Universiteit Utrecht

Department of Mathematics

Master Thesis

The Euler totient function in short intervals

Author:
Tom van Overbeeke

Supervisor:
Prof. Dr. Gunther Cornelissen

Second Reader:
Dr. Damaris Schindler

January, 2017

Abstract

It is known that the two rings $\mathbb{Z}$ and $\mathbb{F}_q[T]$ have a lot of similar properties. Here $\mathbb{F}_q[T]$ denotes the polynomial ring over the finite field $\mathbb{F}_q$ of q elements. In 2014 Keating and Rudnick developed a new technique for calculating the variance of functions in short intervals in $\mathbb{F}_q[T]$. In this thesis we study this technique and apply it to the von Mangoldt function Λ, as Keating and Rudnick did to prove a result analogous to a theorem in $\mathbb{Z}$, due to Goldston and Montgomery. We then apply this technique to the Euler totient function to arrive at some new results concerning its variance in short intervals in $\mathbb{F}_q[T]$. Finally we turn our attention to the Euler totient function in short intervals in $\mathbb{Z}$. Surprisingly, it turns out that the analogue to the statement in $\mathbb{F}_q[T]$ does not hold.

Acknowledgements

I would like to start by expressing my deepest gratitude to my mentor and thesis supervisor Gunther Cornelissen. I have worked on this thesis with great pleasure for almost a year and that would not have been possible without him. He introduced me to the subject of this thesis and he has helped me to write it every step of the way, providing literature, answering questions and giving feedback.

He also introduced me to Zeev Rudnick, whom I would like to thank for taking the time to sit down with me and discuss his papers. His suggestions on what I could work myself were exactly the inspiration I needed halfway through my research.

I would also like to thank Damaris Schindler both for being the second reader on this thesis and for suggesting ways to tackle the problems discussed in chapter 8.

Finally I would like to thank my friends Martijn den Besten and Maxim Faber for proofreading this thesis. Their suggestions on how to improve it were extremely useful. Moreover I would like to thank them for keeping me motivated, listening to my problems and encouraging me to keep trying.

Contents

Notation

We start with an overview of the notations used in this thesis.

- For positive functions $f, g : \mathbb{R} \to \mathbb{R}$, define:

$$\begin{array}{ll} f(x) = o(g(x)) & \text{if } \lim_{x\to\infty} \frac{f(x)}{g(x)} = 0; \\ f(x) = O(g(x)) & \text{if } \exists C \in \mathbb{R}_{\geq 0}, x_0 \in \mathbb{R} \text{ such that } |f(x)| \leq Cg(x) \text{ for all } x \geq x_0; \\ f(x) = \Theta(g(x)) & \text{if } \exists c, C \in \mathbb{R}_{\geq 0}, x_0 \in \mathbb{R} \text{ such that } cg(x) \leq |f(x)| \leq Cg(x) \text{ for all } x \geq x_0; \\ f(x) \sim g(x) & \text{if } \lim_{x\to\infty} \frac{f(x)}{g(x)} = 1. \end{array}$$

- Define the functions $\pi, \psi, M : \mathbb{R} \to \mathbb{R}$ by:

$$\pi(x) = \#\{p \leq x \mid p \text{ prime}\} = \sum_{p\leq x} 1;$$

$$\psi(x) = \sum_{p^m \leq x} \log p;$$

$$M(x) = \sum_{n\leq x} \mu(n).$$

- Define the von Mangoldt function $\Lambda : \mathbb{N} \to \mathbb{C}$ by:

$$\Lambda(n) = \begin{cases} \log p & \text{if } n = p^m, \\ 0 & \text{otherwise.} \end{cases}$$

- Define the Euler totient function $\varphi : \mathbb{N} \to \mathbb{C}$ by

$$\varphi(n) = \#(\mathbb{Z}/n\mathbb{Z})^\times$$

- Define the Entier function $[.] : \mathbb{R} \to \mathbb{Z}$ by

$$[x] = \max\{m \in \mathbb{Z} | m \leq x\}.$$

- Define the fractional part function $\{.\} : \mathbb{R} \to \mathbb{R}$ by

$$\{x\} = x - [x].$$

- For any odd prime q define $\mathbb{F}_q$ to be the finite field of q elements. Define $\mathbb{F}_q[T]$ to be the polynomial ring over this finite field $\mathbb{F}_q$.
- Working in $\mathbb{F}_q[T]$, define the norm function $|.| : \mathbb{F}_q[T] \to \mathbb{Z}$ by $|f| = q^{\deg(f)}$.

- Working in $\mathbb{F}_q[T]$, define the following sets of polynomials:

$$\begin{aligned}\mathcal{P}_n &= \{f \in \mathbb{F}_q[T] : f \text{ has degree } n\};\\ \mathcal{P}_{\leq n} &= \{f \in \mathbb{F}_q[T] : f \text{ has degree less than or equal to } n\};\\ \mathcal{M}_n &= \{f \in \mathbb{F}_q[T] : f \text{ is monic of degree } n\}.\end{aligned}$$

- Given a polynomial $A \in \mathcal{P}_n$ and $0 \leq h \leq n$, define the interval $I(A;h)$ as

$$I(A;h) = \{f \in \mathbb{F}_q[T] : |f - A| \leq q^h\} = A + \mathcal{P}_{\leq h}.$$

- Given a function $\alpha : \mathbb{F}_q[T] \to \mathbb{C}$, denote the average of α over monic polynomials of degree n by

$$\langle \alpha \rangle_n = \frac{1}{q^n} \sum_{f \in \mathcal{M}_n} \alpha(f).$$

- Given a function $\alpha : \mathbb{F}_q[T] \to \mathbb{C}$, denote the sum of α over a short interval $I(A;h)$, where $A \in \mathcal{M}_n$, $0 \leq h \leq n-2$, by

$$\mathcal{N}_\alpha(A;h) = \sum_{f \in I(A;h)} \alpha(f).$$

- Given a function $\alpha : \mathbb{F}_q[T] \to \mathbb{C}$, denote the average of $\mathcal{N}_\alpha(A;h)$ over all monic polynomials A of degree n by

$$\langle \mathcal{N}_\alpha(\bullet;h) \rangle_n = \frac{1}{q^n} \sum_{A \in \mathcal{M}_n} \mathcal{N}_\alpha(A;h).$$

- Given a function $\alpha : \mathbb{F}_q[T] \to \mathbb{C}$, denote the variance of $\mathcal{N}_\alpha(A;h)$ over all monic polynomials A of degree n by

$$\mathrm{Var}_n \mathcal{N}_\alpha(\bullet;h) = \frac{1}{q^n} \sum_{A \in \mathcal{M}_n} \left| \mathcal{N}_\alpha(A;h) - \langle \mathcal{N}_\alpha(\bullet;h) \rangle_n \right|^2.$$

Chapter 1

Introduction

1.1 Two important rings

This master thesis takes place in two rings with similar properties. The first is a ring everybody knows: the ring of integers $\mathbb{Z}$. The second is less known. Let q be an odd prime and denote by $\mathbb{F}_q$ the finite field of q elements. The ring we are interested in is $\mathbb{F}_q[T]$, the polynomial ring with coefficients in $\mathbb{F}_q$. Both rings are principal ideal domains on which we can define a norm function and it turns out that they have a lot more interesting properties in common. It is even possible to define some sort of dictionary between the two rings. This dictionary would then make it possible to translate theorems and conjectures from one ring to another. There are even cases of conjectures we do not know how to prove in $\mathbb{Z}$, but where we are able to prove their translation to $\mathbb{F}_q[T]$! Such a dictionary might look like this.

$\mathbb{Z}$	$\leftrightarrow$	$\mathbb{F}_q[T]$
An integer $n \in \mathbb{Z}$ of norm $\lvert n\rvert = \#(\mathbb{Z}/n\mathbb{Z})$	$\leftrightarrow$	A polynomial $f \in \mathbb{F}_q[T]$ of norm $\lvert f\rvert = \#(\mathbb{F}_q[T]/(f)) = q^{\deg(f)}$
$\log\lvert n\rvert$	$\leftrightarrow$	$\log_q\lvert f\rvert = \deg(f)$
A unit $\pm 1 \in \mathbb{Z}$	$\leftrightarrow$	A unit $c \in \mathbb{F}_q^\times$
A prime $p \in \mathbb{Z}_{\geq 0}$	$\leftrightarrow$	An irreducible monic polynomial $P \in \mathbb{F}_q[T]$
Von Mangoldt function over $\mathbb{Z}$ $\Lambda(n) = \begin{cases} \log(p) & \text{if } n = \pm p^k \\ 0 & \text{otherwise.} \end{cases}$	$\leftrightarrow$	Von Mangoldt function over $\mathbb{F}_q[T]$ $\Lambda(f) = \begin{cases} \deg(P) & \text{if } f = cP^k \\ 0 & \text{otherwise.} \end{cases}$
Euler totient function over $\mathbb{Z}$ $\varphi(n) = \#(\mathbb{Z}/n\mathbb{Z})^\times$	$\leftrightarrow$	Euler totient function over $\mathbb{F}_q[T]$ $\varphi(f) = \#(\mathbb{F}_q[T]/(f))^\times$

For an example of a translation, you could look at the Prime Number Theorem (PNT for short), stating that $\pi(x) \sim \frac{x}{\log x}$ as $x \to \infty$. Here $\pi(x)$ is given by the number of primes smaller than or equal to x. The PNT was proven independently by Hadamard [13] and de la Vallée Poussin [6] in 1896. It is well known that the PNT is equivalent to the statement $\sum_{n \leq x} \Lambda(n) \sim x$, with Λ the von Mangoldt function, see for example [2]. Equivalently you might expect that in $\mathbb{F}_q[T]$ you would have that

$$\sum_{\substack{f \text{ monic} \\ \deg(f) \leq n}} \Lambda(f) \sim q^n.$$

Note that you only look at the monic polynomials, as you take all polynomials up to a unit. This is analogous to the statement in $\mathbb{Z}$, where, by taking only the positive numbers, you also take all numbers up to a unit. It turns out that we even have a stronger theorem in $\mathbb{F}_q[T]$. Define $\mathcal{M}_n$ to be the set of monic polynomials of degree n, then

$$\sum_{f \in \mathcal{M}_n} \Lambda(f) = q^n.$$

We will prove this "Prime Polynomial Theorem" in chapter 5.

