consider the proof of Proposition 2.5.1 and see directly that A – as considered in the proof of Proposition 2.5.1 – has only one strongly connected component. It just remains to note that the automatic sequences generated of A and $\overline{A}$ coincide since adding leading zeros does not change the output. □

Chapter 3

Automatic Sequences fulfill the Sarnak Conjecture

3.1 The Sarnak Conjecture

The Möbius function is defined by

$$\mu(n) = \begin{cases} (-1)^k & \text{if } n \text{ is square-free and} \\ & k \text{ is the number of prime factors} \\ 0 & \text{otherwise.} \end{cases}$$

We say that a sequence $\mathbf{a}$ is orthogonal to $\mu(n)$ if

$$\left| \sum_{n<N} a_n \mu(n) \right| = o\left(\sum_{n<N} |a_n| \right).^1$$

There exists an old - relatively vague - principle (the Möbius Randomness Principle, see for example [28, p. 338]), which states that every "reasonable" bounded complex sequence is orthogonal to the Möbius function. The reasoning behind this principle is that the Möbius function changes signs so randomly that it induces sufficient cancellation. However, it is unclear which sequences are "reasonable". One approach is that "simple" sequences should be "reasonable". "Simple" could be interpreted by the computational complexity of the sequence, but Peter Sarnak proposes another notion of "simpleness" by dynamical systems. A dynamical system (or flow) F is a pair (X, S) where X is a compact metric space and $S : X \to X$ is a continuous map. We call a dynamical system *deterministic* if its topological entropy is zero, which will be the notion of "simpleness" that we use. We say a sequence $(\xi(n))_{n\in\mathbb{N}}$ is realized by a dynamical system (X, S) if there exists a start-point x_0 and a continuous map $f : X \to \mathbb{C}$ such that $\xi(n) = f(S^n(x_0))$ for all $n \in \mathbb{N}$.

In 2009, Peter Sarnak stated the following conjecture [40]:

[1] Here we use the little-o notation. Let $f, g : \mathbb{N} \to \mathbb{R}^+$. We say that $f = o(g)$ if and only if $\lim_{n\to\infty} f(n)/g(n) = 0$.

Conjecture 3.1.1. *Let μ be the Möbius function. For any bounded sequence $\xi(n)$ realized by a deterministic dynamical system* [2] *(X, S), it holds that*

$$\sum_{n\leq N} \xi(n)\mu(n) = o(N).$$

This conjecture has aroused great interest and numerous papers were recently devoted to show the Sarnak conjecture for different classes of dynamical systems [5, 6, 7, 10, 12, 16, 17, 20, 24, 25, 33, 35, 44].

Sarnak stated his conjecture in terms of a deterministic dynamical system (X, S). We are interested in special dynamical systems that origin from a sequence:

Fix a sequence $\mathbf{a} = (a_n)_{n\in\mathbb{N}}$ that takes values in a finite set $\mathcal{A}$. We denote by S the shift operator on the sequences $(\mathcal{A}^{\mathbb{N}})$ in $\mathcal{A}$ and define $X := \overline{\{S^n(\mathbf{a}) : n \in \mathbb{N}_0\}}$. By taking the product topology of the discrete topologies on each copy of $\mathcal{A}$, we find that $\mathcal{A}^{\mathbb{N}}$ is compact and complete. The product topology is induced by the following metric $d(x, y) = \sum_{n=0}^{\infty} 2^{-n-1} d_n(x_n, y_n)$ on X, where d_n denotes the discrete metric on $\mathcal{A}$. The dynamical system (or flow) (X, S) is called the *symbolic dynamical system associated with* $\mathbf{a}$.

Therefore, we say that a sequence $\mathbf{a}$ fulfills the Sarnak conjecture if the symbolic dynamical system associated with $\mathbf{a}$ fulfills the Sarnak conjecture. For more information about symbolic dynamical systems and their complexity see for example [18, 21].

It is easy to compute the topological entropy of a symbolic dynamical system associated to a sequence $\mathbf{a} \in \mathcal{A}^{\mathbb{N}}$. We find (see for example [34]) that the topological entropy is equal to $\lim_{n\to\infty} \frac{p_{\mathbf{a}}(n)}{n}$, where the subword complexity of $\mathbf{a}$ was defined in Section 1.4.3.

A Möbius-randomness principle is often closely related to a Prime Number Theorem (PNT), i.e. an asymptotic formula for the sum $\sum_{n\leq x} \Lambda(n)u_n$, where Λ denotes the von Mangoldt function. That is $\Lambda(n) = \log p$ if $n = p^\ell$ for some $p \in \mathbb{P}$ and $\ell \in \mathbb{N}$ and $\Lambda(n) = 0$ otherwise. An estimate of $\sum_{n\leq x} \Lambda(n)u_n$ can be used to estimate

$$\sum_{\substack{p\leq x \\ p\in\mathbb{P}}} u_p$$

by partial summation. For example by using the following identities one can hope to relate a Möbius-randomness principle to a Prime Number Theorem

$$\Lambda(n) = \sum_{d|n} \mu(d) \log(\frac{n}{d})$$
$$1 = \sum_{d|n} \mu(d)\tau(\frac{n}{d})$$

where τ denotes the divisor function.

Some classes or examples of automatic sequences, which are observed by a deterministic dynamical system, have already been covered:

[2] A dynamical system is a pair (X, S) where X is a compact metric space and $S : X \to X$ is a continuous map.

- The Thue-Morse sequence:
  - A Möbius-randomness principle follows from the work of Indlekofer and Kátai [27, 30]
  - Dartyge and Tenenbaum [9] give additionally an explicit error bound
  - A Möbius-randomness principle and a PNT by Mauduit and Rivat [37]
- The Rudin-Shapiro sequence:
  - Tao suggests a strategy to prove a Möbius-randomness principle in [42]
  - Mauduit and Rivat provide a different approach which allows to prove a Möbius-randomness principle and a PNT for more general functions [38]
- Sequences generated by invertible automata
  - Sarnak conjecture [19]
  - A Möbius-randomness principle and a PNT [13]
- Sequences generated by synchronizing automata
  - The Sarnak conjecture and a PNT [11]
- Some special digital functions (e.g. the parity of occurrences of a word in the digital expansion)
  - A Möbius-randomness principle together with a PNT [26]

The purpose of this chapter is to prove the Sarnak conjecture for all automatic sequences, which covers all results concerning automatic sequences mentioned above.

Theorem 3.1.2. *Let μ be the Möbius function, $(a_n)_{n\in\mathbb{N}}$ be an automatic sequence and let (X, S) be the symbolic dynamical system associated with $(a_n)_{n\in\mathbb{N}}$. Then for all sequences $\xi(n) := f(S^n(x))$, with $x \in X$ and $f \in C(X, \mathbb{C})$, we have*

$$\sum_{n\leq N} \xi(n)\mu(n) = o(N).$$

3.2 Reduction of Theorem 3.1.2

We reduce in this section the main Theorem 3.1.2 of this chapter to statements concerning naturally induced transducers.

First we present a lemma that reduces the Sarnak conjecture for the symbolic dynamical system associated with an automatic sequence to a Möbius randomness law for (possibly different) automatic sequences.

Lemma 3.2.1. *Suppose that for every automatic sequence $(a_n)_{n\in\mathbb{N}}$ with values in $\mathbb{C}$*

$$\sum_{n\leq N} \mu(n)a_{n+r} = o(N), \tag{3.1}$$

uniformly for $r \in \mathbb{N}$. Then Theorem 3.1.2 holds.

Proof. Let us fix one automatic sequence $\mathbf{b} = (b_n)_{n\in\mathbb{N}}$ that takes values in $\mathbb{C}$. Let (X, S) be the symbolic dynamical system associated with $\mathbf{b}$. [11, Lemma 2.1] shows that it is sufficient to show that for every $j \geq 1$ and every $g : \mathbb{C}^j \to \mathbb{C}$ it holds that

$$\sum_{n\leq N} \mu(n)g(b_{n+r}, b_{n+r+1}, \ldots, b_{n+r+j-1}) = o(N)$$

uniformly for $r \in \mathbb{N}$. However, a combination of Theorem 1.2.5,Theorem 1.1.7 and Theorem 1.1.6 shows that $a_n = g(b_n, b_{n+1}, \ldots, b_{n+j-1})$ defines again an automatic sequence and hence the statement follows. □

To further reduce this result we need some ideas of representation theory and a method to detect certain digits of $(n)_k$. We cover the most important definitions and notations briefly. A representation D is a continuous homomorphism $D : G \to U_d$, where U_d denotes the group of unitary $d \times d$ matrices over $\mathbb{C}$. A representation D is called irreducible if there exists no non-trivial subspace $V \subset U_d$ such that $D(g)V \subseteq V$ holds for all $g \in G$.
For $\alpha \in \mathbb{R}$ with $0 \leq \alpha < 1$ we denote by χ_α the characteristic function of the interval $[0, \alpha)$ modulo 1:

$$\chi_\alpha(x) = \lfloor x \rfloor - \lfloor x - \alpha \rfloor .$$

One finds that

$$\chi_{k^\mu}\left(\frac{n - mk^{\lambda-\mu}}{k^\lambda}\right) = 1$$

if and only if $(\varepsilon_{\lambda-1}(n), \ldots, \varepsilon_{\lambda-\mu}(n))$ and $(m)_k$ coincide up to leading zeros.

In Section 3.3, we show the following result.

Proposition 3.2.2. *Let $\mathcal{T}_A$ be a naturally induced transducer by A and suppose that $d(A) = 1, k_0(A) = 1$ holds. Let, furthermore, D be a unitary irreducible representation of G, $r, \lambda_1, \lambda_2 \in \mathbb{N}, 0 < b < k^{\lambda_1}$ and $m < k^{\lambda_2}$. It holds*

$$\left\| \sum_{\substack{n<N \\ n\equiv m \bmod k^{\lambda_2}}} \chi_{1/k^{\lambda_1}}\left(\frac{n + r - bk^{\nu-\lambda_1}}{k^\nu}\right) D(T(q_0, (n + r)_k))\mu(n) \right\|_F = o(N), \tag{3.2}$$

uniformly for r where ν is the unique integer satisfying $k^{\nu-1} \leq N < k^\nu$.

Remark. $n \equiv m \bmod k^{\lambda_2}$ obviously fixes the last λ_2 digits of n. Furthermore, $\chi_{1/k^{\lambda_1}}(.)$ fixes the digits $\nu - \lambda_1, \ldots, \nu - 1$.

Our goal of this section is to show that Proposition 3.2.2 implies Theorem 3.1.2.

We fix any $\varepsilon > 0$ and need to show that there exists N_0 such that for $N \geq N_0(\varepsilon)$

$$|S(r, N)| \leq \varepsilon N$$

holds for all r, where

$$S(r, N) := \sum_{n<N} \mu(n) a_{n+r}.$$

Let us first present the most important ideas for this proof. Obviously, we can only work with naturally induced transducers, when we restrict ourselves to the final components of A. Therefore we fix some of the first digits of $(n + r)$ which allows us to work with a strongly connected component of Q'. Thereafter, we also fix the last digits of $(n+r)$ to fix $\delta(q, (n+r)_k)$. It then remains to find suitable estimates for $T(q, .)$.

More precisely, Proposition 2.5.1 shows that there exists a DFAO $A = (Q', \Sigma, \delta', q'_0, \tau)$ such that $\Sigma = \{0, \ldots, k-1\}, \delta'(q'_0, 0) = q'_0$ and $a_n = \tau(\delta'(q'_0, (n)_k))$. Let $Q'_1, \ldots, Q'_\ell$ be the strongly connected components of Q' that are closed under δ'. Furthermore, let A_i be the restriction of A to Q'_i. Proposition 2.5.1 states that $d(A_i) = k_0(A_i) = 1$ holds.

We define the naturally induced transducers for these components by

$$\mathcal{T}_{A_i} = (Q_i, \Sigma, \delta_i, \Delta_i, \lambda_i)$$

and let $\mathbf{w}_i$ be a synchronizing word of $\mathcal{T}_{A_i}$ (we intentionally avoid to fix the initial state here).

Lemma 3.2.3. *Suppose M is the set of integers n such that there exist $\mathbf{v}_1, \mathbf{v}_2 \in \Sigma^*$ satisfying*

- $(n)_k = \mathbf{v}_1\mathbf{v}_2$
- *For all $q' \in Q' : \delta'(q', \mathbf{v}_1) \in \cup Q'_i$*
- *For all $i : \mathbf{w}_i$ is a subword of $\mathbf{v}_2$, i.e. $\mathbf{v}_2$ is a synchronizing word for all A_i.*

Then M has density one.

The motivation for the set M is that whenever you read a word corresponding to some $n \in M$, you end up in one final component Q'_i – this is achieved by $\mathbf{v}_1$. The purpose of $\mathbf{v}_2$ is not so obvious but will be more apparent shortly.

Remark. Let $n \in M$. By the definition of M it follows that the considered properties hold also for unreduced digital representations of n (i.e. with leading zeros).

As the proof of Lemma 3.2.3 is rather technical, we postpone it to Section 3.7.

We assume for simplicity that $|a_n| \leq 1$ holds for all $n \in \mathbb{N}$. We denote by ν the unique integer satisfying $k^{\nu-1} \leq N < k^\nu$.

The first idea is to remove the digits of $n+r$ of positions $\nu, \nu+1, \dots$ and restrict ourselves to $(n+r) \bmod k^\nu$ as these integers form (at most) two intervals with combined length at least $k^{\nu-1}$. Therefore, we want to detect the digits of $n+r$ at the positions $\nu-\lambda_1, \dots, \nu-1$ for some fixed value λ_1 – only depending on ε – that we choose shortly:

$$\begin{aligned}\left|\sum_{n<N} \mu(n)a_{n+r}\right| &= \left|\sum_{n<N}\sum_{b<k^{\lambda_1}} \chi_{k^{-\lambda_1}}\left(\frac{n+r-bk^{\nu-\lambda_1}}{k^\nu}\right)\mu(n)a_{n+r}\right| \\ &\le \sum_{b<k^{\lambda_1}}\left|\sum_{n<N} \chi_{k^{-\lambda_1}}\left(\frac{n+r-bk^{\nu-\lambda_1}}{k^\nu}\right)\mu(n)a_{n+r}\right| \\ &\le \sum_{\substack{b<k^{\lambda_1}\\ b\in M}}\left|\sum_{n<N} \chi_{k^{-\lambda_1}}\left(\frac{n+r-bk^{\nu-\lambda_1}}{k^\nu}\right)\mu(n)a_{n+r}\right| + \sum_{\substack{b<k^{\lambda_1}\\ b\notin M}} k^{\nu-\lambda_1}.\end{aligned}$$

Choosing λ_1 such that $\left|\{b<k^{\lambda_1} | b\notin M\}\right| \le \frac{\varepsilon}{3k}k^{\lambda_1}$ gives

$$|S(r,N)| \le \sum_{\substack{b<k^{\lambda_1}\\ b\in M}} |S_1(b,r,N)| + \frac{\varepsilon}{3}N$$

where

$$S_1 := S_1(b,r,N) := \sum_{n<N} \chi_{k^{-\lambda_1}}\left(\frac{n+r-bk^{\nu-\lambda_1}}{k^\nu}\right)\mu(n)a_{n+r}.$$

We rewrite $r = r_1 + k^\nu r_2$ with $r_1 < k^\nu$ and $r_1, r_2 \in \mathbb{N}$, and split the sum over n into two parts. We denote by $(n)_k^t$ the unique word $\mathbf{w}$ of length t such that $[\mathbf{w}]_k \equiv n \bmod k^t$ and find

$$\begin{aligned}|S_1| \le &\left|\sum_{\substack{n<N\\ n+r_1<k^\nu}} \chi_{k^{-\lambda_1}}\left(\frac{n+r-bk^{\nu-\lambda_1}}{k^\nu}\right)\mu(n)\tau(\delta'(q_0',(n+r)_k))\right| \\ &+\left|\sum_{k^\nu-r_1\le n<N} \chi_{k^{-\lambda_1}}\left(\frac{n+r-bk^{\nu-\lambda_1}}{k^\nu}\right)\mu(n)\tau(\delta'(q_0',(n+r)_k))\right| \\ =&\left|\sum_{\substack{n<N\\ n<k^\nu-r_1}} \chi_{k^{-\lambda_1}}\left(\frac{n+r-bk^{\nu-\lambda_1}}{k^\nu}\right)\mu(n)\tau(\delta'(q_0',(r_2)_k(b)_k^{\lambda_1}(n+r)_k^{\nu-\lambda_1}))\right| \\ &+\left|\sum_{k^\nu-r_1\le n<N} \chi_{k^{-\lambda_1}}\left(\frac{n+r-bk^{\nu-\lambda_1}}{k^\nu}\right)\mu(n)\tau(\delta'(q_0',(r_2+1)_k(b)_k^{\lambda_1}(n+r)_k^{\nu-\lambda_1}))\right|.\end{aligned}$$

Thus we find

$$\begin{aligned}|S_1| \le &|S_2(b,r_1,\min(N,k^\nu-r_1),\delta'(q_0',(r_2)_k))| + |S_2(b,r_1,N,\delta'(q_0',(r_2+1)_k))| \\ &+|S_2(b,r_1,k^\nu-r_1,\delta'(q_0',(r_2+1)_k))|,\end{aligned} \tag{3.3}$$

where

$$S_2(b,r',N',q') := \sum_{n<N'} \chi_{k^{-\lambda_1}}\left(\frac{n+r'-bk^{\nu-\lambda_1}}{k^\nu}\right)\mu(n)\tau(\delta'(q',(b)_k^{\lambda_1}(n+r')_k^{\nu-\lambda_1}))$$

for $N'+r'<k^\nu$.

We simplify (3.3) to

$$|S_1(b,r,N)| \le 3 \max_{\substack{r',N'\in\mathbb{N}\\ r'+N'<k^\nu}} \max_{q'\in\cup Q_i'} |S_2(b,r',N',q')| \tag{3.4}$$

We work now on an estimate for S_2. We can rewrite $(b)_k^{\lambda_1} = \mathbf{v}_1\mathbf{v}_2$ where $\delta'(q',\mathbf{v}_1)\in Q_i'$ for some i.

We denote by $q_1' = \delta'(q_0',\mathbf{v}_1)$ and $A_{q_1'} = (Q_i',\Sigma,\delta_i',\delta'(q',\mathbf{v}_1),\tau_i)$ the restriction of A to Q_i', i.e. $\delta_i' = \delta_i|_{Q_i'\times\Sigma}, \tau_i = \tau|_{Q_i'}$). Let $\mathcal{T}_{A_{q_1'}} = (Q,\Sigma,\delta,q_0,\Delta,\lambda)$ be a naturally induced transducer of $A_{q'}$. We define M_{λ_2} as the set of synchronizing words of length λ_2 of $\mathcal{T}_{A_{q_1'}}$ and note that for $\mathbf{w}_0\in M_{\lambda_2}, \mathbf{w}\in\Sigma^*$ it holds that $\delta(q_0,\mathbf{w}\mathbf{w}_0) = \delta(q_0,\mathbf{w}_0)$. Lemma 1.4.1 shows that there exist $c,\eta>0$ such that $|(\Sigma^{\lambda_2}\setminus M_{\lambda_2})| \le ck^{(1-\eta)\lambda_2}$ holds for all $\lambda_2\in\mathbb{N}$. Proposition 2.2.2 shows that

$$\begin{aligned}
|S_2(b,r',N',q')| &= \Bigg|\sum_{n<N'} \chi_{k^{-\lambda_1}}\left(\frac{n+r'-bk^{\nu-\lambda_1}}{k^\nu}\right)\mu(n)\\
&\qquad \tau(\pi_1(T(q_0,\mathbf{v}_2(n+r')_k^{\nu-\lambda_1})\delta(q_0,\mathbf{v}_2(n+r')_k^{\nu-\lambda_1})))\Bigg|\\
&= \Bigg|\sum_{m<k^{\lambda_2}} \sum_{\substack{n<N'\\ n+r'\equiv m \bmod k^{\lambda_2}}} \chi_{k^{-\lambda_1}}\left(\frac{n+r'-bk^{\nu-\lambda_1}}{k^\nu}\right)\mu(n)\\
&\qquad \tau(\pi_1(T(q_0,\mathbf{v}_2(n+r')_k^{\nu-\lambda_1})\delta(q_0,\mathbf{v}_2(n+r')_k^{\nu-\lambda_1})))\Bigg|\\
&\le \sum_{m\in M_{\lambda_2}} \Bigg| \sum_{\substack{n<N'\\ n\equiv m-r' \bmod k^{\lambda_2}}} \chi_{k^{-\lambda_1}}\left(\frac{n+r'-bk^{\nu-\lambda_1}}{k^\nu}\right)\mu(n)\\
&\qquad \tau(\pi_1(T(q_0,\mathbf{v}_2(n+r')_k^{\nu-\lambda_1})\delta(q_0,(m)_k^{\lambda_2})))\Bigg| + ck^{(1-\eta)\lambda_2}k^{\nu-\lambda_1-\lambda_2}.
\end{aligned}$$

We choose λ_2 such that $ck^{-\eta\lambda_2} \le \frac{\varepsilon}{9k}$ and find

$$|S_2(b,r',N',q')| \le \sum_{m\le k^{\lambda_2}} \left|S_3(b,r',N',m,\mathcal{T}_{A_{q_1'}})\right| + \frac{\varepsilon N}{9}k^{-\lambda_1},$$

where

$$S_3(b,r',N',m,\mathcal{T}_{A_{q_1'}}) := \sum_{\substack{n<N'\\ n\equiv m \bmod k^{\lambda_2}}} \chi_{k^{-\lambda_1}}\left(\frac{n+r'-bk^{\nu-\lambda_1}}{k^\nu}\right)\mu(n)$$

$$\tau(\pi_1(T(q_0, \mathbf{v}_2(n+r')_k^{\nu-\lambda_1})\delta(q_0,(m)_k^{\lambda_2}))).$$

Note that λ_1 and λ_2 are independent of N and N'.

We define $q_1 := \delta(q_0,(b)_k)$ and find $T(q_0,(b)_k(n+r')_k^{\nu-\lambda_1}) = T(q_0,(b)_k)\circ T(q_1,(n+r')_k^{\nu-\lambda_1})$ as well as $T(q_0,\mathbf{v}_2(n+r')_k^{\nu-\lambda_1}) = T(q_0,\mathbf{v}_2)\circ T(q_1,(n+r')_k^{\nu-\lambda_1})$, since $\mathbf{v}_2$ is synchronizing. This gives in total

$$T(q_0,\mathbf{v}_2(n+r')_k^{\nu-\lambda_1}) = T(q_0,\mathbf{v}_2)\circ T(q_0,(b)_k)^{-1}\circ T(q_1,(b)_k(n+r')_k^{\nu-\lambda_1}).$$

We define for $\mathbf{w}\in\Sigma^*$ a function $f_{\mathbf{w},b}: G\to\mathbb{C}$

$$f_{\mathbf{w},b}(\sigma) := \tau(\pi_1(T(q_0,\mathbf{v}_2)\circ(T(q_0,(b)_k)^{-1}\circ\sigma)\cdot(\delta(q_0,\mathbf{w})))).$$

Thus we find $\tau(\pi_1(T(q_0,\mathbf{v}_2(n+r')_k^{\nu-\lambda_1})\cdot\delta(q_0,(m)_k^{\lambda_2}))) = f_{(m)_k^{\lambda_2},b}(T(q_0,(b)_k(n+r')_k^{\nu-\lambda_1}))$. We denote by $\mathcal{F}(\mathcal{T}_{A_{q_1'}}) := \{f_{\mathbf{w},b} : \mathbf{w}\in\Sigma^*, b\in\mathbb{N}\}$. Note that $\left|\mathcal{F}(\mathcal{T}_{A_{q_1'}})\right| \le |Q|\cdot|G|$ holds, as $\delta(q_0,\mathbf{w})$ can take at most $|Q|$ and $T(q_0,\mathbf{v}_2)\circ T(q_0,(b)_k)^{-1}$ can take at most $|G|$ values. We find

$$S_3(b,r',N',m,\mathcal{T}_{A_{q_1'}}) = \sum_{\substack{n<N'\\ n\equiv m \bmod k^{\lambda_2}}} \chi_{k^{-\lambda_1}}\left(\frac{n+r'-bk^{\nu-\lambda_1}}{k^\nu}\right)\mu(n) f_{(m)_k^{\lambda_2},b}(T(q_0,(n+r')_k))$$

With $T(n) := T(q_0,(n)_k)$ we find

$$\left|S_3(b,r',N',m,\mathcal{T}_{A_{q'}})\right| \le \max_{f\in\mathcal{F}(\mathcal{T}_{A_{q'}})}\left|\sum_{\substack{n<N'\\ n\equiv m \bmod k^{\lambda_2}}} \chi_{k^{-\lambda_1}}\left(\frac{n+r'-bk^{\nu-\lambda_1}}{k^\nu}\right)\mu(n) f(T(n+r'))\right|$$

Consider a finite group G. It is well-known that there only exist finitely many equivalence classes of unitary and irreducible representations of G (see for example [41, Part I, Section 2.5]).

The Peter-Weyl Theorem (see for example [32, Chapter 4, Theorem 1.2]) states that the entry functions of irreducible representations (suitable re-normalized) form an orthonormal basis of $L^2(G)$. Thus we can express any function $f: G\to\mathbb{C}$ by these M_0 entry functions:

Let $D^{(m')} = (d_{ij}^{(m')})_{ij}$ be representations of the equivalence classes mentioned above and $f: G\to\mathbb{C}$. Then, there exist c_ℓ such that

$$f(g) = \sum_{\ell<M_0} c_\ell d_{i_\ell j_\ell}^{(m_\ell)}(g)$$

holds for all $g\in G$. Thus we find

$$|S_3| \le \max_{D\in\mathcal{D}_{q_1'}} c\left\|\sum_{\substack{n<N'\\ n\equiv m \bmod k^{\lambda_2}}} \chi_{k^{-\lambda_1}}\left(\frac{n+r'-bk^{\nu-\lambda_1}}{k^\nu}\right)\mu(n) D(T(n+r'))\right\|_F,$$

where $\|.\|_F$ denotes the Frobenius norm, c is the maximum of all possible values of $c_\ell \cdot M_0$ and $\mathcal{D}_{q_1'}$ is the finite set of all possible appearing representations for $\mathcal{F}(\mathcal{T}_{A_{q_1'}})$.

If we denote by $T_{q_1'}$ the function T for $\mathcal{T}_{A_{q_1'}}$ we find in total

$$\left|\sum_{n<N}\mu(n)a_{n+r}\right| \leq \frac{2}{3}\varepsilon N + c\sum_{b<k^{\lambda_1}}\sup_{\substack{r',N'\in\mathbb{N}\\ N'+r'<k^\nu}}\sum_{m<k^{\lambda_2}} \sum_{q'\in\cup Q_i'}\sum_{D\in\mathcal{D}_{q'}}\left\|\sum_{\substack{n<N'\\ n\equiv m \bmod k^{\lambda_2}}}\chi_{k^{-\lambda_1}}\left(\frac{n+r'-bk^{\nu-\lambda_1}}{k^\nu}\right)\mu(n)D(T_{q'}(n+r'))\right\|_F.$$

Note that $\mathcal{D}_{q_1'}$ is finite and only depends on the automatic sequence $\mathbf{a}$.

$$|S(r,N)| \leq \frac{2}{3}\varepsilon N + c\sum_{b<k^{\lambda_1}}\sum_{m<k^{\lambda_2}}\sum_{q_1'\in\cup Q_i'}\sum_{D\in\mathcal{D}_{q_1'}} \sup_{r'\in\mathbb{N}}\max_{N'<k^\nu}\left\|\sum_{\substack{n<N'\\ n\equiv m \bmod k^{\lambda_2}}}\chi_{k^{-\lambda_1}}\left(\frac{n+r'-bk^{\nu-\lambda_1}}{k^\nu}\right)\mu(n)D(T_{q_1'}(n+r'))\right\|_F,$$

where λ_1, λ_2 and $\mathcal{D}_{q_1'}$ only depend on ε. This finally proves the following proposition, which recapitulates the results of this section.

Proposition 3.2.4. *Assume that Proposition 3.2.2 holds. Then Theorem 3.1.2 follows.*

3.3 A general Möbius Principle

The goal of this section is to establish the setup to show Proposition 3.2.2 . Therefore, we generalize the setup of Mauduit and Rivat which they used to study the Rudin–Shapiro sequence in [38] - we already briefly mentioned this result in Section 1.4.2. We also need that the Discrete Fourier Transform of the sequence is uniformly small and some sort of quantitative statement about the carry propagation. In Section 3.4 we show that for all but very special representations we have a uniformly small Discrete Fourier Transform for $D(T(.))$. In Section 3.5 we show that carry propagations happen rarely. We leave the full proofs of the generalized results of [38] to Section 3.8, as they are technical and very similar to the original proofs.

We consider a function $f : \mathbb{N} \to U_d$ where U_d denotes the set of unitary $d \times d$ matrices. Let $k \in \mathbb{N}$. We denote by f_λ the k^λ-periodic function defined by

$$\forall n \in \{0, \dots, k^\lambda - 1\}, \quad \forall m \in \mathbb{Z}, f_\lambda(n + mk^\lambda) = f(n).$$

Furthermore, we define $f_{\mu,\lambda}(n) := f_\lambda(n)f_\mu(n)^H$ for $\mu \leq \lambda$, where A^H denotes the conjugate transpose of the matrix A.

It is necessary to use matrix valued functions instead of complex valued functions as in [38], as we are working with representations.

Definition 3.3.1. A function $f : \mathbb{N} \to U_d$ has the carry property if there exists $\eta > 0$ such that uniformly for $(\lambda, \alpha, \rho) \in \mathbb{N}^3$ with $\rho < \lambda$, the number of integers $0 \leq \ell < k^\lambda$ such that there exists $(n_1, n_2) \in \{0, \dots, k^\alpha - 1\}^2$ with

$$f(\ell k^\alpha + n_1 + n_2)^H f(\ell k^\alpha + n_1) \neq f_{\alpha+\rho}(\ell k^\alpha + n_1 + n_2)^H f_{\alpha+\rho}(\ell k^\alpha + n_1) \tag{3.5}$$

is at most $O(k^{\lambda-\eta\rho})$ where the implied constant may depend only on k and f.

Remark. One can obviously exchange (3.5) with

$$f(\ell k^\alpha + n_1)^H f(\ell k^\alpha + n_1 + n_2) \neq f_{\alpha+\rho}(\ell k^\alpha + n_1)^H f_{\alpha+\rho}(\ell k^\alpha + n_1 + n_2).$$

We introduce a set of functions with uniformly small Discrete Fourier Transforms as in [38]:

Definition 3.3.2. Given a non decreasing function $\gamma : \mathbb{R} \to \mathbb{R}$ satisfying $\lim_{\lambda\to\infty} \gamma(\lambda) = +\infty$ and $c > 0$ we denote by $F_{\gamma,c}$ the set of functions $f : \mathbb{N} \to U_d$ such that for $(\alpha, \lambda) \in \mathbb{N}^2$ with $\alpha \leq c\lambda$ and $t \in \mathbb{R}$:

$$\left\| k^{-\lambda} \sum_{u<k^\lambda} f(uk^\alpha)\, \mathrm{e}(-ut) \right\|_F \leq k^{-\gamma(\lambda)}. \tag{3.6}$$

One obvious difference to [38] is that we consider matrix valued functions. Moreover, the original definition of the carry property requires $\eta = 1$. Nevertheless, we find similar results as in [38] in this more general setting. The results of [38] have already been generalized to matrix valued functions in [13] and to a weaker carry property (although still more restrictive than Definition 3.3.1) in [26]. We discuss the proof for the following theorems in Section 3.8.

Theorem 3.3.3. *Let $\gamma : \mathbb{R} \to \mathbb{R}$ be a non decreasing function satisfying $\lim_{\lambda\to\infty} \gamma(\lambda) = +\infty$, and $f : \mathbb{N} \to U_d$ be a function satisfying Definition 3.3.1 for some $\eta \in (0, 1]$ and $f \in \mathcal{F}_{\gamma,c}$ for some $c \geq 10$ in Definition 3.3.2. Then for any $\theta \in \mathbb{R}$ we have*

$$\left\| \sum_{n\leq x} \Lambda(n) f(n)\, \mathrm{e}(\theta n) \right\| \ll c_1(k)(\log x)^{c_2(k)} x k^{-\eta\gamma(2\lfloor(\log x)/(80\log k)\rfloor/20}, \tag{3.7}$$

with the same constants as in [38].

Theorem 3.3.4. *Let $\gamma : \mathbb{R} \to \mathbb{R}$ be a non decreasing function satisfying $\lim_{\lambda\to\infty} \gamma(\lambda) = +\infty$, and $f : \mathbb{N} \to U_d$ be a function satisfying Definition 3.3.1 for some $\eta \in (0, 1]$ and $f \in \mathcal{F}_{\gamma,c}$ for some $c \geq 10$ in Definition 3.3.2. Then for any $\theta \in \mathbb{R}$ we have*

$$\left\| \sum_{n\leq x} \mu(n) f(n)\, \mathrm{e}(\theta n) \right\| \ll c_1(k)(\log x)^{\frac{9}{4}+\frac{1}{4}\max(\omega(k),2)} x k^{-\eta\gamma(2\lfloor(\log x)/(80\log k)\rfloor/20}, \tag{3.8}$$

with the same constants as in [38].

We now want to show that Definition 3.3.2 holds for $f(n) = D(T(n))$ for almost all unitary irreducible representations D where $\gamma(\lambda) = \eta'\lambda - c'$ for some $\eta' > 0$ and $c' \in \mathbb{R}$.

3.4 Fourier Estimates

One of the most difficult parts of this approach is to find sufficient bounds for the Fourier terms. The proof is rather technical. However, it justifies some results of Chapter 2.

Let $\mathcal{T}_A$ be a naturally induced transducer of A and suppose that $d(A) = 1, k_0(A) = 1$ holds. We distinguish at this point some special representations, for which we need a different approach.

Lemma 3.4.1. *There exist d' special 1-dimensional representations $D_0, \ldots, D_{d'-1}$ defined by*

$$D_\ell(g) := \mathrm{e}\left(\frac{\ell \cdot s_0(g)}{d'}\right),$$

for $\ell = 0, \ldots d' - 1$, where s_0 is defined by $s_0(T(q, \mathbf{w})) = [\mathbf{w}]_k \bmod d'$.

Proof. The proof follows directly by Theorem 2.4.1 and Theorem 2.3.1 □

Note that $D_\ell(T(q_0, \mathbf{w})) = \mathrm{e}\left((\ell \cdot [\mathbf{w}]_k)/d'\right)$ for $\ell = 0, \ldots, d' - 1, \mathbf{w} \in \Sigma^*$.

We fix $\mathcal{T}_A$ and try to find unitary, irreducible representations D of G for which we have exponentially decreasing bounds for expressions similar to

$$\left\| \frac{1}{k^\lambda} \sum_{u<k^\lambda} D(T(q_0, (u)_k)) \, \mathrm{e}(-ut) \right\|_2,$$

uniformly in t.

We already see that $D = D_\ell$ is a special case for which we are not able to find exponentially decreasing bounds as

$$\frac{1}{k^\lambda} \sum_{u<k^\lambda} D_\ell(T(q_0, u)) \, \mathrm{e}(-ut) = \frac{1}{k^\lambda} \sum_{u<k^\lambda} \mathrm{e}\left(u\left(\frac{\ell}{d'} - t\right)\right)$$

gives 1 for $t = \ell/d'$, which is the trivial bound. Nevertheless, we are able to find exponentially decreasing bounds for all other unitary and irreducible representations of G:

Theorem 3.4.2. *Let D be a unitary and irreducible representation of G different from $D_0, \ldots, D_{d'-1}$. There exists $\eta > 0$ such that*

$$\left\| \frac{1}{k^\lambda} \sum_{u<k^\lambda} D(T(q, (uk^\alpha + r)_k)) \, \mathrm{e}(-ut) \right\|_2 \ll k^{-\eta\lambda} \tag{3.9}$$

holds uniformly for $t \in \mathbb{R}$, $q \in Q, r \in \mathbb{N}$ and $\alpha \in \mathbb{N}$.

The proof is carried out throughout this section. We define

$$\phi^q_{\lambda,\alpha}(t, r) := \frac{1}{k^\lambda} \sum_{u<k^\lambda} D(T(q, (uk^\alpha + r)_k)) \, \mathrm{e}(-ut)$$

and for $r < k^\alpha$

$$\psi^q_{\lambda,\alpha}(t,r) := \frac{1}{k^\lambda} \sum_{u<k^\lambda} D(T(q,(u)^\lambda_k (r)^\alpha_k)) \,\mathrm{e}(-ut).$$

We see that these two definitions look very similar, but it turns out to be much easier to deal with ψ.

Lemma 3.4.3. *Let D be a unitary representation of G. Assume that there exists $\eta > 0$ such that*

$$\left\|\psi^q_{\lambda,\alpha}(t,r)\right\|_2 \ll k^{-\eta\lambda}$$

holds uniformly for $q \in Q, \lambda, \alpha \in \mathbb{N}, t \in \mathbb{R}$ and $r < k^\lambda$. Then Theorem 3.4.2 holds for D.

Proof. We postpone this proof to Section 3.7. □

Therefore, it is sufficient to prove the following proposition.

Proposition 3.4.4. *Let D be a unitary and irreducible representation of G different from $D_0, \dots, D_{d'-1}$. There exists $\eta > 0$ such that*

$$\frac{1}{k^\lambda} \left\| \sum_{u<k^\lambda} D(T(q,(u)^\lambda_k (r)^\alpha_k)) \,\mathrm{e}(-ut) \right\|_2 \ll k^{-\eta\lambda}$$

holds uniformly for $q \in Q, \lambda, \alpha \in \mathbb{N}, t \in \mathbb{R}$ and $r < k^\alpha$.

We use the recursive structure of transducers to find recurrences for these Fourier terms.

Lemma 3.4.5. *Let $q \in Q$, $\lambda, \alpha \in \mathbb{N}, t \in \mathbb{R}$ and $r < k^\alpha$. It holds for all $m < \lambda$*

$$\psi^q_{\lambda,\alpha}(t,r) = \frac{1}{k^m} \sum_{\varepsilon<k^m} D(T(q,(\varepsilon)^m_k)) \,\mathrm{e}(-\varepsilon(k^{\lambda-m}t)) \psi^{\delta(q,\varepsilon)}_{\lambda-m,\alpha}(t,r). \tag{3.10}$$

Proof. We know that $T(q,(\varepsilon)^m_k (u')^{\lambda-m}_k (r)^\alpha_k) = T(q,(\varepsilon)^m_k) \cdot T(\delta(q,(\varepsilon)^m_k),(u')^{\lambda-m}_k (r)^\alpha_k)$ for $u' < k^{\lambda-m}$. By distinguishing the m most significant digits of u, we find

$$\begin{aligned}
\psi^q_{\lambda,\alpha}(t,r) &= \frac{1}{k^\lambda} \sum_{\varepsilon<k^m} \sum_{u'<k^{\lambda-m}} D(T(q,(\varepsilon)^m_k (u')^{\lambda-m}_k (r)^\alpha_k)) \,\mathrm{e}(-(\varepsilon k^{\lambda-m} + u')t) \\
&= \frac{1}{k^m} \sum_{\varepsilon<k^m} D(T(q,(\varepsilon)^m_k)) \,\mathrm{e}(-\varepsilon(k^{\lambda-m}t)) \\
&\quad \frac{1}{k^{\lambda-m}} \sum_{u'<k^{\lambda-m}} D(T(\delta(q,(\varepsilon)^m_k),(u')^{\lambda-m}_k (r)^m_k)) \,\mathrm{e}(-u't).
\end{aligned}$$

□

The main idea is to find $m \in \mathbb{N}, \varepsilon_1, \varepsilon_2 < k^m$ such that two or more terms on the right hand side of (3.10) cancel – at least partially. This means that we want to find $\varepsilon_1, \varepsilon_2$ such that $\delta(q, \varepsilon_1) = \delta(q, \varepsilon_2)$ and

$$\left\| D(T(q,\varepsilon_1))\, \mathrm{e}(-\varepsilon_1 k^{\lambda-m} t) + D(T(q,\varepsilon_2))\, \mathrm{e}(-\varepsilon_2 k^{\lambda-m} t) \right\|_2 \leq 2 - \xi,$$

for some $\xi > 0$. We split the proof into two parts – depending on the dimension of D.

Lemma 3.4.6. *Let D be a unitary and irreducible representation of G of dimension at least 2. Then Proposition 3.4.4 holds for D.*

Proof. We need to show that there exists $\eta > 0$ such that Proposition 3.4.4 holds. Theorem 2.3.1 shows that there exists for every $g \in G$ a path $\mathbf{w}_g \in \Sigma^{m_0}$ such that $\delta(q, \mathbf{w}_g) = q, T(q, \mathbf{w}_g) = g$ holds. We find using Lemma 3.4.5 that

$$\begin{aligned}
\left\|\psi^q_{\lambda,\alpha}(t,r)\right\|_2 &= \frac{1}{k^{m_0}} \left\| \sum_{\varepsilon < k^{m_0}} D(T(q,(\varepsilon)_k^{m_0}))\, \mathrm{e}(-\varepsilon(k^{\lambda-m_0}t)) \psi^{\delta(q,(\varepsilon)_k^{m_0})}_{\lambda-m_0,\alpha}(t,r) \right\|_2 \\
&\leq \frac{1}{k^{m_0}} \left\| \sum_{g\in G} D(T(q,\mathbf{w}_g))\, \mathrm{e}(-[\mathbf{w}_g]_k k^{\lambda-m_0}t) \psi^q_{\lambda-m_0,\alpha}(t,r) \right\|_2 \\
&\quad + \frac{1}{k^{m_0}} \sum_{\substack{\varepsilon < k^{m_0} \\ \varepsilon \neq [\mathbf{w}_g]_k \forall g \in G}} \left\| D(T(q,(\varepsilon)_k^{m_0}))\, \mathrm{e}(-\varepsilon(k^{\lambda-m_0}t)) \psi^{\delta(q,(\varepsilon)_k^{m_0})}_{\lambda-m_0,\alpha}(t,r) \right\|_2 \\
&\leq \frac{1}{k^{m_0}} \left\| \sum_{g\in G} D(T(q,\mathbf{w}_g))\, \mathrm{e}(-[\mathbf{w}_g]_k k^{\lambda-m_0}t) \right\|_2 \cdot \left\|\psi^q_{\lambda-m_0,\alpha}(t,r)\right\|_2 \\
&\quad + \frac{k^{m_0} - |G|}{k^{m_0}} \max_{q'\in Q} \left\|\psi^{q'}_{\lambda-m_0,\alpha}(t,r)\right\|_2 \\
&\leq \frac{1}{k^{m_0}} \left(k^{m_0} - |G| + \left\| \sum_{g\in G} D(g)\, \mathrm{e}(-[\mathbf{w}_g]_k k^{\lambda-m_0}t) \right\|_2 \right) \max_{q'\in Q} \left\|\psi^{q'}_{\lambda-m_0,\alpha}(t,r)\right\|_2 .
\end{aligned}$$

Thus it is sufficient to show that for all $t' \in \mathbb{R}$ holds

$$\left\| \sum_{g\in G} D(g)\, \mathrm{e}(-[\mathbf{w}_g]_k t') \right\|_2 < |G|. \tag{3.11}$$

Since the left hand side is a periodic and continuous function in t', this implies that there exists $\eta' > 0$ such that

$$\left\| \sum_{g\in G} D(g)\, \mathrm{e}(-[\mathbf{w}_g]_k t') \right\|_2 \leq |G| - \eta'$$

for all $t' \in \mathbb{R}$. This gives in total

$$\left\|\psi^q_{\lambda,\alpha}(t,r)\right\|_2 \leq \left(1 - \frac{\eta'}{k^{m_0+\ell_0}}\right) \max_{q'\in Q} \left\|\psi^{q'}_{\lambda-m_0-\ell_0,\alpha}(t,r)\right\|_2$$

and the statement follows easily as η' only depends on G and D (but not on λ, α, t' or r). What follows is a variation of the proof of [15, Lemma 4]:

Let us assume – on the contrary – that there exists $t' \in \mathbb{R}$ such that

$$\left\| \sum_{g \in G} D(g)\, \mathrm{e}(-[\mathbf{w}_g]_k t') \right\|_2 = |G|.$$

This holds if and only if there exists $\mathbf{0} \neq \mathbf{y} \in \mathbb{C}^d$ such that

$$\left\| \sum_{g \in G} D(g)\, \mathrm{e}(-[\mathbf{w}_g]_k t')\mathbf{y} \right\|_2^2 = \sum_{g_1, g_2 \in G} \langle D(g_1)\, \mathrm{e}(-[\mathbf{w}_{g_1}]_k t')\mathbf{y}, D(g_2)\, \mathrm{e}(-[\mathbf{w}_{g_2}]_k t')\mathbf{y} \rangle = |G|^2 \, \|\mathbf{y}\|_2^2$$

However, the Cauchy-Schwarz inequality implies

$$\begin{aligned} &|\langle D(g_1)\, \mathrm{e}(-[\mathbf{w}_{g_1}]_k t')\mathbf{y}, D(g_2)\, \mathrm{e}(-[\mathbf{w}_{g_2}]_k t')\mathbf{y} \rangle| \\ &\qquad \leq \|D(g_1)\, \mathrm{e}(-[\mathbf{w}_{g_1}]_k t')\mathbf{y}\|_2^2 \, \|D(g_2)\, \mathrm{e}(-[\mathbf{w}_{g_2}]_k t')\mathbf{y}\|_2^2 = \|\mathbf{y}\|_2^2 . \end{aligned} \tag{3.12}$$

For equality to hold in (3.12) it is necessary that the $D(g_i)\, \mathrm{e}(-[\mathbf{w}_{g_i}]_k t')$ are linearly dependent. Since we find for $g_1 = g_2 = id$ the summand $\langle \mathrm{e}(-[\mathbf{w}_{id}]_k t')\mathbf{y}, \mathrm{e}(-[\mathbf{w}_{id}]_k t')\mathbf{y} \rangle$ we obtain for all $g \in G$

$$D(g)\, \mathrm{e}(-[\mathbf{w}_g]_k t')\mathbf{y} = \mathrm{e}(-[\mathbf{w}_{id}]_k t')\mathbf{y},$$

i.e. $\mathbf{y}$ is an eigen-vector of all $D(g), g \in G$. We define $W = span(\mathbf{y})$ and find $D(g)W \subseteq W$ for all $g \in G$. This means that D would be reducible which yields a contradiction.

Therefore (3.11) holds, which concludes this proof. □

Lemma 3.4.7. *Let D be a one-dimensional, unitary and irreducible representation of G different from $D_0, \ldots, D_{d'-1}$. Then Proposition 3.4.4 holds for D.*

Proof. Our goal is to show that there exists some $\eta' > 0$ (only depending on G and D) such that

$$\left\| \psi^q_{\lambda,\alpha}(t,r) \right\|_2 \leq (1-\eta') \max_{\widetilde{q} \in Q} \left\| \psi^{\widetilde{q}}_{\lambda - m'_0 - \ell_0, \alpha}(t,r) \right\|_2 \tag{3.13}$$

holds for all $t \in \mathbb{R}$. This implies again Proposition 3.4.4 for some $\eta > 0$ as in the proof of the previous case.

We need two different estimates for this step and start to work on the first estimate.
We start by using the properties of G, i.e. $\{g \in G : s_0(g) = \ell\} = G_0 \cdot (g'_0)^\ell$, to find restrictions for D.
Assume that $D(g) = 1$ for all $g \in G_0$. It follows that $D(g'_0) = \mathrm{e}(\ell/d')$ for some $\ell < d'$ since $(g'_0)^{d'} \in G_0$ and therefore we see directly that $D = D_\ell$. Thus we know that there exists $g \in G_0$ such that $D(g) \neq 1$.
We use Lemma 2.3.7 to see that there exist $\mathbf{w}_{id}, \mathbf{w}_g \in \Sigma^{m'_0}$ such that $\delta(q, \mathbf{w}_{id}) = \delta(q, \mathbf{w}_g) = q$

and $T(q,\mathbf{w}_{id})=id, T(q,\mathbf{w}_g)=g$ holds. By using Lemma 3.4.5 (with $m=m_0'$), we find (as in the proof for representations of dimension ≥ 2)

$$\begin{aligned}\left\|\psi^q_{\lambda,\alpha}(t,r)\right\|_2 \leq \frac{1}{k^{m_0'}} \Big\| & D(T(q,\mathbf{w}_{id}))\,\mathrm{e}([\mathbf{w}_{id}]_k k^{\lambda-m_0'}t)\psi^{\delta(q,\mathbf{w}_{id})}_{\lambda-m_0',\alpha}(t,r) \\ & + D(T(q,\mathbf{w}_g))\,\mathrm{e}([\mathbf{w}_g]_k k^{\lambda-m_0'}t)\psi^{\delta(q,\mathbf{w}_g)}_{\lambda-m_0',\alpha}(t,r)\Big\|_2 \\ & + \frac{k^{m_0'}-2}{k^{m_0'}} \max_{q'\in Q}\left\|\psi^{q'}_{\lambda-m_0',\alpha}(t,r)\right\|_2.\end{aligned}$$

We see that the first term of the right hand side equals

$$\begin{aligned}&\frac{1}{k^{m_0'}}\left\|1\cdot\mathrm{e}([\mathbf{w}_{id}]_k k^{\lambda-m_0'}t)+D(g)\cdot\mathrm{e}([\mathbf{w}_g]_k k^{\lambda-m_0'}t)\right\|_2\left\|\psi^q_{\lambda-m_0',\alpha}(t,r)\right\|_2 \\ &\quad=\frac{1}{k^{m_0'}}\left\|1+D(g)\,\mathrm{e}(([\mathbf{w}_g]_k-[\mathbf{w}_{id}]_k)k^{\lambda-m_0'}t)\right\|_2\left\|\psi^q_{\lambda-m_0',\alpha}(t,r)\right\|_2.\end{aligned}$$

Now we use Lemma 3.4.5 again with $m=\ell_0$ to find in total

Thus we find the first estimate:

$$\left\|\psi^q_{\lambda,\alpha}(t,r)\right\|_2 \leq \frac{k^{m_0'}-2+\left|1+D(g)\,\mathrm{e}(([\mathbf{w}_g]_k-[\mathbf{w}_{id}]_k)k^{\lambda-m_0'}t)\right|}{k^{m_0'}}\max_{\widetilde{q}\in Q}\left\|\psi^{\widetilde{q}}_{\lambda-\ell_0-m_0',\alpha}(t,r)\right\|_2. \quad (3.14)$$

To find the second estimate, we use Lemma 3.4.5 with $m=\ell_0$ and find

$$\left\|\psi^q_{\lambda,\alpha}(t,r)\right\|_2 \leq \max_{\widetilde{q}\in Q}\left\|\psi^{\widetilde{q}}_{\lambda-\ell_0,\alpha}(t,r)\right\|_2.$$

For convenience we assume that the maximal value of the right hand side is attained at q. Theorem 2.4.1 shows $d'\cdot d''(q,q)=\gcd\{[\mathbf{w}_1]_k-[\mathbf{w}_2]_k : \mathbf{w}_i\in\Sigma^{m_0'} \text{ with } \delta(q,\mathbf{w}_i)=q, T(q,\mathbf{w}_i)=id\}$. Therefore, there exist $N\in\mathbb{N}, \mu_{i,j}\in\mathbb{Z}^N\times\mathbb{Z}^N$ such that $d'\cdot d''(q,q)=\sum_{i,j<N}\mu_{i,j}([\mathbf{w}_i]_k-[\mathbf{w}_j]_k)$ where $\mathbf{w}_i,\mathbf{w}_j\in\Sigma^{m_0'}$ such that $\delta(q,\mathbf{w}_i)=\delta(q,\mathbf{w}_j)=q$ and $T(q,\mathbf{w}_i)=T(q,\mathbf{w}_j)=id$. By Lemma 3.4.5 and the same arguments we used above, we find

$$\begin{aligned}\left\|\psi^q_{\lambda-\ell_0,\alpha}(t,r)\right\|_2 &\leq \frac{1}{k^{m_0'}}\left\|\sum_{i<N}D(T(q,\mathbf{w}_i))\,\mathrm{e}([\mathbf{w}_i]_k k^{\lambda-\ell_0-m_0'}t)\psi^q_{\lambda-\ell_0-m_0',\alpha}(t,r)\right\|_2 \\ &\quad+\frac{k^{m_0'}-N}{k^{m_0'}}\max_{\widetilde{q}\in Q}\left\|\psi^{\widetilde{q}}_{\lambda-\ell_0-m_0',\alpha}(t,r)\right\|_2 \\ &=\frac{1}{k^{m_0'}}\left\|\sum_{i<N}\mathrm{e}([\mathbf{w}_i]_k k^{\lambda-\ell_0-m_0'}t)\right\|\left\|\psi^q_{\lambda-\ell_0-m_0',\alpha}(t,r)\right\|_2 \\ &\quad+\frac{k^{m_0'}-N}{k^{m_0'}}\max_{\widetilde{q}\in Q}\left\|\psi^{\widetilde{q}}_{\lambda-\ell_0-m_0',\alpha}(t,r)\right\|_2.\end{aligned}$$

Thus, we find in total

$$\left\|\psi^q_{\lambda,\alpha}(t,r)\right\|_2 \leq \frac{k^{m_0'}-N+\left|\sum_{i<N}\mathrm{e}([\mathbf{w}_i]_k k^{\lambda-\ell_0-m_0'}t)\right|}{k^{m_0'}}\max_{\widetilde{q}\in Q}\left\|\psi^{\widetilde{q}}_{\lambda-\ell_0-m_0',\alpha}(t,r)\right\|_2. \quad (3.15)$$

We combine (3.14) and (3.15) and find

$$\left\|\psi^q_{\lambda,\alpha}(t,r)\right\|_2 \leq \left(1 - \frac{2 - \left|1 + D(g)\,\mathrm{e}(([\mathbf{w}_g]_k - [\mathbf{w}_{id}]_k)k^{\lambda-m_0'}t)\right| + N - \left|\sum_{i<N}\mathrm{e}([\mathbf{w}_i]_k k^{\lambda-\ell_0-m_0'}t)\right|}{2k^{m_0'}}\right) \max_{\widetilde{q}\in Q}\left\|\psi^{\widetilde{q}}_{\lambda-\ell_0-m_0',\alpha}(t,r)\right\|_2.$$

It remains to show that

$$\left|1 + D(g)\,\mathrm{e}(([\mathbf{w}_g]_k - [\mathbf{w}_{id}]_k)k^{\lambda-m_0'}t)\right| + \left|\sum_{i<N}\mathrm{e}([\mathbf{w}_i]_k k^{\lambda-\ell_0-m_0'}t)\right| < N+2 \tag{3.16}$$

holds for all $t \in \mathbb{R}$ as the left hand side is a periodic and continuous function in t. We distinguish the following two cases.
At first let us assume $d'tk^{\lambda-m_0'} \in \mathbb{Z}$:
Since $g \in G_0$ we know that $[\mathbf{w}_{id}] \equiv [\mathbf{w}_g] \bmod d'$ and, therefore, the first term of the left hand side of equation (3.16) simplifies to $|1 + D(g)|$. By the definition of g, we find that $|1 + D(g)| < 2$ and, therefore, equation (3.16) holds.
The remaining case is $d'tk^{\lambda-m_0'} \notin \mathbb{Z}$:
Let us assume

$$\left|\sum_{i<N}\mathrm{e}([\mathbf{w}_i]_k k^{\lambda-\ell_0-m_0'}t)\right| = N.$$

This implies $([\mathbf{w}_i]_k - [\mathbf{w}_j]_k)k^{\lambda-\ell_0-m_0'}t \in \mathbb{Z}$ for all $i,j < N$. As $d' \cdot d''(q_0,q_0)$ is a linear combination of $([\mathbf{w}_i]_k - [\mathbf{w}_j]_k)$, we find $d' \cdot d''(q,q)k^{\lambda-\ell_0-m_0'}t \in \mathbb{Z}$. This yields a contradiction since $d''(q,q)|k^{\ell_0}$. Therefore, we have

$$\left|\sum_{i<N}\mathrm{e}([\mathbf{w}_i]_k k^{\lambda-m_0'}t)\right| < N$$

and, consequently, equation (3.16) holds, which finishes the proof. □

3.5 Carry Lemma

We fix a strongly connected automaton A satisfying $d(A) = 1$ - see for example Proposition 2.5.1. Let $\mathcal{T}_A = (Q, \Sigma, \delta, q_0, G, \lambda)$ be a naturally induced transducer. Thus, $\mathcal{T}_A$ is synchronizing and we find that there exists $\eta > 0$ such that all but $O(k^{n(1-\eta)})$ words of length n are synchronizing - see Lemma 1.4.1. We show in this section that the function $f(n) = D(T(n+r))$ has the carry property – uniformly in r.

Lemma 3.5.1. *Definition 3.3.1 holds – uniformly in r – for $f(n) = D(T(n+r))$ where D is a unitary and irreducible representation of G, η is given by Lemma 1.4.1 and the implied constant does not depend on r.*

Proof. We fix λ, ρ and α. We rewrite $r = r_1 k^\alpha + r_2$ where $r_1, r_2 \in \mathbb{N}, r_2 < k^\alpha$. We want to distinguish the form of $\ell + r_1$, i.e. the maximal number t such that the digits at position $\alpha, \dots, \alpha+t$ of $(\ell+r_1)k^\alpha+n_1+n_2+r_2$ can be affected by the carry of $n_1+n_2+r_2$. As $n_1+n_2+r_2 \leq 3k^\alpha - 3$ limits the carry to $0, 1$, or 2, we define $t := \max(\nu_k(\ell + r_1 + 1), \nu_k(\ell + r_1 + 2))$. Thus we know that the digits at position $\alpha+t+i$ of $(\ell k^\alpha + n_1 + n_2 + r)$ and $(\ell k^\alpha + n_1 + r)$ coincide for $i = 1, 2, \dots$ and are equal to the digits at position $t + i$ of $\ell + r_1$. We fix t such that $t < \rho$ and count the number of integers ℓ such that (3.5) holds for some n_1, n_2. As the terms corresponding to the digits of index $\alpha + t + i$ cancel for $i \geq 1$, we find

$$\begin{aligned} &D(T(q_0, (\ell k^\alpha + n_1 + n_2 + r)_k))^H D(T(q_0, (\ell k^\alpha + n_1 + r)_k)) \\ &\quad = D(T(\delta(q_0, (\lfloor (\ell + r_1)/k^t \rfloor)_k), (\ell k^\alpha + n_1 + n_2 + r)_k^{\alpha+t}))^H \\ &\qquad D(T(\delta(q_0, (\lfloor (\ell + r_1)/k^t \rfloor)_k), (\ell k^\alpha + n_1 + r)_k^{\alpha+t})) \end{aligned}$$

and

$$\begin{aligned} &D(T_{\alpha+\rho}(q_0, (\ell k^\alpha + n_1 + n_2 + r)_k))^H D(T_{\alpha+\rho}(q_0, (\ell k^\alpha + n_1 + r)_k)) \\ &\quad = D(T(\delta(q_0, (\lfloor (\ell + r_1)/k^t \rfloor)_k^{\rho-t}, (\ell k^\alpha + n_1 + n_2 + r)_k^{\alpha+t}))^H \\ &\qquad D(T(\delta(q_0, (\lfloor (\ell + r_1)/k^t \rfloor)_k^{\rho-t}), (\ell k^\alpha + n_1 + r)_k^{\alpha+t})). \end{aligned}$$

Thus, the number of integers ℓ (with fixed t and r) for which (3.5) holds for some n_1, n_2 can be bound by the number of integers ℓ such that $((\ell + r_1)/k^t)_k^{\rho-t}$ is not synchronizing – otherwise $\delta(q_0, (\lfloor (\ell + r_1)/k^t \rfloor)_k) = \delta(q_0, (\lfloor (\ell + r_1)/k^t \rfloor)_k^{\rho-t})$ would imply that (3.5) does not hold. By Lemma 1.4.1, we know that this number is bound by $O(k^{(\rho-t)(1-\eta)} k^{\lambda-\rho})$ for some $\eta > 0$ where the constant only depends on k and A. Note that there are only two possibilities for the digits with index $0, \dots, t-1$. By summing over t we find that the number of such ℓ is bound by

$$\begin{aligned} c \sum_{t \leq \rho} k^{\lambda-\rho} k^{(\rho-t)(1-\eta)} + k^{\lambda-\rho} &\leq c k^{\lambda-\eta\rho} \sum_{t \leq \rho} k^{-t(1-\eta)} + k^{\lambda-\rho} \\ &\leq c k^{\lambda-\eta\rho} \frac{1}{1 - k^{-(1-\eta)}} + k^{\lambda-\rho} \end{aligned}$$

and the result follows easily as c does not depend on λ, ρ, α and r. □

3.6 Proof of Proposition 3.2.2

We are now ready to show Proposition 3.2.2.

The proof for Proposition 3.2.2 differs when $D = D_\ell$ and we start by considering this specific case.

Lemma 3.6.1. *For all $\ell < d'$, $\lambda_1, \lambda_2, r \in \mathbb{N}$ and $b < k^{\lambda_1}, m < k^{\lambda_2}$ we have*

$$\left\| \sum_{\substack{n<N \\ n \equiv m \bmod k^{\lambda_2}}} \chi_{k^{-\lambda_1}}\left(\frac{n + r - bk^{\nu-\lambda_1}}{k^\nu}\right) D_\ell(T(q_0, (n+r)_k))\mu(n) \right\|_F = o(N)$$

uniformly in $r \in \mathbb{N}$.

Proof. We find by the characterizing property of D_ℓ (i.e. $D_\ell(T(q_0,(n)_k)) = \mathrm{e}(\ell n/d')$)

$$\begin{aligned}
&\left\| \sum_{\substack{n<N \\ n\equiv m \bmod k^{\lambda_2}}} \chi_{k^{-\lambda_1}}\left(\frac{n+r-bk^{\nu-\lambda_1}}{k^\nu}\right) D_\ell(T(q_0,(n+r)_k))\mu(n) \right\|_F \\
&\quad = \left| \sum_{\substack{n<N \\ n\equiv m \bmod k^{\lambda_2}}} \chi_{k^{-\lambda_1}}\left(\frac{n+r-bk^{\nu-\lambda_1}}{k^\nu}\right) \mathrm{e}\left((n+r)\frac{\ell}{d'}\right)\mu(n) \right| \\
&\quad \le \sum_{\substack{m'<k^{\lambda_2}d' \\ m'\equiv m \bmod k^{\lambda_2}}} \left|\mathrm{e}\left(m'\frac{\ell}{d'}\right)\right| \left| \sum_{n<N} \chi_{k^{-\lambda_1}}\left(\frac{n+r-bk^{\nu-\lambda_1}}{k^\nu}\right) \mathbf{1}_{[n+r\equiv m' \bmod k^\lambda d']}\,\mu(n) \right|.
\end{aligned}$$

As

$$\sum_{n<N} \chi_{k^{-\lambda_1}}\left(\frac{n+r-bk^{\nu-\lambda_1}}{k^\nu}\right)$$

is indeed a sum over at most two intervals included in $[1,\dots,N-1]$, the result follows from the well know result

$$\sum_{\substack{n<N \\ n\equiv r \bmod s}} \mu(n) = o(N).$$

□

Proof of Proposition 3.2.2: For $D = D_\ell$ Lemma 3.6.1 gives the desired result. Suppose from now on $D \neq D_\ell$ for all $\ell < d'$. We rewrite the left hand side of (3.2) using exponential sums and obtain using Vaaler-approximation - Theorem 3.7.2

$$\begin{aligned}
&\left\| \sum_{n<N} \frac{1}{k^{\lambda_2}} \sum_{h<k^{\lambda_2}} \mathrm{e}\left(h\frac{n-m}{k^{\lambda_2}}\right) \sum_{n<N} \chi_{k^{-\lambda_1}}\left(\frac{n+r-bk^{\nu-\lambda_1}}{k^\nu}\right) D(T(n+r))\mu(n) \right\|_F \\
&\quad \le \frac{1}{k^{\lambda_2}} \sum_{h<k^{\lambda_2}} \left\| \sum_{n<N} \mathrm{e}\left(n\frac{h}{k^{\lambda_2}}\right) \sum_{|h'|\le H} a_{h'}(k^{-\lambda_1},H)\,\mathrm{e}\left(h'\frac{n+r-bk^{\nu-\lambda_1}}{k^\nu}\right) D(T(n+r))\mu(n) \right\|_F \\
&\qquad + \frac{1}{k^{\lambda_2}} \sum_{h<k^{\lambda_2}} \left\| \sum_{n<N} \mathrm{e}\left(n\frac{h}{k^{\lambda_2}}\right) \sum_{|h'|\le H} b_{h'}(k^{-\lambda_1},H)\,\mathrm{e}\left(h'\frac{n+r-bk^{\nu-\lambda_1}}{k^\nu}\right) D(T(n+r))\mu(n) \right\|_F.
\end{aligned}$$

We choose $H = k^{\lambda_1}$ and therefore $\left|a_{h'}(k^{-\lambda_1},k^{\lambda_1})\right| \le k^{-\lambda_1}$, $\left|b_{h'}(k^{-\lambda_1},k^{\lambda_1})\right| \le k^{-\lambda_1}$ holds. Thus, the sum above is bound by

$$2\frac{1}{k^{\lambda_2}} \sum_{h<k^{\lambda_2}} \frac{1}{k^{\lambda_1}} \sum_{|h'|<k^{\lambda_1}} \left\| \sum_{n<N} \mathrm{e}\left(n\left(\frac{h}{k^{\lambda_2}}+\frac{h'}{k^\nu}\right)\right) D(T(n+r))\mu(n) \right\|_F$$

$$\leq 4 \sup_{\theta \in \mathbb{R}} \left\| \sum_{n<N} \mathrm{e}(\theta n) D(T(n+r))\mu(n) \right\|_F$$

However, as $D(T(n+r))$ fulfills Definition 3.3.1 and Definition 3.3.2 uniformly in r, we can apply Theorem 3.3.4. This finishes the proof. □

3.7 Technical and Auxiliary Results

In this section, we will provide the proofs for some results mentioned earlier, where the proof is technical and seemed to disturb the flow of reading.

3.7.1 Technical Results

The first result stated that some properties occur "often" in the digital representation of numbers. More explicitly, we need to show that the set M described in Lemma 3.2.3 has density 1.

Proof of Lemma 3.2.3. One sees easily that it is sufficient to show that

$$\lim_{\nu \to \infty} \frac{M \cap [0, k^{2\nu})}{k^{2\nu}} = 1$$

holds. We define $M_1 := \{n \in \mathbb{N} | \forall q' \in Q' : \delta'(q', (n)_k) \in \cup Q'_i\}$ and $M_{2,\nu} := \{n < k^\nu | \mathbf{w}_1 \mathbf{w}_2 \dots \mathbf{w}_\ell \text{ is a subword of } (n)_k^\nu\}$. The idea is to show that both sets have density one.

Noting that $|(M \cap [0, k^{2\nu}))| \geq |(M_1 \cap [0, k^\nu))| \cdot |(M_{2,\nu})|$ gives then immediately the desired result.

Thus, it just remains to show that M_1 and M_2 have density 1.
Therefore we define $M_{1,q'} := \{n \in \mathbb{N} | \delta'(q', (n)_k) \in \cup Q'_i\}$ and see that $M_1 = \cap M_{1,q'}$. Thus it is sufficient to show that $M_{1,q'}$ has density 1.

One finds easily that for each $q'_1 \in Q'$ there exists $\mathbf{w}_{q'_1} \in \{0, \dots, k-1\}^*$ such that $\delta'(q'_1, \mathbf{w}_{q'_1})$ belongs to Q'_i for some i. We take m to be the maximum of these lengths. As each Q'_i is closed under δ' we may assume that each of these paths $\mathbf{w}_{q'}$ is exactly of length m. We split Q' into two different parts, i.e. $\overline{Q'} = \cup Q_i, \widetilde{Q'} = Q' \setminus \overline{Q'}$. We show that

$$\left| \{\mathbf{w} \in \{0, \dots, k-1\}^n : \delta'(q', \mathbf{w}) \in \widetilde{Q'}\} \right| \leq k^m k^{n(1-\eta)} \tag{3.17}$$

where η is defined by $\eta = 1 - \log(k^m - 1)/\log(k^m) > 0$ such that

$$k^m - 1 = k^{m(1-\eta)}.$$

We show (3.17) by induction on n. The statement is trivial for $n \leq m$. We find for $n > m$ (as $\overline{Q'}$ is closed under δ')

$$|\{\mathbf{w} \in \Sigma^n : \delta'(q', \mathbf{w}) \in \widetilde{Q'}\}| =$$

$$
\begin{aligned}
&= \sum_{q_1' \in \widetilde{Q'}} |\{\mathbf{w}_1 \in \Sigma^{n-m} : \delta'(q', \mathbf{w}) = q_1'\}| \cdot |\{\mathbf{w}_2 \in \Sigma^m : \delta'(q_1', \mathbf{w}_2) \in \widetilde{Q'}\}| \\
&\leq \sum_{q_1' \in \widetilde{Q'}} |\{\mathbf{w}_1 \in \Sigma^{n-m} : \delta'(q', \mathbf{w}) = q_1'\}| \cdot (k^m - 1) \\
&= |\{\mathbf{w} \in \Sigma^{n-m} : \delta'(q', \mathbf{w}) \in \widetilde{Q'}\}| \cdot k^{m(1-\eta)} \leq k^m k^{(n-m)(1-\eta)} k^{m(1-\eta)}.
\end{aligned}
$$

Lemma 1.4.1 shows

$$\lim_{\nu \to \infty} \frac{M_{2,\nu}}{k^\nu} = 1,$$

which finishes the proof. □

We prove the following relation between very similarly defined terms that appeared when we estimated the Fourier terms. More explicitly, we show that a uniform estimate for

$$\left\|\psi_{\lambda,\alpha}^q(t,r)\right\|_2 \ll k^{-\eta\lambda}$$

implies Theorem 3.4.2.