1.2 Short intervals and the major theorems of this thesis

For $X \in \mathbb{Z}$, we can define an interval around X of size H. It does not really matter how you define this interval. For example you can take the interval $[X - \frac{H}{2}, X + \frac{H}{2}]$, but also just $[X, X + H]$. More important is the size of the interval H. If $H = \Theta(X^\delta)$ for some $0 < \delta < 1$, we speak of a *short interval*. The size of the interval tends to infinity as $X \to \infty$, but slower than X. It is then possible to consider (the sum of) functions in short intervals. Define

$$\Psi(X; H) = \sum_{n \in [X - \frac{H}{2}, X + \frac{H}{2}]} \Lambda(n).$$

Not much is known about this function $\Psi(X; H)$. The Riemann Hypothesis implies that $\Psi(X; H) \sim H$ as long as $\delta > \frac{1}{2} + o(1)$. This is due to the fact that

$$\Psi(X; H) = \psi(X + \frac{H}{2}) - \psi(x - \frac{H}{2}),$$

where $\psi(x) = \sum_{n \leq x} \Lambda(n)$ and the fact that Riemann hypothesis implies that $\psi(x) = x + O\left(x^{\frac{1}{2}} (\ln x)^2\right)$, as is shown in [5]. The pair correlation conjecture, stated by Montgomery [22] in 1973, says that the corellation between the zeroes of the Riemann zeta function, normalized to have unit avarage spacing, is given by $1 - \left(\frac{\sin(\pi u)}{\pi u}\right)^2 - \delta(u)$. Assuming the Riemann Hypothesis and the pair correlation conjecture, Goldston and Montgomery [12] showed, that for $H = \Theta(X^\delta)$ for some $0 < \delta < 1$, the following expression for the variance of $\Psi(X; H)$ holds:

$$\frac{1}{X} \int_2^X |\Psi(x; H) - H|^2 \, dx \sim H(\log X - \log H). \tag{1.1}$$

In $\mathbb{F}_q[T]$, the Riemann hypothesis is proven. Therefore it might be possible to prove stronger statements than it is in the ring of integers. This is indeed the case. In this thesis we study a technique introduced by Keating and Rudnick in [20] to prove some of these statements. Let $A \in \mathcal{P}_n$ be a polynomial of degree n. For any $0 \leq h \leq n$ the interval around A is defined as

$$I(A; h) = \{f \in \mathbb{F}_q[T] : |f - A| \leq q^h\}.$$

The size of this interval is given by q^{h+1}. Hence for $0 \leq h \leq n - 2$ we speak of a short interval, because as $q \to \infty$, $|A| = q^n$ and $\#I(A, h) = q^{h+1}$ both tend to infinity, but the latter is again slower. That is $\lim_{q \to \infty} \frac{\#I(A,h)}{|A|} = 0$. Note that we take the limit $q \to \infty$, while we fix A and h. As we want the size of the interval to tend to infinity, it would also be possible to let n and

h tend to infinity, while we keep q fixed. This is another subject entirely, which we will not consider in this thesis. Now fix $0 \leq h \leq n-2$. Denote by $\mathcal{M}_n$ the set of monic polynomials of degree n and for $A \in \mathcal{M}_n$, define

$$\mathcal{N}_\Lambda(A;h) = \sum_{f \in I(A;h)} \Lambda(f).$$

Define $\langle \mathcal{N}_\Lambda(\bullet;h) \rangle_n$ to be the average of $\mathcal{N}_\Lambda(A;h)$ as A runs over $\mathcal{M}_n$ for some fixed n. Finally define the variance of $\mathcal{N}_\Lambda$ to be

$$\mathrm{Var}_n \mathcal{N}_\Lambda(\bullet;h) = \frac{1}{q^n} \sum_{A \in \mathcal{M}_n} \left| \mathcal{N}_\Lambda(A;h) - \langle \mathcal{N}_\Lambda(\bullet;h) \rangle_n \right|^2.$$

Theorem 1.1 (Keating and Rudnick, [20], theorem 2.1)**.** *Fix $0 < h < n-3$. As $q \to \infty$,*

$$\mathrm{Var}_n \mathcal{N}_\Lambda(\bullet;h) \sim q^{h+1}(n-h-2).$$

Note that this is an exact analogue of relation (1.1). For this proof, Keating and Rudnick developed a new technique, which they later used in [18] and [19] to calculate the variance of the Möbius function and the divisor function in short intervals. They were able to transform the short intervals into short arithmetic progressions. They could now apply Dirichlet characters to pick out the progression. Taking the limit $q \to \infty$ and applying some major work by Katz in [15] and [16] to transform a sum over characters into a matrix integral, they could finally calculate the variance. In chapter 6 we will prove theorem 1.1. In chapter 7 we will prove a similar theorem for the Euler totient function, (see the dictionary in the previous section for the definition of this function φ).

Theorem 1.2 (This thesis, theorem 7.16)**.** *Fix $0 < h < n-3$. As $q \to \infty$,*

$$\mathrm{Var}_n \mathcal{N}_{\frac{\varphi(f)}{|f|}}(\bullet;h) \sim q^{-h-3}.$$

Finally in chapter 8 we study the Euler function in $\mathbb{Z}$. Analogous to theorem 1.2, you might expect the variance to be inversely proportional to the size of the interval. It turns out this is not the case. We cannot prove this definitively, but we can show that it follows from some assumptions.

Theorem 1.3 (This thesis, theorem 8.25)**.** *Let $H = \Theta(X^\delta)$, $0 < \delta \leq 1$. Assuming 8.21 and 8.24, we find*

$$\frac{1}{X} \sum_{x=1}^{X} \left[\left(\sum_{x<n<x+H} \frac{\varphi(n)}{n} - \frac{H}{\zeta(2)} \right)^2 \right] \xrightarrow{X \to \infty} \frac{1}{6\zeta(2)} - \frac{1}{6\zeta(2)^2}.$$

What these assumptions exactly are, we will see in chapter 8.

In the four preceding chapters 2-5 we give background information and an introduction in the subjects needed for the proofs in chapters 6-8. We specifically look at theorems and notions needed in the later chapters. We also give references for a more thorough introduction and for the theorems we cannot prove, due to a lack of space.

Chapter 2

Introduction to analytic number theory

2.1 Introduction

In this chapter we give a brief introduction to analytic number theory. As with all of the four introductory chapters of this master thesis, we will only focus on the theorems required for our research in later chapters. In this chapter the notions we need later on are the definitions and theorems in section 2.2, together with theorem 2.31. We will also prove some first properties of the Euler totient function in section 2.6.

For some standard works about analytic number theory, one could look at [1] or [2]. We will follow the same books for proving the theorems in this chapter.

2.2 Convolution product and Möbius inversion

In this section we give a quick introduction into arithmetic functions.

Definition 2.1. *An* arithmetic function *f is a function $f : \mathbb{N} \to \mathbb{C}$.*

Definition 2.2. *An arithmetic function f is called* multiplicative *if $f(mn) = f(m)f(n)$ for any coprime $m, n \in \mathbb{N}$.*

As for any multiplicative function f we have $f(p_1^{k_1} \dots p_n^{k_n}) = f(p_1^{k_1}) \dots f(p_n^{k_n})$, the following lemma follows trivially.

Lemma 2.3. *Any multiplicative arithmetic function f is uniquely defined by its values $f(p^k)$ for every prime power p^k.* □

Definition 2.4. *Given two arithmetic functions f, g, define their* convolution product *$f * g : \mathbb{N} \to \mathbb{C}$ by*

$$(f * g)(n) = \sum_{a \cdot b = n} f(a)g(b) = \sum_{d|n} f(d)g\left(\frac{n}{d}\right).$$

Example 2.5. *Two examples of arithmetic functions are the functions $e, E : \mathbb{N} \to \mathbb{C}$ defined by $e(1) = 1$, $e(n) = 0$ for $n \neq 1$ and $E(n) = 1$ for all $n \in \mathbb{N}$. Now for any arithmetic function f, we have that $e * f(n) = f(n)$ and*

$$E * f(n) = \sum_{d|n} f(d).$$

Lemma 2.6. *The convolution product makes the set of arithmetic functions f with $f(1) \neq 0$ into an Abelian group with identity e.*

Proof. To prove this, we first show the commutativity and associativity of the convolution product. Commutativity is clear, as for any two arithmetic functions f, g we have

$$(f * g)(n) = \sum_{a \cdot b = n} f(a)g(b) = \sum_{b \cdot a = n} f(b)g(a) = (g * f)(n).$$

Associativity is slightly less obvious, but this follows from

$$((f * g) * h)(n) = \sum_{a \cdot b = n} h(b) \sum_{x \cdot y = a} f(x)g(y) = \sum_{a \cdot b \cdot c = n} f(a)g(b)h(c)$$

and

$$(f * (g * h))(n) = \sum_{a \cdot b = n} f(a) \sum_{x \cdot y = b} g(x)h(y) = \sum_{a \cdot b \cdot c = n} f(a)g(b)h(c).$$

for any arithmetic f, g, h. Now we already noted that we have an identity element e, so we only need to prove the existence of an inverse f^{-1} for any arithmetic function f. Fix f with $f(1) \neq 0$ and suppose that we are given an f^*, such that $(f * f^*)(n) = e(n)$. Then

$$1 = e(1) = \sum_{a \cdot b = 1} f(a)f^*(b) = f(1)f^*(1)$$

and hence $f^*(1) = \frac{1}{f(1)}$. We now prove that for any $n \geq 2$, $f^*(n)$ is uniquely defined by its values $f^*(m)$ for all $m < n$. Suppose we know al these values, then

$$0 = e(n) = \sum_{d|n} f\left(\frac{n}{d}\right) f^*(d) = f^*(n)f(1) + \sum_{\substack{d|n \\ d<n}} f\left(\frac{n}{d}\right) f^*(d).$$

Hence

$$f^*(n) = \frac{-1}{f(1)} \sum_{\substack{d|n \\ d<n}} f\left(\frac{n}{d}\right) f^*(d).$$

It follows that for any f there exists some f^{-1} s.t. $(f * f^{-1})(n) = e(n)$ and that this inverse is unique. This proves the lemma. □

Lemma 2.7. *The set of multiplicative arithmetic functions is a subgroup of the group described in the previous lemma.*

Proof. Not that if a function is multiplicative, then $f(1) = 1$, so the set of multiplicative functions is indeed a subset of the set described in the previous lemma. Clearly the identity function is multiplicative. We first prove that the convolution product preserves multiplicativity. Let f, g denote two multiplicative arithmetic functions and let m, n be two coprime integers. Note that

when $a|mn$ for m, n coprime, then we can write a uniquely as the product $a = bc$ with $b|m$ and $c|n$. Furthermore if $b|m$ and $c|n$, then $bc|mn$ if m, n are coprime. Hence

$$\begin{aligned}(f*g)(mn) &= \sum_{a|mn} f(a)g\left(\frac{mn}{a}\right)\\ &= \sum_{b|m}\sum_{c|n} f(bc)g\left(\frac{mn}{bc}\right)\\ &= \sum_{b|m}\sum_{c|n} f(b)f(c)g\left(\frac{m}{b}\right)g\left(\frac{n}{c}\right)\\ &= \left(\sum_{b|m} f(b)g\left(\frac{m}{b}\right)\right)\left(\sum_{c|n} f(c)g\left(\frac{n}{c}\right)\right)\\ &= (f*g)(m)\cdot(f*g)(n).\end{aligned}$$