Proof of Lemma 3.4.3. We start by proving

$$\left\|\phi_{\lambda,\alpha}^q(t,r)\right\|_2 \leq \sum_{1 \leq j \leq \lambda} \max_{q' \in Q} \left\|\psi_{\lambda-j,\alpha}^{q'}(t, r \bmod k^\alpha)\right\|_2 \frac{1}{k^j} + \frac{1}{k^\lambda}. \tag{3.18}$$

We rewrite $r = r_1 k^{\alpha+\lambda} + r_2 k^\alpha + r_3$ where $r_3 < k^\alpha$ and $r_2 < k^\lambda$, such that $uk^\alpha + r = r_1 k^{\alpha+\lambda} + (u + r_2)k^\alpha + r_3$. Let us first consider the case $r_1 > 0$. We find by distinguishing $u + r_2 < k^\lambda$ and $u + r_2 \geq k^\lambda$,

$$
\begin{aligned}
\phi_{\lambda,\alpha}^q(t,r) = &\frac{1}{k^\lambda} \sum_{\substack{u < k^\lambda \\ u + r_2 < k^\lambda}} D(T(q, (r_1)_k (u + r_2)_k^\lambda (r_3)_k^\alpha)) \, \mathrm{e}(-ut) \\
&+ \frac{1}{k^\lambda} \sum_{\substack{u < k^\lambda \\ u + r_2 \geq k^\lambda}} D(T(q, (r_1 + 1)_k (u + r_2 - k^\lambda)_k^\lambda (r_3)_k^\alpha)) \, \mathrm{e}(-ut) \\
= &\frac{1}{k^\lambda} D(T(q, (r_1)_k)) \sum_{r_2 \leq u' < k^\lambda} D(T(\delta(q, (r_1)_k), (u')_k^\lambda (r_3)_k^\alpha)) \, \mathrm{e}(-(u' - r_2)t) \\
&+ \frac{1}{k^\lambda} D(T(q, (r_1 + 1)_k)) \sum_{u' < r_2} D(T(\delta(q, (r_1 + 1)_k), (u')_k^\lambda (r_3)_k^\alpha)) \, \mathrm{e}(-(u' - r_2 + k^\lambda)t).
\end{aligned}
$$

This gives

$$\left\|\phi_{\lambda,\alpha}^q(t,r)\right\|_2 \leq \max_{q' \in Q} \frac{1}{k^\lambda} \left\| \sum_{r_2 \leq u' < k^\lambda} D(T(q', (u')_k^\lambda (r_3)_k^\alpha)) \, \mathrm{e}(-u't) \right\|_2$$

$$+ \max_{q' \in Q} \frac{1}{k^\lambda} \left\| \sum_{u' < r_2} D(T(q', (u')_k^\lambda (r_3)_k^\alpha)) \, \mathrm{e}(-u't) \right\|_2 .$$

The case $r_1 = 0$ can be treated similarly and we find in this case

$$\begin{aligned} \left\| \phi_{\lambda,\alpha}^q(t,r) \right\|_2 \leq & \max_{q' \in Q} \frac{1}{k^\lambda} \left\| \sum_{r_2 \leq u' < k^\lambda} D(T(q', (u')_k (r_3)_k^\alpha)) \, \mathrm{e}(-u't) \right\|_2 \\ & + \max_{q' \in Q} \frac{1}{k^\lambda} \left\| \sum_{u' < r_2} D(T(q', (u')_k^\lambda (r_3)_k^\alpha)) \, \mathrm{e}(-u't) \right\|_2 . \end{aligned} \tag{3.19}$$

We work from now on just with the case $r_1 = 0$ as the case $r_1 > 0$ works similarly. We rewrite $r_2 = r_2' k^{\lambda-1} + r_2''$ with $r_2' < k, r_2'' < k^{\lambda-1}$ and find by distinguishing the most significant digits of u' the following upper bound for the right hand side of (3.19):

$$\begin{aligned} & \frac{1}{k^\lambda} \sum_{r_2'+1 \leq u_1' < k} \max_{q' \in Q} \left\| \sum_{u_2' < k^{\lambda-1}} D(T(q', (u_2')_k^{\lambda-1} (r_3)_k^\alpha)) \, \mathrm{e}(-u_2' t) \right\|_2 \\ & + \frac{1}{k^\lambda} \max_{q' \in Q} \left\| \sum_{r_2'' \leq u_2' < k^{\lambda-1}} D(T(q'(u_2')_k (r_3)_k^\alpha)) \, \mathrm{e}(-u_2' t) \right\|_2 \\ & + \frac{1}{k^\lambda} \sum_{u_1' < r_2'} \max_{q' \in Q} \left\| \sum_{u_2' < k^{\lambda-1}} D(T(q', (u_2')_k^{\lambda-1} (r_3)_k^\alpha)) \, \mathrm{e}(-u_2' t) \right\|_2 \\ & + \frac{1}{k^\lambda} \max_{q' \in Q} \left\| \sum_{u_2' < r_2''} D(T(q'(u_2')_k^{\lambda-1} (r_3)_k^\alpha)) \, \mathrm{e}(-u_2' t) \right\|_2 . \end{aligned}$$

The first and third line give

$$\frac{k}{k-1} \psi_{\lambda-1,\alpha}^q(t, r_3).$$

By applying this step inductively we find (3.18). Thus it just remains to use the bound of $\psi_{\lambda,\alpha}^q(t,r)$

$$\begin{aligned} \left\| \phi_{\lambda,\alpha}^q(t,r) \right\|_2 & \leq \sum_{1 \leq j \leq \lambda} \max_{q' \in Q} \left\| \psi_{\lambda-j,\alpha}^{q'}(t, r \bmod k^\alpha) \right\|_2 \frac{1}{k^j} + \frac{1}{k^\lambda} \\ & \leq c \sum_{j \leq \lambda} k^{-(\lambda-j)\eta} k^{-j} + k^{-\lambda} \\ & \leq c k^{-\lambda\eta} \frac{k^{-(1-\eta)}}{1 - k^{-(1-\eta)}} + k^{-\lambda}. \end{aligned}$$

□

3.7.2 Van-der-Corput's inequality

The following lemma is a generalization of Van-der-Corput's inequality.

Lemma 3.7.1. *Let N be a positive integer and $Z(n) \in \mathbb{C}^{d\times d}$ for all $0 \le n \le N$. Then we have for any real number $S \ge 1$ and any integer $k \ge 1$ the estimate*

$$\left\| \sum_{0\le n\le N} Z(n) \right\|_F^2 \le \frac{N+k(S-1)+1}{S} \sum_{|s|<S} \left(1-\frac{|s|}{S}\right) \sum_{0\le n\le N-ks} \operatorname{tr}\left(Z(n+ks)Z(n)^H\right) \quad (3.20)$$

where $\operatorname{tr}(Z)$ *denotes the trace of* Z.

Proof. See [15]. □

3.7.3 Vaaler's method

The following theorem is a classical method to detect real numbers in an interval modulo 1 by means of exponential sums. For $\alpha \in \mathbb{R}$ with $0 \le \alpha < 1$, we denote by χ_α the characteristic function of the interval $[0,\alpha)$ modulo 1:

$$\chi_\alpha(x) = \lfloor x \rfloor - \lfloor x-\alpha \rfloor . \quad (3.21)$$

The following theorem is due to Vaaler [43].

Theorem 3.7.2. *For all $\alpha \in \mathbb{R}$ with $0 \le \alpha < 1$ and all integer $H \ge 1$, there exist real-valued trigonometric polynomials $A_{\alpha,H}(x)$ and $B_{\alpha,H}(x)$ such that for all $x \in \mathbb{R}$*

$$|\chi_\alpha(x) - A_{\alpha,H}(x)| \le B_{\alpha,H}(x). \quad (3.22)$$

The trigonometric polynomials are defined by

$$A_{\alpha,H}(x) = \sum_{|h|\le H} a_h(\alpha,H)\,\mathrm{e}(hx),\ B_{\alpha,H}(x) = \sum_{|h|\le H} b_h(\alpha,H)\,\mathrm{e}(hx), \quad (3.23)$$

with coefficients $a_h(\alpha,H)$ and $b_h(\alpha,H)$ satisfying

$$a_0(\alpha,H) = \alpha,\ |a_h(\alpha,H)| \le \min\left(\alpha, \tfrac{1}{\pi|h|}\right),\ |b_h(\alpha,H)| \le \tfrac{1}{H+1}. \quad (3.24)$$

Using this method we can detect points in a d-dimensional box (modulo 1) - which will be useful in Chapter 5:

Lemma 3.7.3. *For $(\alpha_1,\dots,\alpha_d) \in [0,1)^d$ and $(H_1,\dots,H_d) \in \mathbb{N}^d$ with $H_1 \ge 1,\dots,\ H_d \ge 1$, we have for all $(x_1,\dots,x_d) \in \mathbb{R}^d$*

$$\left| \prod_{j=1}^d \chi_{\alpha_j}(x_j) - \prod_{j=1}^d A_{\alpha_j,H_j}(x_j) \right| \le \sum_{\emptyset\ne J\subseteq\{1,\dots,d\}} \prod_{j\notin J} \chi_{\alpha_j}(x_j) \prod_{j\in J} B_{\alpha_j,H_j}(x_j) \quad (3.25)$$

where $A_{\alpha,H}(.)$ and $B_{\alpha,H}(.)$ are the real valued trigonometric polynomials defined by (3.23).

Proof. See [14]. □

Let $(U_1, \ldots, U_d) \in \mathbb{N}^d$ with $U_1 \geq 1, \ldots, U_d \geq 1$ and define $\alpha_1 = 1/U_1, \ldots, \alpha_d = 1/U_d$. For $j = 1, \ldots, d$ and $x \in \mathbb{R}$ we have

$$\sum_{0 \leq u_j < U_j} \chi_{\alpha_j}\left(x - \frac{u_j}{U_j}\right) = 1. \tag{3.26}$$

Let $N \in \mathbb{N}$ with $N \geq 1$, $f : \{1, \ldots, N\} \to \mathbb{R}^d$ and $g : \{1, \ldots, N\} \to \mathbb{C}$ such that $|g| \leq 1$. If $f = (f_1, \ldots, f_d)$, we can express the sum

$$S = \sum_{n=1}^{N} g(n)$$

as

$$S = \sum_{n=1}^{N} g(n) \sum_{0 \leq u_1 < U_1} \chi_{\alpha_1}\left(f_1(n) - \frac{u_1}{U_1}\right) \cdots \sum_{0 \leq u_d < U_d} \chi_{\alpha_d}\left(f_d(n) - \frac{u_d}{U_d}\right).$$

We now define $(H_1, \ldots, H_d) \in \mathbb{N}^d$ with $H_1 \geq 1, \ldots, H_d \geq 1$,

$$\widetilde{S} = \sum_{n=1}^{N} g(n) \sum_{0 \leq u_1 < U_1} A_{\alpha_1, H_1}\left(f_1(n) - \frac{u_1}{U_1}\right) \cdots \sum_{0 \leq u_d < U_d} A_{\alpha_d, H_d}\left(f_d(n) - \frac{u_d}{U_d}\right).$$

Lemma 3.7.4. *With the notations from above, we have*

$$\begin{aligned} \left|S - \widetilde{S}\right| \leq \sum_{\ell=1}^{d-1} \sum_{1 \leq j_1 < \cdots < j_\ell} \frac{U_{j_1} \cdots U_{j_\ell}}{H_{j_1} \cdots H_{j_\ell}} \sum_{|h_{j_1}| \leq H_{j_1}/U_{j_1}} \cdots \sum_{|h_{j_\ell}| \leq H_{j_\ell}/U_{j_\ell}} \\ \left|\sum_{n=1}^{N} \mathrm{e}\left(h_{j_1} U_{j_1} f_{j_1}(n) + \cdots + h_{j_\ell} U_{j_\ell} f_{j_\ell}(n)\right)\right|. \end{aligned} \tag{3.27}$$

Proof. See again [14]. □

3.8 Proof of Theorem 3.3.3 and Theorem 3.3.4

The proof of Theorem 3.3.3 and Theorem 3.3.4 is completely analogous to the corresponding proof in [38]. One of the main ideas is to use the Vaughan method. It follows from the following identity concerning the von Mangoldt function. Let $U, V \geq 1$. Then

$$\Lambda(n) = \Lambda(n)\, \mathbf{1}_{[1,U]}(n) + \sum_{\substack{ab=n \\ a \leq V}} \mu(a) \log(b) - \sum_{\substack{ab=n \\ a \leq UV}} \sum_{\substack{cd=a \\ c \leq V, d \leq U}} \mu(c)\Lambda(d) - \sum_{\substack{ab=n \\ a > U, b > V}} \Lambda(a) \sum_{\substack{c \mid b \\ c \leq V}} \mu(c).$$

Using this identity, one can find estimates for

$$\sum_{n<N} \Lambda(n) f(n)$$

by estimating sums of type I

$$S_I(\theta) = \sum_{\frac{M}{q}<m\leq M} \left| \sum_{\substack{n \\ mn\in I(M,N)}} f(mn)\, \mathrm{e}(\theta mn) \right|,$$

where $I(M,N) \subset [0, MN]$ is an interval, and sums of type II

$$S_{II}(\theta) = \sum_{\frac{M}{q}<m\leq M} \sum_{\frac{N}{q}<n\leq N} a_m b_n f(mn)\, \mathrm{e}(\theta mn)$$

for any $a_m, b_n \in \mathbb{C}$ with $|a_m| \leq 1, |b_n| \leq 1$. A very similar result is true when considering

$$\sum_{n<N} \mu(n) f(n).$$

This is a classical method used for complex valued functions f. However, one can easily generalize this method for matrix valued functions f.

We will only comment on how to adopt the original proof in [38] to our situation. We only describe the important changes briefly (and assume that the reader is familiar with [38]). First we comment on how to adopt the auxiliary results of [38].

For a generalization of [38, Lemma 3] see [15]. [38, Lemma 6] as well as the Cauchy-Schwarz inequality can be easily adapted to matrix-valued functions, where we use the Frobenius norm instead of the absolute value. These results change at most by a factor $\sqrt{d}$.
[38, Lemma 8] can be adopted to our case where the proof stays unchanged:

Lemma 3.8.1. *Let $f : \mathbb{N} \to U_d$ satisfies Definition 3.3.1 with $\eta > 0$. For $(\mu, \nu, \rho) \in \mathbb{N}^3$ with $2\rho < \nu$ the set $\mathcal{E}$ of pairs $(m,n) \in \{k^{\mu-1}, \dots, k^\mu - 1\} \times \{k^{\nu-1}, \dots, k^\nu - 1\}$ such that there exists $\ell < k^{\mu+\rho}$ with $f(mn+k)^H f(mn) \neq f_{\mu+2\rho}(mn+k)^H f_{\mu+2\rho}(mn)$ satisfies*

$$\operatorname{card} \mathcal{E} \ll (\log k) k^{\mu+\nu-\eta\rho}. \tag{3.28}$$

[38, Lemma 9] can be adopted as well:

Lemma 3.8.2. *Let $f : \mathbb{N} \to U_d$ satisfying Definition 3.3.1 and $(\mu, \nu, \mu_0, \mu_1, \mu_2) \in \mathbb{N}^5$ with $\mu_0 \leq \mu_1 \leq \mu \leq \mu_2$, $\mu \leq \nu$ and $2(\mu_2 - \mu) \leq \mu_0$. For $(a,b,c) \in \mathbb{N}^3$ the set $\mathcal{E}(a,b,c)$ of pairs $(m,n) \in \{k^{\mu-1}, \dots, k^\mu - 1\} \times \{k^{\nu-1}, \dots, k^\nu - 1\}$ such that*

$$\begin{aligned} &f_{\mu_2}(mn+am+bn+c)^H f_{\mu_2}(k^{\mu_0} \mathbf{r}_{\mu_0,\mu_2}(mn+am+bn+c)) \\ &\quad \neq f_{\mu_1}(mn+am+bn+c)^H f_{\mu_1}(k^{\mu_0} \mathbf{r}_{\mu_0,\mu_2}(mn+am+bn+c)) \end{aligned}$$

satisfies

$$\operatorname{card} \mathcal{E}(a,b,c) \ll \max(\tau(k), \log k) \mu_2^{\omega(k)} k^{\mu+\nu+\eta(\mu_0-\mu_1)}. \tag{3.29}$$

Sums of Type I

We take a non decreasing function $\gamma : \mathbb{R} \to \mathbb{R}$ satisfying $\lim_{\lambda\to\infty} \gamma(\lambda) = +\infty$, $c \geq 2$ and $f : \mathbb{N} \to U_d$ be a function satisfying Definition 3.3.1 and $f \in \mathcal{F}_{\gamma,c}$ according to Definition 3.3.2. Let

$$1 \leq M \leq N \text{ such that } M \leq (MN)^{1/3}. \tag{3.30}$$

Let μ and ν be the unique integers such that

$$k^{\mu-1} \leq M < k^{\mu} \text{ and } k^{\nu-1} \leq N < k^{\nu}.$$

Let $\theta \in \mathbb{R}$, an interval $I(M,N) \subseteq [0, MN]$ and

$$S_I(\theta) = \sum_{\frac{M}{q} < m \leq M} \left\| \sum_{n : mn \in I(M,N)} f(mn)\, \mathrm{e}(\theta mn) \right\|_F. \tag{3.31}$$

Proposition 3.8.3. *Assuming* (3.30) *and with $c \geq 2$, we have – uniformly for $\theta \in \mathbb{R}$ – that*

$$S_I(\theta) \ll (\log k)^{5/2} (\mu+\nu)^2 k^{\mu+\nu-\frac{\eta}{2}\gamma\left(\frac{\mu+\nu}{3}\right)}. \tag{3.32}$$

Proof (Sketch): The proof proceeds as in [38] up to [38, equation (32)]. We find

$$\widehat{f_{\mu+\nu}}(t) = \frac{1}{k^{\mu+\nu-\alpha}} \sum_{v < k^{\mu+\nu-\alpha}} f(vk^{\alpha})\, \mathrm{e}\left(-\frac{vt}{k^{\mu+\nu-\alpha}}\right) \frac{1}{k^{\alpha}} \sum_{u \leq k^{\alpha}} f(vk^{\alpha})^H f(u+vk^{\alpha})\, \mathrm{e}\left(-\frac{ut}{k^{\mu+\nu}}\right).$$

Here we needed to ensure the correct order of the terms. From now on the proof does not change and we just have to take the factor η into account when using Definition 3.3.1. We find for example

$$\operatorname{card} \widetilde{\mathcal{W}_{\alpha}} \ll k^{\mu+\nu-\eta\rho_1}.$$

Thereafter one only needs to keep the order of the terms as the values of f are now matrices and thus the order of the terms is important. However, this does not change any important properties and all arguments of [38] still hold. By taking η into account, we find

$$\left\| S''_{I,2}(M,d) \right\|_F \ll (\log k)^{1/2} k^{\mu+\nu-\eta\rho_1/2}$$

$$S'_{I,2}(\theta') \ll (\log k)^{1/2} \sum_{1 \leq d \leq M} \frac{k^{-\eta\rho_1/2}}{d} \ll \mu (\log k)^{3/2} k^{-\eta\rho_1/2}.$$

Choosing

$$\rho_1 = \gamma\left(\frac{\mu+\nu}{3}\right) \frac{2}{1+\eta}$$

then gives the desired result. □

Sums of Type II

We take $\gamma : \mathbb{R} \to \mathbb{R}$ a non decreasing function satisfying $\lim_{\lambda\to\infty} \gamma(\lambda) = +\infty, c \geq 10$ and $f : \mathbb{N} \to U_d$ be a function satisfying Definition 3.3.1 and $f \in \mathcal{F}_{\gamma,c}$ in Definition 3.3.2. Let $1 \leq M \leq N$. We denote by μ and ν the unique integers such that

$$k^{\mu-1} \leq M < k^{\mu} \text{ and } k^{\nu-1} \leq N < k^{\nu}.$$

Let us assume that

$$\frac{1}{4}(\mu+\nu) \leq \mu \leq \nu \leq \frac{3}{4}(\mu+\nu). \tag{3.33}$$

We assume – as in [38] – that the multiplicative dependency of the variables in the type II sums has been removed by the classical method described (for example) in [37, Section 5]. Let $\theta \in \mathbb{R}, a_m \in \mathbb{C}, b_n \in \mathbb{C}$ with $|a_m| \leq 1, |b_n| \leq 1$ and

$$S_{II}(\theta) = \sum_m \sum_n a_m b_n f(mn)\, \mathrm{e}(\theta mn)$$

where we sum over $m \in (M/k, M]$ and $n \in (N/k, N]$. We will prove

Proposition 3.8.4. *Assuming* (3.33) *and* $c \geq 10$, *uniformly for* $|a_m| \leq 1, |b_n| \leq 1$ *and* $\theta \in \mathbb{R}$, *we have*

$$\|S_{II}(\theta)\|_F \ll \max(\tau(k)\log k, \log^3 k)^{1/4}(\mu+\nu)^{\frac{1}{4}(1+\max(\omega(k),2))} k^{\mu+\nu-\frac{\eta}{20}\gamma(2\lfloor\mu/15\rfloor)}. \tag{3.34}$$

Proof. The proof of the corresponding result in [38] is the most difficult part – quite long and complicated. We will try to focus only on the necessary changes and keep it as short as possible, as no important new ideas are needed.
We find by using the corresponding results:

$$\|S_{II}(\theta)\|_F^2 \ll \frac{M^2N^2}{R} + \frac{MN}{R}\sum_{1\leq r<R}\left(1-\frac{r}{R}\right)\mathrm{tr}(S_1(r))$$

with

$$S_1(r) = \sum_m \sum_{n\in I(N,r)} b_{n+r}\overline{b_n} f(mn+mr)f(mn)^H\, \mathrm{e}(\theta mr),$$

where $I(N,r) = (N/k, N-r]$. Let

$$\mu_2 = \mu + 2\rho. \tag{3.35}$$

As we are only interested in $\mathrm{tr}(S_1(r))$, we can – by well-known properties of the trace – exchange the order of the matrices

$$\mathrm{tr}(S_1(r)) = \sum_m \sum_{n\in I(N,r)} b_{n+r}\overline{b_n}\,\mathrm{tr}(f(mn+mr)f(mn)^H)\,\mathrm{e}(\theta mr) \tag{3.36}$$

$$= \sum_m \sum_{n \in I(N,r)} b_{n+r}\overline{b_n} \operatorname{tr}(f(mn)^H f(mn+mr)) \operatorname{e}(\theta mr) \tag{3.37}$$

$$= \operatorname{tr}\left(\sum_m \sum_{n \in I(N,r)} b_{n+r}\overline{b_n} f(mn)^H f(mn+mr) \operatorname{e}(\theta mr) \right). \tag{3.38}$$

We denote from now on by $\widetilde{S}$ the corresponding sum where we exchange the order of the matrices, e.g. :

$$\widetilde{S}_1(r) = \sum_m \sum_{n \in I(N,r)} b_{n+r}\overline{b_n} f(mn)^H f(mn+mr) \operatorname{e}(\theta mr). \tag{3.39}$$

If f satisfies the carry property described in Definition 3.3.1, then by Lemma 3.8.1 the number of pairs (m,n) for which $f(mn)^H f(mn+mr) \neq f_{\mu_2}(mn)^H f_{\mu_2}(mn+mr)$ is $O(k^{\mu+\nu-\eta\rho})$. Hence

$$\operatorname{tr}(S_1(r)) = \operatorname{tr}(\widetilde{S}_1(r)) = \operatorname{tr}(S_1'(r)) + O(k^{\mu+\nu-\eta\rho}) \tag{3.40}$$

where

$$S_1'(r) = \sum_m \sum_{n \in I(N,r)} b_{n+r}\overline{b_n} f_{\mu_2}(mn)^H f_{\mu_2}(mn+mr) \operatorname{e}(\theta mr).$$

Note that it was important to change the order of the matrices first to apply Lemma 3.8.1. Using again the Cauchy-Schwarz inequality for the summation over r and $|\operatorname{tr}(A)| \leq \|A\|_F \sqrt{d}$, leads to

$$\|S_{II}(\theta)\|_F^4 \ll \frac{M^4N^4}{R^2} + \frac{M^2N^2}{R^2} R \sum_{1 \leq r < R} \|S_1'(r)\|_F^2 . \tag{3.41}$$

When applying the Van-der-Corput inequality for the summation over m (see for example [15]) we need again to keep the correct order of terms. We find

$$\sum_{1 \leq r < R} \|S_1'(r)\|_F^2 \ll \frac{M^2N^2R}{S} + \frac{MN}{S} \operatorname{tr}(\widetilde{S}_2)$$

with

$$\widetilde{S}_2 = \sum_{1 \leq r < R} \sum_{1 \leq s < S} \left(1 - \frac{s}{S}\right) \operatorname{e}(\theta k^{\mu_1} rs) \widetilde{S}_2'(r,s)$$

where

$$\begin{aligned}
\widetilde{S}_2'(r,s) &= \sum_m \sum_n f_{\mu_2}((m+sk^{\mu_1})n)^H f_{\mu_2}((m+sk^{\mu_1})(n+r)) f_{\mu_2}(m(n+r))^H f_{\mu_2}(mn) \\
&= \sum_m \sum_n f_{\mu_1}((m+sk^{\mu_1})n)^H f_{\mu_1,\mu_2}((m+sk^{\mu_1})n)^H f_{\mu_1,\mu_2}((m+sk^{\mu_1})(n+r)) \\
&\quad f_{\mu_1}((m+sk^{\mu_1})(n+r)) f_{\mu_1}(m(n+r))^H f_{\mu_1,\mu_2}(m(n+r))^H f_{\mu_1,\mu_2}(mn) f_{\mu_1}(mn).
\end{aligned}$$

However, we find that $\operatorname{tr}(\widetilde{S}_2'(r,s)) = \operatorname{tr}(S_2'(r,s))$ for

$$S_2'(r,s) = \sum_m \sum_n f_{\mu_1,\mu_2}(mn) f_{\mu_1,\mu_2}((m+sk^{\mu_1})n)^H f_{\mu_1,\mu_2}((m+sk^{\mu_1})(n+r)) f_{\mu_1,\mu_2}(m(n+r))^H$$

and, therefore,

$$\sum_{1\le r<R} \|S_1'(r)\|_F^2 \ll \frac{M^2N^2R}{S} + \frac{MN}{S}\operatorname{tr}(S_2)$$

where

$$S_2 = \sum_{1\le r<R} \sum_{1\le s<S} \left(1-\frac{s}{S}\right) \mathrm{e}(\theta k^{\mu_1} rs) S_2'(r,s).$$

We choose the order of the terms such that we are able to use the Cauchy-Schwarz inequality later in a sufficient way. Whenever we use Definition 3.3.1 (at least implicitly), we find a different error term e.g. : instead of [38, (60)] we can use

$$\operatorname{card} \mathcal{E}_{\mu_0,\mu_1,\mu_2}(r,s) \ll \max(\tau(k), \log k)(\mu+\nu)^{\omega(k)} k^{\mu+\nu-2\eta\rho'}.$$

Thus we find

$$S_2'(r,s) = S_3(r,s) + O(\max(\tau(k), \log k)(\mu+\nu)^{\omega(k)} k^{\mu+\nu-2\eta\rho'}),$$

where the order of the terms in S_3 has to be changed:

$$\begin{aligned} S_3(r,s) = \sum_m \sum_n \sum_{0\le u_0<k^{\mu_2-\mu_0}} \sum_{0\le u_1<k^{\mu_2-\mu_0}} & \chi_{k^{\mu_0-\mu_2}}\left(\frac{mn}{k^{\mu_2}} - \frac{u_0}{k^{\mu_2-\mu_0}}\right) \chi_{k^{\mu_0-\mu_2}}\left(\frac{mn+mr}{k^{\mu_2}} - \frac{u_1}{k^{\mu_2-\mu_0}}\right) \\ & g(u_0) g(u_0 + k^{\mu_1-\mu_0} sn)^H g(u_1 + k^{\mu_1-\mu_0} sn + k^{\mu_1-\mu_0} sr) g(u_1)^H. \end{aligned}$$

This impact of η carries through the rest of the work and we will only comment on the specific form of some intermediate results, e.g. S_4:

$$\begin{aligned} S_4(r,s) = k^{2(\mu_2-\mu_0)} \sum_{|h_0|\le H} \sum_{|h_1|\le H} a_{h_0}(k^{\mu_0-\mu_2}, H) a_{h_1}(k^{\mu_0-\mu_2}, H) \sum_{0\le h_2<k^{\mu_2-\mu_0}} \sum_{0\le h_3<k^{\mu_2-\mu_0}} \mathrm{e}\left(\frac{h_3 sr}{k^{\mu_2-\mu_1}}\right) \\ \widehat{g}(h_0-h_2)\widehat{g}(-h_2)^H \widehat{g}(h_3)\widehat{g}(h_3-h_1)^H \\ \sum_m \sum_n \mathrm{e}\left(\frac{(h_0+h_1)mn + h_1 mr + (h_2+h_3)k^{\mu_1} sn}{k^{\mu_2}}\right). \end{aligned}$$

The definitions of S_6 and S_7 have to be adopted as well, e.g. :

$$S_7(h_1) = \sum_{0\le h'<k^{\mu_2-\mu_0}} \left\| \widehat{g}(h'-h_1)\widehat{g}(h')^H \right\|_F^2.$$

The following arguments of [38] carry over to this generalization.

We find the following lemma, which is the analogon to [38, Lemma 10]

Lemma 3.8.5. *If*

$$\mu \leq \left(2 + \frac{4}{3}c\right)\rho$$

then we have – uniformly for $\lambda \in \mathbb{N}$ *with* $\frac{1}{3}(\mu_2 - \mu_0) \leq \lambda \leq \frac{4}{5}(\mu_2 - \mu_0)$ *–*

$$\sum_{0 \leq h < k^{\mu_2 - \mu_0}} \sum_{0 \leq k < k^{\mu_2 - \mu_0 - \lambda}} \left\| \widehat{g}(h+k)\widehat{g}(h)^H \right\|_F^2 \ll k^{-\gamma_1(\lambda, \mu_1 - \mu_0)} (\log k^{\mu_2 - \mu_1})^2$$

where

$$\gamma_1(\lambda, \mu_1 - \mu_0) = \frac{\gamma(\lambda) - \mu_1 + \mu_0}{2}\eta. \tag{3.42}$$

Proof. The proof works as in [38], where one has to be careful to keep the order of the terms. However, this does not change the proof substantially and by using the new estimate given by Definition 3.3.1 one finds the desired result. □

The rest of the proof does not change and it just remains to balance the error terms differently. One finds in total, uniformly for $\theta \in \mathbb{R}$

$$\begin{aligned}
\|S_{II}(\theta)\|_F^4 \ll\ & k^{4\mu+4\nu+\mu_1-\mu_0}(k^{-\gamma_1(\mu_2-\mu_0-2\rho,\mu_1-\mu_0)} + k^{-\rho}\log k^{\rho}) \\
& (\tau\left(k^{\mu_2-\mu_1}\right) + k^{\mu_2-\mu_1-\nu\log k^{\mu_2-\mu_1}}) \\
& + (\log k)^3(\mu+\nu)^3 k^{4\mu+4\nu+3(\mu_2-\mu_0)+2\rho}(k^{-\mu_2} + k^{-\nu}) \\
& + \max(\log k^{\mu_0}, \tau(k^{\mu_0}))k^{4\mu+4\nu-2\rho} \\
& + \max(\tau(k), \log k)(\mu+\nu)^{\omega(k)} k^{4\mu+4\nu-2\eta\rho'}.
\end{aligned}$$

Note that only the first and last error terms have changed compared to [38]. As in [38], we assumed

$$\begin{aligned}
\mu_2 &= \mu + 2\rho \\
\mu_1 &= \mu - 2\rho \\
\mu_0 &= \mu_1 - 2\rho' \\
\mu &\leq \nu \leq 3\mu.
\end{aligned}$$

In total, we find

$$\begin{aligned}
\|S_{II}(\theta)\|_F^4 \ll\ & \tau(k)(\mu_2-\mu_1)^{\omega(k)} k^{4\mu+4\nu+\frac{2+\eta}{2}(\mu_1-\mu_0)-\frac{\eta}{2}\gamma(2\rho)}\log k^{\rho} \\
& + (\log k)^3(\mu+\nu)^3 k^{4\mu+4\nu+3(\mu_1-\mu_0)+14\rho-\mu} \\
& + \max(\log k^{\mu_0}, \tau(k^{\mu_0}))k^{4\mu+4\nu-2\rho} \\
& + \max(\tau(k), \log k)(\mu+\nu)^{\omega(k)} k^{4\mu+4\nu-2\eta\rho'}.
\end{aligned}$$

Taking

$$\rho' = \lfloor \eta\gamma(2\rho)/10 \rfloor,$$

we have $\mu_1 - \mu_0 = 2\rho' \leq \eta\gamma(2\rho)/5 \leq \eta\rho/5$, and thus

$$\frac{2+\eta}{2}(\mu_1 - \mu_0) - \frac{\eta}{2}\gamma(2\rho) \leq \frac{3}{10}\eta\gamma(2\rho) - \frac{\eta}{2}\gamma(2\rho) = -\eta\frac{\gamma(2\rho)}{5}.$$

Furthermore, we choose

$$\rho = \lfloor \mu/15 \rfloor,$$

which yields by the same arguments as in [38] that – finally –

$$\|S_{II}(\theta)\|_F^4 \ll \max(\tau(k)\log k, \log^3 k)(\mu+\nu)^{1+\max(\omega(k),2)} k^{4\mu+4\nu-\eta\gamma(2\lfloor \mu/15 \rfloor)/5},$$

which completes the proof of Proposition 3.8.4. □

Proof of Theorem 3.3.4 and Theorem 3.3.3. It just remains to use the estimates of sums of type I and type II to find the corresponding estimates, see for example [38]. □

Chapter 4

Subsequences of Automatic sequences

The setup we have developed in Chapters 2 and 3 can also be used to work with linear subsequences and also to derive a Prime Number Theorem for a large class of automatic sequences – which covers almost all mentioned results at the beginning of Chapter 3.

This chapter is dedicated to show the following two results.

Theorem 4.0.1. *Let $A = (Q', \Sigma, \delta', q_0', \tau)$ be a strongly connected deterministic finite automaton with output (DFAO) with $\Sigma = \{0, \dots, k-1\}$, $\delta'(q_0', 0) = q_0'$ and automatic sequence $\mathbf{u}$. Then the frequencies of letters in $(\mathbf{u}_{an+b})_{n\in\mathbb{N}}$ exists for every $a, b \in \mathbb{N}$.*

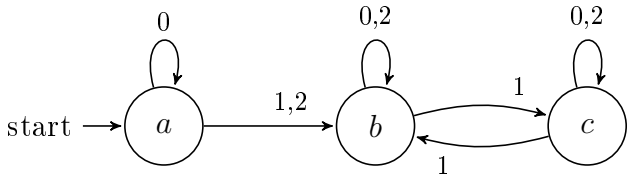
Theorem 4.0.2. *Let $A = (Q', \Sigma, \delta', q_0', \tau)$ be a strongly connected deterministic finite automaton with output (DFAO) with $\Sigma = \{0, \dots, k-1\}$, $\delta'(q_0', 0) = q_0'$ and automatic sequence $\mathbf{a}$. Then the frequencies of the letters for the prime subsequence $(\mathbf{a}_p)_{p\in\mathcal{P}}$ exist.*

Remark. The proofs allow us to determine these frequencies.
All block-additive, i.e. digital, functions[1] are covered by Theorem 4.0.2. For the residue of any block-additive function $f \bmod m$ satisfying $(k-1, m) = 1$ and $(\gcd(f(n)_{n\in\mathbb{N}}), m) = 1$ one finds that all letters appear with the same frequencies along the primes.

Let $\mathbf{a}$ be an automatic sequence. We have observed in Chapter 1 that every linear subsequence is again an automatic sequence. However, it is not clear if the existence of densities for the original sequence $\mathbf{a}$ implies the existence of densities for its linear subsequence and how these densities are related.

Remark. The requirements in Theorem 4.0.1 and Theorem 4.0.2 for the DFAO A are sufficient to ensure frequencies for the automatic sequence itself. On the one hand, one might relax these conditions without changing the theorem itself. On the other hand, we would like to mention that frequencies of the automatic sequence do not ensure frequencies for linear subsequences or the prime subsequence:

[1] For the definition of digital functions see Chapter 1.



Note that $a_n = b$ holds in exactly two cases:

- the sum of digits of n in base 3 is even and the first digit of n in base 3 is 2
- the sum of digits of n in base 3 is odd and the first digit of n in base 3 is 1.

We are able to reformulate this characterization as the sum of digits of n in base 3 is even if and only if n is even: we have $a_n = b$ if and only if n is even and the first digit of n in base 3 is 2 or n is odd and the first digit of n in base 3 is 1.

The corresponding automatic sequence is

$$abbbcbbcbbcbcbcbcbbcbcbcbcbbcbcbcbcbcbcbcbcbcbcbcbcb$$

One finds easily that the automatic sequence is equally distributed on $\{b, c\}$, as the generating automaton has only one final component which is primitive and symmetric. But, if we consider only odd integers we are detecting the first digit of n in base 3: the subsequence of odd integers of the automatic sequence is constant on all intervals of the form $[3^m, 2 \cdot 3^m), [2 \cdot 3^m, 3^{m+1})$. This shows that no densities exist for the subsequence of odd integers. This consideration together with the prime number theorem also shows that no densities exist along the primes.

4.1 Linear subsequences of automatic sequences

We have mentioned that it is unclear under what conditions we can expect that the frequencies of letters in subsequences of automatic sequences exist. By only considering automata that fulfill a - rather weak - technical condition, we are able to show that the frequencies of the letters for automatic subsequences along arithmetic progressions exist. We are also able to express these frequencies in terms of the naturally induced transducer and it is given by (4.2).

We use Corollary 2.5.2 and find that there exists a possible different automaton $\overline{A}$ with the same automatic sequence as A and $d(\overline{A}) = k_0(\overline{A}) = 1$. We denote by $\overline{\Sigma} = \{0, \ldots, k-1\}$ the input alphabet of $\overline{A}$ and by $\mathcal{T} := \mathcal{T}_{\overline{A}} = (Q, \overline{\Sigma}, \delta, q_0, G, \lambda)$ the naturally induced transducer corresponding to $\overline{A}$.