This shows that the convolution product of two multiplicative functions is again multiplicative. Finally we show that the inverse of a multiplicative function is again multiplicative. Suppose that it is not. Then there exists a multiplicative function f, s.t. f^{-1} is not multiplicative. Let $m, n \in \mathbb{N}$ be coprime integers such that mn denotes the smallest integer such that $f^{-1}(mn) \neq f^{-1}(m)f^{-1}(n)$. Note that $m, n \neq 1$, as multiplicativity implies that $f(1) = 1$ and hence $f^{-1}(1) = 1$. Now

$$\begin{aligned}0 &= (f*f^{-1})(mn)\\ &= \sum_{a|mn} f(a)f^{-1}\left(\frac{mn}{a}\right)\\ &= f(1)f^{-1}(mn) + \sum_{\substack{a|mn\\ a<mn}} f(a)f^{-1}\left(\frac{mn}{a}\right)\\ &= f^{-1}(mn) + \sum_{\substack{b|m,c|n\\ bc<mn}} f(bc)f^{-1}\left(\frac{mn}{bc}\right)\\ &= f^{-1}(mn) + \sum_{\substack{b|m,c|n\\ bc<mn}} f(b)f(c)f^{-1}\left(\frac{m}{b}\right)f^{-1}\left(\frac{n}{c}\right)\\ &= f^{-1}(mn) + \left(\sum_{b|m} f(b)f^{-1}\left(\frac{m}{b}\right)\right)\left(\sum_{c|n} f(c)f^{-1}\left(\frac{n}{c}\right)\right) - f(m)f(n)\\ &= f(mn) - f(m)f(n) + (f*f^{-1})(m)\cdot(f*f^{-1})(n)\\ &= f(mn) - f(m)f(n).\end{aligned}$$

We have found a contradiction and conclude that f^{-1} is multiplicative. We conclude that the lemma holds. □

Definition 2.8. *Define the* Möbius function $\mu : \mathbb{N} \to \mathbb{C}$ *by*

$$\mu(n) = \begin{cases} 0 & \text{if } n \text{ is not square free,}\\ (-1)^t & \text{is } n = p_1 \dots p_t \text{ for } t \text{ distinct primes.}\end{cases}$$

Lemma 2.9. *The Möbius function μ is the inverse of the arithmetic function $E(n) = 1$.*

Proof. If $n = 1$, then

$$(E * \mu)(1) = \mu(1) = 1 = e(1).$$

Note that μ is multiplicative. It is hence sufficient to show that $(E * \mu)(p^k) = 0$ for all primes p and integers $k \geq 1$, as $E * \mu$ is multiplicative by lemma 2.7. We see that

$$(E * \mu)(p^k) = \sum_{d|p^k} \mu(d) = \mu(1) + \mu(p) + \mu(p^2) + \cdots + \mu(p^k) = 1 - 1 = 0,$$

applying the definition of μ. □

Corollary 2.10 (Möbius inversion)**.** *Let g be an arithmetic function and let $f : \mathbb{N} \to \mathbb{C}$ be defined by $f(n) = \sum_{d|n} g(d)$. Then $g(n) = \sum_{d|n} \mu(d) f\left(\frac{n}{d}\right)$.*

Proof. As $f = E * g$, we see that

$$g = e * g = (\mu * E) * g = \mu * (E * g) = \mu * f.$$

□

Example 2.11. *Let $\sigma_0(n) = \#\{d \in \mathbb{N} \text{ such that } d|n\}$ denote the divisor function of n. Then $\sigma_0(n) = \sum_{d|n} 1 = \sum_{d|n} E(d)$. Applying Möbius inversion, we see that*

$$1 = E(n) = \sum_{d|n} \mu(d)\sigma_0\left(\frac{n}{d}\right)$$

for any n.

Definition 2.12. *Given an arithmetic function f, we define its L-*series *as a formal power series by*

$$L_f(s) = \sum_{n \geq 1} f(n) n^{-s}.$$

Lemma 2.13. *For any two arithmetic functions f, g, we have that*

$$L_f \cdot L_g = L_{f*g}.$$

Proof. Let f, g be any two arithmetic functions, then

$$L_f \cdot L_g = \left(\sum_{n \geq 1} f(n) n^{-s}\right)\left(\sum_{m \geq 1} g(m) m^{-s}\right) = \sum_{m,n \geq 1} f(n) g(m) (nm)^{-s} = \sum_{n \geq 1} \sum_{d|n} f(d) g\left(\frac{n}{d}\right) n^{-s}.$$

By definition, the right hand side equals L_{f*g}. □

Corollary 2.14. *For the Möbius function μ we have that $L_\mu(s) = \frac{1}{\zeta(s)}$.*

Proof. Note that by definition

$$\zeta(s) = \sum_{n \geq 1} n^{-s} = L_E(s).$$

Now $E * \mu = e$ and $L_e = \sum_{n \geq 1} e(n) n^{-s} = 1^{-s} = 1$. Therefore

$$1 = L_e(s) = L_{E*\mu}(s) = L_E(s) \cdot L_\mu(s) = \zeta(s) \cdot L_\mu(s),$$

from which we conclude that $L_\mu(s) = \frac{1}{\zeta(s)}$. □

2.3 Another Möbius inversion

In this section we will look at another use of Möbius inversion. We are especially interested in its corollaries 2.16 and 2.17 as we need them later on in the chapter.

Theorem 2.15. *Suppose $f, g : \mathbb{R}_{\geq 1} \to \mathbb{R}$ are real functions. If*

$$f(x) = \sum_{n \leq x} g\left(\frac{x}{n}\right), \tag{2.1}$$

then

$$g(x) = \sum_{n \leq x} \mu(n) f\left(\frac{x}{n}\right). \tag{2.2}$$

Conversely, if $g(x)$ is defined by the lower equation, then $f(x)$ is given by the upper equation.

Proof. First suppose that relation (2.1) holds. Then

$$\begin{aligned}\sum_{n \leq x} \mu(n) f\left(\frac{x}{n}\right) &= \sum_{n \leq x} \mu(n) \sum_{m \leq \frac{x}{n}} g\left(\frac{x}{mn}\right) = \sum_{\substack{n \leq x \\ m \leq \frac{x}{n}}} \mu(n) g\left(\frac{x}{mn}\right) \\ &\overset{(a)}{=} \sum_{k \leq x} \sum_{n|k} \mu(n) g\left(\frac{x}{k}\right) \overset{(b)}{=} \sum_{k \leq x} e(k) g\left(\frac{x}{k}\right) = g(x).\end{aligned}$$

Here we substituted $k = mn$ at (a) and at (b) we used that $e = \mu * E$, implying that $e(k) = \sum_{n|k} \mu(n)$. Next suppose that relation (2.2) holds. Then we use the same reasoning to see that

$$\begin{aligned}\sum_{n \leq x} g\left(\frac{x}{n}\right) &= \sum_{n \leq x} \sum_{m \leq \frac{x}{n}} \mu(m) f\left(\frac{x}{mn}\right) = \sum_{\substack{n \leq x \\ m \leq \frac{x}{n}}} \mu(m) f\left(\frac{x}{mn}\right) \\ &\overset{(a)}{=} \sum_{k \leq x} \sum_{m|k} \mu(m) f\left(\frac{x}{k}\right) \overset{(b)}{=} \sum_{k \leq x} e(k) f\left(\frac{x}{k}\right) = f(x).\end{aligned}$$

□

Corollary 2.16.

$$\sum_{n \leq x} \mu(n) \left[\frac{x}{n}\right] = 1$$

Proof. Apply theorem 2.15 with $g(x) = 1$ for all x. Then $f(x) = \sum_{n \leq x} 1 = [x]$, so

$$\sum_{n \leq x} \mu(n) \left[\frac{x}{n}\right] = g(x) = 1$$

□

Corollary 2.17. *For any $x \in \mathbb{R}_{\geq 1}$ we have that*

$$\left| \sum_{n \leq x} \frac{\mu(n)}{n} \right| < 1.$$

Proof. Fix $x \in \mathbb{R}_{\geq 1}$. Define the fractional part of any real x as $\{x\} = x - [x]$, (so $0 \leq \{x\} < 1$). Applying corollary 2.16, we see that

$$\begin{aligned} x\left|\sum_{n\leq x}\frac{\mu(n)}{n}\right| &= \left|\sum_{n\leq x}\mu(n)\frac{x}{n}\right| = \left|\sum_{n\leq x}\mu(n)\left(\left[\frac{x}{n}\right]+\left\{\frac{x}{n}\right\}\right)\right| \\ &\leq \left|\sum_{n\leq x}\mu(n)\left[\frac{x}{n}\right]\right| + \left|\sum_{n\leq x}\mu(n)\left\{\frac{x}{n}\right\}\right| \leq 1+\{x\}+\left|\sum_{2\leq n\leq x}\left\{\frac{x}{n}\right\}\right| \\ &\leq 1+\{x\}+([x]-1) = x. \end{aligned}$$

□

2.4 The Prime Number Theorem in equivalent forms

In this section we will introduce the famous prime number theorem, (PNT for short). We will not prove it, but every book on analytic number theory, such as [1] or [2], contains one or more proofs of this theorem.

Definition 2.18. *Define the functions $\pi, \psi, M : \mathbb{R} \to \mathbb{R}$ by:*

$$\begin{aligned} \pi(x) &= \#\{p \leq x \mid p \text{ prime}\} = \sum_{p\leq x} 1; \\ \psi(x) &= \sum_{p^m\leq x} \log p; \\ M(x) &= \sum_{n\leq x}\mu(n). \end{aligned}$$

Definition 2.19. *Define the von Mangoldt function $\Lambda : \mathbb{N} \to \mathbb{C}$ by:*

$$\Lambda(n) := \begin{cases} \log p & \text{if } n = p^m, \\ 0 & \text{otherwise.} \end{cases}$$

Note that with this function we have that $\psi(x) = \sum_{n\leq x}\Lambda(n)$. Also note that

$$\psi(x) = \sum_{p^m\leq x}\log p = \sum_{p\leq x}\left[\log_p(x)\right]\log p = \sum_{p\leq x}\left[\frac{\log x}{\log p}\right]\log p.$$

We can now state the PNT, which was proven independently by Hadamard [13] and de la Vallée Poussin [6] in 1896.

Theorem 2.20 (Prime Number Theorem)**.**

$$\pi(x) \sim \frac{x}{\log x}$$

The prime number theorem is of course a famous theorem in analytic number theory. It has a lot of equivalent forms. One of these forms is given by the statement

Theorem 2.21.

$$\psi(x) - x = o(x).$$

Usually a course on Analytic Number Theory will first prove the statement above and then prove the two statements to be equivalent. We will not do this here. For a rigorous proof, see [2, Chapter 2]. Theorem 2.21 will be the form of the Prime number theorem that we will use in this thesis. The following statement is also equivalent to the Prime number theorem.

Theorem 2.22.

$$M(x) = o(x).$$

Again we will not prove that the equivalence, but because this statement is important to us, we will prove that the prime number theorem implies it. Hence we will prove that

Theorem 2.23. *If $\psi(x) - x = o(x)$, then $M(x) = o(x)$.*

Here we follow the proof of Apostol in [1]. We'll first need two lemmas.

Lemma 2.24 (Partial summation)**.** *Let $a_n \in \mathbb{C}$ for every integer $n \geq 1$. Define $A(t) := \sum_{n \leq t} a_n$ and suppose $g : [1, \infty) \to \mathbb{C}$ is a differentiable function. Finally let $x \in \mathbb{R}$. Then*

$$\sum_{n \leq x} a_n g(n) = A(x)g(x) - \int_1^x A(t)g'(t)dt.$$

Proof. Note that according to the definition $A(0) = 0$, so we find that

$$\begin{aligned}\sum_{n \leq x} a_n g(n) &= \sum_{n \leq x} \left(A(n) - A(n-1)\right) g(n) = \sum_{n \leq x} A(n)g(n) - \sum_{n \leq x} A(n-1)g(n) \\ &= \sum_{n \leq x} A(n)g(n) - \sum_{n \leq x-1} A(n)g(n+1) = A([x])g([x]) - \sum_{n \leq x-1} A(n)(g(n+1) - g(n)).\end{aligned}$$

Now for fixed n it holds that

$$A(n)(g(n+1) - g(n)) = A(n) \int_n^{n+1} g'(t)dt = \int_n^{n+1} A(t)g'(t)dt,$$

as $A(t) = A(n)$ for all $n \leq t < n + 1$. Analogous $A(x)g(x) - A([x])g([x]) = \int_{[x]}^x A(t)g'(t)dt$. This implies that

$$\sum_{n \leq x} a_n g(n) = A([x])g([x]) - \sum_{n \leq x-1} \int_n^{n+1} A(t)g'(t)dt = A(x)g(x) - \int_1^x A(t)g'(t)dt.$$

□

Defining

$$H(x) = \sum_{n \leq x} \mu(n) \log n,$$

we have

Lemma 2.25.

$$\lim_{x \to \infty} \left(\frac{M(x)}{x} - \frac{H(x)}{x \log x} \right) = 0.$$

Proof. Applying partial summation with $g(t) = \log t$, $a_n = \mu(n)$, we see that

$$H(x) = \sum_{n \leq x} \mu(n) \log n = M(x) \log x - \int_1^x \frac{M(t)}{t} dt.$$

Hence

$$\begin{aligned} \left| \frac{M(x)}{x} - \frac{H(x)}{x \log x} \right| &= \left| \frac{1}{x \log x} \int_1^x \frac{M(t)}{t} dt \right| \leq \frac{1}{x \log x} \int_1^x \left| \frac{M(t)}{t} \right| dt \\ &\leq \frac{1}{x \log x} \int_1^x dt = \frac{1}{\log x}. \end{aligned}$$

Here we used the obvious inequality $|M(t)| \leq \sum_{n \leq t} |\mu(n)| \leq t$. The lemma is proven. □

We're now able to prove theorem 2.23.

Proof of theorem 2.23. We first note that

$$\log n = \sum_{d|n} \Lambda(d)$$

for each n. This follows from the fact that if $n = p_1^{k_1} \dots p_m^{k_m}$, then

$$\begin{aligned} \sum_{d|n} \Lambda(d) &= \sum_{j=1}^{k_1} \Lambda(p_1^j) + \dots + \sum_{j=1}^{k_n} \Lambda(p_n^j) = \sum_{j=1}^{k_1} \log p_1 + \dots + \sum_{j=1}^{k_m} \log p_m \\ &= k_1 \log p_1 + \dots + k_m \log p_m = \log n. \end{aligned}$$

Applying Möbius inversion, we see that

$$\begin{aligned} \Lambda(n) &= \sum_{d|n} \mu(d) \log \frac{n}{d} = \log n \sum_{d|n} \mu(d) - \sum_{d|n} \mu(d) \log(d) \\ &= e(n) \log(n) - \sum_{d|n} \mu(d) \log(d) = - \sum_{d|n} \mu(d) \log(d). \end{aligned}$$

The last equality follows from the fact that $e(n) = 0$, except when $n = 1$, in which case $\log n = 0$. Now we apply Möbius inversion again, to see that

$$-\mu(n) \log n = \sum_{d|n} \mu(d) \Lambda\left(\frac{n}{d}\right).$$

It follows that

$$-H(x) = -\sum_{n \leq x} \mu(n) \log n = \sum_{n \leq x} \sum_{d|n} \mu(d) \Lambda\left(\frac{n}{d}\right) = \sum_{d \leq x} \mu(d) \sum_{n \leq \frac{x}{d}} \Lambda(n) = \sum_{d \leq x} \mu(d) \psi\left(\frac{x}{d}\right).$$

Fix $\epsilon > 0$. We will show that there exists some $B \in \mathbb{R}$ s.t. $x > B$ implies that $\left| \frac{H(x)}{x \log x} \right| < \epsilon$. By assumption $\psi(x) - x = o(x)$, so there exists an $A \in \mathbb{R}$ s.t. $x > A$ implies that $|\psi(x) - x| \leq \frac{\epsilon}{2} x$.

Suppose $x > A$ and define $y = \left[\frac{x}{A}\right]$. Now if $n \leq y \leq \frac{x}{A}$, then $\frac{x}{n} \geq A$. Hence

$$\begin{aligned}\left|\sum_{n\leq y}\mu(n)\psi\left(\frac{x}{n}\right)\right| &= \left|\sum_{n\leq y}\mu(n)\left(\frac{x}{n}+\psi\left(\frac{x}{n}\right)-\frac{x}{n}\right)\right| \\ &\leq x\left|\sum_{n\leq y}\frac{\mu(n)}{n}\right| + \sum_{n\leq y}\left|\psi\left(\frac{x}{n}\right)-\frac{x}{n}\right| \\ &\overset{(a)}{\leq} x+\frac{\epsilon}{2}\sum_{n\leq y}\frac{x}{n} \\ &< x+\frac{\epsilon}{2}x(1+\log y) < x+\frac{\epsilon}{2}x+\frac{\epsilon}{2}x\log x.\end{aligned}$$

At (a) we used corollary 2.17. Next suppose n is such that $\frac{x}{A} < y+1 \leq n \leq x$. Then $A > \frac{x}{y+1} \geq \frac{x}{n}$, so $\psi(A) \geq \psi\left(\frac{x}{n}\right)$, as ψ is increasing. Now

$$\left|\sum_{y<n\leq x}\mu(n)\psi\left(\frac{x}{n}\right)\right| \leq \sum_{y<n\leq x}|\mu(n)|\psi(A) \leq x\psi(A).$$

We find that

$$|H(x)| = \left|\sum_{d\leq x}\mu(d)\psi\left(\frac{x}{d}\right)\right| \leq x+\frac{\epsilon}{2}x+\frac{\epsilon}{2}x\log x + x\psi(A) < (2+\psi(A))x+\frac{\epsilon}{2}x\log x.$$

Now choosing B such that $x > B$ implies $\frac{2+\psi(A)}{\log x} < \frac{\epsilon}{2}$, we see that for $x > A$, $x > B$

$$\left|\frac{H(x)}{x\log x}\right| < \frac{2+\psi(A)}{\log x}+\frac{\epsilon}{2} < \epsilon.$$

We conclude that $\lim_{x\to\infty}\frac{H(x)}{x\log x} = 0$ and hence by lemma 2.25 $\lim_{x\to\infty}\frac{M(x)}{x} = 0$. □

2.5 The summation of $\frac{\mu(n)}{n}$

In this section we will see another theorem and a very important corollary.

Definition 2.26. *Let $f : [a,b] \to \mathbb{R}$ be a real-valued function. The* total variation *of f on a interval $[a',b'] \subseteq [a,b]$ is given by*

$$V_{a'}^{b'}(f) = \sup_{P\in\mathcal{P}}\sum_{i=0}^{n_P-1}\left|f(x_{i+1})-f(x_i)\right|.$$

Here $\mathcal{P}$ denotes the set of all partitions $P = \{x_0,\ldots,x_{n_P}\}$ of $[a',b']$.

Definition 2.27. *A real-valued function f is of* bounded variation *on an interval $[a,b]$ if $V_a^b(f)$ is finite.*

Example 2.28. *If $[a,b]$ is a finite interval, (that is $a,b \in \mathbb{R}$), and if a function $f : [a,b] \to \mathbb{R}$ is an absolutely continuous function except on a finite number of points, then f is of bounded variation.*

Theorem 2.29. *Suppose $B(x) : \mathbb{R}_{\geq 1} \to \mathbb{R}$ is a real function and $a(x)$ an arithmetic function, such that*

1. *$B(x) = O(1)$;*
2. *$B(x)$ is of bounded variation in every finite interval;*
3. *$\sum_{n\leq x} a(n) = o(x)$ for all x;*
4. *$\sum_{n\leq x} |a(n)| = O(x)$ for all x.*

Then

$$\sum_{n\leq x} a(n)B\left(\frac{x}{n}\right) = o(x).$$

Proof. Let $\epsilon > 0$. We show that there exists an X, s.t. $x > X$ implies that

$$\left|\sum_{n\leq x} a(n)B\left(\frac{x}{n}\right)\right| < \epsilon x.$$

For some $0 < \delta < 1$, we define

$$S_1 = \sum_{n\leq \delta x} a(n)B\left(\frac{x}{n}\right) \text{ and } S_2 = \sum_{\delta x < n\leq x} a(n)B\left(\frac{x}{n}\right).$$

Then by properties 1 and 4, we have that

$$|S_1| \leq \sum_{n\leq \delta x} |a(n)| \left|B\left(\frac{x}{n}\right)\right| = O\left(\sum_{n\leq \delta x} |a(n)|\right) = O(\delta x).$$

We can choose δ small enough, such that $|S_1| < \frac{\epsilon}{2}x$. Defining $A(x) = \sum_{n\leq x} a(n)$, we see that

$$\begin{aligned}
S_2 &= \sum_{\delta x < n\leq x} a(n)B\left(\frac{x}{n}\right) = \sum_{\delta x < n\leq x} (A(n) - A(n-1))\, B\left(\frac{x}{n}\right) \\
&= \sum_{\delta x < n\leq x} A(n)B\left(\frac{x}{n}\right) - \sum_{\delta x < n\leq x} A(n-1)B\left(\frac{x}{n}\right) \\
&= \sum_{\delta x < n\leq x} A(n)B\left(\frac{x}{n}\right) - \sum_{\delta x - 1 < n\leq x-1} A(n)B\left(\frac{x}{n+1}\right) \\
&= A([x])B\left(\frac{x}{[x]}\right) - A([\delta x])B\left(\frac{x}{[\delta x]}\right) - \sum_{n=[\delta x]+1}^{[x]-1} A(n)\left(B\left(\frac{x}{n}\right) - B\left(\frac{x}{n+1}\right)\right).
\end{aligned}$$