We need to analyze $\delta(q_0, (an+b)_k)$ and $T(q_0, (an+b)_k)$[2] originating from $\mathcal{T}$:

$$\sum_{n<N} \mathbf{1}_{[u_{an+b}=\alpha]} = \sum_{n<N} \mathbf{1}_{[\tau(\pi_1(T(q_0,(an+b)_k)\cdot\delta(q_0,(an+b)_k)))=\alpha]}$$

[2]The function T was defined at the beginning of Chapter 2.

$$= \sum_{\substack{n<aN \\ n\equiv b \bmod a}} \mathbf{1}_{[\tau(\pi_1(T(q_0,(n)_k)\cdot\delta(q_0,(n)_k)))=\alpha]} + O(b)$$
$$= \sum_{q\in Q}\sum_{\sigma\in G} \mathbf{1}_{[\tau(\pi_1(\sigma\cdot q))=\alpha]} \sum_{\substack{n<aN \\ n\equiv b \bmod a}} \mathbf{1}_{[\delta(q_0,(n)_k)=q]}\,\mathbf{1}_{[T(q_0,(n)_k)=\sigma]} + O(b).$$

We expect that the restriction for n effects $\delta(.)$ as well as $T(.)$. We rewrite $a = a'\cdot a''$ where $\gcd(a'',k)=1$ and $a' \mid k^{\lambda_1}$ for a minimal λ_1. Of course this implies that $\gcd(a',a'')=1$.

Lemma 1.4.1 shows that there exist $1>\eta>0$ such that at most $O(k^{\lambda(1-\eta)})$ words of length λ are not synchronizing. Thus, we find

$$\sum_{n<N} \mathbf{1}_{[u_{an+b}=\alpha]} = \sum_{q\in Q}\sum_{\sigma\in G} \mathbf{1}_{[\tau(\pi_1(\sigma\cdot q))=\alpha]} \sum_{m<k^\lambda} \sum_{\substack{m_1<k^{\lambda_1} \\ m_1\equiv b \bmod a'}}$$
$$\sum_{\substack{n<aN \\ n\equiv b \bmod a'' \\ n\equiv mk^{\lambda_1}+m_1 \bmod k^{\lambda+\lambda_1}}} \mathbf{1}_{[\delta(q_0,(n)_k)=q]}\,\mathbf{1}_{[T(q_0,(n)_k)=\sigma]} + O(b)$$
$$= \sum_{q\in Q}\sum_{\sigma\in G} \mathbf{1}_{[\tau(\pi_1(\sigma\cdot q))=\alpha]} \sum_{m<k^\lambda} \sum_{\substack{m_1<k^{\lambda_1} \\ m_1\equiv b \bmod a'}} \mathbf{1}_{[\delta(q_0,(mk^{\lambda_1}+m_1)_k)=q]}$$
$$\sum_{\substack{n<aN \\ n\equiv b \bmod a'' \\ n\equiv mk^{\lambda_1}+m_1 \bmod k^{\lambda+\lambda_1}}} \mathbf{1}_{[T(q_0,(n)_k)=\sigma]} + O(Nk^{-\eta\lambda}),$$

where we assume that $k^\lambda \ll N$. We start to evaluate the inner most sum,

$$\sum_{\substack{n<aN \\ n\equiv b \bmod a'' \\ n\equiv m' \bmod k^{\lambda'}}} \mathbf{1}_{[T(q_0,(n)_k)=\sigma]},$$

where we expect the frequencies to exist.

We want to use the following Lemma from representation theory.

Lemma 4.1.1. *Let G be a compact group and $\boldsymbol{\nu}$ a regular normed Borel measure in G. Then a sequence $(x_n)_{n\geq 0}$ is $\boldsymbol{\nu}$-uniformly distributed in G, i.e., $\frac{1}{N}\sum_{n<N}\delta_{x_n} \to \boldsymbol{\nu}$, if and only if*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n<N} D(x_n) = \int_G D d\boldsymbol{\nu} \tag{4.1}$$

holds for all irreducible unitary representations D of G.

A proof for Lemma 4.1.1 is given for example in [32]. Since we are working here with a finite group, we find that the error term for

$$\frac{1}{N}\sum_{n<N}\delta_{x_n} \to \boldsymbol{\nu}$$

is bounded - up to constants - by the maximal error term of (4.1). To apply Lemma 4.1.1 we start to evaluate the left side of (4.1). Therefore we have to distinguish some special representations, which we already encountered in Chapter 3.

Lemma 3.4.1. *There exist d' special 1-dimensional representations $D_0, \dots, D_{d'-1}$ defined by*

$$D_\ell(g) := \mathrm{e}\left(\frac{\ell \cdot s_0(g)}{d'}\right),$$

for $\ell = 0, \dots d'-1$, where s_0 is defined by $s_0(T(q, \mathbf{w})) = [\mathbf{w}]_k \bmod d'$.

Remark. These representations play usually a different role than the other representations.

Note that $D_\ell(T(q_0, \mathbf{w})) = \mathrm{e}\left((\ell \cdot [\mathbf{w}]_k)/d'\right)$ for $\ell = 0, \dots, d'-1, \mathbf{w} \in \Sigma^*$.

We evaluate the left-hand side of (4.1) for $D = D_\ell$:

$$\begin{aligned}
\sum_{\substack{n<N \\ n\equiv b \bmod a'' \\ n\equiv m' \bmod k^{\lambda'}}} D_\ell(T(q_0,(n)_k)) &= \sum_{\substack{n<N \\ n\equiv b \bmod a'' \\ n\equiv m' \bmod k^{\lambda'}}} \mathrm{e}\left(\frac{n\ell}{d'}\right) \\
&= \sum_{\substack{b'<\mathrm{lcm}(a'',d') \\ b'\equiv b \bmod a''}} \sum_{\substack{n<N \\ n\equiv b' \bmod lcm(a'',d') \\ n\equiv m' \bmod k^{\lambda'}}} \mathrm{e}\left(\frac{b'\ell}{d'}\right) \\
&= \sum_{\substack{b'<lcm(a'',d') \\ b'\equiv b \bmod a''}} \mathrm{e}\left(\frac{b'\ell}{d'}\right) \frac{N}{lcm(a'',d')k^{\lambda'}} + O\left(k^{\lambda'} \frac{lcm(a'',d')^2}{\gcd(a'',d')}\right) \\
&= \sum_{x<lcm(a'',d')/a''} \mathrm{e}\left(\frac{(b+x\cdot a'')\ell}{d'}\right) \frac{N}{lcm(a'',d')k^{\lambda'}} \\
&\quad + O\left(k^{\lambda'} \frac{lcm(a'',d')^2}{\gcd(a'',d')}\right) \\
&= \mathrm{e}\left(\frac{b\ell}{d'}\right) \sum_{x<d'/\gcd(a'',d')} \mathrm{e}\left(\frac{x\ell(a''/\gcd(a'',d'))}{d'/\gcd(a'',d')}\right) \frac{N}{lcm(a'',d')k^{\lambda'}} \\
&\quad + O\left(k^{\lambda'} \frac{lcm(a'',d')^2}{\gcd(a'',d')}\right) \\
&= \mathbf{1}_{[\ell\equiv 0 \bmod d'/\gcd(a'',d')]}\, \mathrm{e}\left(\frac{b\ell}{d'}\right) \frac{N}{a''k^{\lambda'}} + O\left(k^{\lambda'} \frac{lcm(a'',d')^2}{\gcd(a'',d')}\right).
\end{aligned}$$

For $D \notin \{D_0, \dots, D_{d'-1}\}$ we find a good bound by applying Theorem 3.4.2.

Let ν be the unique integer such that $k^{\nu-1} \le x < k^\nu$. We find (see for example [39, Lemma 3.7])

$$\left|\sum_{\substack{n<N \\ n\equiv b \bmod a'' \\ n\equiv m' \bmod k^{\lambda'}}} D(T(q_0,(n)_k))\right| = \left|\sum_{\substack{n<N/k^{\lambda'} \\ nk^{\lambda'}+m'\equiv b \bmod a}} D(T(q_0,(nk^{\lambda'}+m')_k))\right| + O(k^{\lambda'})$$

$$
\begin{aligned}
&= \left| \sum_{n<N/k^{\lambda'}} \frac{1}{a''} \sum_{h<a''} \mathrm{e}\left(\frac{h(nk^{\lambda'}+m'-b)}{a''}\right) D(T(q_0,(nk^{\lambda'}+m')_k)) \right| \\
&\quad + O(k^{\lambda'}) \\
&\leq \frac{1}{a''} \sum_{h<a''} (\nu-\lambda') \sup_{\theta\in\mathbb{R}} \left| \sum_{n<k^{\nu-\lambda'}} \mathrm{e}(n\theta) D(T(q_0,(nk^{\lambda'}+m')_k)) \right| \\
&\quad + O(k^{\lambda'}) \\
&\ll (\nu-\lambda')k^{(\nu-\lambda')(1-\eta)} + O(k^{\lambda'}).
\end{aligned}
$$

In total, we find

$$
\lim_{N\to\infty} \frac{a''k^{\lambda'}}{N} \sum_{\substack{n<N \\ n\equiv b \bmod a'' \\ n\equiv m' \bmod k^{\lambda'}}} D(T(q_0,(n)_k)) = \begin{cases} \mathrm{e}\left(\frac{b\ell}{d'}\right), & \text{for } D=D_\ell \text{ and } \ell \equiv 0 \bmod d'/\gcd(d',a'') \\ 0, & \text{otherwise} \end{cases}
$$

It remains to find the correct measure $\boldsymbol{\nu}$. Therefore, we define a function $f : G \to \mathbb{C}$ as follows

$$
f(g) = \sum_{x<\gcd(d',a'')} \mathrm{e}\left(-\frac{bx}{\gcd(d',a'')}\right) D_{x(d'/\gcd(d',a''))}(g).
$$

We compute $f(g)$ now for some $g \in G_\ell = \{g \in G : s_0(g) = \ell\}$.

$$
\begin{aligned}
f(g) &= \sum_{x<\gcd(d',a'')} \mathrm{e}\left(-\frac{bx}{\gcd(d',a'')}\right) \mathrm{e}\left(\frac{\ell x}{\gcd(d',a'')}\right) \\
&= \sum_{x<\gcd(d',a'')} \mathrm{e}\left(\frac{x(\ell-b)}{\gcd(d',a'')}\right) \\
&= \gcd(d',a'')\, \mathbf{1}_{[\ell\equiv b \bmod \gcd(d',a'')]}.
\end{aligned}
$$

Thus we can define

$$
d\boldsymbol{\nu} = f d\boldsymbol{\mu},
$$

where $\boldsymbol{\mu}$ denotes the *Haar*-measure of G.

Let $\{D^\alpha = (d^\alpha_{ij})_{i,j\leq n_\alpha}, \alpha \in \mathcal{A}\}$ be a complete set of pairwise inequivalent irreducible unitary representations and set $e^\alpha_{ij}(g) = \sqrt{n_\alpha} d^\alpha_{ij}(g)$. Note that $\mathcal{A}$ is finite because G is a finite group. Recall that the set $\{e^\alpha_{ij}\}$ forms a complete orthonormal system in the Hilbert space $L^2(G)$. We obtain for $D_{x(d'/\gcd(d',a''))}$ that

$$
\int_G D_{x(d'/\gcd(d',a''))} f d\boldsymbol{\mu} = \sum_{x'<\gcd(d',a'')} \mathrm{e}\left(\frac{bx'}{\gcd(d',a'')}\right) < D_{x(d'/\gcd(d',a''))} | \overline{D_{x'(d'/\gcd(d',a''))}} >
$$

$$= \mathrm{e}\left(\frac{bx}{\gcd(d', a'')}\right).$$

For all other representations $D^\alpha = (d^\alpha_{ij})_{1\le i,j\le n_\alpha}$, we find

$$\int_G d^\alpha_{ij} f d\boldsymbol{\nu} = \sum_{x'<\gcd(d',a'')} \mathrm{e}\left(\frac{bx'}{\gcd(d', a'')}\right) < d^\alpha_{ij} | \overline{D_{x(d'/\gcd(d',a''))}} >= 0.$$

This proves that $(T(q_0,(n)_k))_{n\in\mathbb{N}(m',k^{\lambda_1})\cap\mathbb{N}(b,a'')}$[3] is $\boldsymbol{\nu}$-uniformly distributed. $\boldsymbol{\nu}$ does not depend on m' and the error term is of form $O((N/k^{\lambda'})^{1-\eta_1})$.

Consequently, we find

$$\begin{aligned}
\sum_{n<N} \mathbf{1}_{[u_{an+b}=\alpha]} &= \sum_{q\in Q}\sum_{\sigma\in G} \mathbf{1}_{[\tau(\pi_1(\sigma\cdot q))=\alpha]} \sum_{m<k^\lambda} \sum_{\substack{m_1<k^{\lambda_1}\\ m_1\equiv b \bmod a'}} \mathbf{1}_{[\delta(q_0,(mk^{\lambda_1}+m_1)_k)=q]} \\
&\qquad \sum_{\substack{n<aN\\ n\equiv b \bmod a''\\ n\equiv mk^{\lambda_1}+m_1 \bmod k^{\lambda+\lambda_1}}} \mathbf{1}_{[T(q_0,(n)_k)=\sigma]} + O(Nk^{-\eta\lambda}) \\
&= \sum_{q\in Q}\sum_{\sigma\in G} \mathbf{1}_{[\tau(\pi_1(\sigma\cdot q))=\alpha]} \sum_{m<k^\lambda} \sum_{\substack{m_1<k^{\lambda_1}\\ m_1\equiv b \bmod a'}} \mathbf{1}_{[\delta(q_0,(mk^{\lambda_1}+m_1)_k)=q]} \\
&\qquad \frac{\gcd(d', a'')}{|G|} \mathbf{1}_{[s_0(\sigma)\equiv b \bmod \gcd(d',a'')]} \frac{aN}{a''k^{\lambda+\lambda_1}} \\
&\qquad + O(k^{\lambda+\lambda_1}(N/k^{\lambda+\lambda_1})^{1-\eta_1}) + O(Nk^{-\eta\lambda}) \\
&= N\sum_{q\in Q} \frac{\gcd(d', a'')}{|G|} \sum_{\substack{\sigma\in G\\ s_0(\sigma)\equiv b \bmod \gcd(d',a'')}} \mathbf{1}_{[\tau(\pi_1(\sigma\cdot q))=\alpha]} \sum_{\overline{q}\in Q} \frac{1}{k^\lambda} \sum_{m<k^\lambda} \mathbf{1}_{[\delta(q_0,(m)_k)=\overline{q}]} \\
&\qquad \frac{a'}{k^{\lambda_1}} \sum_{\substack{m_1<k^{\lambda_1}\\ m_1\equiv b \bmod a'}} \mathbf{1}_{[\delta(\overline{q},(m_1)_k)=q]} + O(k^{\lambda+\lambda_1}(N/k^{\lambda+\lambda_1})^{1-\eta_1}) + O(Nk^{-\eta\lambda})
\end{aligned}$$

Theorem 1.4.2 shows that the frequencies of every letter in $\delta(q_0,(m)_k)$ exists - where no explicit error term is given. However, we use the stronger version (1.2).

Furthermore, we denote by

$$f^{b,a'}_{\overline{q},q} = \frac{a'}{k^{\lambda_1}} \sum_{\substack{m_1<k^{\lambda_1}\\ m_1\equiv b \bmod a'}} \mathbf{1}_{[\delta(\overline{q},(m_1)_k)=q]}$$

the densities of the transition restricted to $m_1 \equiv b \bmod a'$. We find

$$\sum_{n<N} \mathbf{1}_{[u_{an+b}=\alpha]} = N\frac{\gcd(d', a)}{|G|} \sum_{\substack{\sigma\in G\\ s_0(\sigma)\equiv b \bmod \gcd(d',a)}} \sum_{q\in Q} \mathbf{1}_{[\tau(\pi_1(\sigma\cdot q))=\alpha]} \sum_{\overline{q}\in Q} f_{\overline{q}} f^{b,a'}_{\overline{q},q}$$

[3] $\mathbb{N}(x,y)$ denotes the set $\{n\in\mathbb{N} : n\equiv x \bmod y\}$.

$$+ O(k^{\lambda+\lambda_1}(N/k^{\lambda+\lambda_1})^{1-\eta_1}) + O(Nk^{-\eta\lambda}).$$

This identity is not particularly surprising. We have already seen that $n \bmod d'$ influences which elements of G can appear. If we restrict ourselves to $n \equiv b \bmod a$, we also induce restrictions for $n \bmod \gcd(a, d')$. This corresponds to the term $\frac{\gcd(d',a)}{|G|} \sum_{\substack{\sigma\in G \\ s_0(\sigma)\equiv b \bmod \gcd(d',a)}}$.

The condition for $n \bmod a'$ gives us information for the last few steps, which accounts for the term $\sum_{\overline{q}\in Q} f_{\overline{q}} f_{\overline{q},q}^{b,a'}$.

If we choose now $\lambda = \lfloor \log_k(N)/2 \rfloor$ we find in total

$$\sum_{n<N} \mathbf{1}_{[u_{an+b}=\alpha]} = N\frac{\gcd(d',a)}{|G|} \sum_{\substack{\sigma\in G \\ s_0(\sigma)\equiv b \bmod \gcd(d',a)}} \sum_{q\in Q} \mathbf{1}_{[\tau(\pi_1(\sigma\cdot q))=\alpha]} \sum_{\overline{q}\in Q} f_{\overline{q}} f_{\overline{q},q}^{b,a'} + O(N^{1-\eta'}) \tag{4.2}$$

for some $\eta' > 0$.

4.2 Prime Number Theorem for Automatic Sequences

In this section we develop a Prime Number Theorem for a large class of automatic sequences - i.e. we estimate the sum $\sum_{n<N} \Lambda(n)a_n$. This gives then - by partial summation - information about the prime subsequence of $\mathbf{a}$, i.e. $(a_p)_{p\in\mathbb{P}}$. More precisely, we show that the frequencies of letters in this subsequence exist and are also able to derive them. We use the same notations as in the previous section and some similar ideas. We have seen that representations play an important role. For the prime number theorem we have to distinguish again the special representations D_ℓ. We recall the most important facts about them.
We have defined in Chapter 2, $G_\ell := \{g \in G : s_0(g) = \ell\}$ for $\ell = 0, \dots, d'(A) - 1$ and found the following d' special 1-dimensional representations $D_0, \dots, D_{d'}$ defined by

$$D_\ell(g) := \mathrm{e}\left(\frac{\ell \cdot s_0(g)}{d'}\right),$$

for $\ell = 0, \dots, d' - 1$ and s_0 is defined by $s_0(T(q, \mathbf{w})) = [\mathbf{w}]_k \bmod d'$.

Note that $D_\ell(T(q_0, \mathbf{w})) = \mathrm{e}\left((\ell \cdot [\mathbf{w}]_k)/d'\right)$ for $\ell = 0, \dots, d' - 1, \mathbf{w} \in \Sigma^*$.

In Section 4.3, we will prove the following Proposition.

Proposition 4.2.1. *Let D be a unitary, irreducible representations of G different from $D_\ell, 0 \le \ell < d'$. There exists some $\eta > 0$ such that*

$$\left\| \frac{1}{k^\nu} \sum_{p<k^\nu} D(T(q_0, (p)_k)) \,\mathrm{e}(-pt) \right\|_2 \ll k^{-\nu\eta}$$

holds uniformly in $t \in \mathbb{R}$.

The rest of this section is devoted to prove

Proposition 4.2.2. *Suppose Proposition 4.2.1 holds. Then Theorem 4.0.2 holds.*

Proof. We start as in the previous section and find by using Corollary 2.5.2 and Proposition 2.2.2 that

$$\frac{1}{\pi(x)}\sum_{p\le x}\mathbf{1}_{[a_p=b]} = \frac{1}{\pi(x)}\left(\sum_{\substack{a\le k^\lambda\\(a,k^\lambda)=1}}\sum_{\substack{p\le x\\ p\equiv a \bmod k^\lambda}}\mathbf{1}_{[a_p=b]}+O(k^\lambda)\right)$$
$$= \frac{1}{\pi(x)}\left(\sum_{\substack{a\le k^\lambda\\(a,k^\lambda)=1}}\sum_{\substack{p\le x\\ p\equiv a \bmod k^\lambda}}\mathbf{1}_{[\tau(\pi_1(T(q_0,(p)_k)\cdot\delta(q_0,(p)_k)))=b]}+O(k^\lambda)\right).$$

We note that $(a,k^\lambda)=1$ holds if and only if $(\varepsilon_0(a),k)=1$ where $\varepsilon_0(x)$ denotes the least significant digit of x in base k. We denote again by $M_\lambda := \{n<k^\lambda : (n,k)=1 \text{ and } (n)_k \text{ is synchronizing}\}$. One finds easily that $|\{n\le k^\lambda : (n,k)=1, n\notin M_\lambda\}| = O(k^{\lambda(1-\eta)})$ for some $\eta>0$, as in the previous section.

We fix λ for now to find the necessary estimates. Thereafter we let λ grow. This then gives the desired result.

$$\frac{1}{\pi(x)}\sum_{p\le x}\mathbf{1}_{[a_p=b]} = \frac{1}{\pi(x)}\sum_{a\in M_\lambda}\sum_{\substack{p\le x\\ p\equiv a \bmod k^\lambda}}\mathbf{1}_{[\tau(\pi_1(T(q_0,(p)_k)\cdot\delta(q_0,(p)_k)))=b]}$$
$$+\frac{\pi(x;1,k^\lambda)}{\pi(x)}O(k^{\lambda(1-\eta)})+O\left(\frac{k^\lambda}{\pi(x)}\right)$$
$$= \frac{1}{\pi(x)}\sum_{a\in M_\lambda}\sum_{\substack{p\le x\\ p\equiv a \bmod k^\lambda}}\mathbf{1}_{[\tau(\pi_1(T(q_0,(p)_k)\cdot\delta(q_0,(a)_k)))=b]}+O(k^{-\lambda\eta})+O\left(\frac{k^\lambda}{\pi(x)}\right)$$
$$= \frac{1}{\pi(x)}\sum_{q\in Q}\sum_{\substack{a\in M_\lambda\\ \delta(q_0,(a)_k)=q}}\sum_{\substack{p\le x\\ p\equiv a \bmod k^\lambda}}\mathbf{1}_{\tau([\pi_1(T(q_0,(p)_k)\cdot q))=b]}+O(k^{-\lambda\eta})+O\left(\frac{k^\lambda}{\pi(x)}\right)$$
$$= \frac{1}{\pi(x)}\sum_{q\in Q}\sum_{\substack{a<k^\lambda\\(a,k)=1}}\mathbf{1}_{[\delta(q_0,(a)_k)=q]}\sum_{\substack{p\le x\\ p\equiv a \bmod k^\lambda}}\mathbf{1}_{[\tau(\pi_1(T(q_0,(p)_k)\cdot q))=b]}$$
$$+O(k^{-\lambda\eta})+O\left(\frac{k^\lambda}{\pi(x)}\right).$$

Thus we are interested in the distribution of $T(q_0,(p)_k)$ along primes in arithmetic progressions. We use the following result which is rather technical. We postpone the proof to Section 4.3.

Lemma 4.2.3. *Let $\lambda \in \mathbb{N}$, $\mathcal{T}$ be a naturally induced transducer with function T. For every $g \in G$ exists f_g such that*

$$\frac{1}{\pi(x;a,k^\lambda)} \sum_{\substack{p<x \\ p\equiv a \bmod k^\lambda}} \mathbf{1}_{[T(q_0,(p)_k)=g]} = f_g + o(1) \text{ for } x \to \infty$$

holds (uniformly) for all $a < k^\lambda$ such that $(a,k) = 1$.

By writing

$$\sum_{\substack{p\le x \\ p\equiv a \bmod k^\lambda}} \mathbf{1}_{[\tau(\pi_1(T(q_0,(p)_k)\cdot q))=b]} = \sum_{g\in G} \mathbf{1}_{[\pi_1(g\cdot q)=b]} \sum_{\substack{p\le x \\ p\equiv a \bmod k^\lambda}} \mathbf{1}_{[T(q_0,(p)_k)=g]}$$

we find that

$$\sum_{\substack{p\le x \\ p\equiv a \bmod k^\lambda}} \mathbf{1}_{[\tau(\pi_1(T(q_0,(p)_k)\cdot q))=b]} = \pi(x;1,k^\lambda) f_{q,b} + o(\pi(x;1,k^\lambda))$$

holds, where

$$f_{q,b} = \sum_{g\in G} f_g \, \mathbf{1}_{[\tau(\pi_1(g\cdot q))=b]} \,.$$

Therefore, we are interested in the quantity

$$\frac{1}{\varphi(k^\lambda)} \left|\{a < k^\lambda : (\varepsilon_0(a),k) = 1, \delta(q_0,(a)_k) = q\}\right|.$$

We want to show that it has a limit (for $\lambda \to \infty$). We start by rewriting it as

$$\begin{aligned}
&\frac{1}{\varphi(k)} \sum_{\substack{a_0<k \\ (a_0,k)=1}} \frac{1}{k^{\lambda-1}} \left|\{a < k^{\lambda-1} : \delta(q_0,(a\cdot k + a_0)_k) = q\}\right| \\
&\quad = \frac{1}{\varphi(k)} \sum_{\substack{a_0<k \\ (a_0,k)=1}} \sum_{q'\in Q} \mathbf{1}_{[\delta(q',a_0)=q]} \frac{1}{k^{\lambda-1}} \left|\{a < k^{\lambda-1} : \delta(q_0,(a)_k) = q'\}\right|.
\end{aligned}$$

Thus it is sufficient to show that $\frac{1}{k^{\lambda-1}}|\{a < k^{\lambda-1} : \delta(q_0,(a)_k) = q'\}|$ has a limit. However, $\mathcal{T}$ is synchronizing and strongly connected and therefore primitive. It is a well-known result that the densities exist in this setting, see for example Theorem 1.3.5. This allows us to write

$$|\{a < k^\lambda : (\varepsilon_0(a),k) = 1, \delta(q_0,(a)_k) = q\}| = \varphi(k^\lambda) f_q + o(\varphi(k^\lambda)),$$

where f_q does not depend on λ. Thus we find in total

$$\frac{1}{\pi(x)} \sum_{p\le x} \mathbf{1}_{[a_p=b]} = \frac{1}{\pi(x)} \sum_{q\in Q} (\varphi(k^\lambda) f_q + o(\varphi(k^\lambda)))(\pi(x;1,k^\lambda) f_{q,b} + o(\pi(x;1,k^\lambda))) + O(k^{-\lambda\eta})$$

$$= \sum_{q \in Q} f_q f_{q,b} + \frac{1}{\pi(x;1,k^\lambda)} o(\pi(x;1,k^\lambda)) + \frac{1}{\varphi(k^\lambda)} o(\varphi(k^\lambda)) + O(k^{-\lambda\eta}).$$

To show that the sum of the three error terms is smaller than ε, we choose λ such that the second and third error term are bounded by $\varepsilon/3$. Then we find for x large enough that the first error term (for this given λ) is also bounded by $\varepsilon/3$ (note that $k^{\lambda-1} \leq \varphi(k^\lambda) \leq k^\lambda$). Consequently we find

$$\lim_{x\to\infty} \frac{1}{\pi(x)} \sum_{p \leq x} \mathbf{1}_{[a_p = b]} = \sum_{q \in Q} f_q f_{q,b},$$

which – finally – finishes the proof of Proposition 4.2.2. □

Remark. The error terms can be made explicit. We find for example that the error term in Lemma 4.2.3 is actually of the form $O(k^{-\lambda\eta'})$. The dominant error term seems to correspond to the distribution of primes in arithmetic progressions.

Remark. As mentioned earlier, Theorem 4.0.2 covers block-additive functions. However, to find the corresponding result for block-additive functions the group structure degenerates to an additive structure and many arguments simplify (e.g. one does not need to use group representations).

4.3 Technical Results

Proof of Proposition 4.2.1: By classical partial summation one finds the following well known result (see for example [37, Lemma 11]).

If $f : \mathbb{N} \to \mathbb{C}$ is such that $|f(n)| \leq 1$ for all $n \in \mathbb{N}$ then

$$\left| \sum_{p \leq x} f(p) \right| \leq \frac{2}{\log x} \max_{t \leq x} \left| \sum_{n \leq t} \Lambda(n) f(n) \right| + O(\sqrt{x}). \tag{4.3}$$

This result can easily be adopted to matrix valued functions. Thus we find

$$\left\| \frac{1}{\pi(k^\nu)} \sum_{p < k^\nu} D(T(q_0,(p)_k)) \, \mathrm{e}(pt) \right\|_F \ll \frac{1}{\log k^\nu} \frac{1}{\pi(k^\nu)} \max_{t \leq k^\nu} \left\| \sum_{n \leq t} \Lambda(n) D(T(q_0,(n)_k)) \, \mathrm{e}(nt) \right\|_F + O(k^{-\nu/2})$$

As $D(T(n))$ fulfills Definition 3.3.1 and Definition 3.3.2, we can apply Theorem 3.3.3 and find that the right hand side is asymptotically bounded by

$$\max_{t \leq k^\nu} t^{-\eta'} + O(k^{-\nu/2})$$

for some $\eta' > 0$ – if we only choose η' strictly smaller than η to remove the other terms. This gives directly the desired result. □

4.3.1 Frequency of T along primes

We show in this subsection that the frequencies of the sequence $(T(q_0,(p)_k))_{p\in\mathcal{P}(a,k^\lambda)}$ exist and are independent of a and λ.

Lemma 4.2.3. *Let $\lambda\in\mathbb{N}$, $\mathcal{T}$ be a naturally induced transducer with function T. For every $g\in G$ exists f_g such that*

$$\frac{1}{\pi(x;a,k^\lambda)}\sum_{\substack{p<x\\ p\equiv a \bmod k^\lambda}}\mathbf{1}_{[T(q_0,(p)_k)=g]}=f_g+o(1)\ \textit{for}\ x\to\infty$$

holds (uniformly) for all $a<k^\lambda$ such that $(a,k)=1$.

Proof. We use again Lemma 4.1.1.

The sequence we are concerned about is $(T(q_0,(p)_k))_{p\in\mathcal{P}(a,k^\lambda)}$.

We start by computing the left hand side of (4.1) for $D=D_\ell$ for our specific sequence. Recall that $d'|k-1$ holds.

$$\begin{aligned}
&\lim_{x\to\infty}\frac{1}{\pi(x;a,k^\lambda)}\sum_{\substack{p<x\\ p\equiv a \bmod k^\lambda}}D_\ell(T(q_0,(p)_k))\\
&\quad=\lim_{x\to\infty}\frac{1}{\pi(x;a,k^\lambda)}\sum_{\substack{p<x\\ p\equiv a \bmod k^\lambda}}\mathrm{e}(\ell p/d')\\
&\quad=\lim_{x\to\infty}\frac{1}{\pi(x;a,k^\lambda)}\sum_{b<d'}\mathrm{e}(\ell b/d')\sum_{\substack{p<x\\ p\equiv a \bmod k^\lambda\\ p\equiv b \bmod d'}}1\\
&\quad=\frac{1}{\varphi(d')}\sum_{\substack{b<d'\\ (b,d')=1}}\mathrm{e}(\ell b/d')\\
&\quad=\frac{1}{\varphi(d')}S(\ell,0;d').
\end{aligned}$$

Here $S(a,b;c)$ denotes the Kloosterman (or Ramanujan) sum, defined by

$$S(a,b;c):=\sum_{\substack{x<c\\ (x,c)=1}}\mathrm{e}\left(\frac{ax+b\bar{x}}{c}\right),$$

where $\bar{x}$ denotes the multiplicative inverse of x modulo c.

Furthermore, we want to show that

$$\lim_{x\to\infty}\frac{1}{\pi(x;a,k^\lambda)}\sum_{\substack{p<x\\ p\equiv a \bmod k^\lambda}}D(T(q_0,(p)_k))=0$$

holds for $D \notin \{D_\ell : \ell = 0, \dots, d'-1\}$. We use a standard technique of analytic number theory to restrict ourselves to sums with $x = k^\nu$. Let ν be the unique integer such that $k^{\nu-1} \le x < k^\nu$. We find (see for example [39, Lemma 3.7])

$$\left\| \frac{1}{\pi(x;a,k^\lambda)} \sum_{\substack{p\le x \\ p\equiv a \bmod k^\lambda}} D(T(q_0,(p)_k)) \right\|_F \ll \nu \sup_{\theta\in\mathbb{R}} \left\| \frac{1}{\pi(x;a,k^\lambda)} \sum_{\substack{p\le k^\nu \\ p\equiv a \bmod k^\lambda}} \mathrm{e}(\theta p) D(T(q_0,(p)_k)) \right\|_F .$$

We can rewrite the condition $p \equiv a \bmod k^\lambda$ using exponential sums and find that the right hand side is bounded by

$$\begin{aligned}
&\nu \frac{1}{\pi(k^\nu;a,k^\lambda)} \sup_{\theta\in\mathbb{R}} \left\| \frac{1}{k^\lambda} \sum_{h<k^\lambda} \sum_{p\le k^\nu} \mathrm{e}\left(h\frac{p-a}{k^\lambda}\right) \mathrm{e}(\theta p) D(T(q_0,(p)_k)) \right\|_F \\
&\quad \le \nu \frac{1}{\pi(k^\nu;a,k^\lambda)} \sup_{\theta\in\mathbb{R}} \max_{h<k^\lambda} \left\| \sum_{p\le k^\nu} \mathrm{e}\left(p\left(\frac{h}{k^\lambda}+\theta\right)\right) D(T(q_0,(p)_k)) \right\|_F \\
&\quad \le \nu \frac{1}{\pi(k^\nu;a,k^\lambda)} \sup_{\theta'\in\mathbb{R}} \left\| \sum_{p\le k^\nu} \mathrm{e}(p\theta') D(T(q_0,(p)_k)) \right\|_F
\end{aligned}$$

Thus we find – by using Proposition 4.2.1 – that

$$\lim_{x\to\infty} \frac{1}{\pi(x;a,k^\lambda)} \sum_{\substack{p<x \\ p\equiv a \bmod k^\lambda}} D(T(q_0,(p)_k)) = \begin{cases} \frac{1}{\varphi(d')} S(\ell,0;d'), & \text{for } D = D_\ell \\ 0, & \text{otherwise,} \end{cases}$$

We define the function f by

$$\begin{aligned}
f(g) &:= \frac{d'}{\varphi(d')} + \frac{d'}{\varphi(d')} \overline{S(1,0;d')} D_1(g) + \ldots + \frac{d'}{\varphi(d')} \overline{S(d'-1,0;d')} D_{d'-1}(g) \\
&= \sum_{\ell=0}^{d'-1} \frac{d'}{\varphi(d')} \overline{S(\ell,0;d')} D_\ell(g).
\end{aligned}$$

For $v < d'$, we find:

$$\begin{aligned}
f(T(nd'+v)) &= \frac{d'}{\varphi(d')} \sum_{\ell=0}^{d'-1} \overline{S(\ell,0;d')}\, \mathrm{e}\left(\frac{\ell v}{d'}\right) \\
&= \frac{d'}{\varphi(d')} \sum_{\ell=0}^{d'-1} \sum_{\substack{x<d' \\ (x,d')=1}} \mathrm{e}\left(-\frac{\ell x}{d'}\right) \mathrm{e}\left(\frac{\ell v}{d'}\right) \\
&= \frac{d'}{\varphi(d')} \sum_{\substack{x<d' \\ (x,d')=1}} \sum_{\ell=0}^{d'-1} \mathrm{e}\left(\ell\frac{v-x}{d'}\right)
\end{aligned}$$

$$= \frac{d'}{\varphi(d')} \sum_{\substack{x<d' \\ (x,d')=1}} \mathbf{1}_{[x=v]}\,.$$

In total we find

$$f(T(nd'+v)) = \begin{cases} \frac{d'}{\varphi(d')}, & \text{for } (v,d')=1 \\ 0, & \text{otherwise,} \end{cases}$$

i.e. f is a positive function. We can actually rewrite this by using Theorem 2.4.1:

$$f(g) = \begin{cases} \frac{d'}{\varphi(d')}, & \text{for } (s_0(g),d')=1 \\ 0, & \text{otherwise,} \end{cases}\quad .$$

Thus we can define

$$d\boldsymbol{\nu} = f d\boldsymbol{\mu},$$

where $\boldsymbol{\mu}$ denotes the Haar-measure of G.