It follows that

$$|S_2| \leq |A([x])| \left|B\left(\frac{x}{[x]}\right)\right| + |A([\delta x])| \left|B\left(\frac{x}{[\delta x]}\right)\right| + \sum_{n=[\delta x]+1}^{[x]-1} |A(n)| \left|\left(B\left(\frac{x}{n}\right) - B\left(\frac{x}{n+1}\right)\right)\right|.$$

Now by property 3 $|A(x)| = o(x)$, and by property 2 $\left|B\left(\frac{x}{[x]}\right)\right|$, $\left|B\left(\frac{x}{[\delta x]}\right)\right|$ are bounded by some constant. Finally by property 2

$$\sum_{n=[\delta x]+1}^{[x]-1} \left|\left(B\left(\frac{x}{n}\right) - B\left(\frac{x}{n+1}\right)\right)\right|$$

is bounded by a function f depending on δ, not on x. Hence $|S_2| = o(x)f(\delta)$. Choosing x large enough, we find that $|S_2| < \frac{\epsilon}{2}x$. Finally we find that

$$\left|\sum_{n\leq x} a(n)B\left(\frac{x}{n}\right)\right| \leq |S_1| + |S_2| < \epsilon x.$$

□

Corollary 2.30. *The statement $M(x) = o(x)$ implies that*

$$\sum_{n\leq x} \frac{\mu(n)}{n} = o(1)$$

as $x \to \infty$.

Proof. We prove that $M(x) = o(x)$ implies that $\sum_{n\leq x} \mu(n)\frac{x}{n} = o(x)$ as $x \to \infty$. Recalling the fractional part of any real x as $\{x\} = x - [x]$, we see that

$$\begin{aligned}\sum_{n\leq x} \mu(n)\frac{x}{n} &= \sum_{n\leq x} \mu(n)\left(\left[\frac{x}{n}\right] + \left\{\frac{x}{n}\right\}\right) \\ &= \sum_{n\leq x} \mu(n)\left[\frac{x}{n}\right] + \sum_{n\leq x} \mu(n)\left\{\frac{x}{n}\right\}.\end{aligned}$$

By corollary 2.16 the first sum is equal to 1. For the second sum we apply theorem 2.29 with $a(n) = \mu(n)$ and $B(x) = \{x\}$. Clearly $\sum_{n\leq x} |\mu(n)| = O(x)$ and by assumption $\sum_{n\leq x} \mu(n) = o(x)$. Furthermore $B(x) = O(1)$, as $B(x) < 1$ for all x. Finally note that $B(x)$ is absolutely continuous, except at the integers. As in every finite interval, there are only finitely many integers, $B(x)$ is on every finite interval absolutely continuous, except at a finite number of points. Hence $B(x)$ is of bounded variation in every finite interval. We can then apply theorem 2.29 and see that $\sum_{n\leq x} \mu(n)\left\{\frac{x}{n}\right\} = o(x)$. We conclude that $\sum_{n\leq x} \mu(n)\frac{x}{n} = o(x)$ and hence $\sum_{n\leq x} \frac{\mu(n)}{n} = o(1)$. □

By the theorem 2.23 and corollary 2.30 it hence follows that the PNT implies

Theorem 2.31.

$$\sum_{n\leq x} \frac{\mu(n)}{n} = o(1).$$

We will use this theorem a lot in chapter 8.

2.6 The Euler totient function over $\mathbb{Z}$

In this section we look at the first properties of the Euler totient function in $\mathbb{Z}$, cf. [1]. In chapter 7 we will see that the Euler totient function in $\mathbb{F}_q[T]$ has similar properties.

Definition 2.32. *Define*

$$\begin{aligned}\varphi : \mathbb{Z} &\to \mathbb{Z} \\ n &\mapsto \#(\mathbb{Z}/n\mathbb{Z})^\times\end{aligned}$$

Hence $\varphi(n)$ is given by the number of integers that are both smaller than n and coprime to n.

Lemma 2.33. *For all $n \in \mathbb{N}$, the following statements hold:*

1. $n = \sum_{d|n} \varphi(d)$.
2. $\varphi = I_1 * \mu$, *where I_1 is the arithmetic function given by $I_1(n) = n$.*
3. $$\varphi(n) = n \prod_{\substack{p|n \\ p \text{ prime}}} \left(1 - \frac{1}{p}\right).$$

Proof.

1. Let d be a divisor of n. Denote
$$V_d = \{m \in \mathbb{Z} |\ 1 \leq m \leq n,\ \gcd(m, n) = d\}.$$
By dividing every element of V_d by d, you see that $\#V_d = \varphi\left(\frac{n}{d}\right)$. Now every $m \in \{1, \ldots, n\}$ occurs exactly in one V_d, so $\{1, \ldots, n\} = \cup_{d|n} V_d$. Hence
$$n = \#\{1, \ldots, n\} = \sum_{d|n} \#V_d = \sum_{d|n} \varphi\left(\frac{n}{d}\right).$$

2. By 1. we know that $I_1 = \varphi * E$. Applying Möbius inversion, we see that $\varphi = I_1 * \mu$.

3. In the proof of lemma 2.7 we saw that the convolution product preserves multiplicativity. Hence φ is a multiplicative function and it is sufficient to calculate $\varphi(p^k)$ for all primes p. Now
$$\varphi(p^k) = \sum_{l=0}^{k} \mu\left(p^l\right) I_1\left(p^{k-l}\right) = p^k - p^{k-1} = p^k \left(1 - \frac{1}{p}\right).$$
The statement then follows.

□

Corollary 2.34.
$$\frac{\varphi(n)}{n} = \sum_{d|n} \frac{\mu(d)}{d}$$

Proof. By lemma 2.33.2
$$\varphi(n) = \sum_{d|n} \mu(d) \frac{n}{d}.$$
Dividing by n yields the required result. □

We see that $\varphi(n) = n \prod_{\substack{p|n \\ p \text{ prime}}} \left(1 - \frac{1}{p}\right)$. We are actually interested in the factor $\prod_{\substack{p|n \\ p \text{ prime}}} \left(1 - \frac{1}{p}\right) = \frac{\varphi(n)}{n}$, as this gives the fluctuations in $\varphi(n)$. The following theorem tells us what this factor is on average, for $n \to \infty$.

Theorem 2.35. *For all $x \geq 1$, we have*

$$\frac{1}{x}\sum_{n\leq x}\frac{\varphi(n)}{n} = \frac{1}{\zeta(2)} + O\left(\frac{\log x}{x}\right).$$

Proof. We will prove this by showing that

$$\sum_{n\leq x}\frac{\varphi(n)}{n} = \sum_{n\leq x}\frac{\mu(n)}{n}\left[\frac{x}{n}\right] = \frac{x}{\zeta(2)} + O(\log x).$$

For the first equality we apply corollary 2.34 to see that

$$\sum_{n\leq x}\frac{\varphi(n)}{n} = \sum_{n\leq x}\sum_{d|n}\frac{\mu(d)}{d} = \sum_{d\leq x}\frac{\mu(d)}{d}\left[\frac{x}{d}\right].$$

Here we used that for a given d, there are $\left[\frac{x}{d}\right]$ integers $n \leq x$, such that d divides n. Now

$$\sum_{n\leq x}\frac{\mu(n)}{n}\left[\frac{x}{n}\right] = \sum_{n\leq x}\frac{\mu(n)}{n}\left(\frac{x}{n} - \left\{\frac{x}{n}\right\}\right) = x\sum_{n\leq x}\frac{\mu(n)}{n^2} - \sum_{n\leq x}\frac{\mu(n)}{n}\left\{\frac{x}{n}\right\} = x\sum_{n\leq x}\frac{\mu(n)}{n^2} + O\left(\sum_{n\leq x}\frac{1}{n}\right).$$

We know that $\frac{1}{\zeta(s)} = \sum_{n\geq 0}\frac{\mu(n)}{n^s}$ by corollary 2.14. Hence

$$\sum_{n\leq x}\frac{\mu(n)}{n^2} = \frac{1}{\zeta(2)} - \sum_{n>x}\frac{\mu(n)}{n^2}.$$

Finally we bound $\sum_{n\leq x}\frac{1}{n}$ and $\sum_{n>x}\frac{\mu(n)}{n^2}$ by

$$\sum_{n\leq x}\frac{1}{n} \leq 1 + \int_1^x \frac{1}{t}dt = 1 + \log t|_1^x = 1 + \log x = O(\log x)$$

and

$$\left|\sum_{n>x}\frac{\mu(n)}{n^2}\right| \leq \sum_{n>x}\frac{1}{n^2} \leq \frac{1}{\lceil x\rceil^2} + \int_x^\infty \frac{1}{t^2}dt = \frac{1}{\lceil x\rceil^2} - \frac{1}{t}\Big|_x^\infty = \frac{1}{x} + \frac{1}{\lceil x\rceil^2} = O\left(\frac{1}{x}\right).$$

We conclude that

$$\sum_{n\leq x}\frac{\mu(n)}{n}\left[\frac{x}{n}\right] = \frac{x}{\zeta(2)} + O(1) + O(\log x) = \frac{x}{\zeta(2)} + O(\log x).$$

□

Proving this theorem, we took a very rough estimation for the sum of fractional parts $\sum_{n\leq x}\frac{\mu(n)}{n}\left\{\frac{x}{n}\right\}$. When we look at the variance of $\frac{\varphi(n)}{n}$, we will need to take a far more precise approach in estimating this term. This actually proves to be the main difficulty in calculating the variance. As we're interested in the Euler function in (short) intervals, we prove one final theorem this chapter. As a reminder, we let H be the size of the interval, which should be $\Theta(x^\delta)$ for some $0 < \delta \leq 1$. If $0 < \delta < 1$, then we say it is a short interval.

Theorem 2.36. *Let $x \geq 1$ and let $H = \Theta(x^\delta)$ for some $0 < \delta \leq 1$. Then*

$$\frac{1}{H} \sum_{x < n \leq x+H} \frac{\varphi(n)}{n} = \frac{1}{\zeta(2)} + O\left(\frac{\log(x+H)}{H}\right).$$

Proof. The proof of this theorem is almost the same as that of theorem 2.35, including the same rough estimate. We know that

$$\begin{aligned}
&\sum_{x < n \leq x+H} \frac{\varphi(n)}{n} = \sum_{n \leq x+H} \frac{\varphi(n)}{n} - \sum_{n \leq x} \frac{\varphi(n)}{n} \\
&= \sum_{n \leq x+H} \sum_{d \mid n} \frac{\mu(d)}{d} - \sum_{n \leq x} \sum_{d \mid n} \frac{\mu(d)}{d} \\
&= \sum_{n \leq x+H} \frac{\mu(n)}{n} \left[\frac{x+H}{n}\right] - \sum_{n \leq x} \frac{\mu(n)}{n} \left[\frac{x}{n}\right] \\
&= \sum_{n \leq x+H} \frac{\mu(n)}{n} \left(\frac{x+H}{n} - \left\{\frac{x+H}{n}\right\}\right) - \sum_{n \leq x} \frac{\mu(n)}{n} \left(\frac{x}{n} - \left\{\frac{x}{n}\right\}\right) \\
&= H \sum_{n \leq x+H} \frac{\mu(n)}{n^2} + x \sum_{x < n \leq x+H} \frac{\mu(n)}{n^2} - \left(\sum_{n \leq x+H} \frac{\mu(n)}{n} \left\{\frac{x+H}{n}\right\} - \sum_{d \leq n} \frac{\mu(n)}{n} \left\{\frac{x}{n}\right\}\right) \\
&= \frac{H}{\zeta(2)} - H \sum_{n > x+H} \frac{\mu(n)}{n^2} + x \sum_{x < n \leq x+H} \frac{\mu(n)}{n^2} - \left(\sum_{n \leq x+H} \frac{\mu(n)}{n} \left\{\frac{x+H}{n}\right\} - \sum_{n \leq x} \frac{\mu(n)}{n} \left\{\frac{x}{n}\right\}\right).
\end{aligned}$$

Now we again apply the estimates

$$\left|\sum_{n > x} \frac{\mu(n)}{n^2}\right| = O\left(\frac{1}{x}\right)$$

and

$$\left(\sum_{n \leq x+H} \frac{\mu(n)}{n} \left\{\frac{x+H}{n}\right\} - \sum_{n \leq x} \frac{\mu(n)}{n} \left\{\frac{x}{n}\right\}\right) = O\left(\sum_{n \leq x+H} \frac{1}{n}\right) = O(\log(x+H))$$

to prove the statement. □

Chapter 3

Introduction to random matrix theory

3.1 Introduction

In this chapter we will see some definitions and basic proofs in random matrix theory. The main result of this chapter will be the Weyl integration formula, which we need to prove theorem 3.7. We need the matrix integral in this theorem in chapter 6.

In this chapter we introduce the basic notions as Miller, Takloo-Bighash did in [21]. For the proof of theorem 3.5 we follow Gamburd in [10] and Fulton and Harris in [9], as well as a series of lectures by Keating in [17]. We follow the same series of lectures in the proof of theorem 3.7.

Random matrix theory is a relatively new theory, with applications in various fields of science. Physicists first used it in nuclear physics and statistical mechanics to estimate the behaviour of particles in a closed system, but it turned out that the same theory could be used by mathematicians to estimate the behaviour of the prime numbers! Of course this last application is what interests us most in this thesis.

Let us first make the definition of a matrix ensemble.

Definition 3.1. *A* matrix ensemble *is a collection of matrices, together with a probability measure over these matrices.*

We list a few ensembles to get a feeling of what a matrix ensemble might look like. For any ensemble we let g denote an element in this ensemble.

Example 3.2.

1. Real Wigner ensemble. *For a fixed N, the set of real symmetric $N \times N$-matrices with matrix coefficients x_{ij} together with some probability measure $\mathbb{P}_{ij}(x_{ij})dx_{ij}$ for each $1 \leq i, j \leq N$. Note that x_{ij} determines x_{ji}. Hence if $g = [x_{ij}]_{i,j=1,\dots,N}$, then*

$$\mathbb{P}(g)dg = \prod_{1 \leq i \leq j \leq N} \mathbb{P}_{ij}(x_{ij})dx_{ij}.$$

2. Complex Wigner ensemble. *For a fixed N, the set of complex hermitian $N \times N$-matrices with matrix coefficients $x_{ij} + iy_{ij}$, together with some probability measures $\mathbb{P}_{x_{ij}}(x_{ij})dx_{ij}, \mathbb{P}_{y_{ij}}(y_{ij})dy_{ij}$.*

3. Unitary invariant ensemble. *For a fixed N, the set of complex hermitian $N \times N$-matrices with some probability measure satisfying $\mathbb{P}(U^\dagger gU) = \mathbb{P}(g)$ for any $N \times N$ unitary matrix U.*

4. Gaussian unitary ensemble (GUE). *The unique unitary invariant ensemble of complex Wigner random matrices.*

5. The ensemble of unitary matrices, *consisting of the set of unitary matrices $U(N)$ for some fixed N and a unitary invariant probability measure.*

Given such an ensemble and a function depending on matrices in this ensemble, our goal is to calculate the expectation value of this function in this ensemble. An example of a function depending on matrices in the ensemble could be the distribution of the eigenvalues of a matrix or the distribution of the differences of the eigenvalues of a matrix. You will see that eigenvalues are very important in this theory.

3.2 Class functions

In this specific case, we're interested in the ensemble of unitary matrices $U(N)$. In the whole chapter N will be arbitrary, but fixed. This ensemble has a measure, which we will write as $d\mu(g) = \mathbb{P}(g)dg$ for each $g \in G$, which is unitary invariant. This means that $\mathbb{P}(g) = \mathbb{P}(U^\dagger gU)$ for all unitary matrices U, g. Note that as U is unitary, we have that $U^\dagger = U^{-1}$, so this condition is equivalent to $\mathbb{P}(g) = \mathbb{P}(U^{-1}gU)$ for all unitary matrices U, g. It is hence a Haar-measure, as the unitary matrices we translate by are in fact part of our group $U(N)$. For any unitary matrix g, we know that its eigenvalues have norm 1. We can hence write them as $e^{i\theta_1}, \ldots, e^{i\theta_N}$. This enables us to define class functions.

Definition 3.3. *Let $\tilde{f} : U(N) \to \mathbb{C}$ be a function. $\tilde{f}$ is called a* class function *if there exists some function $f : [0, 2\pi[^N \to \mathbb{C}$ symmetric in its arguments such that for each matrix $g \in U(N)$ with eigenvalues $e^{i\theta_1}, \ldots, e^{i\theta_N}$ we have $\tilde{f}(g) = f(\theta_1, \ldots, \theta_N)$.*

For the ease of notation we will write $\tilde{f} = f$, (so $f(g) = f(\theta_1, \ldots, \theta_N)$).

Example 3.4. *The function $h \mapsto \mathrm{Tr}(h^k)$ is a class function for each $k \in \mathbb{Z}$, as for each unitary matrix h with eigenvalues $e^{i\theta_1}, \ldots, e^{i\theta_N}$ we have an unitary matrix g, s.t. $h = gag^{-1}$, where a is the diagonal matrix consisting of its eigenvalues. Now*

$$\mathrm{Tr}(h^k) = \mathrm{Tr}((gag^{-1})^k) = \mathrm{Tr}(ga^kg^{-1}) = \mathrm{Tr}(a^kg^{-1}g) = \mathrm{Tr}(a^k) = \sum_{j=1}^{N} e^{ik\theta_j},$$

using the fact that the trace function is invariant under cyclic permutations. Hence $\mathrm{Tr}(h^k)$ only depends on the eigenvalues of h. Furthermore it is symmetric in arguments: interchanging the order of the θ_j does not change $\mathrm{Tr}(h^k)$.

For any integrable class function f on $U(N)$, we can calculate its expectation value by

$$\mathop{\mathbb{E}}_{g \in U(N)}[f(g)] = \int_{U(N)} f(g)d\mu(g) = \int_{U(N)} f(g)\mathbb{P}(g)dg.$$

3.3 The Weyl integration formula

In this section we prove a very important formula to calculate the expectation value of a class function. It is given by

Theorem 3.5 (Weyl integration formula)**.**
For a class function f on $U(N)$, we have

$$\mathop{\mathbb{E}}_{g\in U(N)}[f(g)] = \frac{1}{(2\pi)^N N!}\int_0^{2\pi}\cdots\int_0^{2\pi} f(\theta_1,\ldots,\theta_N)\prod_{1\le j<k\le N}\left|e^{i\theta_j}-e^{i\theta_k}\right|^2 d\theta_1\ldots d\theta_N.$$

This is a special case of a general theorem in Lie groups, applied to the compact, connected group $U(N)$. This general theorem, the Weyl character formula, was proven by Weyl [27, 28, 29] in 1926. We will prove it specifically for $U(N)$, using as little Lie algebra theory as possible, but it turns out we will need some of it.

Proof. Define A to be the subgroup of $U(N)$ consisting of all unitary diagonal matrices. Since a diagonal matrix has its eigenvalues on its diagonal, we know that A consists exactly of all matrices of the form $\begin{pmatrix} e^{i\theta_1} & & \\ & \ddots & \\ & & e^{i\theta_N}\end{pmatrix} = \mathrm{diag}(e^{i\theta_1},\ldots,e^{i\theta_N})$ with $\theta_1,\ldots,\theta_N\in[0,2\pi[$. Hence A is isomorphic to $(S^1)^N$, where S^1 is the complex unit circle. As we know, every unitary matrix is diagonalizable. Hence for every unitary matrix $h\in U(N)$, there exist some $a\in A$, $g\in U(N)$, such that $h=gag^{-1}$. Now define the map

$$\begin{aligned}\rho : A\times U(N) &\to U(N)\\ (a,g) &\mapsto gag^{-1}.\end{aligned}$$

Note that this function is surjective. Moreover, as A is commutative, we know that $\rho(a,g)=\rho(a,ga')$ for every $a,a'\in A$ and $g\in U(N)$, so we can factor out this group A to get a map

$$\begin{aligned}\bar\rho : A\times(U(N)/A) &\to U(N)\\ (a,\bar g) &\mapsto gag^{-1}.\end{aligned}$$

This function is the key to the proof. It allows us to make a transformation of variables, whose jacobian will give us the factor $\prod\left|e^{i\theta_j}-e^{i\theta_k}\right|^2$. First we use this function $\bar\rho$ to lift the Haar measure $d\mu$ on $U(N)$ to a (Haar-)measure $d\bar\mu$ on $A\times(U(N)/A)$. This measure will be zero on the set where $\bar\rho$ is singular. $d\bar\mu$ gives us a first glance at the transformation at hand. For a function φ on $A\times(U(N)/A)$, we have:

$$\int_{A\times(U(N)/A)}\varphi(a,\bar g)d\bar\mu(a,\bar g) = \int_{U(N)}\left(\sum_{(a,\bar g)\in\bar\rho^{-1}(g)}\varphi(a,\bar g)\right)d\mu(g).$$