Our goal is to show that $(T(q_0,(p)_k))_{p\in\mathcal{P}(a,k^\lambda)}$ is $\boldsymbol{\nu}$-uniformly distributed in G.

Let $\{D^\alpha = (d^\alpha_{ij})_{i,j\le n_\alpha}, \alpha\in\mathcal{A}\}$ be a complete set of pairwise inequivalent irreducible unitary representations and set $e^\alpha_{ij}(g) = \sqrt{n_\alpha} d^\alpha_{ij}(g)$. Note that A is finite as we are actually considering a finite group G. Recall that the set $\{e^\alpha_{ij}\}$ forms a complete orthonormal system in the Hilbert space $L^2(G)$. We obtain for D_ℓ that

$$\int_G D_\ell f d\boldsymbol{\nu} = \frac{1}{\varphi(d')} \sum_{m<d'} S(m,0;d') < D_\ell | \overline{D_m} >= \frac{1}{\varphi(d')} S(\ell,0;d').$$

For all other representations $D^\alpha = (d^\alpha_{ij})_{1\le i,j\le n_\alpha}$, we find

$$\int_G d^\alpha_{ij} f d\boldsymbol{\nu} = \frac{1}{\varphi(d')} \sum_{m<d'} S(m,0;d') < d^\alpha_{ij} | \overline{D_m} >= 0.$$

This proves that $(T(q_0,(p)_k))_{p\in\mathcal{P}(a,k^\lambda)}$ is $\boldsymbol{\nu}$-uniformly distributed, where $\boldsymbol{\nu}$ does not depend on a. □

Chapter 5

Normality of digital sequences along squares

The goal of this chapter is to give new examples of deterministic sequences - actually even almost periodic sequences - whose subsequence along the squares is normal. The result also gives many new examples of normal numbers in any given base.

The results of this chapter are a continuation of a work by Drmota, Mauduit and Rivat [14], which we discussed in Section 1.4.

We have already discussed in Chapter 1 that the subword complexity of automatic sequences, i.e. the number of different blocks of length n that appear within the sequence, grows at most linearly (this was Theorem 1.4.7).

For a random sequence $\mathbf{u} \in \{0,1\}^{\mathbb{N}}$ one finds that $p_{\mathbf{u}}(n) = 2^n$. Thus, we see that automatic sequences are far from being random.

However, the situation changes completely when one considers the subsequence along squares. Drmota, Mauduit and Rivat gave a first example for that phenomenon. They considered the Thue–Morse sequence $(t_n)_{n\geq 0}$ and showed that not only $p_{(t_{n^2})_{n\geq 0}}(L) = 2^L$, but were able to prove how often such a block appears.

$$\lim_{N\to\infty} \frac{1}{N} \#\{i < N : t_{i^2} = b_0, \ldots, t_{(i+L-1)^2} = b_{L-1}\} = \frac{1}{2^L},$$

where $(t_n)_{n\in\mathbb{N}}$ denotes the classical Thue-Morse sequence. We go here a step further and show a similar result for digital sequences by the same method.

We use here a slightly different definition of digital functions.

Definition 5.0.1. We call a function $b : \mathbb{N} \to \mathbb{N}$ a *strongly block-additive q-ary function* or also *digital function* if and only if there exist $m \in \mathbb{N}_{>0}$ and $F : \{0, \ldots, q-1\}^m \to \mathbb{N}$ such that

$$b(n) = \sum_{j\in\mathbb{Z}} F(\varepsilon_{j+m-1}(n), \ldots, \varepsilon_j(n))^1.$$

A quite prominent examples of strongly block-additive functions is probably the sum of digits function in base q, $s_q(n)$. This is a strongly block-additive function with $m = 1$ and $F(x) = x$. Another prominent example is the Rudin-Shapiro sequence which is given by $q = 2, m = 2$ and $F(x, y) = x \cdot y$ - i.e. it counts the number of blocks of the form "11" in the digital expansion in base 2.

5.1 Outline

The goal of this chapter is to give a proof of the following theorem.

Theorem 5.1.1. *Let b be a strongly block-additive function and $m' \in \mathbb{N}$ with $\gcd(q-1, m') = 1$. Then $(b(n^2) \bmod m')_{n\in\mathbb{N}}$ is normal i.e. every sub-sequence of length k appears with asymptotic frequency $(m')^{-k}$.*

Further, we also fix an arbitrary strongly block-additive function b - and therefore q and m - and m' satisfying $\gcd(q-1, m') = 1$.

In order to prove our main result, we will work with exponential sums. We present here the main theorem on exponential sums which we will prove throughout this chapter and further show its connection to Theorem 5.1.1.

Theorem 5.1.2. *For any integer $k \geq 1$ and $(\alpha_0, \dots, \alpha_{k-1}) \in \{\frac{0}{m'}, \dots, \frac{m'-1}{m'}\}^k$ such that $(\alpha_0, \dots, \alpha_{k-1}) \neq (0, \dots, 0)$, there exists $\eta > 0$ such that*

$$S_0 = \sum_{n<N} \mathrm{e}\left(\sum_{\ell=0}^{k-1} \alpha_\ell b((n+\ell)^2)\right) \ll N^{1-\eta}. \tag{5.1}$$

Lemma 5.1.3. *Theorem 5.1.2 implies Theorem 5.1.1.*

Proof. Let $(c_0, \dots, c_{k-1}) \in \{0, \dots, m'-1\}^k$ be an arbitrary sequence of length k. We count the number of occurrences of this sequence in $(b(n^2) \bmod m')_{n\leq N}$. Assuming that (5.1) holds we obtain by using the well known identity – $\sum_{n=0}^{m'-1} \mathrm{e}(\frac{n}{m'}\ell) = m'$ for $\ell \equiv 0 \bmod m'$ and 0 otherwise – that

$$\begin{aligned}
&\left|\{n < N : (b(n^2) \bmod m', \dots, b((n+k-1)^2) \bmod m') = (c_0, \dots, c_{k-1})\}\right| \\
&= \sum_{n<N} \mathbf{1}_{[b(n^2)\equiv c_0 \bmod m']} \cdots \mathbf{1}_{[b((n+k-1)^2)\equiv c_{k-1} \bmod m']} \\
&= \sum_{n<N} \prod_{\ell=0}^{k-1} \frac{1}{m'} \sum_{\alpha'_\ell=0}^{m'-1} \mathrm{e}\left(\frac{\alpha'_\ell}{m'}\left(b((n+\ell)^2) - c_\ell\right)\right)
\end{aligned}$$

$$= \frac{1}{(m')^k} \sum_{(\alpha'_0,\dots,\alpha'_{k-1})\in\{0,\dots,m'-1\}^k} \mathrm{e}\left(-\frac{\alpha'_0 c_0 + \cdots + \alpha'_{k-1} c_{k-1}}{m'}\right) \sum_{n<N} \mathrm{e}\left(\sum_{\ell=0}^{k-1} \underbrace{\frac{\alpha'_\ell}{m'}}_{=:\alpha_\ell} b((n+\ell)^2)\right)$$

[1] We define $\varepsilon_{-i}(n) = 0$ for all $i \geq 1$.

$$= \frac{N}{(m')^k} + \mathcal{O}\left(N^{1-\eta}\right)$$

with the same $\eta > 0$ as in Theorem 5.1.2. To obtain the last equality we separate the term with $(\alpha'_0, \ldots, \alpha'_{k-1}) = (0, \ldots, 0)$. □

It remains to show Theorem 5.1.2.

The structure of the full proof of Theorem 5.1.2 is presented below.

In Section 5.3, we derive the main ingredients of the proof of Theorem 5.1.2 which are upper bounds on the Fourier terms

$$H^I_\lambda(h,d) = \frac{1}{q^\lambda} \sum_{0 \le u < q^\lambda} \mathrm{e}\left(\sum_{\ell=0}^{k-1} \alpha_\ell b_\lambda(u + \ell d + i_\ell) - hq^{-\lambda} \right),$$

where $I = (i_0, \ldots, i_{k-1}) \in \mathbb{N}^k$ with some special properties defined in Section 5.3.2 and b_λ is a truncated version of b which is properly defined in Definition 5.2.1.

The main results of Section 5.3 are Propositions 5.3.7 and 5.3.8. Proposition 5.3.7 yields a bound on averages of Fourier transforms and Proposition 5.3.8 yields a uniform bound on Fourier transforms which is harder to prove.
In Section 5.4, we present some auxiliary results also used in [14].

In Section 5.5, we complete the proof for Theorem 5.1.2. We use Van-der-Corput-like inequalities in order to reduce our problem to sums depending only on few digits of $n^2, (n+1)^2, \ldots, (n+k-1)^2$. By detecting these few digits, we are able to remove the quadratic terms, which allows a proper Fourier analytic treatment. After the Fourier analysis, the remaining sum is split into two sums. The first sum involves quadratic exponential sums. The result from Section 5.4.2 allows us to find a proper bound here.

The Fourier terms $H^I_\lambda(h,d)$ appear in the second sum and Propositions 5.3.7 and 5.3.8 will provide the necessary bounds.

For the proof of the main theorem we have to distinguish the cases $K = \alpha_0 + \cdots + \alpha_{k-1} \in \mathbb{Z}$ and $K \notin \mathbb{Z}$. Sections 5.5.1 and 5.5.2 tackle one of these cases each. In Section 5.5.1, we prove that – if $K \in \mathbb{Z}$ – we deduce Theorem 5.1.2 from Proposition 5.3.7. For $K \notin \mathbb{Z}$, Section 5.5.2 shows that we can deduce Theorem 5.1.2 from Proposition 5.3.8.

5.2 Digital Functions

Definition 5.2.1. we define the truncated function b_λ and the two-fold restricted function $b_{\mu,\lambda}$ by

$$b_\lambda(n) = \sum_{j<\lambda} F(\varepsilon_{j+m-1}(n), \ldots, \varepsilon_j(n)) \text{ and } b_{\mu,\lambda}(n) = b_\lambda(n) - b_\mu(n).$$

We see directly that $b_\lambda(.) : \mathbb{N} \to \mathbb{N}$ is a $q^{\lambda+m-1}$ periodic function and we extend it to a ($q^{\lambda+m-1}$ periodic) function $\mathbb{Z} \to \mathbb{N}$ which we also denote by $b_\lambda(.) : \mathbb{Z} \to \mathbb{N}$.

For any $n \in \mathbb{N}$ we define $F(n) := F(\varepsilon_{m-1}(n), \dots, \varepsilon_0(n))$.

Thus we can rewrite $b(n)$ as follows

$$b(n) = \sum_{j \geq 0} F\left(\left\lfloor \frac{q^{m-1}n}{q^j} \right\rfloor\right)$$

since $F(0) = 0$.

Now, we want to show that for any block-additive function, we can choose F without loss of generality such that it fulfills a nice property.

Lemma 5.2.2. *Let $b' : \mathbb{N} \to \mathbb{N}$ be a strongly block-additive function corresponding to F'. Then, there exists another strongly block-additive function $b : \mathbb{N} \to \mathbb{N}$ corresponding to F such that*

$$b(n) = b'(n) \tag{5.2}$$

$$\sum_{j=1}^{m-1} F(nq^j) = 0 \tag{5.3}$$

holds for all $n \in \mathbb{N}$.

Proof. We start by defining a new function

$$G(n) := \sum_{j=1}^{m-1} F'(nq^j).$$

This already allows us to define the function F:

$$F(n) := F'(n) + G(n) - G(\lfloor n/q \rfloor).$$

We find directly that $G(0) = F(0) = 0$. It remains to show (5.2) and (5.3), which are simple computations:

$$\begin{aligned} b(n) &= \sum_{j \geq 0} F\left(\left\lfloor \frac{q^{m-1}n}{q^j} \right\rfloor\right) \\ &= \sum_{j \geq 0} F'\left(\left\lfloor \frac{q^{m-1}n}{q^j} \right\rfloor\right) + \sum_{j \geq 0} G\left(\left\lfloor \frac{q^{m-1}n}{q^j} \right\rfloor\right) - \sum_{j \geq 0} G\left(\left\lfloor \frac{q^{m-1}n}{q^{j+1}} \right\rfloor\right) \\ &= b'(n) + G(0) = b'(n). \end{aligned}$$

Furthermore, we find

$$\sum_{j=1}^{m-1} F(nq^j) = \sum_{j=1}^{m-1} F'(nq^j) + \sum_{j=1}^{m-1} G(nq^j) - \sum_{j=1}^{m-1} G(nq^{j-1})$$

$$= \sum_{j=1}^{m-1} F'(nq^j) + G(nq^{m-1}) - G(n)$$
$$= \sum_{j=1}^{m-1} F'(nq^j) + 0 - \sum_{j=1}^{m-1} F'(nq^j) = 0.$$

□

We assume from now on that for any strongly block-additive function b (5.3) holds. This allows us to find an easier expression for b:

Corollary 5.2.3. *Let $b(n)$ be a digital function fulfilling (5.3). Then*

$$b(n) = \sum_{j\geq 0} F\left(\left\lfloor \frac{n}{q^j} \right\rfloor\right) \tag{5.4}$$

and

$$b_\lambda(n) = \sum_{j=0}^{\lambda-1} F\left(\left\lfloor \frac{n}{q^j} \right\rfloor\right)$$

holds for all $n, \lambda \in \mathbb{N}$.

We easily find the following recursion.

Lemma 5.2.4. *Let $\alpha \in \mathbb{N}, n_1 \in \mathbb{N}$ and $n_2 < q^\alpha$. Then*

$$b_\lambda(n_1 q^\alpha + n_2) = b_{\lambda-\alpha}(n_1) + b_\alpha(n_1 q^\alpha + n_2) \tag{5.5}$$

holds for all $\lambda > \alpha$ and

$$b(n_1 q^\alpha + n_2) = b(n_1) + b_\alpha(n_1 q^\alpha + n_2). \tag{5.6}$$

Proof. We compute $b_\lambda(n_1 q^\alpha + n_2)$:

$$\begin{aligned} b_\lambda(n_1 q^\alpha + n_2) &= \sum_{j=0}^{\lambda-1} F\left(\left\lfloor \frac{n_1 q^\alpha + n_2}{q^j} \right\rfloor\right) \\ &= \sum_{j=\alpha}^{\lambda-1} F\left(\left\lfloor \frac{n_1 q^\alpha + n_2}{q^j} \right\rfloor\right) + \sum_{j=0}^{\alpha-1} F\left(\left\lfloor \frac{n_1 q^\alpha + n_2}{q^j} \right\rfloor\right) \\ &= \sum_{j=0}^{\lambda-\alpha-1} F\left(\left\lfloor \frac{n_1}{q^j} \right\rfloor\right) + \sum_{j=0}^{\alpha-1} F\left(\left\lfloor \frac{n_1 q^\alpha + n_2}{q^j} \right\rfloor\right) \\ &= b_{\lambda-\alpha}(n_1) + b_\alpha(n_1 q^\alpha + n_2). \end{aligned}$$

The second case can be treated analogously. □

As we are dealing with the distribution of digital functions along a special subsequence, we will start discussing some distributional result for digital functions.

Lemma 5.2.5. *Let b be a strongly block-additive function and $m' > 1$. Then the following three statements are equivalent.*

(i) $\exists n \in \mathbb{N} : m' \nmid b(n)$

(ii) $\exists n < q^m : m' \nmid F(n)$

(iii) $\exists n < q^m : m' \nmid b(n)$

Proof. Obviously $(iii) \implies (i)$. Next we show $(i) \implies (ii)$:
Let n_0 be the smallest natural number > 0 such that $m' \nmid b(n_0)$. By Lemma 5.2.4 holds

$$b(n_0) = b(\lfloor n_0/q \rfloor) + F(n_0).$$

By the definition of n_0 holds $m' \mid b(\lfloor n_0/q \rfloor)$ and therefore $m' \nmid F(n_0) = F(n_0 \bmod q^m)$.

Therefore, it just remains to prove $(ii) \implies (iii)$:
Let n_0 be the smallest natural number > 0 such that $m' \nmid b(n_0)$. By (ii), we have $n_0 < q^m$. We compute $b(n_0) \bmod m'$:

$$b(n_0) = \sum_{j \geq 0} F\left(\left\lfloor \frac{n_0}{q^j} \right\rfloor\right) \equiv F(n_0) \neq 0 \bmod m'$$

as $\left\lfloor \frac{n_0}{q^j} \right\rfloor < n_0$ for $j \geq 1$ implies that $F\left(\left\lfloor \frac{n_0}{q^j} \right\rfloor\right) \equiv 0 \bmod m'$. □

Remark. We can not replace $m' \nmid .$ by $\gcd(m', .) = 1$ in Lemma 5.2.5 as the following example shows:
Let $m = 1, q = 3, m' = 6$ and $F(0) = 0, F(1) = 2, F(2) = 3$. We see that $\gcd(m', F(n)) > 1$ for all $n < q^m = 3$ and also $\gcd(m', b(n)) > 1$ for all $n < q^m = 3$. However, $b(5) = F(1) + F(2) = 5$ and $\gcd(m', b(5)) = 1$.

Next, we show a technical result concerning block-additive functions, which will be useful later on.

Lemma 5.2.6. *Let b be a strongly block-additive function in base q and $k > 1$ such that $\gcd(k, q-1) = 1$ and $\gcd(k, \gcd(\{b(n) : n \in \mathbb{N}\})) = 1$. Then there exist integers $\mathbf{e}_1, \mathbf{e}_2 < q^{2m-1}$ such that*

$$b(q^{m-1}(\mathbf{e}_1 + 1) - 1) - b(q^{m-1}(\mathbf{e}_1 + 1)) \neq b(q^{m-1}(\mathbf{e}_2 + 1) - 1) - b(q^{m-1}(\mathbf{e}_2 + 1)) \bmod k \tag{5.7}$$

holds.

Proof. Without loss of generality we can restrict ourselves to the case $p \in \mathbb{P}$ where $p \mid k$. Let us assume on the contrary that there exists c such that

$$b(q^{m-1}(\mathbf{e}+1)-1) - b(q^{m-1}(\mathbf{e}+1)) \equiv c \bmod p$$

holds for all $\mathbf{e} < q^{2m-1}$. We find under this assumption a new expression for $b(n) \bmod p$, where $n < q^m$:

$$\begin{aligned} n \cdot q^{m-1}c &\equiv \sum_{\mathbf{e}<nq^{m-1}} b(q^{m-1}(\mathbf{e}+1)-1) - b(q^{m-1}(\mathbf{e}+1)) \\ &\equiv \sum_{\mathbf{e}<nq^{m-1}} b(\mathbf{e}) + b_{m-1}(q^{m-1}\mathbf{e}+q^{m-1}-1) - b(\mathbf{e}+1) \\ &\equiv -b(nq^{m-1}) + \sum_{\mathbf{e}<nq^{m-1}} b_{m-1}(q^{m-1}\mathbf{e}+q^{m-1}-1) \\ &\equiv -b(nq^{m-1}) + n\sum_{\mathbf{e}<q^{m-1}} b_{m-1}(q^{m-1}\mathbf{e}+q^{m-1}-1), \end{aligned}$$

where the last equality holds since $b_{m-1}(q^{m-1}\mathbf{e}+q^{m-1}-1)$ is a q^{m-1} periodic function in $\mathbf{e}$. This gives

$$b(n) = b(nq^{m-1}) \equiv n\left(\sum_{\mathbf{e}<q^{m-1}} b_{m-1}(q^{m-1}\mathbf{e}+q^{m-1}-1) - q^{m-1}c\right) \bmod p. \tag{5.8}$$

By comparing this expression for $b(1) = b(q)$ we find

$$\begin{aligned} (q-1)\left(\sum_{\mathbf{e}<q^{m-1}} b_{m-1}(q^{m-1}\mathbf{e}+q^{m-1}-1) - q^{m-1}c\right) &\equiv 0 \bmod p \\ \sum_{\mathbf{e}<q^{m-1}} b_{m-1}(q^{m-1}\mathbf{e}+q^{m-1}-1) - q^{m-1}c &\equiv 0 \bmod p \end{aligned}$$

as $\gcd(p, q-1) = 1$.

Together with (5.8) this implies that $p \mid b(n)$ for all $n < q^m$. This is a contradiction to $\gcd(p, \gcd(\{b(n) : n \in \mathbb{N}\})) = 1$ by Lemma 5.2.5 □

We will use this result in a different form, given by the following corollary.

Corollary 5.2.7. *Let b be a strongly block-additive function in base q and $m' > 1$ such that $\gcd(m', q-1) = 1$ and $\gcd(m', \gcd(\{b(n) : n \in \mathbb{N}\})) = 1$. For every $\alpha \in \{\frac{1}{m'}, \ldots, \frac{m'-1}{m'}\}$ exist $\mathbf{e}_1, \mathbf{e}_2 < q^{2m-1}$ and $d \in \mathbb{N}$ such that $d\alpha \notin \mathbb{Z}$ and*

$$b(q^{m-1}(\mathbf{e}_1+1)-1) - b(q^{m-1}(\mathbf{e}_1+1)) - b(q^{m-1}(\mathbf{e}_2+1)-1) + b(q^{m-1}(\mathbf{e}_2+1)) = d.$$

Proof. Let $\alpha = \frac{x}{y}$ where $\gcd(x, y) = 1$ and $1 < y \mid m'$. We apply now Lemma 5.2.6 for $k = y$ and find $\mathbf{e}_1, \mathbf{e}_2$ such that

$$b(q^{m-1}(\mathbf{e}_1+1)-1) - b(q^{m-1}(\mathbf{e}_1+1)) - b(q^{m-1}(\mathbf{e}_2+1)-1) + b(q^{m-1}(\mathbf{e}_2+1)) = d,$$

where

$$d \not\equiv 0 \bmod y.$$

This implies

$$d\alpha = \frac{dx}{y} \not\equiv 0 \bmod 1.$$

□

5.3 Bounds on Fourier Transforms

The goal of this section is to prove Propositions 5.3.7 and 5.3.8. To find the necessary bounds we first need to state one important result on the norm of matrix products.

Afterward, we deal with Fourier estimates and formulate Proposition 5.3.7 and Proposition 5.3.8. The following Sections 5.3.3 and 5.3.4 give proofs of Proposition 5.3.7 and Proposition 5.3.8 respectively.

5.3.1 Auxiliary Results for the Bounds of the Fourier Transforms

In this section we find necessary conditions under which the product of matrices decreases exponentially with respect to the matrix row-sum norm.

Lemma 5.3.1. *Let $\mathbf{M}_\ell$, $\ell \in \mathbb{N}$, be $N \times N$-matrices with complex entries $M_{\ell;i,j}$, for $1 \le i, j \le N$, and absolute row sums*

$$\sum_{j=1}^{N} |M_{\ell;i,j}| \le 1 \text{ for } 1 \le i \le N.$$

Furthermore, we assume that there exist integers $m_0 \ge 1$ and $m_1 \ge 1$ and constants $c_0 > 0$ and $\eta > 0$ such that

1. *every product $\mathbf{A} = (A_{i,j})_{(i,j)\in\{1,\dots,N\}^2}$ of m_0 consecutive matrices $\mathbf{M}_\ell$ has the property that,*

$$|A_{i,1}| \ge c_0 \quad \text{or} \quad \sum_{j=1}^{N} |A_{i,j}| \le 1 - \eta \text{ for every row } i; \tag{5.9}$$

2. *every product $\mathbf{B} = (B_{i,j})_{(i,j)\in\{1,\dots,N\}^2}$ of m_1 consecutive matrices $\mathbf{M}_\ell$ has the property*

$$\sum_{j=1}^{N} |B_{1,j}| \le 1 - \eta. \tag{5.10}$$

Then there exist constants $C > 0$ and $\delta > 0$ such that

$$\left\| \prod_{\ell=r}^{r+k-1} \mathbf{M}_\ell \right\|_\infty \leq C q^{-\delta k} \tag{5.11}$$

uniformly for all $r \geq 0$ and $k \geq 0$ (where $\|\cdot\|_\infty$ denotes the matrix row-sum norm).

Proof. See [14]. □

Lemma 5.3.2. *Let $x_1, x_2, \xi_1, \xi_2 \in \mathbb{R}$. Then*

$$|\mathrm{e}(x_1) + \mathrm{e}(x_1 + \xi_1)| + |\mathrm{e}(x_2) + \mathrm{e}(x_2 + \xi_2)| \leq 4 - 8 \left(\sin \left(\frac{\pi \, \|\xi_1 - \xi_2\|}{4} \right) \right)^2 .$$

Proof. The proof can be found in [38]. □

5.3.2 Fourier estimates

In this section, we discuss some general properties of the occurring Fourier terms. Therefore, we need some more definitions.

For any $k \in \mathbb{N}$, we denote by $\mathcal{I}_k$ the set of integer vectors $I = (i_0, \ldots, i_{k-1})$ with $i_0 < q^{m-1}$ and $i_{\ell-1} \leq i_\ell \leq i_{\ell-1} + q^{m-1}$ for $1 \leq \ell \leq k-1$.
Furthermore, we denote by $\mathcal{I}'_k$ the set of integer vectors $I' = (i'_0, \ldots, i'_{k-1})$ with $i'_0 = 0$ and $i'_{\ell-1} \leq i'_\ell \leq i'_{\ell-1} + 1$.
This set $\mathcal{I}_k$ obviously consists of $q^{m-1}(q^{m-1}+1)^{k-1}$ elements. For any $I \in \mathcal{I}'_k$, $h \in \mathbb{Z}$ and $(d, \lambda) \in \mathbb{N}^2$, we define

$$H^I_\lambda(h, d) = \frac{1}{q^{\lambda+m-1}} \sum_{0 \leq u < q^{\lambda+m-1}} \mathrm{e}\left(\sum_{\ell=0}^{k-1} \alpha_\ell b_\lambda(u + \ell d + i_\ell) - huq^{-\lambda-m+1} \right), \tag{5.12}$$

for fixed coefficients $\alpha_\ell \in \{\frac{0}{m'}, \ldots, \frac{m'-1}{m'}\}$. We sum up $u < q^{\lambda+m-1}$ because b_λ is a $q^{\lambda+m-1}$ periodic function. This sum $H^I_\lambda(\,.\,, d)$ can then be seen as the discrete Fourier transform of the function

$$u \mapsto \mathrm{e}\left(\sum_{\ell=0}^{k-1} \alpha_\ell b_\lambda(u + \ell d + i_\ell) \right),$$

which is $q^{\lambda+m-1}$ periodic.

Furthermore, we define the important parameter

$$K := \alpha_0 + \cdots + \alpha_{k-1}.$$

We would like to find a simple recursion for H_λ in terms of $H_{\lambda-1}$. Instead we relate it to a different function for which the recursion is much simpler:

$$G_\lambda^I(h,d) = \frac{1}{q^\lambda} \sum_{u<q^\lambda} \mathrm{e}\left(\sum_{\ell=0}^{k-1} \alpha_\ell b_\lambda(q^{m-1}(u+\ell d)+i_\ell) - huq^{-\lambda} \right).$$

This sum $G_\lambda^I(\,.\,,d)$ can then be seen as the discrete Fourier transform of the function

$$u \mapsto \mathrm{e}\left(\sum_{\ell=0}^{k-1} \alpha_\ell b_\lambda(q^{m-1}(u+\ell d)+i_\ell) \right),$$

which is q^λ periodic. We show now how G and H are related.

Lemma 5.3.3. *Let $I \in \mathcal{I}'_k, h \in \mathbb{Z}, (d,\lambda) \in \mathbb{N}^2$ and $\delta \in \{0,\dots,q^{m-1}-1\}$. It holds*

$$H_\lambda^I(h, q^{m-1}d+\delta) = \frac{1}{q^{m-1}} \sum_{\varepsilon=0}^{q^{m-1}-1} \mathrm{e}\left(\sum_{\ell<k} \alpha_\ell \cdot F(\ell\delta+\varepsilon+i_\ell) - \frac{h\varepsilon}{q^{\lambda+m-1}} \right) \mathrm{G}_\lambda^{J_{\varepsilon,\delta}}(h,d), \quad (5.13)$$

where

$$J_{\varepsilon,\delta} = J_{\varepsilon,\delta}(I) = (i_\ell + \ell\delta + \varepsilon)_{\ell\in\{0,\dots,k-1\}} \in \mathcal{I}_k.$$

Proof. One checks easily that $J_{\varepsilon,\delta}(I) \in \mathcal{I}_k$. We evaluate $H_\lambda^I(h, q^{m-1}d+\delta)$:

$$\begin{aligned}
H_\lambda^I(h, q^{m-1}d+\delta) &= \frac{1}{q^{\lambda+m-1}} \sum_{0\le u<q^{\lambda+m-1}} \mathrm{e}\left(\sum_{\ell=0}^{k-1} \alpha_\ell b_\lambda(u+\ell(q^{m-1}d+\delta)+i_\ell) - huq^{-\lambda+m-1} \right) \\
&= \frac{1}{q^{\lambda+m-1}} \sum_{\varepsilon<q^{m-1}} \sum_{0\le u<q^\lambda} \mathrm{e}\left(-\frac{h(q^{m-1}u)}{q^{\lambda+m-1}} \right) \mathrm{e}\left(-\frac{h\varepsilon}{q^{\lambda+m-1}} \right) \\
&\qquad \cdot \mathrm{e}\left(\sum_{\ell=0}^{k-1} \alpha_\ell b_\lambda(q^{m-1}u+\varepsilon+\ell(q^{m-1}d+\delta)+i_\ell) \right) \\
&= \frac{1}{q^{\lambda+m-1}} \sum_{\varepsilon<q^{m-1}} \sum_{u<q^\lambda} \mathrm{e}\left(-\frac{hu}{q^\lambda} \right) \mathrm{e}\left(-\frac{h\varepsilon}{q^{\lambda+m-1}} \right) \\
&\qquad\qquad \cdot \mathrm{e}\left(\sum_{\ell=0}^{k-1} \alpha_l b_\lambda\left((u+\ell d)q^{m-1} + (\ell\delta + i_\ell + \varepsilon) \right) \right) \\
&= \frac{1}{q^{m-1}} \sum_{\varepsilon<q^{m-1}} \mathrm{e}\left(-\frac{h\varepsilon}{q^{\lambda+m-1}} \right) \mathrm{G}_\lambda^{J_{\varepsilon,\delta}}(h,d)
\end{aligned}$$

□

Next we define a transformation on $\mathcal{I}_k$ and a weight function v.

Definition 5.3.4. Let $j \geq 1$ and $\varepsilon, \delta \in \{0, \dots, q^j - 1\}$. Then, we define for $I \in \mathcal{I}_k$

$$T^j_{\varepsilon,\delta}(I) := \left(\left\lfloor \frac{i_\ell + q^{m-1}(\varepsilon + \ell\delta)}{q^j} \right\rfloor \right)_{\ell \in \{0,\dots,k-1\}}$$
$$v^j(I, \varepsilon, \delta) := \mathrm{e}\left(\sum_{\ell<k} \alpha_\ell \cdot b_j(i_\ell + q^{m-1}(\varepsilon + \ell\delta)) \right).$$

Furthermore, we extend the definition of T^j for arbitrary ε, δ by

$$T^j_{\varepsilon,\delta}(I) := T^j_{\varepsilon \bmod q^j, \delta \bmod q^j}(I).$$

The next Lemma shows some basic properties of these functions.

Lemma 5.3.5. *Let $\lambda, j, j_1, j_2 \in \mathbb{N}$ and $\varepsilon_i, \delta_i \in \{0, \dots, q^{j_i} - 1\}$. Then, the following facts hold.*

- $T^j_{\varepsilon,\delta}(I) \in \mathcal{I}_k$
- $T^{j_2}_{\varepsilon_2\delta_2} \circ T^{j_1}_{\varepsilon_1\delta_1} = T^{j_1+j_2}_{\varepsilon_2 q^{j_1}+\varepsilon_1, \delta_2 q^{j_1}+\delta_1}$
- $G^I_\lambda(h, d) = \sum_{u<q^\lambda} v^\lambda(I, u, d)\, \mathrm{e}(-huq^{-\lambda})$.

Proof. These facts are direct consequences of basic properties of the floor function. □

Now we can find a nice recursion for the Fourier transform G.

Lemma 5.3.6. *Let $I \in \mathcal{I}_k, h \in \mathbb{Z}, d, \lambda \in \mathbb{N}$ and $1 \leq j \leq \lambda, \delta \in \{0, \dots, q^j - 1\}$. We have*

$$G^I_\lambda(h, q^j d + \delta) = \frac{1}{q^j} \sum_{\varepsilon<q^j} \mathrm{e}(-h\varepsilon q^{-\lambda}) v^j(I, \varepsilon, \delta) \cdot G^{T^j_{\varepsilon,\delta}(I)}_{\lambda-j}(h, d).$$

Proof. We evaluate $G^I_\lambda(h, q^j d + \delta)$ and use (5.5):

$$\begin{aligned} G^I_\lambda(h, q^j d + \delta) &= \frac{1}{q^\lambda} \sum_{u<q^\lambda} \mathrm{e}\left(\sum_{\ell=0}^{k-1} \alpha_\ell b_\lambda(q^{m-1}(u + \ell(q^j d + \delta)) + i_\ell) - huq^{-\lambda} \right) \\ &= \frac{1}{q^j} \sum_{\varepsilon<q^j} \frac{1}{q^{\lambda-j}} \sum_{u<q^{\lambda-j}} \mathrm{e}\left(\sum_{\ell=0}^{k-1} \alpha_\ell b_\lambda(q^{m-1+j}(u + \ell d) + q^{m-1}(\varepsilon + \ell\delta) + i_\ell) \right) \\ &\quad \cdot \mathrm{e}(-h(uq^j + \varepsilon)q^{-\lambda}) \\ &= \frac{1}{q^j} \sum_{\varepsilon<q^j} \mathrm{e}\left(\sum_{\ell=0}^{k-1} \alpha_\ell b_j(q^{m-1}(\varepsilon + \ell\delta) + i_\ell) \right) \mathrm{e}(-h\varepsilon q^{-\lambda}) \end{aligned}$$

$$\cdot \frac{1}{q^{\lambda-j}} \sum_{u<q^{\lambda-j}} \mathrm{e}\left(\sum_{\ell=0}^{k-1} b_{\lambda-j}\left(q^{m-1}(u+\ell d)+\left\lfloor \frac{\varepsilon q^{m-1}+\ell\delta q^{m-1}+i_\ell}{q^j}\right\rfloor\right) - huq^{\lambda-j}\right)$$
$$= \frac{1}{q^j}\sum_{\varepsilon<q^j} v^j(I,\varepsilon,\delta)\,\mathrm{e}(h\varepsilon q^{-\lambda})\cdot G_{\lambda-j}^{T^j_{\varepsilon,\delta}(I)}(h,d).$$

□

The following propositions are crucial for our proof of the main Theorem 5.1.2.

Proposition 5.3.7. *If $K \equiv 0 \bmod 1$ and $\frac{1}{2}\lambda \le \lambda' \le \lambda$, then there exists $\eta > 0$ such that for any $I \in \mathcal{I}'_k$*

$$\frac{1}{q^{\lambda'}} \sum_{0\le d<q^{\lambda'}} \left|H_\lambda^I(h,d)\right|^2 \ll q^{-\eta\lambda}$$

holds uniformly for all integers h.

Proposition 5.3.8. *If $K \not\equiv 0 \bmod 1$, then there exists $\eta > 0$ such that for any $I \in \mathcal{I}'_k$*

$$\left|H_\lambda^I(h,d)\right| \ll q^{-\eta L} \max_{J\in\mathcal{I}_k} \left|G_{\lambda-L}^J(h,\lfloor d/q^L\rfloor)\right|$$

holds uniformly for all non-negative integers h, d and L.

Proofs for Proposition 5.3.7 and 5.3.8 are given in the following sections.

5.3.3 Proof of Proposition 5.3.7

This section is dedicated to prove Proposition 5.3.7. We start by reducing the problem from $H_\lambda^I(h,d)$ to $G_\lambda^I(h,d)$ for which we have found a nice recursion.