The question that now arises is how many pairs $(a,\bar g)$ there are s.t. $(a,\bar g)\in\bar\rho^{-1}(g)$. That is, given an $h\in U(N)$, how many pairs $(a,\bar g)$ are there such that $\bar\rho(a,\bar g)=h$? If h has N distinct eigenvalues, the answer turns out to be $N!$. You can see this as follows: if we have a unitary matrix h with N distinct eigenvalues we can permute these values in $N!$ different ways. Once you've chosen such a permutation, (so you've fixed $\theta_1,\ldots,\theta_N$), there is only one $\bar g$, s.t. $h=gag^{-1}$ with $a=\mathrm{diag}(e^{i\theta_1},\ldots,e^{i\theta_N})$. If there exist $g,g'\in U(N)$ s.t. $gag^{-1}=h=g'a(g')^{-1}$,

then $(g')^{-1}ga = a(g')^{-1}g$. So a and $g(g')^{-1}$ commute, implying that $g(g')^{-1} \in A$. Here the need for distinct eigenvalues arises. If h, (and hence a), does not have N distinct eigenvalues, there exist elements $g \in U(N) \setminus A$ s.t. g and a commute. We conclude that for each permutation of eigenvalues, there is exactly one equivalence class $\bar{g}$ s.t. $\rho(a, \bar{g}) = h$. Since there are $N!$ such permutations, we conclude that there are $N!$ pairs $(a, \bar{g}) \in \bar{\rho}^{-1}(h)$. If h doesn't have N distinct eigenvalues, then there infinitely many pairs $(a, \bar{g})$ s.t. $\bar{\rho}(a, \bar{g}) = h$, but luckily for us in that case $\bar{\rho}$ is singular so the measure will be zero. Now for f a class function on $U(N)$, (so $f(a) = f(gag^{-1})$ as this do not change the eigenvalues), we can take

$$\varphi(a, \bar{g}) = f(\bar{\rho}(a, \bar{g})) = f(gag^{-1}) = f(a) = f(\theta_1, \ldots, \theta_N).$$

From the above we then conclude that

$$\begin{aligned}\mathop{\mathbb{E}}_{g \in U(N)}[f(g)] &= \int_{U(N)} f(g) d\mu(g) = \int_{U(N)} \left(\sum_{(a,\bar{g}) \in \bar{\rho}^{-1}(g)} \varphi(a, \bar{g}) \right) d\mu(g) \\ &= \frac{1}{N!} \int_{A \times (U(N)/A)} \varphi(a, \bar{g}) d\bar{\mu}(a, \bar{g}) = \frac{1}{N!} \int_{A \times (U(N)/A)} f(a) d\bar{\mu}(a, \bar{g}).\end{aligned}$$

The next step is to consider the measure $d\bar{\mu}(a, \bar{g})$. As this measure is a lift from our Haar measure on $U(N)$, using $\bar{\rho}(a, \bar{g}) = gag^{-1}$, we know that a change in $\bar{g}$ does not result in a change in $d\bar{\mu}(a, \bar{g})$, (as $d\mu(a) = d\mu(gag^{-1})$ in our original measure). Hence $d\bar{\mu}$ is a function only depending on a. We can write $d\bar{\mu}(a, \bar{g}) = d\nu(a) d\bar{g}$. Furthermore note that all of our groups are smooth manifolds and all our functions are smooth maps, so $d\nu$ is absolutely continuous. This means that

$$d\nu(a) = \nu(a) da,$$

with $\nu(a)$ still to be determined. We find that

$$\mathop{\mathbb{E}}_{g \in U(N)}[f(g)] = \frac{1}{N!} \int_{A \times (U(N)/A)} f(a) \nu(a) da dg = \frac{1}{N!} \int_A f(a) \nu(a) da \int_{(U(N)/A)} d\bar{g},$$

Now $\nu(a)$ is actually given by the determinant of the Jacobian of our transformation, so the logical next step is to calculate the Jacobian of the map $\bar{\rho}(a, \bar{g}) = gag^{-1}$. This step actually requires some Lie algebra, but we won't get into it in too much detail. To calculate the Jacobian at point $(a_0, \bar{g}_0) \in A \times (U(N)/A)$ we need to consider the tangent space at this point. As we're working in a product space, it is enough to consider the tangent space of $a_0 \in A$ and of $g_0 A = \bar{g}_0 \in (U(N)/A)$. These are Lie algebras. It turns out that $a_0 e^{ta}$ and $g_0 e^{tg} A$ are parametrizations of these tangent spaces. Hence $(a_0 e^{ta}, g_0 e^{tg} A)$ is a parametrization of the product space. Note that $\bar{\rho}$ maps the point $(a_0 e^{ta}, g_0 e^{tg} A)$ to $g_0 e^{tg} a_0 e^{ta} e^{-tg} g_0^{-1}$. Now both da and $d\bar{g}$ are Haar measures on A and $U(N)/A$ respectively. This implies that a translation on the left by $a_0^{-1} g_0^{-1}$ and a translation on the right by g_0 do not change the measure. It is hence sufficient to calculate the Jacobian of the map

$$\begin{aligned}T_{a_0} A \times T_{\bar{g}_0}(U(N)/A) &\to U(N) \\ (a_0 e^{ta}, g_0 e^{tg} A) &\mapsto a_0^{-1} e^{tg} a_0 e^{ta} e^{-tg}\end{aligned}$$

at the point $t = 0$ as this map has the same Jacobian of our map $\bar{\rho}$ at the point $t = 0$. Now

$$\frac{d}{dt} a_0^{-1} e^{tg} a_0 e^{ta} e^{-tg} \Big|_{t=0} = a_0^{-1} g a_0 + a_0^{-1} a_0 a - a_0^{-1} a_0 g = a - g + a_0^{-1} g a_0.$$

If we denote for $h \in U(N)$ the function $\mathrm{Ad}_h : U(N) \to U(N), g \mapsto h^{-1}gh$, (as is standard notation in Lie groups), we conclude that the Jacobian of the transformation is given by

$$J = \begin{pmatrix} \mathbb{I}_A & 0 \\ 0 & \mathrm{Ad}_{a_0}|_{U(N)/A} - \mathbb{I}_{U(N)/A} \end{pmatrix}$$

Now to calculate the determinant $|J|$, some authors use what is standard knowledge about the Ad-function. We take a different approach: we know that the space $U(N)$ has dimension N^2 as a vector space over $\mathbb{C}$. Hence the dimension of $U(N)/A$ is $N^2 - N = N(N-1)$. This is also seen by looking at the Lie algebra $\mathfrak{u}(N)$ associated to $U(N)$. As we have seen the function e^{tX} parametrizes $U(N)$ if $X \in \mathfrak{u}(N)$. Now unitary matrices are matrices M s.t. $M^\dagger M = I$. Equivalently these are the matrices M s.t. $\langle Mu, Mv\rangle = \langle u, v\rangle$ for all $u, v \in \mathbb{C}^N$. Here $\langle u, v\rangle$ denotes the standard inner product on $\mathbb{C}^N$, given by $u^\dagger v$. Next we want to find a defining condition for X to be in the Lie algebra $\mathfrak{u}(N)$, (the tangent space of $U(N)$). Then $e^{tX} \in U(N)$, so

$$\langle e^{tX}u, e^{tX}v\rangle = \langle u, v\rangle.$$

Differentiating this with respect to t gives us

$$0 = \langle Xe^{tX}u, e^{tX}v\rangle + \langle e^{tX}u, Xe^{tX}v\rangle = u^\dagger(e^{tX})^\dagger(X + X^\dagger)e^{tX}v$$

for all $u, v \in \mathbb{C}^n$. We conclude $X + X^\dagger = 0$ and it turns out that this is sufficient as a defining property of $\mathfrak{u}(N)$. Hence $X \in \mathfrak{u}(N)$ if and only if X is skew-symmetric. Note that the dimension of skew-symmetric $N \times N$ matrices is N^2 as you have one degree of freedom for all N diagonal entries and two degrees of freedom for all $\frac{N(N-1)}{2}$ upper triangle (complex) entries. This coincides with the dimension of $U(N)$, as a tangent space has the same dimension as the group it is tangent with. Now the Lie algebra $\mathfrak{a}$, or the tangent space of A, is given by all $N \times N$ diagonal matrices with real entries, so this has dimension N. The orthogonal complement of this Lie algebra, $\mathfrak{a}^\perp$, or the tangent space of $U(N)/A$, is given by the skew-symmetric $N \times N$ matrices with zeroes on the diagonal. Of course this has dimension $N(N-1)$. Hence we expect the determinant $|J|$ to be a polynomial of at most degree $N(N-1)$. Next we ask ourself for which $a_0 \in A$ this determinant is zero. This is exactly the case when $a_0^{-1}ga_0 = g$ for some $a_0 \in A$ and $g \in \mathfrak{a}^\perp$. Writing

$$a_0 = \mathrm{diag}(e^{i\theta_1}, \ldots, e^{i\theta_N}),\ g = \begin{pmatrix} 0 & g_{12} & g_{13} & \ldots \\ -g_{12} & 0 & g_{23} & \ldots \\ -g_{13} & -g_{23} & 0 & \ldots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

we see that

$$a_0^{-1}ga_0 = \begin{pmatrix} 0 & e^{i\theta_2 - i\theta_1}g_{12} & e^{i\theta_3 - i\theta_1}g_{13} & \ldots \\ -e^{i\theta_1 - i\theta_2}g_{12} & 0 & e^{i\theta_3 - i\theta_2}g_{23} & \ldots \\ -e^{i\theta_1 - i\theta_3}g_{13} & -e^{i\theta_2 - i\theta_3}g_{23} & 0 & \ldots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

This is equal to g if $\theta_j = \theta_k$ for some $j < k$ and if g_{jk} is the only non-zero element of g. Note that the dimension of the subset $\{g \in \mathfrak{a}^\perp | g_{jk} \text{ is the only non-zero element of } g\}$ is two. Therefore there is a zero of order two at each a with $\theta_j = \theta_k$ for some $j < k$. It hence follows that $\prod_{j<k} \left|e^{i\theta_j} - e^{i\theta_k}\right|^2$ divides $|J|$. As there are $\frac{N(N-1)}{2}$ pairs $j < k$, the degree of this polynomial is $N(N-1)$, so there cannot be any more factors. We conclude that $|J| = c\prod_{j<k} \left|e^{i\theta_j} - e^{i\theta_k}\right|^2$

for some constant c. Hence

$$\mathop{\mathbb{E}}_{g\in U(N)}[f(g)] = \frac{c}{N!}\int_A f(a)|J|da = \frac{c}{N!}\int_0^{2\pi}\dots\int_0^{2\pi} f(\theta_1,\dots,\theta_N)\prod_{1\le j<k\le N}\left|e^{i\theta_j}-e^{i\theta_k}\right|^2 d\theta_1\dots d\theta_N.$$

Here we pulled the factor $\int_{(U(N)/A)} d\bar{g}$ into the constant c. The next and final step is to calculate this factor c. For this we will fix f to be the constant function 1, giving

$$1 = \frac{c}{N!}\int_0^{2\pi}\dots\int_0^{2\pi}\prod_{1\le j<k\le N}\left|e^{i\theta_j}-e^{i\theta_k}\right|^2 d\theta_1\dots d\theta_N.$$

To calculate c, we calculate this integral. We first prove the following lemma.