Proposition 5.3.9. *For $K \in \mathbb{Z}$ and $\frac{1}{2}\lambda \le \lambda' \le \lambda$, we find $\eta > 0$ such that for any $I \in \mathcal{I}_k$*

$$\frac{1}{q^{\lambda'}} \sum_{0\le d<q^{\lambda'}} \left|G_\lambda^I(h,d)\right|^2 \ll q^{-\eta\lambda}$$

holds uniformly for all integers h.

Lemma 5.3.10. *Proposition 5.3.9 implies Proposition 5.3.7.*

Proof. We see by (5.13) that

$$\left|H_\lambda^I(h,d)\right| \le \max_{J\in\mathcal{I}_k}\left|G_\lambda^J(h,\lfloor d/q^m\rfloor)\right| \le \sum_{J\in\mathcal{I}_k}\left|G_\lambda^J(h,\lfloor d/q^m\rfloor)\right|$$

Thus we find

$$\frac{1}{q^{\lambda'}} \sum_{0\le d<q^{\lambda'}} \left|H_\lambda^I(h,d)\right|^2 \le \sum_{J\in\mathcal{I}_k}\frac{1}{q^{\lambda'}}\sum_{0\le d<q^{\lambda'}}\left|G_\lambda^J(h,\lfloor d/q^m\rfloor)\right|^2 \ll q^{-\eta\lambda}.$$

□

Using Lemma 5.3.6, it is easy to establish a recursion for

$$\Phi^{I,I'}_{\lambda,\lambda'}(h) = \frac{1}{q^{\lambda'}} \sum_{0\le d<q^{\lambda'}} \mathrm{G}^I_\lambda(h,d)\overline{\mathrm{G}^{I'}_\lambda(h,d)}.$$

where $h \in \mathbb{Z}$, $(\lambda,\lambda') \in \mathbb{N}^2$ and $(I,I') \in \mathcal{I}_k^2$. For $\lambda,\lambda' \ge 1$ and $1 \le j \le \min(\lambda,\lambda')$ we yield

$$\Phi^{I,I'}_{\lambda,\lambda'}(h) = \frac{1}{q^{3j}} \sum_{\delta<q^j} \sum_{\varepsilon_1<q^j} \sum_{\varepsilon_2<q^j} \mathrm{e}\left(-\frac{(\varepsilon_1-\varepsilon_2)h}{q^\lambda}\right) v^j(I,\varepsilon_1,\delta)\overline{v^j(I,\varepsilon_2,\delta)}\Phi^{T^j_{\varepsilon_1,\delta}(I)T^j_{\varepsilon_2,\delta}(I')}_{\lambda-j,\lambda'-j}(h).$$

To find this recursion, one has to split up the sum over $0 \le d < q^{\lambda'}$ into the equivalence classes modulo q^j.

This identity gives rise to a vector recursion for $\Psi_{\lambda,\lambda'}(h) = \left(\Phi^{I,I'}_{\lambda,\lambda'}(h)\right)_{(I,I')\in\mathcal{I}_k^2}$. We use the recursion for $j = 1$:

$$\Psi_{\lambda,\lambda'}(h) = \mathbf{M}(h/q^\lambda) \cdot \Psi_{\lambda-1,\lambda'-1}(h)$$

where the $2^{2(k-1)} \times 2^{2(k-1)}$-matrix $\mathbf{M}(\beta) = (M_{(I,I'),(J,J')}(\beta))_{((I,I'),(J,J'))\in\mathcal{I}_k^2\times\mathcal{I}_k^2}$ is independent of λ and λ'. By construction, all absolute row sums of $\mathbf{M}(\beta)$ are equal to 1.

It is useful to interpret these matrices as weighted directed graphs. The vertices are the pairs $(I,I') \in \mathcal{I}_k^2$ and, starting from each vertex, there are q^3 directed edges to the vertices $(\mathrm{T}_{\varepsilon_1,\delta}(I), \mathrm{T}_{\varepsilon_2,\delta}(I'))$ - where $(\delta,\varepsilon_1,\varepsilon_2) \in \{0,\dots,q-1\}^3$ - with corresponding weights

$$\frac{1}{q^3}\,\mathrm{e}\left(-\frac{(\varepsilon_1-\varepsilon_2)h}{q^\lambda}\right) v^1(I,\varepsilon_1,\delta)\overline{v^1(I',\varepsilon_2,\delta)}.$$

Products of j such matrices correspond to oriented paths of length j in these graphs, which are weighted with the corresponding products. The entries at position $((I,I'),(J,J'))$ of such product matrices correspond to the sum of weights along paths from (I,I') to (J,J'). Lemma 5.3.6 allows us to describe this product of matrices directly.

Lemma 5.3.11. *The entry* $((I,I'),(J,J'))$ *of* $\mathbf{M}(h/q^\lambda)\cdot\mathbf{M}(h/q^{\lambda-1})\cdot\ldots\cdot\mathbf{M}(h/q^{\lambda-j+1})$ *equals*

$$\frac{1}{q^{3j}} \sum_{\delta<q^j} \sum_{\varepsilon_1,\varepsilon_2<q^j} \mathbf{1}_{[T^j_{\varepsilon_1\delta}(I)=J]}\,\mathbf{1}_{[T^j_{\varepsilon_2\delta}(I')=J']}\, v^j(I,\varepsilon_1,\delta)\overline{v^j(I',\varepsilon_2,\delta)}\,\mathrm{e}\left(-\frac{(\varepsilon_1-\varepsilon_2)h}{q^\lambda}\right).$$

Proof. Follows directly by Lemma 5.3.6. □

This product of matrices corresponds to oriented paths of length j. They can be encoded by the triple $(\varepsilon_1,\varepsilon_2,\delta)$ and they correspond to a path from (I,I') to $(T^j_{\varepsilon_1\delta}(I), T^j_{\varepsilon_2\delta}(I'))$ with weight $v^j(I,\varepsilon_1,\delta)\overline{v^j(I',\varepsilon_2,\delta)}\,\mathrm{e}\left(-\frac{(\varepsilon_1-\varepsilon_2)h}{q^\lambda}\right)$.

In order to prove Proposition 5.3.7, we will use Lemma 5.3.1 uniformly for h with $\mathbf{M}_l = \mathbf{M}(h/q^l)$. Therefore, we need to check Conditions (5.9) and (5.10).

Note that, since $\frac{1}{2}\lambda \le \lambda' \le \lambda$, we have

$$\Psi_{\lambda,\lambda'}(h) = \mathbf{M}(h/q^\lambda)\cdots\mathbf{M}(h/q^{\lambda-\lambda'+1})\Psi_{\lambda-\lambda',0}(h).$$

Lemma 5.3.12. *The matrices M_l defined above fulfill Condition* (5.9) *of Lemma 5.3.1.*

Proof. We need to show that there exists an integer $m_0 \geq 1$ such that every product

$$\mathbf{A} = (A_{(I,I'),(J,J')})_{((I,I'),(J,J'))\in\mathcal{I}_k^2\times\mathcal{I}_k^2}$$

of m_0 consecutive matrices $\mathbf{M}_l = \mathbf{M}(h/q^l)$ verifies Condition (5.9) of 5.3.1.

We define $m_0 = m-1+\lceil \log_q(k+1)\rceil$. It follows directly from the definition, that $T_{0,0}^{m_0}(I) = \mathbf{0}$ for all $I \in \mathcal{I}_k$. In the graph interpretation this means that for every vertex (I, I') there is a path of length m_0 from (I, I') to $(\mathbf{0}, \mathbf{0})$. Fix a row indexed by (I, I') in the matrix $\mathbf{A}$. We already showed that the entry $A_{(I,I'),(\mathbf{0},\mathbf{0})}$ is the sum of at least one term of absolute value q^{-3m_0}.

There are two possible cases. If the absolute row sum is at most

$$\leq 1-\eta$$

with $\eta \leq q^{-3m_0}$ then we are done.

In case the absolute row sum is strictly greater than $1-\eta$, we show that $|A_{(I,I'),(\mathbf{0},\mathbf{0})}| \geq q^{-3m_0}/2$: The inequality $|A_{(I,I'),(\mathbf{0},\mathbf{0})}| < q^{-3m_0}/2$ implies that $A_{(I,I'),(\mathbf{0},\mathbf{0})}$ is the sum of at least two terms of absolute value q^{-3m_0}. Thus the absolute row sum would be bounded by

$$\sum_{(J,J')} |A_{(I,I'),(J,J')}| < \frac{1}{2}q^{-3m_0} + \left(1 - 2\cdot q^{-3m_0}\right) = 1 - \frac{3}{2}q^{-3m_0} < 1 - q^{-3m_0}.$$

This contradicts the assumption that the absolute row sum is strictly greater than

$$1-\eta \geq 1-q^{-3m_0}.$$

Consequently, we yield

$$|A_{(I,I'),(\mathbf{0},\mathbf{0})}| \geq c_0 \text{ for } c_0 = q^{-3m_0}/2.$$

□

Lemma 5.3.13. *The matrices M_l fulfill Condition* (5.10) *of Lemma 5.3.1.*

Proof. We need to show that there exists an integer $m_1 \geq 1$ such that for every product

$$\mathbf{B} = (B_{(I,I'),(J,J')})_{((I,I'),(J,J'))\in\mathcal{I}_k^2\times\mathcal{I}_k^2}$$

of m_1 consecutive matrices $\mathbf{M}_l = \mathbf{M}(h/q^l)$ the absolute row-sum of the first row is bounded by $1-\eta$. We concentrate on the entry $B_{(\mathbf{0},\mathbf{0}),(\mathbf{0},\mathbf{0})}$, i.e. we consider all possible paths from $(\mathbf{0}, \mathbf{0})$ to $(\mathbf{0}, \mathbf{0})$ of length m_1 in the corresponding graph and show that a positive saving for the absolute row sum is just due to the structure of this entry.

Since $T_{00}^m(\mathbf{0}) = T_{10}^m(\mathbf{0}) = \mathbf{0}$, we have at least two paths from $(\mathbf{0}, \mathbf{0})$ to $(\mathbf{0}, \mathbf{0})$ and it follows that the entry $B_{(\mathbf{0},\mathbf{0}),(\mathbf{0},\mathbf{0})}$ is certainly a sum of $k_0 = k_0(m_1) \geq 2$ terms of absolute value q^{-3m_1}

(for every $m_1 \geq m$). This means that there are $k_0 \geq 2$ paths from $(\mathbf{0},\mathbf{0})$ to $(\mathbf{0},\mathbf{0})$ of length m_1 in the corresponding graph.

Our goal is to construct two paths $(\varepsilon_1^i, \varepsilon_2^i, \delta^i)$ from $(\mathbf{0},\mathbf{0})$ to $(\mathbf{0},\mathbf{0})$ such that

$$\left|\sum_{i=1}^{2} v^{m_1}(\mathbf{0},\varepsilon_1^i,\delta^i)\overline{v^{m_1}(\mathbf{0},\varepsilon_2^i,\delta^i)}\,\mathrm{e}\left(-\frac{(\varepsilon_1^i-\varepsilon_2^i)h}{q^\lambda}\right)\right| \leq 2-\eta$$

holds for all $h \in \mathbb{Z}$.

At first, we construct a path from $\mathbf{0}$ to $(q^{m-1}-1,\dots,q^{m-1}-1,q^{m-1},\dots,q^{m-1}) =: I_0 \in \mathcal{I}_k$ with exactly n_0+1 times $q^{m-1}-1$ (where $n_0 = \min\{n \in \mathbb{N} : \alpha_n \neq 0\}$). Therefore, let $n_1 = \lfloor \log_q(k-1) \rfloor + m$.

Lemma 5.3.14. *Let n_0, n_1 and I_0 be as above. Then*

$$T^{n_1}_{q^{n_1}-n_0-1,1}(\mathbf{0}) = I_0.$$

Proof. This follows directly by the definitions and simple computations. □

By applying Lemma 5.3.14 we find a transformation from $\mathbf{0}$ to I_0. This gives a path from $(\mathbf{0},\mathbf{0})$ to (I_0,I_0) by applying this transformation component-wise. We concatenate this path with another path $(\mathbf{e}_1,\mathbf{e}_2,0)$ of length $n_2 = 3m-1$ where $\mathbf{e}_i < q^{2m-1}$. The weight of the concatenation of these two paths equals

$$\begin{aligned} &v^{n_1}(\mathbf{0},q^{n_1}-n_0-1,1)v^{n_2}(I_0,\mathbf{e}_1,0)\overline{v^{n_1}(\mathbf{0},q^{n_1}-n_0-1,1)v^{n_2}(I_0,\mathbf{e}_2,0)}\,\mathrm{e}\left(-\frac{(\mathbf{e}_1-\mathbf{e}_2)h}{q^{\lambda-n_1}}\right)\\ &\quad = v^{n_2}(I_0,\mathbf{e}_1,0)\overline{v^{n_2}(I_0,\mathbf{e}_2,0)}\,\mathrm{e}\left(-\frac{(\mathbf{e}_1-\mathbf{e}_2)h}{q^{\lambda-n_1}}\right). \end{aligned}$$

We see that

$$\begin{aligned} T^{3m-1}_{\mathbf{e}_i,0}(I_0) &= \left(\left\lfloor \frac{I_{0|\ell}+q^{m-1}\mathbf{e}_i}{q^{3m-1}} \right\rfloor\right)_{\ell\in\{0\dots k-1\}}\\ &\leq \left(\left\lfloor \frac{q^{m-1}+q^{m-1}(q^{2m-1}-1)}{q^{3m-1}} \right\rfloor\right)_{\ell\in\{0\dots k-1\}}\\ &= \left(\left\lfloor \frac{q^{m-1}\cdot q^{2m-1}}{q^{3m-1}} \right\rfloor\right)_{\ell\in\{0\dots k-1\}} = \mathbf{0} \end{aligned}$$

Thus, we have found for each $\mathbf{e}_1,\mathbf{e}_2 < q^{2m-1}$ a path from $(\mathbf{0},\mathbf{0})$ to $(\mathbf{0},\mathbf{0})$.

We can use the special structure of I_0 to make the weight of this path more explicit: At first, we note that

$$\sum_{\ell=0}^{n_0} \alpha_\ell = \alpha_{n_0}$$

by the definition of n_0. Furthermore as $\sum_\ell \alpha_\ell \in \mathbb{Z}$

$$\sum_{\ell=n_0+1}^{k-1} \alpha_\ell \equiv -\alpha_{n_0} \bmod 1.$$

We find by the definition of v that for each $\mathbf{e} < q^{2m-1}$,

$$\begin{aligned} v^{3m-1}(I_0, \mathbf{e}, 0) &= \mathrm{e}\left(\sum_{\ell=0}^{k-1} \alpha_\ell b_{3m-1}(q^{m-1}\mathbf{e} + I_{0|\ell})\right) \\ &= \mathrm{e}\left(\alpha_{n_0}\left(b_{3m-1}(q^{m-1}\mathbf{e} + q^{m-1} - 1) - b_{3m-1}(q^{m-1}\mathbf{e} + q^{m-1})\right)\right) \\ &= \mathrm{e}\left(\alpha_{n_0}\left(b(q^{m-1}\mathbf{e} + q^{m-1} - 1) - b(q^{m-1}(\mathbf{e}+1)\right)\right). \end{aligned}$$

We find by Corollary 5.2.7 that there exist $\mathbf{e}_1, \mathbf{e}_2 < q^{2m-1}$ such that

$$b(q^{m-1}(\mathbf{e}_1+1)-1) - b(q^{m-1}(\mathbf{e}_1+1)) - b(q^{m-1}(\mathbf{e}_2+1)-1) + b(q^{m-1}(\mathbf{e}_2+1)) = d$$

and $\alpha_{n_0} d \notin \mathbb{Z}$.

We now compare the following two paths from $(\mathbf{0},\mathbf{0})$ to $(\mathbf{0},\mathbf{0})$ of length $m_1 = n_1 + n_2 = \lfloor \log_q(k-1) \rfloor + 4m - 1$:

- $(\mathbf{e}_1 q^{n_1} + q^{n_1} - n_0 - 1, \mathbf{e}_2 q^{n_1} + q^{n_1} - n_0 - 1, 1)$: We split up this path into the path of length n_1 from $(\mathbf{0},\mathbf{0})$ to (I_0, I_0) and the path of length n_2 from (I_0, I_0) to $(\mathbf{0},\mathbf{0})$: The first path can be described by the triple $(q^{n_1} - n_0 - 1, q^{n_1} - n_0 - 1, 1)$ and its weight is obviously 1.
  The second path - i.e. the path from (I_0, I_0) to $(\mathbf{0},\mathbf{0})$ - can be described by the triple $(\mathbf{e}_1, \mathbf{e}_2, 0)$ and its weight equals

$$\begin{aligned} & v^{n_2}(I_0, \mathbf{e}_1, 0)\overline{v^{n_2}(I_0, \mathbf{e}_2, 0)}\,\mathrm{e}\left(-\frac{(\mathbf{e}_1 - \mathbf{e}_2)h}{q^{\lambda - n_1}}\right) \\ &\quad = \mathrm{e}\left(\alpha_{n_0}\left(b(q^{m-1}(\mathbf{e}_1+1)-1) - b(q^{m-1}(\mathbf{e}_1+1))\right)\right) \\ &\qquad \overline{\mathrm{e}\left(\alpha_{n_0}\left(b(q^{m-1}(\mathbf{e}_2+1)-1) - b(q^{m-1}(\mathbf{e}_2+1))\right)\right)}\,\mathrm{e}\left(-\frac{(\mathbf{e}_1 - \mathbf{e}_2)h}{q^{\lambda - n_1}}\right) \\ &\quad = \mathrm{e}(\alpha_{n_0} d)\,\mathrm{e}\left(-\frac{(\mathbf{e}_1 - \mathbf{e}_2)h}{q^{\lambda - n_1}}\right). \end{aligned}$$

  Thus, the overall weight of the path from $(\mathbf{0},\mathbf{0})$ to $(\mathbf{0},\mathbf{0})$ has weight

$$\mathrm{e}(\alpha_{n_0} d)\,\mathrm{e}\left(-\frac{(\mathbf{e}_1 - \mathbf{e}_2)h}{q^{\lambda - n_1}}\right).$$

- $(\mathbf{e}_1 q^{n_1}, \mathbf{e}_2 q^{n_1}, 0)$: we compute directly the weight of this path:

$$v^{m_1}(\mathbf{0}, \mathbf{e}_1 q^{n_1}, 0)\overline{v^{m_1}(\mathbf{0}, \mathbf{e}_2 q^{n_1}, 0)}\,\mathrm{e}\left(-\frac{(\mathbf{e}_1 - \mathbf{e}_2)h}{q^{\lambda - n_1}}\right)$$

$$
\begin{aligned}
&= \mathrm{e}\left(\sum_{\ell=0}^{k-1}\alpha_\ell b_{m_1}(\mathbf{e}_1 q^{n_1}) - \sum_{\ell=0}^{k-1}\alpha_\ell b_{m_1}(\mathbf{e}_2 q^{n_1})\right)\mathrm{e}\left(-\frac{(\mathbf{e}_1-\mathbf{e}_2)h}{q^{\lambda-n_1}}\right)\\
&= \mathrm{e}\left(K(b_{m_1}(\mathbf{e}_1 q^{n_1}) - b_{m_1}(\mathbf{e}_2 q^{n_1}))\right)\mathrm{e}\left(-\frac{(\mathbf{e}_1-\mathbf{e}_2)h}{q^{\lambda-n_1}}\right)\\
&= \mathrm{e}\left(-\frac{(\mathbf{e}_1-\mathbf{e}_2)h}{q^{\lambda-n_1}}\right).
\end{aligned}
$$

We finally see that

$$
\begin{aligned}
|B_{(\mathbf{0},\mathbf{0}),(\mathbf{0},\mathbf{0})}| &\le \left(k_0 - 2 + \left|\mathrm{e}(\alpha_{n_0} d)\,\mathrm{e}\left(-\frac{(\mathbf{e}_1-\mathbf{e}_2)h}{q^{\lambda-n_1}}\right) + \mathrm{e}\left(-\frac{(\mathbf{e}_1-\mathbf{e}_2)h}{q^{\lambda-n_1}}\right)\right|\right) q^{-3m_1}\\
&= (k_0 - 2 + |1 + \mathrm{e}(\alpha_{n_0} d)|) q^{-3m_1}\\
&= (k_0 - 2 + 2\,|\cos(\pi\alpha_{n_0} d)|) q^{-3m_1}\\
&= \left(k_0 - 2 + 2\left|1 - 2\left(\sin\left(\frac{\pi\alpha_{n_0} d}{2}\right)\right)^2\right|\right) q^{-3m_1}\\
&\le \left(k_0 - 4\left(\sin\left(\frac{\pi}{2m'}\right)\right)^2\right) q^{-3m_1}.
\end{aligned}
$$

Thus we have

$$
\begin{aligned}
\sum_{(J,J')} |B_{(0,0),(J,J')}| &\le \left(k_0 - 4\left(\sin\left(\frac{\pi}{2m'}\right)\right)^2\right) q^{-3m_1} + (1 - k_0 q^{-3m_1})\\
&\le 1 - 4\left(\sin\left(\frac{\pi}{2m'}\right)\right)^2 \cdot q^{-3m_1}.
\end{aligned}
$$

Therefore condition (5.10) of Lemma 5.3.1 is verified with $m_1 = \lfloor \log_q(k-1)\rfloor + 4m - 1$ and $\eta = 4\left(\sin\left(\frac{\pi}{2m'}\right)\right)^2 q^{-3m_1} \ge 4\left(\sin\left(\frac{\pi}{2m'}\right)\right)^2 (k-1) q^{-12m+3} > 0$. □

At the end of this section, we want to recall the important steps of the proof of Proposition 5.3.7. At first we observe that

$$
\frac{1}{q^{\lambda'}} \sum_{0 \le d < q^{\lambda'}} |G_\lambda^I(h,d)|^2 = \Phi_{\lambda,\lambda'}^{I,I}(h).
$$

Thus Proposition 5.3.7 is equivalent to $\Phi_{\lambda,\lambda'}^{I,I}(h) \ll q^{-\eta\lambda}$. Next we considered the vector $\Psi_{\lambda,\lambda'}(h) = \left(\Phi_{\lambda,\lambda'}^{I,I'}(h)\right)_{(I,I')\in\mathcal{I}_k^2}$ and find the recursion

$$
\Psi_{\lambda,\lambda'}(h) = M(h/q^\lambda)\cdots M(h/q^{\lambda-\lambda'+1})\Psi_{\lambda-\lambda',0}(h)
$$

Then we defined $M_\ell := M(h/q^\ell)$ and showed that we can apply Lemma 5.3.1. Therefore we know that – since $\left|\Phi_{\lambda-\lambda'+1,0}^{I,I'}(h)\right| \le 1$

$$
|\Phi_{\lambda,\lambda'}^{I,I'}(h)| \le \|M_\lambda \cdots M_{\lambda-\lambda'+1}\|_\infty \le C q^{-\delta\lambda'} \le C q^{-\delta\lambda/2}
$$

with C and δ obtained by Lemma 5.3.1. Thus we know that $\Phi_{\lambda,\lambda'}^{I,I'}(h) \ll q^{-\eta\lambda}$ with $\eta = \delta/2$ uniformly for all h. This concludes the proof of Proposition 5.3.7.

5.3.4 Proof of Proposition 5.3.8

We start again by reducing the problem from $H_\lambda^{I'}(h,d)$ to $G_\lambda^I(h,d)$.

Proposition 5.3.15. *For $K \not\equiv 0 \bmod 1$ there exists $\eta > 0$ such that for any $I \in \mathcal{I}_k$*

$$\left|G_\lambda^I(h,d)\right| \ll q^{-\eta L} \max_{J \in \mathcal{I}_k} \left|G_{\lambda-L}^J(h, \lfloor d/q^L \rfloor)\right| \tag{5.14}$$

holds uniformly for all non-negative integers h, d and L.

Lemma 5.3.16. *Proposition 5.3.15 implies Proposition 5.3.8.*

Proof. Follows directly by (5.13). □

We assume from now on that $K \notin \mathbb{Z}$ holds.

We formulate Lemma 5.3.6 as a matrix vector multiplication:

$$G_\lambda(h, q^j d + \delta) = \frac{1}{q^j} M_\delta^j \left(\mathrm{e}\left(-\frac{h}{q^\lambda}\right)\right) G_{\lambda-j}(h,d)$$

where for any $\delta \in \{0, \dots, q^j - 1\}$ and $z \in \mathbb{U}$ we have

$$M_\delta^j(z) = \sum_{\varepsilon=0}^{q^j-1} (\mathbf{1}_{[J=T_{\varepsilon,\delta}^j(I)]} v^j(I,\varepsilon,\delta) z^\varepsilon)_{(I,J)\in\mathcal{I}_k^2}.$$

To prove Proposition 5.3.15 we aim to show that

$$\exists m_1 \in \mathbb{N}, \eta' \in \mathbb{R}^+ \text{ such that } \forall \delta < q^{m_1}, z \in \mathbb{U} \text{ holds } \left\|M_\delta^{m_1}(z)\right\|_\infty \le q^{m_1} - \eta'. \tag{5.15}$$

Indeed, we find that this is already sufficient to show Proposition 5.3.15.

Lemma 5.3.17. (5.15) *implies Proposition 5.3.15.*

Proof. We first note that

$$\left\|M_\delta^j(z)\right\|_\infty \le q^j$$

holds for all j and $\delta < q^j$ by definition.
Next we split the digital expansion of $d \bmod q^L$ - read from right to left - into $\lfloor L/m_1 \rfloor$ parts of length m_1 and possible one part of length $L \bmod m_1$. We denote the first parts by $\delta_1, \dots, \delta_{\lfloor L/m_1 \rfloor}$ and the last part by δ_0. Thus we find

$$\begin{aligned}\max_{I\in\mathcal{I}_k} \left|G_\lambda^I(h,d)\right| &= \|G_\lambda(h,d)\|_\infty \\ &\le \frac{1}{q^L} \max_z \left\|M_d^L(z)\right\|_\infty \cdot \left\|G_{\lambda-L}(h, \lfloor d/q^L \rfloor)\right\|_\infty\end{aligned}$$

$$\begin{aligned}
&\leq \frac{1}{q^L} \prod_{j=1}^{\lfloor L/m_1 \rfloor} \max_z \left\| M_{\delta_j}^{m_1}(zq^{m_1(j-1)}) \right\|_\infty \cdot q^{(L \bmod m_1)} \cdot \left\| G_{\lambda-L}(h, \lfloor d/q^L \rfloor) \right\|_\infty \\
&\leq \frac{1}{q^L} (q^{m_1} - \eta')^{\lfloor L/m_1 \rfloor} q^{(L \bmod m_1)} \cdot \left\| G_{\lambda-L}(h, \lfloor d/q^L \rfloor) \right\|_\infty \\
&\ll q^{-L\eta} \cdot \left\| G_{\lambda-L}(h, \lfloor d/q^L \rfloor) \right\|_\infty
\end{aligned}$$

where $\eta = \frac{\eta'}{q^{m_1} \log(q^{m_1})} > 0$. □

Throughout the rest of this section, we aim to prove (5.15).

Therefore, we try to find for each $I \in \mathcal{I}_k$ and $\delta < q^{m_1}$ a pair $(\varepsilon_1, \varepsilon_2)$ and $m_1' \leq m_1$ such that for all $z \in \mathbb{U}$ holds

$$\begin{aligned}
&T_{\varepsilon_i,\delta}^{m_1'}(I) = T_{\varepsilon_i+1,\delta}^{m_1'}(I) \\
&\left| v^{m_1'}(I, \varepsilon_1, \delta) + z v^{m_1'}(I, \varepsilon_1 + 1, \delta) \right| + \left| v^{m_1'}(I, \varepsilon_2, \delta) + z v^{m_1'}(I, \varepsilon_2 + 1, \delta) \right| \leq 4 - \eta'.
\end{aligned} \tag{5.16}$$

Let us assume for now that (5.16) holds. Indeed we find

$$\left\| M_\delta^{m_1'}(z) \right\|_\infty = \max_{I \in \mathcal{I}_k} \max_{z \in \mathbb{U}} \sum_{J \in \mathcal{I}_k} \left| \sum_{\varepsilon < q^{m_1'}} \mathbf{1}_{[T_{\varepsilon,\delta}^{m_1'}(I) = J]} z^\varepsilon v^{m_1'}(I, \varepsilon, \delta) \right|$$

However, we find for each I some $\varepsilon_1, \varepsilon_2$ fulfilling (5.16). This gives

$$\begin{aligned}
\max_{z \in \mathbb{U}} \sum_{J \in \mathcal{I}_k} \left| \sum_{\varepsilon < q^{m_1'}} \mathbf{1}_{[T_{\varepsilon,\delta}^{m_1'}(I) = J]} z^\varepsilon v^{m_1'}(I, \varepsilon, \delta) \right| &\leq \left(q^{m_1'} - 4 \right) + \sum_{i=1}^{2} \left| \sum_{j=0}^{1} z^{\varepsilon_i + j} v^{m_1'}(I, \varepsilon_i + j, \delta) \right| \\
&\leq q^{m_1'} - \eta'.
\end{aligned}$$

Thus, we find in total

$$\left\| M_\delta^{m_1}(z) \right\|_\infty \leq q^{m_1 - m_1'} (q^{m_1'} - \eta') \leq q^{m_1} - \eta'.$$

It just remains to find $\varepsilon_1, \varepsilon_2, m_1'$ fulfilling (5.16) and this turns out to be a rather tricky task.

We fix now some arbitrary $I \in \mathcal{I}_k$ and $\delta \in \mathbb{N}$. We start by defining for $0 \leq x \leq (4m-2)k$ and $c \in \mathbb{N}$

$$M_{x,c} = M_{x,(c \bmod q^x)} := \{\ell < k : \lfloor i_\ell / q^{m-1} \rfloor + d\ell \equiv c \bmod q^x\}$$

and show some basic properties of $M_{x,c}$.

Lemma 5.3.18. *For every $x < q^{(4m-2)k}$ exists c_0 such that*

$$\sum_{\ell \in M_{x,c_0}} \alpha_\ell \notin \mathbb{Z}.$$

Proof. One finds easily that

$$\{0,\dots,k-1\} = \bigcup_{c<q^x} M_{x,c},$$

which means that $\{M_{x,c} : c < q^x\}$ is a partition of $\{0,\dots,k-1\}$ for each x. Thus, we find for every x

$$\sum_c \sum_{\ell\in M_{x,c}} \alpha_\ell = \sum_{\ell<k} \alpha_\ell = K \notin \mathbb{Z}$$

and the proof follows easily. □

Lemma 5.3.19. *Let $d < q^{(4m-2)k}$ and $I \in \mathcal{I}_k$.*

Then, there exists $0 \le x_0 \le (4m-2)(k-1)$ such that for each $c < q^{x_0}$ exists $c^+ < q^{x_0+(4m-2)}$ such that

$$M_{x_0,c} = M_{x_0+(4m-2),c^+}.$$

Remark. This is equivalent to the statement that

$$\left\lfloor i_{\ell_1}/q^{m-1}\right\rfloor + d\ell_1 \equiv \left\lfloor i_{\ell_2}/q^{m-1}\right\rfloor + d\ell_2 \bmod q^{x_0}$$

implies

$$\left\lfloor i_{\ell_1}/q^{m-1}\right\rfloor + d\ell_1 \equiv \left\lfloor i_{\ell_2}/q^{m-1}\right\rfloor + d\ell_2 \bmod q^{x_0+4m-2}$$

Proof. We have already seen that $\{M_{x,c} : c < q^x\}$ is a partition of $\{0,\dots,k-1\}$. Furthermore, we find for $0 \le x \le (4m-2)k$ and $c < q^x$ that

$$M_{x,c} = \bigcup_{c'<q^{4m-2}} M_{x+(4m-2),c+q^x c'}.$$

This implies that $\{M_{x+4m-2,c} : c < q^{x+4m-2}\}$ is a refinement of $\{M_{x,c} : c < q^x\}$ and we find

$$\{M_{(4m-2)\cdot 0,c} : c < 1\} \ge \{M_{(4m-2)\cdot 1,c} : c < q^{4m-2}\} \ge \ldots \ge \{M_{(4m-2)k,c} : c < q^{(4m-2)k}\}.$$

It is well known that the maximal length of a chain in the set of partitions of $\{0,\dots,k-1\}$ is k. This means that there exists x_0' such that $\{M_{(4m-2)x_0',c} : c < q^{(4m-2)x_0'}\} = \{M_{(4m-2)(x_0'+1),c'} : c' < q^{(4m-2)(x_0'+1)}\}$. □

Furthermore, we define

$$\beta_{x,c} := \sum_{\ell\in M_{x,c}} \alpha_\ell.$$

We can now choose $m_1 := (4m-2)k$, $m_1' := x_0+(4m-2)$ where x_0 is given by Lemma 5.3.19. We consider $c_0 < q^{x_0}$ and c_0^+ provided by Lemma 5.3.18 and Lemma 5.3.19 and know that $\beta_{x,c_0} \notin \mathbb{Z}$. Therefore we apply Corollary 5.2.7 and find $\mathbf{e}_1, \mathbf{e}_2 < q^{2m-1}$ such that

$$b(q^{m-1}(\mathbf{e}_1+1)-1) - b(q^{m-1}(\mathbf{e}_1+1)) - b(q^{m-1}(\mathbf{e}_2+1)-1) + b(q^{m-1}(\mathbf{e}_2+1)) = d$$

and $d\beta_{x,c_0} \notin \mathbb{Z}$.

We are now able to define

$$\varepsilon_1 = (q^{x_0+m-1}(\mathbf{e}_1+1) - c_0^+ - 1) \bmod q^{x_0+4m-2}$$
$$\varepsilon_2 = (q^{x_0+m-1}(\mathbf{e}_2+1) - c_0^+ - 1) \bmod q^{x_0+4m-2}.$$

It just remains to check (5.16) which we split up into the following two lemmata.

Lemma 5.3.20. *Let x_0, ε_i be defined as above. Then*

$$T^{x_0+4m-2}_{\varepsilon_i,d}(I) = T^{x_0+4m-2}_{\varepsilon_i+1,d}(I)$$

holds.

Proof. We need to show that

$$\left\lfloor \frac{i_\ell + q^{m-1}(\ell d + \varepsilon_i)}{q^{x_0+4m-2}} \right\rfloor = \left\lfloor \frac{i_\ell + q^{m-1}(\ell d + \varepsilon_i + 1)}{q^{x_0+4m-2}} \right\rfloor \tag{5.17}$$

holds for all $\ell < k$ and $i = 1, 2$. We know that ℓ belongs to M_{x_0+4m-2,c^+} for some $c < q^{x_0}$. Thus, we find for $j = 0, 1$

$$\begin{aligned}\left\lfloor \frac{i_\ell + q^{m-1}(\ell d + \varepsilon_i + j)}{q^{x_0+4m-2}} \right\rfloor &= \left\lfloor \frac{(i_\ell \bmod q^{m-1}) + q^{m-1}(c^+ + \varepsilon_i + j)}{q^{x_0+4m-2}} \right\rfloor \\ &= \left\lfloor \frac{c^+ + \varepsilon_i + j}{q^{x_0+3m-1}} \right\rfloor\end{aligned}$$

Therefore, (5.17) does hold, unless

$$c^+ + \varepsilon_i + 1 \equiv 0 \bmod q^{x_0+3m-1}.$$

We find

$$c^+ + \varepsilon_i + 1 \equiv c^+ + q^{x_0+m-1}(\mathbf{e}_i + 1) - c_0^+ \bmod q^{x_0+3m-1}.$$

We first consider the case $c \neq c_0$:

$$c^+ + \varepsilon_i + 1 \equiv c - c_0 \not\equiv 0 \bmod q^{x_0}$$

For $c = c_0$:

$$c_0^+ + \varepsilon_i + 1 \equiv q^{x_0+m-1}(\mathbf{e}_i + 1) \bmod q^{x_0+3m-1}$$

However

$$\mathbf{e}_i + 1 \not\equiv 0 \bmod q^{2m}$$

as $\mathbf{e}_i < q^{2m-1}$. Thus, (5.17) holds. □

Lemma 5.3.21. *There exists $\eta' > 0$ only depending on m' such that for x_0 and ε_i defined as above holds*

$$\sum_{i=1}^{2} \left| v^{x_0+4m-2}(I, \varepsilon_i, \delta) + z \cdot v^{x_0+4m-2}(I, \varepsilon_i + 1, \delta) \right| \leq 4 - \eta' \tag{5.18}$$

for all $z \in \mathbb{U}$.

Proof. We start by computing the weights $v^{x_0+4m-2}(I, \varepsilon_i + j, \delta)$. For arbitrary $\varepsilon < q^{\lambda_0+4m-2}$, we find:

$$\begin{aligned} v^{x_0+4m-2}(I, \varepsilon, d) &= \prod_{\ell<k} \mathrm{e}(\alpha_\ell b_{x_0+4m-2}(i_\ell + q^{m-1}(\varepsilon + \ell d))) \\ &= \prod_{\ell<k} \mathrm{e}(\alpha_\ell b_{m-1}(i_\ell + q^{m-1}(\varepsilon + \ell d)))\, \mathrm{e}\left(\alpha_\ell b_{x_0+3m-1}\left(\left\lfloor i_\ell / q^{m-1} \right\rfloor + \varepsilon + \ell d\right)\right) \\ &= \mathrm{e}(g(\varepsilon)) \cdot \prod_{\ell<k} \mathrm{e}\left(\alpha_\ell b_{x_0+3m-1}\left(\left\lfloor i_\ell / q^{m-1} \right\rfloor + \varepsilon + \ell d\right)\right). \end{aligned}$$

where

$$g(\varepsilon) := \sum_{\ell<k} \alpha_\ell b_{m-1}(i_\ell + q^{m-1}(\varepsilon + \ell d)).$$

Note that $g(\varepsilon)$ only depends on $\varepsilon \bmod q^{m-1}$.

We can describe this product by using the weights β defined above.

$$v^{x_0+4m-2}(I, \varepsilon, d) = \mathrm{e}(g(\varepsilon)) \cdot \prod_{c'<q^{x_0+4m-2}} \mathrm{e}\left(\beta_{x_0+4m-2,c'} \cdot b_{x_0+3m-1}\left(c' + \varepsilon\right)\right).$$

Furthermore, we can rewrite every $c' < q^{x_0+4m-2}$ for which $\beta_{x_0+4m-2,c'} \neq 0$ as some c^+ where $c < q^{x_0}$. This gives then

$$\begin{aligned} v^{x_0+4m-2}(I, \varepsilon, d) &= \mathrm{e}(g(\varepsilon)) \cdot \prod_{c<q^{x_0}} \mathrm{e}\left(\beta_{x_0,c} \cdot b_{x_0+3m-1}\left(c^+ + \varepsilon\right)\right) \\ &= g(\varepsilon) \cdot \prod_{c<q^{x_0}} \mathrm{e}\left(\beta_{x_0,c} \cdot b_{x_0}(c^+ + \varepsilon)\right) \cdot \prod_{c<q^{x_0}} \mathrm{e}\left(\beta_{x_0,c} \cdot b_{3m-1}\left(\left\lfloor \frac{c^+ + \varepsilon}{q^{x_0}} \right\rfloor\right)\right) \end{aligned}$$

Thus we find for $\varepsilon = \varepsilon_i + j$ that:

$$\begin{aligned} v^{x_0+4m-2}(I, \varepsilon_i + j, d) &= \mathrm{e}(g(\varepsilon_i + j)) \cdot \prod_{c<q^{x_0}} \mathrm{e}\left(\beta_{x_0,c} \cdot b_{x_0}(c^+ + \varepsilon_i + j)\right) \\ &\qquad \cdot \prod_{c<q^{x_0}} \mathrm{e}\left(\beta_{x_0,c} \cdot b_{3m-1}\left(\left\lfloor \frac{c^+ + \varepsilon_i + j}{q^{x_0}} \right\rfloor\right)\right) \\ &= \mathrm{e}(g(-c_0^+ - 1 + j)) \cdot \prod_{c<q^{x_0}} \mathrm{e}\left(\beta_{x_0,c} \cdot b_{x_0}(c^+ - c_0^+ - 1 + j)\right) \end{aligned}$$

$$\begin{aligned}
&\cdot \prod_{c<q^{x_0}} \mathrm{e}\left(\beta_{x_0,c}\cdot b_{3m-1}\left(q^{m-1}(\mathbf{e}_i+1)+\left\lfloor\frac{c^+-c_0^+-1+j}{q^{x_0}}\right\rfloor\right)\right)\\
&=\mathrm{e}(g(-c_0^+-1+j))\cdot\prod_{c<q^{x_0}}\mathrm{e}\left(\beta_{x_0,c}\cdot b_{x_0}(c^+-c_0^+-1+j)\right)\\
&\quad\cdot\prod_{\substack{c<q^{x_0}\\ c\neq c_0}}\mathrm{e}\left(\beta_{x_0,c}\cdot b_{3m-1}\left(q^{m-1}(\mathbf{e}_i+1)+\left\lfloor\frac{c^+-c_0^+-1+j}{q^{x_0}}\right\rfloor\right)\right)\\
&\quad\cdot\mathrm{e}\left(\beta_{x_0,c_0}\cdot b_{3m-1}(q^{m-1}(\mathbf{e}_i+1)-1+j)\right).
\end{aligned}$$

For $c\neq c_0$, we find

$$\left\lfloor\frac{c^+-c_0^+-1}{q^{x_0}}\right\rfloor=\left\lfloor\frac{c^+-c_0^+}{q^{x_0}}\right\rfloor$$

as $c^+\equiv c\not\equiv c_0\equiv c_0^+ \mod q^{x_0}$.

Consequently, we find

$$\begin{aligned}
v^{x_0+4m-2}(I,\varepsilon_i,d)&=\mathrm{e}(x_i)\\
v^{x_0+4m-2}(I,\varepsilon_i+1,d)&=\mathrm{e}(x_i+\xi_i)
\end{aligned}$$

where

$$\begin{aligned}
x_i=g(-c_0^+-1)&+\sum_{c<q^{x_0}}\beta_{x_0,c}\cdot b_{x_0}(c^+-c_0^+-1)\\
&+\sum_{\substack{c<q^{x_0}\\ c\neq c_0}}\beta_{x_0,c}\cdot b_{3m-1}\left(q^{m-1}(\mathbf{e}_i+1)+\left\lfloor\frac{c^+-c_0^+}{q^{x_0}}\right\rfloor\right)\\
&+\beta_{x_0,c_0}\cdot b_{3m-1}(q^{m-1}(\mathbf{e}_i+1)-1)
\end{aligned}$$

and

$$\begin{aligned}
\xi_i=g(-c_0^+)-g(-c_0^+-1)&+\sum_{c<q^{x_0}}\beta_{x_0,c}\cdot b_{x_0}(c^+-c_0^+)-\sum_{c<q^{x_0}}\beta_{x_0,c}\cdot b_{x_0}(c^+-c_0^+-1)\\
&+\beta_{x_0,c_0}\cdot b_{3m-1}(q^{m-1}(\mathbf{e}_i+1))-\beta_{x_0,c_0}\cdot b_{3m-1}(q^{m-1}(\mathbf{e}_i+1)-1).
\end{aligned}$$

Also, we find

$$\xi_1-\xi_2=\beta_{x_0,c_0}d\notin\mathbb{Z},$$

where

$$d=b(q^{m-1}(\mathbf{e}_1+1))-b(q^{m-1}(\mathbf{e}_1+1)-1)-b(q^{m-1}(\mathbf{e}_2+1))+b(q^{m-1}(\mathbf{e}_2+1)-1).$$

This implies

$$\|\xi_1-\xi_2\|\geq\frac{1}{m'}.$$

It remains to apply Lemma 5.3.2 to find that (5.18) holds with $\eta'=8\left(\sin\left(\frac{\pi}{4m'}\right)\right)^2$. □

At the end of this section, we recall the important steps of the proof of Proposition 5.3.15. We started to rewrite our recursion for G_λ^I into a matrix vector multiplication

$$G_\lambda(h, q^L d + \delta) = \frac{1}{q^L} M_\delta^L \left(\mathrm{e}\left(-\frac{h}{q^\lambda} \right) \right) G_{\lambda - L}(h, d).$$

We then split up this matrix $M_\delta^L(.)$ into a product of many matrices $M_{\delta_j}^{m_1}(.)$, where $m_1 = (4m-2)k$. Thereafter, we showed that $\left\| M_{\delta_j}^{m_1}(.) \right\| \leq q^{m_1} - \eta$, where $\eta = 8\left(\sin\left(\frac{\pi}{4m'}\right)\right)^2$. This implies then Proposition 5.3.15.
To show that $\left\| M_{\delta_j}^{m_1} \right\| \leq q^{m_1} - \eta$, we found two different ε_i such that

$$\begin{aligned} &T_{\varepsilon_i,\delta}^{m_1'}(I) = T_{\varepsilon_i+1,\delta}^{m_1'}(I) \\ &\left| v^{m_1'}(I, \varepsilon_1, \delta) + z v^{m_1'}(I, \varepsilon_1 + 1, \delta) \right| + \left| v^{m_1'}(I, \varepsilon_2, \delta) + z v^{m_1'}(I, \varepsilon_2 + 1, \delta) \right| \leq 4 - \eta' \end{aligned}$$

holds for all $z \in \mathbb{U}$.

5.4 Auxiliary Results

In this section, we present some auxiliary results which are used in Section 5.5, to prove the main theorem. For this proof, it is crucial to approximate characteristic functions of the intervals $[0, \alpha)$ mod 1 where $0 \leq \alpha < 1$ by trigonometric polynomials. This is done by using Vaaler's method - see Section 3.7.3. As we deal with exponential sums we also use a generalization of Van-der-Corput's inequality which we have already seen in Section 3.7.2. In Section 5.4.1, we acquire some results dealing with sums of geometric series which we use to bound linear exponential sums. Section 5.4.2 is dedicated to one classic result on Gauss sums and allows us to find appropriate bounds on the occurring quadratic exponential sums in Section 5.5. The last part of this section deals with carry propagation. We find a quantitative statement that carry propagation along several digits is rare, i.e. exponentially decreasing.

We would like to note that all these auxiliary results have already been presented in [14].

5.4.1 Sums of geometric series

We will often make use of the following upper bound for geometric series with ratio $\mathrm{e}(\xi), \xi \in \mathbb{R}$ and $L_1, L_2 \in \mathbb{Z}$, $L_1 \leq L_2$:

$$\left| \sum_{L_1 < \ell \leq L_2} \mathrm{e}(\ell \xi) \right| \leq \min(L_2 - L_1, |\sin \pi \xi|^{-1}), \tag{5.19}$$

which is obtained from the formula for finite geometric series.

The following results allow us to find useful estimates for special double and triple sums involving geometric series.

Lemma 5.4.1. *Let $(a,m) \in \mathbb{Z}^2$ with $m \geq 1$, $\delta = \gcd(a,m)$ and $b \in \mathbb{R}$. For any real number $U > 0$, we have*

$$\sum_{0 \leq n \leq m-1} \min\left(U, \left|\sin\left(\pi \tfrac{an+b}{m}\right)\right|^{-1}\right) \leq \delta \min\left(U, \left|\sin\left(\pi \tfrac{\delta \|b/\delta\|}{m}\right)\right|^{-1}\right) + \frac{2\,m}{\pi}\log(2\,m). \quad (5.20)$$

Proof. See for example [36] or [14]. □

Lemma 5.4.2. *Let $m \geq 1$ and $A \geq 1$ be integers and $b \in \mathbb{R}$. For any real number $U > 0$, we have*

$$\frac{1}{A}\sum_{1 \leq a \leq A}\sum_{0 \leq n < m} \min\left(U, \left|\sin\left(\pi \tfrac{an+b}{m}\right)\right|^{-1}\right) \ll \tau(m)\; U + m\log m \quad (5.21)$$

and, if $|b| \leq \frac{1}{2}$, we have an even sharper bound

$$\frac{1}{A}\sum_{1 \leq a \leq A}\sum_{0 \leq n < m} \min\left(U, \left|\sin\left(\pi \tfrac{an+b}{m}\right)\right|^{-1}\right) \ll \tau(m)\min\left(U, \left|\sin\left(\pi \tfrac{b}{m}\right)\right|^{-1}\right) + m\log m, \quad (5.22)$$

where $\tau(m)$ denotes the number of divisors of m.

Proof. See [14]. □

5.4.2 Gauss sums

In the proof of the main theorem, we will meet quadratic exponential sums. We first consider Gauss sums $\mathrm{G}(a,b;m)$ which are defined by:

$$\mathrm{G}(a,b;m) := \sum_{n=0}^{m-1} \mathrm{e}\left(\frac{an^2+bn}{m}\right).$$

In this chapter, we want to prove one classic result on Gauss sums, namely Theorem 5.4.3.

Theorem 5.4.3. *For all $(a,b,m) \in \mathbb{Z}^3$ with $m \geq 1$,*

$$\left|\sum_{n=0}^{m-1} \mathrm{e}\left(\tfrac{an^2+bn}{m}\right)\right| \leq \sqrt{2m\gcd(a,m)} \quad (5.23)$$

holds.

Proof. For a detailed proof see [23]. □

Consequently we obtain the following result for incomplete quadratic Gauss sums.

Lemma 5.4.4. *For all $(a,b,m,N,n_0) \in \mathbb{Z}^5$ with $m \geq 1$ and $N \geq 0$, we have*

$$\left|\sum_{n=n_0+1}^{n_0+N} \mathrm{e}\left(\tfrac{an^2+bn}{m}\right)\right| \leq \left(\tfrac{N}{m} + 1 + \tfrac{2}{\pi}\log\tfrac{2m}{\pi}\right)\sqrt{2m\gcd(a,m)}. \quad (5.24)$$

Proof. This is Lemma 9 of [14]. □

5.4.3 Carry Lemmas

As mentioned before, we want to find a quantitative statement on how rare carry propagation along several digits is.

Lemma 5.4.5. *Let $(\nu, \lambda, \rho) \in \mathbb{N}^3$ such that $\nu + \rho \leq \lambda \leq 2\nu$. For any integer r with $0 \leq r \leq q^\rho$, the number of integers $n < q^\nu$ for which there exists an integer $j \geq \lambda$ with $\varepsilon_j((n+r)^2) \neq \varepsilon_j(n^2)$ is $\ll q^{2\nu+\rho-\lambda}$. Hence, we find for any block-additive function b, that the number of integers $n < q^\nu$ with*

$$b_{\lambda-m+1}((n+r)^2) - b_{\lambda-m+1}(n^2) \neq b((n+r)^2) - b(n^2)$$

is also $\ll q^{2\nu+\rho-\lambda}$.

Proof. We follow the idea of [14] with some minor changes to suit our case better.

First we suppose that $\lambda \geq \nu + \rho + 3$; otherwise we know that the number of all integers $n < q^\nu$ is bounded by $q^\nu \leq q^{\nu-\lambda+\nu+\rho+2} = q^2 \cdot q^{2\nu+\rho-\lambda}$.

We know that $2nr + r^2 < 2q^{\rho+\nu} + q^{2\rho} \leq 3q^{\rho+\nu} < q^{\rho+\nu+2}$. In order to affect the j-th digit for $j \geq \lambda$, it is necessary to transfer a carry for the digits $\rho + \nu + 2$ to j. Therefore, for $\rho + \nu + 2 \leq j' < \lambda$, $a_{j'} = q - 1$ must hold. Hence there exists $t \in \mathbb{N}$ such that $\lfloor n^2/q^{\rho+\nu+2} \rfloor = q^{\lambda-\nu-\rho-2}t - 1$. In other words:

$$q^{\lambda-\nu-\rho-2}t - 1 \leq \frac{n^2}{q^{\nu+\rho+2}} < q^{\lambda-\nu-\rho-2}t.$$

Therefore, we can bound $t \in \mathbb{N}$

$$\frac{n^2}{q^\lambda} < t \leq \left\lfloor \frac{q^{2\nu}}{q^\lambda} + \frac{1}{q^{\lambda-\nu-\rho-2}} \right\rfloor = q^{2\nu-\lambda}.$$

For fixed t, there are at most $\sqrt{q^\lambda t} - \sqrt{q^\lambda t - q^{\nu+\rho+2}} = \sqrt{q^\lambda t}\left(1 - \sqrt{1 - \frac{1}{tq^{\lambda-\nu-\rho-2}}}\right)$ integers n such that $\lfloor n^2/q^{\nu+\rho+2} \rfloor = q^{\lambda-\nu-\rho-2}t - 1$.

For $0 \leq u \leq 1$ it holds that $1 - \sqrt{1-u} \leq u$. Since $tq^{\lambda-\nu-\rho-2} \geq 1$, we know that the number of integers $n < q^\nu$ for which there exists an integer $j \geq \lambda$ with $\varepsilon_j((n+r)^2) \neq \varepsilon_j(n^2)$ is bounded by

$$\sum_{t=1}^{q^{2\nu-\lambda}} \sqrt{q^\lambda t}\left(1 - \sqrt{1 - \frac{1}{tq^{\lambda-\nu-\rho-2}}}\right) \leq \sum_{t=1}^{q^{2\nu-\lambda}} \frac{\sqrt{q^\lambda t}}{q^{\lambda-\nu-\rho-2}t} = q^{\nu+\rho+2-\lambda/2} \sum_{t=1}^{q^{2\nu-\lambda}} \frac{1}{\sqrt{t}}$$
$$\overset{(*)}{\leq} q^{5/2} q^{2\nu+\rho-\lambda}.$$

The last inequality $(*)$ holds since

$$\sum_{t=1}^{q^n} \frac{1}{\sqrt{t}} = q^{-\frac{n}{2}} + \sum_{\ell=1}^{n} \sum_{t=q^{\ell-1}}^{q^\ell - 1} \frac{1}{\sqrt{t}} \leq 1 + \sum_{\ell=1}^{n} (q^\ell - q^{\ell-1}) \frac{1}{\sqrt{q^{\ell-1}}}$$

$$\leq 1+\sum_{\ell=1}^{n}\left(q^{\frac{\ell+1}{2}}-q^{\frac{\ell-1}{2}}\right)=1+q^{\frac{n+1}{2}}-1=q^{\frac{1}{2}}q^{\frac{n}{2}}.$$

This completes the proof. □

The next lemma helps to replace quadratic exponential sums depending only on few digits.

Lemma 5.4.6. *Let $(\lambda,\mu,\nu,\rho')\in\mathbb{N}^4$ such that $0<\mu<\nu<\lambda$, $2\rho'\leq\mu\leq\nu-\rho'$ and $\lambda-\nu\leq 2(\mu-\rho')$and set $\mu'=\mu-\rho'$. For integers $n<q^\nu$, $s\geq 1$ and $1\leq r\leq q^{(\lambda-\nu)/2}$ we set*

$$\begin{aligned}
n^2 &\equiv u_1q^{\mu'}+w_1 \bmod q^{\lambda+m-1} && (0\leq w_1<q^{\mu'},\ 0\leq u_1<q^{\lambda+m-1-\mu+\rho'})\\
(n+r)^2 &\equiv u_2q^{\mu'}+w_2 \bmod q^{\lambda+m-1} && (0\leq w_2<q^{\mu'},\ 0\leq u_2<q^{\lambda+m-1-\mu+\rho'})\\
2n &\equiv u_3q^{\mu'}+w_3 \bmod q^{\lambda+m-1} && (0\leq w_3<q^{\mu'},\ 0\leq u_3<q^{\nu+1-\mu+\rho'})\\
2sq^{m-1}n &\equiv v \bmod q^{\lambda-\mu+m-1}, && (0\leq v<q^{\lambda-\mu+m-1})
\end{aligned}\tag{5.25}$$

where the integers $u_1=u_1(n)$, $u_2=u_2(n)$,$u_3=u_3(n)$, $v=v(n)$,$w_1=w_1(n)$, $w_2=w_2(n)$ and $w_3=w_3(n)$ satisfy the above conditions. Then for any integer $\ell\geq 1$ the number of integers $n<q^\nu$ for which one of the following conditions

$$\begin{aligned}
\mathrm{s}_{\mu,\lambda}((n+\ell)^2) &\neq \mathrm{s}_{\rho',\lambda-\mu+\rho'}(u_1+\ell u_3)\\
\mathrm{s}_{\mu,\lambda}((n+\ell+sq^{\mu+m-1})^2)) &\neq \mathrm{s}_{\rho',\lambda-\mu+\rho'}(u_1+\ell u_3+vq^{\rho'}+2\ell sq^{m-1}q^{\rho'})\\
\mathrm{s}_{\mu,\lambda}((n+r+\ell)^2) &\neq \mathrm{s}_{\rho',\lambda-\mu+\rho'}(u_2+\ell u_3)\\
\mathrm{s}_{\mu,\lambda}((n+r+\ell+sq^{\mu+m-1})^2)) &\neq \mathrm{s}_{\rho',\lambda-\mu+\rho'}(u_2+\ell u_3+vq^{\rho'}+2(\ell+r)sq^{m-1}q^{\rho'})
\end{aligned}\tag{5.26}$$

is satisfied is $\ll q^{\nu-\rho'}$.

Proof (by [14]). We first consider the case $(n+\ell)^2$. The other cases are similar and we will comment on them at the end of the proof. We find that

$$(n+\ell)^2\equiv(u_1+\ell u_3)q^{\mu'}+w_1+\ell w_3+\ell^2 \bmod q^{\lambda+m-1}.$$

If $w_1+\ell w_3+\ell^2<q^{\mu'}$ and $0\leq j<\lambda-\mu'+m-1$, we have $\varepsilon_{\mu'+j}((n+\ell)^2)=\varepsilon_j(u_1+\ell u_3)$. For $w_1+\ell w_3+\ell^2\geq q^{\mu'}$, there is a carry propagation. We show that there are only few exceptions where more than ρ' digits are changed. The proof is split into the following two steps:

1. If the digits block $(\varepsilon_j((n+\ell)^2))_{\mu\leq j<\lambda+m-1}$ differ from the digits block $(\varepsilon_j(u_1+\ell u_3))_{\rho'\leq j<\lambda+m-1-\mu+\rho'}$, where $u_1=u_1(n)$ and $u_3=u_3(n)$ are defined by (5.25), it follows that

$$\frac{(n+\ell)^2}{q^\mu}-\left\lfloor\frac{(n+\ell)^2}{q^\mu}\right\rfloor\leq\frac{C}{q^{\rho'}}\quad\text{or}\quad\frac{(n+\ell)^2}{q^\mu}-\left\lfloor\frac{(n+\ell)^2}{q^\mu}\right\rfloor\geq 1-\frac{C}{q^{\rho'}},\tag{5.27}$$

   for some constant $C=C(\ell)$.

2. The number of integers $n<q^\nu$ fulfilling (5.27) is $\ll q^{\nu-\rho'}$.

Obviously these two properties are sufficient to prove Lemma 5.4.6.

We start with the proof of the first property. As mentioned above we just have to consider the case $w_1 + \ell w_3 + \ell^2 \geq q^{\mu'} = q^{\mu-\rho'}$. Since $w_1, w_3 < q^{\mu'}$ the carry

$$\widetilde{w} := \left\lfloor q^{-\mu'} \left(w_1 + \ell w_3 + \ell^2\right) \right\rfloor$$

is bounded and, thus, can only attain finitely many values $\{1, 2, \ldots, D\}$ (where D is a constant depending on ℓ). These values of $\widetilde{w}$ will certainly affect some digits (of lower order) of $u_1 + \ell u_3$. Let $\widetilde{v} := u_1 + \ell u_3 \bmod q^{\rho'}$ with $0 \leq \widetilde{v} < q^{\rho'}$. The digits $\varepsilon_j(u_1 + \ell u_3)$, $\rho' \leq j < \lambda + m - 1 - \mu'$ might be affected by this carry if $\widetilde{v} \in \{q^{\rho'} - 1, q^{\rho'} - 2, \ldots, q^{\rho'} - D\}$. Since

$$\begin{aligned} \frac{(n+\ell)^2}{q^\mu} &\equiv \frac{u_1 + \ell u_3}{q^{\rho'}} + \frac{w_1 + \ell w_3 + \ell^2}{q^{\mu'+\rho'}} \bmod 1 \\ &\equiv \frac{\widetilde{v}}{q^{\rho'}} + \frac{w_1 + \ell w_3 + \ell^2}{q^{\mu'+\rho'}} \bmod 1, \end{aligned}$$

it immediately follows that (5.27) holds with $C = D + 1$. This completes the proof of the first part.

Next, let Z denote the number of integers $n < q^\nu$ with (5.27). By Theorem 3.7.2, we have

$$\begin{aligned} Z &= \sum_{n<q^\nu} \left(\chi_\alpha\left(q^{-\mu}(n+\ell)^2\right) + \chi_\alpha\left(-q^{-\mu}(n+\ell)^2\right)\right) \\ &\leq 2 \sum_{|h|\leq H} \left(\alpha + \frac{1}{H}\right) \left| \sum_{n<q^\nu} \mathrm{e}\left(h\frac{(n+\ell)^2}{q^\mu}\right)\right| \end{aligned}$$

with $\alpha = Cq^{-\rho'}$. We can set $H = q^{\rho'}$.

It is clear that the main contribution comes from the term corresponding to $h = 0$ which gives an upper bound of form $\mathcal{O}(q^{\nu-\rho'})$. Each $h \neq 0$ with $|h| \leq H = q^{\rho'}$ can be written as $h = h'd$, where $d \mid q^\mu$ and $\gcd(h', q) = 1$. Therefore, we have by Lemma 5.4.4

$$\sum_{n<q^\nu} \mathrm{e}\left(h\frac{(n+\ell)^2}{q^\mu}\right) = \mathcal{O}\left(q^{\nu-\mu/2}\sqrt{d} + \mu q^{\mu/2}\sqrt{d}\right)$$

and, consequently,

$$q^{-\rho'} \sum_{0\neq|h|\leq q^{\rho'}} \left| \sum_{n<q^\nu} \mathrm{e}\left(h\frac{(n+\ell)^2}{q^\mu}\right)\right| = \mathcal{O}\left((q^{-\rho'}q^{\nu-\mu/2} + \mu q^{\mu/2}) \sum_{\substack{d\mid q^\mu \\ d\leq q^{\rho'}}} \frac{q^{\rho'}}{d}\sqrt{d}\right).$$

This equals $\mathcal{O}\left(q^{\nu-\mu/2} + \mu q^\mu\right)$ since

$$\sum_{d\mid q^\mu} d^{-1/2} \leq \prod_{j=1}^{\omega(q)} \frac{1}{1 - \frac{1}{\sqrt{p_j}}}.$$

where $p_1, \dots, p_{\omega(q)}$ are exactly the prime divisors of q. Since $2\rho' \le \mu \le \nu - \rho'$, all contributions are $\ll q^{\nu-\rho'}$. This completes the proof of the second part.

Finally, we comment on the other cases. First, there is no change for $(n+\ell+sq^{\mu+m-1})^2$ since the term $sq^{\mu+m-1}$ does not affect the discussed carry propagation. For $(n+\ell+r)^2$, we have

$$(n+\ell+r)^2 = (u_2+\ell u_3)q^{\mu'} + w_2 + \ell w_3 + \ell^2 + 2r\ell.$$

Here we have to assure that $q^{-\mu'}(w_2 + \ell w_3 + \ell^2 + 2r\ell)$ remains bounded. However, this is ensured by the assumption $\lambda - \nu \le 2(\mu - \rho')$. The same argument applies for the final case $(n+\ell+sq^{\mu+m-1}+r)^2$. □

5.5 Proof of the Main Theorem

In this section, we complete the proof of Theorem 5.1.2 following the ideas and structure of [14]. As the proof is very similar, we only outline it briefly and comment on the important changes.

The structure of the proof is similar for both cases: At first we want to substitute the function b by $b_{\mu,\lambda}$. This can be done by applying Lemma 5.4.5 and Lemma 3.7.1 in the case $K \in \mathbb{Z}$. For the case $K \notin \mathbb{Z}$ we have to use Lemma 3.7.1 first.

Thereafter, we apply Lemma 5.4.6 to detect the digits between μ and λ. Next, we use characteristic functions to detect suitable values for $u_1(n), u_2(n), u_3(n)$. Lemma 3.7.3 allows us to replace the characteristic functions by exponential sums. We split the remaining exponential sum into a quadratic and a linear part and find that the quadratic part is negligibly small. For the remaining sum, we apply Proposition 5.3.7 or 5.3.8 – depending on whether $K \in \mathbb{Z}$. The case $K \notin \mathbb{Z}$ needs more effort to deal with.

5.5.1 The case $K \in \mathbb{Z}$

In this section, we show that, if $K = \alpha_0 + \cdots + \alpha_{k-1} \in \mathbb{Z}$, Proposition 5.3.7 provides an upper bound for the sum

$$S_0 = \sum_{n<N} \mathrm{e}\left(\sum_{\ell=0}^{k-1} \alpha_\ell b((n+\ell)^2)\right).$$

Let ν be the unique integer such that $q^{\nu-1} < N \le q^\nu$ and we choose all appearing exponents - i.e. λ, μ, ρ, etc. - as in [14].

By using Lemma 5.4.5, and the same arguments as in [14], we find

$$S_0 = S_1 + \mathcal{O}\left(q^{\nu-(\lambda-\nu)}\right), \tag{5.28}$$

where

$$S_1 = \sum_{n<N} \mathrm{e}\left(\sum_{\ell=0}^{k-1} \alpha_\ell b_\lambda((n+\ell)^2)\right).$$

Now we use Lemma 3.7.1 - with $Q = q^{\mu+m-1}$ and $S = q^{\nu-\mu}$ - to relate S_1 to a sum in terms of $b_{\mu,\lambda}$:

$$|S_1|^2 \ll \frac{N^2}{S} + \frac{N}{S}\Re(S_2), \tag{5.29}$$

where

$$S_2 = \sum_{1\le s<S} \left(1 - \frac{s}{S}\right) S_2'(s)$$

and

$$S_2'(s) = \sum_{n\in I(N,s)} \mathrm{e}\left(\sum_{\ell=0}^{k-1} \alpha_\ell (b_{\mu,\lambda}((n+\ell)^2) - b_{\mu,\lambda}((n+\ell+sq^{\mu+m-1})^2))\right),$$

where $I(N,s)$ is an interval included in $[0, N-1]$ (which we do not specify).

Next we use Lemma 5.4.6 to detect the digits of $(n+\ell)^2$ and $(n+\ell+sq^{m-1}q^{\mu})^2$ between μ and $\lambda+m-1$ - with a negligible error term. Therefore, we have to take the digits between $\mu' = \mu - \rho'$ and μ into account, where $\rho' > 0$ will be chosen later.

We set the integers $u_1 = u_1(n)$, $u_3 = u_3(n)$, $v = v(n)$, $w_1 = w_1(n)$, and $w_3 = w_3(n)$ to satisfy the conditions of Lemma 5.4.6 and detect them by characteristic functions. Thus, we find

$$S_2'(s) = S_3'(s) + \mathcal{O}(q^{\nu-\rho'}), \tag{5.30}$$

where

$$\begin{aligned} S_3'(s) = \sum_{0\le u_1<U_1} \sum_{0\le u_3<U_3} \sum_{n\in I(N,s)} & \\ \mathrm{e}\left(\sum_{\ell=0}^{k-1} \alpha_\ell (b_{\rho',\lambda-\mu+\rho'}(u_1+\ell u_3) - b_{\rho',\lambda-\mu+\rho'}(u_1+\ell u_3+v(n)q^{\rho'}+2\ell s q^{m-1}q^{\rho'})\right) & \\ \chi_{q^{\mu'-\lambda-m+1}}\left(\frac{n^2}{q^{\lambda+m-1}} - \frac{u_1}{U_1}\right) \chi_{q^{\mu'-\nu-1}}\left(\frac{2n}{q^{\nu+1}} - \frac{u_3}{U_3}\right), & \end{aligned}$$

where χ_α is defined by (3.21) and $U_1 = q^{\lambda+m-1-\mu'}, U_3 = q^{\nu-\mu'+1}$. Lemma 3.7.3 allows us to replace the characteristic functions χ by trigonometric polynomials. More precisely, using (3.27) with $H_1 = U_1 q^{\rho''}$ and $H_3 = U_3 q^{\rho''}$ for some suitable $\rho'' > 0$ (which is chosen later and again as a fraction of ν), we have

$$S_3'(s) = S_4(s) + \mathcal{O}(E_1) + \mathcal{O}(E_3) + \mathcal{O}(E_{1,3}), \tag{5.31}$$

where E_1, E_3 and $E_{1,3}$ are the error terms specified in (3.27) and

$$\begin{aligned} S_4(s) \;=\; & \sum_{0\le u_1<U_1} \sum_{0\le u_3<U_3} \sum_{0\le v<q^{\lambda-\mu+m-1}} \\ & \sum_{n\in I(N,s)} \mathrm{e}\left(\sum_{\ell=0}^{k-1} \alpha_\ell (b_{\rho',\lambda-\mu+\rho'}(u_1+\ell u_3) - b_{\rho',\lambda-\mu+\rho'}(u_1+\ell u_3+vq^{\rho'}+2\ell s q^{m-1}q^{\rho'}))\right) \end{aligned}$$

$$A_{U_1^{-1},H_1}\left(\frac{n^2}{q^{\lambda+m-1}}-\frac{u_1}{U_1}\right)A_{U_3^{-1},H_3}\left(\frac{2n}{q^{\nu+1}}-\frac{u_3}{U_3}\right)$$
$$\frac{1}{q^{\lambda-\mu+m-1}}\sum_{0\le h<q^{\lambda-\mu+m-1}}\mathrm{e}\left(h\frac{2sq^{m-1}n-v}{q^{\lambda-\mu+m-1}}\right),$$

where we use the last sum to detect the correct value of $v=v(n)$.

The error terms E_1, E_3, $E_{1,3}$ can easily be estimated with the help of Lemma 5.4.4, just as in [14]. By using the representations of $A_{U_1^{-1},H_1}$ and $A_{U_3^{-1},H_3}$, we obtain

$$\begin{aligned}S_4(s)=&\frac{1}{q^{\lambda-\mu+m-1}}\sum_{|h_1|\le H_1}\sum_{|h_3|\le H_3}\sum_{0\le h<q^{\lambda-\mu+m-1}}a_{h_1}(U_1^{-1},H_1)\,a_{h_3}(U_3^{-1},H_3)\\&\sum_{0\le u_1<U_1}\sum_{0\le u_3<U_3}\sum_{0\le v<q^{\lambda-\mu+m-1}}\mathrm{e}\left(-\frac{h_1u_1}{U_1}-\frac{h_3u_3}{U_3}-\frac{hv}{q^{\lambda-\mu+m-1}}\right)\\&\mathrm{e}\left(\sum_{\ell=0}^{k-1}\alpha_\ell(b_{\rho',\lambda-\mu+\rho'}(u_1+\ell u_3)-b_{\rho',\lambda-\mu+\rho'}(u_1+\ell u_3+vq^{\rho'}+2\ell sq^{m-1}q^{\rho'}))\right)\\&\cdot\sum_n\mathrm{e}\left(\frac{h_1n^2}{q^{\lambda+m-1}}+\frac{h_3n}{q^\nu}+\frac{2hsn}{q^{\lambda-\mu}}\right).\end{aligned}$$

We now distinguish the cases $h_1=0$ and $h_1\neq 0$. For $h_1\neq 0$, we can estimate the exponential sum by using Lemma 5.4.4 and the following estimate

$$\sum_{1\le h_1\le H_1}\sqrt{\gcd(h_1,q^\lambda)}\ll_q H_1. \tag{5.32}$$

Thus, we find

$$\sum_{0<|h_1|\le H_1}\sum_{|h_3|\le H_3}\sum_{0\le h<q^{\lambda-\mu+m-1}}\left|\sum_n\mathrm{e}\left(\frac{h_1n^2}{q^{\lambda+m-1}}+\frac{h_3n}{q^\nu}+\frac{2hsn}{q^{\lambda-\mu}}\right)\right|\ll\lambda H_1H_3q^{\lambda/2+\lambda-\mu}.$$

This gives then

$$S_4(s)=S_5(s)+\mathcal{O}(\lambda q^{3\lambda/4}), \tag{5.33}$$

where $S_5(s)$ denotes the part of $S_4(s)$ with $h_1=0$.

We set $u_1=u_1''+q^{\rho'}u_1'$ and $u_3=u_3''+q^{\rho'}u_3'$ (where $0\le u_1'',u_3''<q^{\rho'}$). Furthermore, we define $i_\ell=\lfloor(u_1''+\ell u_3'')/q^{\rho'}\rfloor$. As $I=(i_\ell)_{0\le\ell<k}=(\lfloor(u_1''+\ell u_3'')/q^{\rho'}\rfloor)_{0\le\ell<k}$ is contained in $\mathcal{I}_k'$, we have - by the same arguments as in [14] -

$$\begin{aligned}S_5(s)\le&\sum_{|h_3|\le H_3}\sum_{0\le h<q^{\lambda-\mu+m-1}}\frac{1}{q^{\nu+1-\mu}}\sum_{0\le u_3'<q^{\nu-\mu+1}}\sum_{I\in\mathcal{I}_k}\left|H_{\lambda-\mu}^I(h,u_3')\overline{H_{\lambda-\mu}^I(h,u_3'+2sq^{m-1})}\right|\\&\cdot\min\left(N,\left|\sin\left(\pi\left(\frac{h_3}{q^\nu}+\frac{2hs}{q^{\lambda-\mu}}\right)\right)\right|^{-1}\right).\end{aligned}$$

Using the estimate $\left|H^I_{\lambda-\mu}(h,u'_3+2sq^{m-1})\right|\le 1$ and the Cauchy-Schwarz inequality, we yield

$$\sum_{0\le u'_3<q^{\nu-\mu+1}}\left|H^I_{\lambda-\mu}(h,u'_3)\overline{H^I_{\lambda-\mu}(h,u'_3+2sq^{m-1})}\right|\le q^{(\nu-\mu+1)/2}\left(\sum_{0\le u'_3<q^{\nu-\mu+1}}\left|H^I_{\lambda-\mu}(h,u'_3)\right|^2\right)^{1/2}.$$

We now replace λ by $\lambda-\mu+m-1$, λ' by $\nu-\mu+1$ and apply Proposition 5.3.7.

$$S_5(s)\ll q^{-\eta(\lambda-\mu)/2}\sum_{|h_3|\le H_3}\sum_{0\le h<q^{\lambda-\mu+m-1}}\min\left(N,\left|\sin\left(\pi\left(\frac{h_3}{q^\nu}+\frac{2hs}{q^{\lambda-\mu+m-1}}\right)\right)\right|^{-1}\right).$$

Then it just remains to average over s and combine all the estimates as in [14]. This gives then

$$|S_0|\ll q^{\nu-(\lambda-\nu)}+\nu^{(\omega(q)+1)/2}q^\nu q^{-\eta(\lambda-\nu)/2}+q^{\nu-\rho'/2}+q^{\nu-\rho''/2}+\lambda^{1/2}q^{\nu/2+3\lambda/8}$$

– provided that the following conditions hold

$$\begin{aligned}&2\rho'\le\mu\le\nu-\rho',\quad \rho''<\mu'/2,\quad \mu'\ll 2^{\nu-\mu'},\quad 2\mu'\ge\lambda,\\&(\nu-\mu)+2(\lambda-\mu)+2(\rho'+\rho'')\le\lambda/4,\quad \nu-\mu'+\rho''+\lambda-\mu\le\nu.\end{aligned}$$

For example, the choice

$$\lambda=\nu+\left\lfloor\frac{\nu}{20}\right\rfloor\ \text{ and }\ \rho'=\rho''=\left\lfloor\frac{\nu}{200}\right\rfloor$$

ensures that the above conditions are satisfied.

Summing up we proved that for $\eta'<\min(1/200,\eta/40)$ - where η is given by 5.3.7 - holds

$$S_0\ll q^{\nu(1-\eta')}\ll N^{1-\eta'}$$

which is precisely the statement of Theorem 5.1.2.

5.5.2 The case $K\notin\mathbb{Z}$

In this section, we show that, for $K=\alpha_0+\cdots+\alpha_{k-1}\notin\mathbb{Z}$, Proposition 5.3.8 provides an upper bound for the sum

$$S_0=\sum_{n<N}\mathrm{e}\left(\sum_{\ell=0}^{k-1}\alpha_\ell b((n+\ell)^2)\right).$$

Let μ, λ, ρ and ρ_1 be integers satisfying

$$0\le\rho_1<\rho<\mu=\nu-2\rho<\nu<\lambda=\nu+2\rho<2\nu \tag{5.34}$$

to be chosen later - just as in [14]. Since $K \notin \mathbb{Z}$ we can not use Lemma 5.4.5 directly. Therefore, we apply Lemma 3.7.1 with $Q = 1$ and $R = q^{\rho}$. Summing trivially for $1 \leq r \leq R_1 = q^{\rho_1}$ yields

$$|S_0|^2 \ll \frac{N^2 R_1}{R} + \frac{N}{R} \sum_{R_1 < r < R} \left(1 - \frac{r}{R}\right) \Re(S_1(r)),$$

where

$$S_1(r) = \sum_{n \in I_1(r)} \mathrm{e}\left(\sum_{\ell=0}^{k-1} \alpha_\ell \left(b((n+\ell)^2) - b((n+r+\ell)^2)\right)\right)$$

and $I_1(r)$ is an interval included in $[0, N-1]$. By Lemma 5.4.5 we conclude that $b_{\lambda,\infty}((n+\ell)^2) = b_{\lambda,\infty}((n+r+\ell)^2)$ for all but $\mathcal{O}(Nq^{-(\lambda-\nu-\rho)})$ values of n. Therefore, we see that

$$S_1(r) = S_1'(r) + \mathcal{O}(q^{\nu-(\lambda-\nu-\rho)}),$$

with

$$S_1'(r) = \sum_{n \in I_1(r)} \mathrm{e}\left(\sum_{\ell=0}^{k-1} \alpha_\ell \left(b_\lambda((n+\ell)^2) - b_\lambda((n+r+\ell)^2)\right)\right).$$

This leads to

$$|S_0|^2 \ll q^{2\nu-\rho+\rho_1} + q^{3\nu+\rho-\lambda} + \frac{q^\nu}{R} \sum_{R_1 < r < R} |S_1'(r)|$$

and, by using the Cauchy-Schwarz inequality to

$$|S_0|^4 \ll q^{4\nu-2\rho+2\rho_1} + q^{6\nu+2\rho-2\lambda} + \frac{q^{2\nu}}{R} \sum_{R_1 < r < R} |S_1'(r)|^2.$$

For $|S_1'(r)|^2$ we can use Lemma 3.7.1 again: Let $\rho' \in \mathbb{N}$ to be chosen later such that $1 \leq \rho' \leq \rho$. After applying Lemma 3.7.1 with $Q = q^{\mu+m-1}$ and

$$S = q^{2\rho'} \leq q^{\nu-\mu}, \tag{5.35}$$

we observe that for any $\widetilde{n} \in \mathbb{N}$ we have

$$b_\lambda((\widetilde{n} + sq^{\mu+m-1})^2) - b_\lambda(\widetilde{n}^2) = b_{\mu,\lambda}((\widetilde{n} + sq^{\mu+m-1})^2) - b_{\mu,\lambda}(\widetilde{n}^2),$$

and thus

$$|S_0|^4 \ll q^{4\nu-2\rho+2\rho_1} + q^{6\nu+2\rho-2\lambda} + \frac{q^{4\nu}}{S} + \frac{q^{3\nu}}{RS} \sum_{R_1 < r < R} \sum_{1 \leq s < S} |S_2(r,s)|, \tag{5.36}$$

with

$$\begin{aligned} S_2(r,s) \;=\; & \sum_{n \in I_2(r,s)} \mathrm{e}\Bigg(\sum_{\ell=0}^{k-1} \alpha_\ell \big(b_{\mu,\lambda}((n+\ell)^2) - b_{\mu,\lambda}((n+r+\ell)^2) \\ & - b_{\mu,\lambda}((n+sq^{\mu+m-1}+\ell)^2) + b_{\mu,\lambda}((n+sq^{\mu+m-1}+r+\ell)^2)\big)\Bigg), \end{aligned}$$

where $I_2(r,s)$ is an interval included in $[0, N-1]$.

We now make a Fourier analysis similar to the case $K \equiv 0 \bmod 1$ - as in [14]. We set $U = q^{\lambda+m-1-\mu'}, U_3 = q^{\nu-\mu'+1}$ and $V = q^{\lambda-\mu+m-1}$. We apply Lemma 5.4.6 and detect the correct values of u_1, u_2, u_3 by characteristic functions. This gives

$$\begin{aligned} S_2(r,s) \;=\; & \sum_{0\le u_1<U}\sum_{0\le u_2<U}\sum_{0\le u_3<U_3} \\ & \sum_{n\in I_2(r,s)} \mathrm{e}\Bigg(\sum_{\ell=0}^{k-1}\alpha_\ell\big(b_{\rho',\lambda-\mu+\rho'}(u_1+\ell u_3) - b_{\rho',\lambda-\mu+\rho'}(u_2+\ell u_3) \\ & \qquad -b_{\rho',\lambda-\mu+\rho'}(u_1+\ell u_3+v(n)q^{\rho'}+2\ell s q^{m-1}q^{\rho'}) \\ & \qquad +b_{\rho',\lambda-\mu+\rho'}(u_2+\ell u_3+v(n)q^{\rho'}+2(\ell+r)sq^{m-1}q^{\rho'})\big)\Bigg) \\ & \chi_{U^{-1}}\left(\frac{n^2}{q^{\lambda+m-1}}-\frac{u_1}{U}\right)\chi_{U^{-1}}\left(\frac{(n+r)^2}{q^{\lambda+m-1}}-\frac{u_2}{U}\right)\chi_{U_3^{-1}}\left(\frac{2n}{q^\nu}-\frac{u_3}{U_3}\right) \\ & +\mathcal{O}(q^{\nu-\rho'}). \end{aligned}$$

Furthermore, we use Lemma 3.7.3 to replace the characteristic functions χ by trigonometric polynomials. Using (3.27) with $U_1 = U_2 = U$, $H_1 = H_2 = Uq^{\rho_2}$ and $H_3 = U_3 q^{\rho_3}$, and integers ρ_2, ρ_3 verifying

$$\rho_2 \le \mu - \rho',\ \rho_3 \le \mu - \rho', \tag{5.37}$$

we obtain

$$\begin{aligned} S_2(r,s) = S_3(r,s) \;+\; & \mathcal{O}(q^{\nu-\rho'}) + \mathcal{O}\left(E_{30}(r)\right) + \mathcal{O}\left(E_{31}(0)\right) + \mathcal{O}\left(E_{31}(r)\right) \\ +\; & \mathcal{O}\left(E_{32}(0)\right) + \mathcal{O}\left(E_{32}(r)\right) + \mathcal{O}\left(E_{33}(r)\right) + \mathcal{O}\left(E_{34}(r)\right), \end{aligned} \tag{5.38}$$

for the error terms obtained by 3.27 and $S_3(r,s)$ obtained by replacing the characteristic function by trigonometric polynomials. We now reformulate $S_3(r,s)$ by expanding the trigonometric polynomials, detecting the correct value of $v = v(n)$ and restructuring the sums:

$$\begin{aligned} S_3(r,s) \;=\; & \frac{1}{q^{\lambda-\mu+m-1}} \sum_{0\le h<q^{\lambda-\mu+m-1}} \sum_{|h_1|\le H_1} a_{h_1}(U^{-1},H_1) \\ & \sum_{|h_2|\le H_2} a_{h_2}(U^{-1},H_2) \sum_{|h_3|\le H_3} a_{h_3}(U_3^{-1},H_3) \\ & \sum_{0\le u_1<U}\sum_{0\le u_2<U}\sum_{0\le u_3<U_3}\sum_{0\le v<V} \mathrm{e}\left(-\frac{h_1u_1+h_2u_2}{U}-\frac{h_3u_3}{U_3}-\frac{hv}{q^{\lambda-\mu+m-1}}\right) \\ & \mathrm{e}\Bigg(\sum_{\ell=0}^{k-1}\alpha_\ell\big(b_{\rho',\lambda-\mu+\rho'}(u_1+\ell u_3) - b_{\rho',\lambda-\mu+\rho'}(u_2+\ell u_3) \\ & \qquad -b_{\rho',\lambda-\mu+\rho'}(u_1+\ell u_3+vq^{\rho'}+2\ell s q^{m-1}q^{\rho'}) \\ & \qquad +b_{\rho',\lambda-\mu+\rho'}(u_2+\ell u_3+vq^{\rho'}+2(\ell+r)sq^{m-1}q^{\rho'})\big)\Bigg) \end{aligned}$$

$$\sum_{n\in I_2(r,s)} \mathrm{e}\left(\frac{h_1n^2+h_2(n+r)^2}{q^{\lambda+m-1}}+\frac{2h_3n}{q^{\nu}}+\frac{2hsn}{q^{\lambda-\mu}}\right).$$

One can estimate the error terms just as in [14] and finds that they are bounded by either $q^{\nu-\rho_3}$ or $q^{\nu-\rho_2}$. In conclusion we deduce that

$$S_2(r,s)=S_3(r,s)+\mathcal{O}(q^{\nu-\rho'})+\mathcal{O}(q^{\nu-\rho_2})+\mathcal{O}(q^{\nu-\rho_3}). \tag{5.39}$$

We now split the sum $S_3(r,s)$ into two parts:

$$S_3(r,s)=S_4(r,s)+S_4'(r,s), \tag{5.40}$$

where $S_4(r,s)$ denotes the contribution of the terms for which $h_1+h_2=0$ while $S_4'(r,s)$ denotes the contribution of the terms for which $h_1+h_2\neq 0$. We can estimate $S_4'(r,s)$ just as in [14] and find

$$S_4'(r,s) \ll \nu^4 q^{\nu+\frac{1}{2}(8\lambda-9\mu+7\rho'+\rho_2)}$$

and it remains to consider $S_4(r,s)$. Setting $u_1=u_1''+q^{\rho'}u_1'$, $u_2=u_2''+q^{\rho'}u_2'$ and $u_3=u_3''+q^{\rho'}u_3'$, (where $0\le u_1'',u_2'',u_3''<q^{\rho'}$) we can replace the two-fold restricted block-additive function by a truncated block-additive function

$$\begin{aligned}
&b_{\rho',\lambda-\mu+\rho'}(u_1+\ell u_3)=b_{\lambda-\mu}\left(u_1'+\ell u_3'+\left\lfloor (u_1''+\ell u_3'')/q^{\rho'}\right\rfloor\right),\\
&b_{\rho',\lambda-\mu+\rho'}(u_2+\ell u_3)=b_{\lambda-\mu}\left(u_2'+\ell u_3'+\left\lfloor (u_2''+\ell u_3'')/q^{\rho'}\right\rfloor\right),\\
&b_{\rho',\lambda-\mu+\rho'}(u_1+\ell u_3+vq^{\rho'}+2\ell sq^{m-1}q^{\rho'})=b_{\lambda-\mu}\left(u_1'+v+\ell(u_3'+2sq^{m-1})+\left\lfloor (u_1''+\ell u_3'')/q^{\rho'}\right\rfloor\right)\\
&b_{\rho',\lambda-\mu+\rho'}(u_2+\ell u_3+vq^{\rho'}+2(\ell+r)sq^{m-1}q^{\rho'})\\
&=b_{\lambda-\mu}\left(u_2'+v+2srq^{m-1}+\ell(u_3'+2sq^{m-1})+\left\lfloor (u_2''+\ell u_3'')/q^{\rho'}\right\rfloor\right).
\end{aligned}$$

Using the periodicity of b modulo $V:=q^{\lambda-\mu+m-1}$, we replace the variable v by v_1 such that $v_1\equiv u_1'+v \bmod q^{\lambda-\mu+m-1}$. Furthermore we introduce a new variable v_2 such that

$$v_2\equiv u_2'+v+2srq^{m-1}\equiv v_1+u_2'-u_1'+2srq^{m-1} \bmod q^{\lambda-\mu+m-1}.$$

We then follow the arguments of [14] and find

$$\begin{aligned}
S_4(r,s) \ll\ & q^{2\lambda-2\mu}\sum_{0\le h<q^{\lambda-\mu+m-1}}\sum_{0\le h'<q^{\lambda-\mu+m-1}}\sum_{|h_2|\le H_2}\min(U^{-2},h_2^{-2})\sum_{|h_3|\le H_3}\min(U_3^{-1},h_3^{-1})\\
&\sum_{0\le u_1''<q^{\rho'}}\sum_{0\le u_2''<q^{\rho'}}\sum_{0\le u_3''<q^{\rho'}}\sum_{0\le u_3'<U_3'}\\
&\left|H_{\lambda-\mu}^{I(u_1'',u_3'')}(h'-h-h_2,u_3')\right|\left|H_{\lambda-\mu}^{I(u_2'',u_3'')}(h'-h_2,u_3')\right|\\
&\left|H_{\lambda-\mu}^{I(u_1'',u_3'')}(h'-h,u_3'+2sq^{m-1})\right|\left|H_{\lambda-\mu}^{I(u_2'',u_3'')}(h',u_3'+2sq^{m-1})\right|
\end{aligned}$$

$$\left| \sum_{n \in I_2(r,s)} \mathrm{e}\left(\frac{2h_2rn}{q^{\lambda+m-1}} + \frac{2h_3n}{q^{\nu}} + \frac{2hsn}{q^{\lambda-\mu}} \right) \right|,$$

with

$$I(u,\tilde{u}) = \left(\left\lfloor \frac{u}{q^{\rho'}} \right\rfloor, \left\lfloor \frac{u+\tilde{u}}{q^{\rho'}} \right\rfloor, \ldots, \left\lfloor \frac{u+(k-1)\tilde{u}}{q^{\rho'}} \right\rfloor \right) \quad \text{for } (u,\tilde{u}) \in \mathbb{N}^2.$$

The next few steps are again so similar to the corresponding ones in [14] that we skip the details. We find

$$\begin{aligned} S_4(r,s) \;\ll\; (\lambda-\mu)\ \gcd(2s,q^{\lambda-\mu})\ q^{2\lambda-2\mu} \sum_{0\le u_1'',u_2'',u_3''<q^{\rho'}} \sum_{|h_2|\le H_2} \min(U^{-2},h_2^{-2}) \\ S_6(h_2,s,u_1'',u_3'')^{1/2} S_6(h_2,s,u_2'',u_3'')^{1/2} \\ \sum_{|h_3|\le H_3} \min(U_3^{-1},h_3^{-1}) \min\left(q^{\nu}, \left| \sin \pi \tfrac{2h_2r+2q^{\lambda-\nu+m-1}h_3}{q^{\lambda+m-1}} \right|^{-1} \right). \end{aligned}$$

where

$$S_6(h_2,s,u'',u_3'') = \sum_{0\le u_3'<U_3'} \sum_{0\le h'<q^{\lambda-\mu+m-1}} \left| H_{\lambda-\mu}^{I(u'',u_3'')}(h'-h_2,u_3') \right|^2 \left| H_{\lambda-\mu}^{I(u'',u_3'')}(h',u_3'+2sq^{m-1}) \right|^2. \tag{5.41}$$

Here we introduce the integers H_2' and κ such that

$$H_2' = q^{\lambda-\nu+m} H_3/R_1 = q^{\lambda-\mu+\rho'+\rho_3-\rho_1+m+1} = q^{\kappa}. \tag{5.42}$$

This leads to

$$S_4(r,s) \ll S_{41}(r,s) + S_{42}(r,s) + S_{43}(r,s),$$

where $S_{41}(r,s)$, $S_{42}(r,s)$ and $S_{43}(r,s)$ denote the contribution of the terms $|h_2| \le H_2'$, $H_2' < |h_2| \le q^{\lambda+m-1-\mu}$ and $q^{\lambda+m-1-\mu} < |h_2| \le H_2$ respectively.

Estimate of $S_{41}(r,s)$ By (5.20) we have

$$\sum_{|h_3|\le H_3} \min\left(q^{\nu}, \left| \sin \pi \tfrac{2h_3+2h_2rq^{\nu-\lambda-m+1}}{q^{\nu}} \right|^{-1} \right) \ll \nu q^{\nu},$$

and, therefore,

$$\begin{aligned} S_{41}(r,s) \;\ll\; \nu\,(\lambda-\mu)\ \gcd(2s,q^{\lambda-\mu})\ q^{\nu+2\lambda-2\mu} U^{-2} U_3^{-1} \\ \sum_{0\le u_1'',u_2'',u_3''<q^{\rho'}} \sum_{|h_2|\le H_2'} S_6(h_2,s,u_1'',u_3'')^{1/2} S_6(h_2,s,u_2'',u_3'')^{1/2}. \end{aligned}$$

By Proposition 5.3.8 (replacing λ by $\lambda-\mu$ and L by $\lambda-\mu-\kappa$), we find some $0<\eta'\le 1$ such that

$$\left| H_{\lambda-\mu}^{I(u'',u_3'')}(h'-h_2',u_3') \right| \ll q^{-\eta'(\lambda-\mu-\kappa)} \max_{J\in\mathcal{I}_k} \left| G_{\kappa}^{J}(h'-h_2',\lfloor u_3'/q^L \rfloor) \right|.$$

By Parseval's equality and recalling that $\#(\mathcal{I}_k) = q^{m-1}(q^{m-1}+1)^{k-1}$, it follows that

$$\sum_{|h_2|\leq H_2'} \max_{J\in\mathcal{I}_k} \left|H_\kappa^J\lfloor(h'-h_2,u_3'/q^L\rfloor)\right|^2 \leq \sum_{J\in\mathcal{I}_k}\sum_{|h_2|\leq H_2'} \left|G_\kappa^J(h'-h_2,\lfloor u_3'/q^L\rfloor)\right|^2 \leq q^{m-1}(q^{m-1}+1)^{k-1}.$$

We obtain

$$\sum_{|h_2|\leq H_2'} \left|H_{\lambda-\mu}^{I(u'',u_3'')}(h'-h_2,u_3')\right|^2 \ll q^{-\eta'(\lambda-\mu-\kappa)} = \left(\frac{H_2'}{q^{\lambda-\mu}}\right)^{\eta'}$$

uniformly in λ, μ, H_2', u_3', u'' and u_3''. The remaining proof is analogue to the corresponding proof in [14]. This yields

$$\frac{1}{RS}\sum_{R_1<r<R}\sum_{1\leq s<S} S_{41}(r,s) \ll \nu\,(\lambda-\mu)^{\omega(q)+1}\; q^{\nu-\eta'(\rho_1-\rho'-\rho_3)}, \tag{5.43}$$

which concludes this part.

Estimate of $S_{42}(r,s)$ and $S_{43}(r,s)$ By following the arguments of [14] and applying the same changes as in the estimate of S_{41} we find

$$\frac{1}{RS}\sum_{R_1<r<R}\sum_{1\leq s<S} S_{42}(r,s) \ll \rho\rho_3(\lambda-\mu)^{1+\omega(q)}\; q^{\nu-\rho_3}. \tag{5.44}$$

and

$$\frac{1}{RS}\sum_{R_1<r<R}\sum_{1\leq s<S} S_{43}(r,s) \ll \rho\;(\lambda-\mu)^{1+\omega(q)}\; q^{\nu-\rho+3\rho'}. \tag{5.45}$$

Combining the estimates for S_4 It follows from (5.43), (5.44) and (5.45) that

$$\frac{1}{RS}\sum_{R_1<r<R}\sum_{1\leq s<S} S_4(r,s) \ll \nu^{3+\omega(q)}q^\nu\left(q^{-2\eta'(\rho_1-\rho'-\rho_3)}+q^{-\rho_3}+q^{-\rho+3\rho'}\right).$$

Choosing

$$\rho_1=\rho-\rho',\ \rho_2=\rho_3=\rho',$$

we obtain

$$\frac{1}{RS}\sum_{R_1<r<R}\sum_{1\leq s<S} S_4(r,s) \ll \nu^{3+\omega(q)}q^\nu\left(q^{-2\eta'(\rho-3\rho')}+q^{-\rho'}+q^{-(\rho-3\rho')}\right).$$

Since $0<\eta'<1$, we obtain using (5.40) and (5.39), that

$$\frac{1}{RS}\sum_{R_1<r<R}\sum_{1\leq s<S} S_2(r,s) \ll \nu^{3+\omega(q)}q^\nu\left(q^{-\eta'(\rho-3\rho')}+q^{-\rho'}+q^{\frac{1}{2}(8\lambda-9\mu+8\rho')}\right).$$

We recall by (5.35) that $S = q^{2\rho'}$ and by (5.34) that $\mu = \nu - 2\rho$, $\lambda = \nu + 2\rho$ and insert the estimation from above in (5.36):

$$|S_0|^4 \ll q^{4\nu-2\rho'} + q^{4\nu-2\rho} + \nu^{3+\omega(q)} q^{4\nu} \left(q^{-\eta'(\rho-3\rho')} + q^{-\rho'} + q^{-\frac{\nu}{2}+17\rho+4\rho'}. \right)$$

For $\rho' = \lfloor \nu/146 \rfloor$ and $\rho = 4\rho'$, we obtain

$$|S_0| \ll \nu^{(3+\omega(q))/4} q^{\nu - \frac{\eta'\rho'}{4}} \ll N^{1-\eta_1},$$

for all $\eta_1 < \eta'/584$. Therefore we have seen that Proposition 5.3.8 implies the case $K \not\equiv 0 \bmod 1$ of Theorem 5.1.2.

Bibliography

[1] J.-P. Allouche and J. Shallit. *Automatic sequences. Theory, applications, generalizations.* Cambridge: Cambridge University Press, 2003.

[2] M. V. Berlinkov. On the probability to be synchronizable, preprint, arXiv:1304.5774, 2013.

[3] M. V. Berlinkov. Testing for synchronization, preprint, arXiv:1401.2553, 2014.

[4] J. Bésineau. Indépendance statistique d'ensembles liés à la fonction "somme des chiffres". *Acta Arith.*, 20:401–416, 1972.

[5] J. Bourgain. Möbius-Walsh correlation bounds and an estimate of Mauduit and Rivat. *J. Anal. Math.*, 119:147–163, 2013.

[6] J. Bourgain. On the correlation of the Moebius function with rank-one systems. *J. Anal. Math.*, 120:105–130, 2013.

[7] J. Bourgain, P. Sarnak, and T. Ziegler. Disjointness of Moebius from horocycle flows. In *From Fourier analysis and number theory to Radon transforms and geometry*, volume 28 of *Dev. Math.*, pages 67–83. Springer, New York, 2013.

[8] A. Cobham. Uniform tag sequences. *Math. Systems Theory*, 6:164–192, 1972.

[9] C. Dartyge and G. Tenenbaum. Sommes des chiffres de multiples d'entiers. *Ann. Inst. Fourier (Grenoble)*, 55(7):2423–2474, 2005.

[10] H. Davenport. On some infinite series involving arithmetical functions (II). *The Quarterly Journal of Mathematics*, os-8(1):313–320, 1937.

[11] J.-M. Deshouillers, M. Drmota, and C. Müllner. Automatic Sequences generated by synchronizing automata fulfill the Sarnak conjecture. *Studia Mathematica*, 231:83–95, 2015.

[12] T. Donarowicz and S. Kasjan. Odometers and toeplitz subshifts revisited in the context of sarnak's conjecture, preprint.

[13] M. Drmota. Subsequences of automatic sequences and uniform distribution. In *Uniform distribution and quasi-Monte Carlo methods*, volume 15 of *Radon Ser. Comput. Appl. Math.*, pages 87–104. De Gruyter, Berlin, 2014.

[14] M. Drmota, C. Mauduit, and J. Rivat. The thue-morse sequence along squares is normal. manuscript.

[15] M. Drmota and J. F. Morgenbesser. Generalized Thue-Morse sequences of squares. *Israel J. Math.*, 190:157–193, 2012.

[16] E. H. El Abdalaoui, S. Kasjan, and M. Lemańczyk. 0-1 sequences of the Thue-Morse type and Sarnak's conjecture. *Proc. Amer. Math. Soc.*, 144(1):161–176, 2016.

[17] E. H. El Abdalaoui, M. Lemańczyk, and T. de la Rue. On spectral disjointness of powers for rank-one transformations and Möbius orthogonality. *J. Funct. Anal.*, 266(1):284–317, 2014.

[18] S. Ferenczi. Complexity of sequences and dynamical systems. *Discrete Math.*, 206(1-3):145–154, 1999. Combinatorics and number theory (Tiruchirappalli, 1996).

[19] S. Ferenczi, J. Kułaga-Przymus, M. Lemanczyk, and C. Mauduit. Substitutions and Möbius disjointness, preprint, arXiv:1507.01123v1, 2015.

[20] S. Ferenczi and C. Mauduit. On sarnak's conjecture and veech's question for interval exchanges, preprint.

[21] N. P. Fogg. *Substitutions in dynamics, arithmetics and combinatorics*, volume 1794 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2002. Edited by V. Berthé, S. Ferenczi, C. Mauduit and A. Siegel.

[22] A. Gelfond. Sur les nombres qui ont des propriétés additives et multiplicatives données. *Acta Arith.*, 13:259–265, 1967/1968.

[23] S. W. Graham and G. Kolesnik. *Van der Corput's method of exponential sums*, volume 126 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1991.

[24] B. Green. On (not) computing the Möbius function using bounded depth circuits. *Combin. Probab. Comput.*, 21(6):942–951, 2012.

[25] B. Green and T. Tao. The mobius function is strongly orthogonal to nilsequences. *Ann. of Math. (2)*, 175(2):541–566, 2012.

[26] G. Hanna. Sur les occurrences des mots dans les nombres premiers. *ArXiv e-prints*, Nov. 2015.

[27] K.-H. Indlekofer and I. Kátai. Investigations in the theory of q-additive and q-multiplicative functions. I. *Acta Math. Hungar.*, 91(1-2):53–78, 2001.

[28] H. Iwaniec and E. Kowalski. *Analytic number theory*, volume 53 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2004.

[29] H. Iwaniec and E. Kowalski. *Analytic number theory*, volume 53 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2004.

[30] I. Kátai. A remark on a theorem of H. Daboussi. *Acta Math. Hungar.*, 47(1-2):223–225, 1986.

[31] D.-H. Kim. On the joint distribution of q-additive functions in residue classes. *J. Number Theory*, 74(2):307–336, 1999.

[32] L. Kuipers and H. Niederreiter. *Uniform distribution of sequences*. Wiley-Interscience [John Wiley & Sons], New York-London-Sydney, 1974. Pure and Applied Mathematics.

[33] J. Kułaga-Przymus and M. Lemańczyk. The Moebius function and continuous extensions of rotations. *ArXiv e-prints*, Oct. 2013.

[34] P. Kurka. *Topological and symbolic dynamics*, volume 11 of *Cours Spécialisés [Specialized Courses]*. Société Mathématique de France, Paris, 2003.

[35] J. Liu and P. Sarnak. The Möbius function and distal flows. *Duke Math. J.*, 164(7):1353–1399, 2015.

[36] C. Mauduit and J. Rivat. La somme des chiffres des carrés. *Acta Math.*, 203(1):107–148, 2009.

[37] C. Mauduit and J. Rivat. Sur un problème de Gelfond: la somme des chiffres des nombres premiers. *Ann. of Math. (2)*, 171(3):1591–1646, 2010.

[38] C. Mauduit and J. Rivat. Prime numbers along Rudin-Shapiro sequences. *J. Eur. Math. Soc. (JEMS)*, 17(10):2595–2642, 2015.

[39] C. Müllner and L. Spiegelhofer. Normality of the Thue–Morse sequence along Piatetski-Shapiro sequences, II. *ArXiv e-prints*, Nov. 2015.

[40] P. Sarnak. Three lectures on the Mobius function randomness and dynamics. http://publications.ias.edu/sites/default/files/MobiusFunctionsLectures

[41] J.-P. Serre. *Linear representations of finite groups*. Springer-Verlag, New York-Heidelberg, 1977. Translated from the second French edition by Leonard L. Scott, Graduate Texts in Mathematics, Vol. 42.

[42] T. Tao. Möbius randomness of the rudin-shapiro sequence. http://mathoverflow.net/questions/97261/m2012.

[43] J. D. Vaaler. Some extremal functions in Fourier analysis. *Bull. Amer. Math. Soc. (N.S.)*, 12(2):183–216, 1985.

[44] W. A. Veech. Möbius orthogonality for generalized morse-kakutani flows, preprint.

Acknowledgements

First I want to thank my advisers Michael Drmota and Joël Rivat. They introduced me to many interesting and challenging problems and provided support whenever needed. They had a lot of trust in my abilities and allowed me to work freely on my own idea. Furthermore, I am very thankful for being given the possibility to write this thesis as a *Co-tutelle de thèse*. I would like to thank the responsible people from both the TU Wien and the Universite d Aix Marseille for this great program and everyone who supported me with it.

Moreover, I thank the Austrian Science Foundation FWF who financed my research (during the course of my studies) and also the Institute for Discrete Mathematics and Geometry at the TU Wien and the Institute de Mathématiques de Luminy at the Université d'Aix-Marseille for providing good working conditions.

In this context, I would also like to thank all my colleagues for many interesting conversations and a good working atmosphere. Specifically, I'd like to mention Lukas Spiegelhofer who provided helpful advise on several occasions.

Furthermore, I would like to thank Cécile Dartye and Christian Elsholtz for not only reviewing my thesis but also being part of the jury for my final exam. I also thank Mariusz Lemánczyk for his interest in my work and for completing the exam jury as an additional member.

Special thanks go to my family and especially my parents for their constant support and encouragement throughout my entire life. They have aroused and stimulated my interest in science from an early stage on. Finally, I want to thank Christina Satzinger for her help and support, particularly for spending her time to proofread this thesis and giving helpful advise on structure and formulations.

Clemens Müllner

Curriculum Vitae

Personal Data

Date of birth	07/12/1990
Hometown	Vienna

Education

11/2014 - 02/2017	PhD, Technische Universität Wien and Aix Marseille
5/2013 - 10/2014	MSc, Technische Universität Wien
8/2009 - 4/2013	BSc, Technische Universität Wien

PhD Thesis

Expected Title	Exponential sum estimates and Fourier analytic methods for digitally based dynamical systems
Supervisors	Professor Michael Drmota, Professor Joël Rivat

Master Thesis

Title	On the normality of generalized Thue-Morse sequences
Supervisor	Professor Michael Drmota

Stay Abroad

11/2014 - 02/2017	Cotutelle - Aix Marseille
8/2013- 1/2014	Erasmus stay at KTH Stockholm

Publications

Automatic sequences generated by synchronizing automata fulfill the Sarnak conjecture (with Jean-Marc Deshouillers and Michael Drmota) in Studia Math. 231 (2015), no. 1, 83-95

Normality of the Thue-Morse sequence along Piatetski-Shapiro sequences, II (with Lukas Spiegelhofer) accepted for publication in Israel Journal of Mathematics

Preprints

Automatic Sequences fulfill the Sarnak Conjecture, submitted

Invited Talks

09/2016 Seminaire Teich (Marseille): „Automatic Sequences fulfill the Sarnak Conjecture“

12/2016 Workshop Ergodic Theory and Möbius Disjointness (Marseille): „Automatic Sequences fulfill the Sarnak Conjecture“

05/2017 Workshop: Bridges between Automatic Sequences and algebra and number theory (Montreal): „Automatic Sequences fulfill the Sarnak Conjecture“

Contributed Talks

09/2015 Number Theory, Numeration Systems, Ergodic Theory (Le Bourget du Lac)

02/2016 Séminaire Dynamique, Arithmétique, Combinatoire (Aix Marseille)

05/2016 MUDERA 2016 (Vienna)

09/2016 Conference on Elementary and Analytic number theory (ELAZ) 2016 (Strobl, Austria)

Mathematical Competitions

International mathematics competition (IMC)

2013 116. place (of 321, 2nd Price)

2012 69. place (of 316, 2nd Price)

International mathematics olympiade (IMO)

2009 354. place (of 550, Honorable Mention)

Middle european mathematical olympiad (MEMO)

2008 27. place (3rd Price)

2007 23. place

Austrian mathematical olympiade

2009 4. place (2nd Price, qualification for IMO)

2008 12. place (qualification for MEMO)

2007 12. place (qualification for MEMO)