Lemma 3.6. *For f a class function, we have*

$$\begin{aligned}&\int_0^{2\pi}\dots\int_0^{2\pi} f(\theta_1,\dots,\theta_N)\prod_{1\le j<k\le N}|e^{i\theta_j}-e^{i\theta_k}|^2 d\theta_1\dots d\theta_N\\ &= N!\int_0^{2\pi}\dots\int_0^{2\pi} f(\theta_1,\dots,\theta_N)\begin{vmatrix} 1 & e^{-i\theta_1} & e^{-2i\theta_1} & \dots & e^{-(N-1)i\theta_1}\\ e^{i\theta_2} & 1 & e^{-i\theta_2} & \dots & e^{-(N-2)i\theta_2}\\ e^{2i\theta_3} & e^{i\theta_3} & 1 & \dots & e^{-(N-3)i\theta_3}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ e^{(N-1)i\theta_N} & e^{(N-2)i\theta_N} & e^{(N-3)i\theta_N} & \dots & 1\end{vmatrix} d\theta_1\dots d\theta_N.\end{aligned}$$

Proof. First recall the Van der Monde-determinant

$$\prod_{1\le j<k\le N}(x_j-x_k) = \begin{vmatrix} 1 & 1 & 1 & \cdots & 1\\ x_1 & x_2 & x_3 & \cdots & x_N\\ x_1^2 & x_2^2 & x_3^2 & \cdots & x_N^2\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ x_1^{N-1} & x_2^{N-1} & x_3^{N-1} & \cdots & x_N^{N-1}\end{vmatrix}.$$

This implies that

$$\begin{aligned}\prod_{1\le j<k\le N}\left|e^{i\theta_j}-e^{i\theta_k}\right|^2 &= \prod_{1\le j<k\le N}\left(e^{i\theta_j}-e^{i\theta_k}\right)\left(e^{-i\theta_j}-e^{-i\theta_k}\right)\\ &= \begin{vmatrix} 1 & 1 & 1 & \cdots & 1\\ e^{i\theta_1} & e^{i\theta_2} & e^{i\theta_3} & \cdots & e^{i\theta_N}\\ e^{2i\theta_1} & e^{2i\theta_2} & e^{2i\theta_3} & \cdots & e^{2i\theta_N}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ e^{(N-1)i\theta_1} & e^{(N-1)i\theta_2} & e^{(N-1)i\theta_3} & \dots & e^{(N-1)i\theta_N}\end{vmatrix}\cdot\begin{vmatrix} 1 & e^{-i\theta_1} & e^{-2i\theta_1} & \dots & e^{-(N-1)i\theta_1}\\ 1 & e^{-i\theta_2} & e^{-2i\theta_2} & \dots & e^{-(N-1)i\theta_2}\\ 1 & e^{-i\theta_3} & e^{-2i\theta_3} & \dots & e^{-(N-1)i\theta_3}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 1 & e^{-i\theta_N} & e^{-2i\theta_N} & \dots & e^{-(N-1)i\theta_N}\end{vmatrix}\\ &= \begin{vmatrix} \sum_{l=1}^N 1 & \sum_{l=1}^N e^{-i\theta_l} & \sum_{l=1}^N e^{-2i\theta_l} & \dots & \sum_{l=1}^N e^{-(N-1)i\theta_l}\\ \sum_{l=1}^N e^{i\theta_l} & \sum_{l=1}^N 1 & \sum_{l=1}^N e^{-i\theta_l} & \dots & \sum_{l=1}^N e^{-(N-2)i\theta_l}\\ \sum_{l=1}^N e^{2i\theta_l} & \sum_{l=1}^N e^{i\theta_l} & \sum_{l=1}^N 1 & \dots & \sum_{l=1}^N e^{-(N-3)i\theta_l}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ \sum_{l=1}^N e^{(N-1)i\theta_l} & \sum_{l=1}^N e^{(N-2)i\theta_l} & \sum_{l=1}^N e^{(N-3)i\theta_l} & \dots & \sum_{l=1}^N 1\end{vmatrix}\end{aligned}$$

For notational convenience, from now on, we will write $\iint d\theta$ for $\int_0^{2\pi} \dots \int_0^{2\pi} d\theta_1 \dots d\theta_N$. Hence we want to calculate

$$\iint f(\theta_1,\dots,\theta_N) \begin{vmatrix} \sum_{l=1}^N 1 & \sum_{l=1}^N e^{-i\theta_l} & \sum_{l=1}^N e^{-2i\theta_l} & \cdots & \sum_{l=1}^N e^{-(N-1)i\theta_l} \\ \sum_{l=1}^N e^{i\theta_l} & \sum_{l=1}^N 1 & \sum_{l=1}^N e^{-i\theta_l} & \cdots & \sum_{l=1}^N e^{-(N-2)i\theta_l} \\ \sum_{l=1}^N e^{2i\theta_l} & \sum_{l=1}^N e^{i\theta_l} & \sum_{l=1}^N 1 & \cdots & \sum_{l=1}^N e^{-(N-3)i\theta_l} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum_{l=1}^N e^{(N-1)i\theta_l} & \sum_{l=1}^N e^{(N-2)i\theta_l} & \sum_{l=1}^N e^{(N-3)i\theta_l} & \cdots & \sum_{l=1}^N 1 \end{vmatrix} d\theta.$$

Since f is a class function, it is symmetric in arguments. That means that the contributions of all θ_l to the sums in the first row, when integrated, are all equal. Hence we can just replace these sums by N times the first term. The expression above then equals

$$\iint f(\theta_1,\dots,\theta_N) \begin{vmatrix} N\cdot 1 & Ne^{-i\theta_1} & Ne^{-2i\theta_1} & \cdots & Ne^{-(N-1)i\theta_1} \\ \sum_{l=1}^N e^{i\theta_l} & \sum_{l=1}^N 1 & \sum_{l=1}^N e^{-i\theta_l} & \cdots & \sum_{l=1}^N e^{-(N-2)i\theta_l} \\ \sum_{l=1}^N e^{2i\theta_l} & \sum_{l=1}^N e^{i\theta_l} & \sum_{l=1}^N 1 & \cdots & \sum_{l=1}^N e^{-(N-3)i\theta_l} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum_{l=1}^N e^{(N-1)i\theta_l} & \sum_{l=1}^N e^{(N-2)i\theta_l} & \sum_{l=1}^N e^{(N-3)i\theta_l} & \cdots & \sum_{l=1}^N 1 \end{vmatrix} d\theta.$$

Now we can subtract the first row $\frac{e^{i\theta_1}}{N}$ times from the second row, $\frac{e^{2i\theta_1}}{N}$ times from the third row, etcetera, to see that this equals

$$\iint f(\theta_1,\dots,\theta_N) \begin{vmatrix} N\cdot 1 & Ne^{-i\theta_1} & Ne^{-2i\theta_1} & \cdots & Ne^{-(N-1)i\theta_1} \\ \sum_{l=2}^N e^{i\theta_l} & \sum_{l=2}^N 1 & \sum_{l=2}^N e^{-i\theta_l} & \cdots & \sum_{l=2}^N e^{-(N-2)i\theta_l} \\ \sum_{l=2}^N e^{2i\theta_l} & \sum_{l=2}^N e^{i\theta_l} & \sum_{l=2}^N 1 & \cdots & \sum_{l=2}^N e^{-(N-3)i\theta_l} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum_{l=2}^N e^{(N-1)i\theta_l} & \sum_{l=2}^N e^{(N-2)i\theta_l} & \sum_{l=2}^N e^{(N-3)i\theta_l} & \cdots & \sum_{l=2}^N 1 \end{vmatrix} d\theta.$$

The function integrated over is symmetric in the arguments $\theta_2,\dots,\theta_N$, so we can replace the sums in the second row by $(N-1)$ times the first term in the sum, which is given by $l=2$. Then we get

$$\iint f(\theta_1,\dots,\theta_N) \begin{vmatrix} N\cdot 1 & Ne^{-i\theta_1} & Ne^{-2i\theta_1} & \cdots & Ne^{-(N-1)i\theta_1} \\ (N-1)e^{i\theta_2} & (N-1)\cdot 1 & (N-1)e^{-i\theta_2} & \cdots & (N-2)e^{-(N-1)i\theta_2} \\ \sum_{l=2}^N e^{2i\theta_l} & \sum_{l=2}^N e^{i\theta_l} & \sum_{l=2}^N 1 & \cdots & \sum_{l=2}^N e^{-(N-3)i\theta_l} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum_{l=2}^N e^{(N-1)i\theta_l} & \sum_{l=2}^N e^{(N-2)i\theta_l} & \sum_{l=2}^N e^{(N-3)i\theta_l} & \cdots & \sum_{l=2}^N 1 \end{vmatrix} d\theta.$$

We can then subtract the second row $\frac{e^{i\theta_1}}{N}$ times from the third row, $\frac{e^{2i\theta_1}}{N}$ times from the fourth row, etcetera. Repeat this process to see that this term equals

$$\iint f(\theta_1,\dots,\theta_N) \begin{vmatrix} N\cdot 1 & Ne^{-i\theta_1} & Ne^{-2i\theta_1} & \cdots & Ne^{-(N-1)i\theta_1} \\ (N-1)e^{i\theta_2} & (N-1)\cdot 1 & (N-1)e^{-i\theta_2} & \cdots & (N-1)\cdot e^{-(N-2)i\theta_2} \\ (N-2)e^{2i\theta_3} & (N-2)e^{i\theta_3} & (N-2)\cdot 1 & \cdots & (N-2)e^{-(N-3)i\theta_3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ e^{(N-1)i\theta_N} & e^{(N-2)i\theta_N} & e^{(N-3)i\theta_N} & \cdots & 1 \end{vmatrix} d\theta.$$

Now take the factor N out of the first row, the factor $(N-1)$ out of the second row, etc. This gives us a factor $N!$ and we find exactly the statement we wanted to prove. □

Remember that we wanted to calculate $\int\int \prod_{1\leq j<k\leq N} |e^{i\theta_j} - e^{i\theta_k}| d\theta$. By the lemma above, this is equal to

$$N! \int\int \begin{vmatrix} 1 & e^{-i\theta_1} & e^{-2i\theta_1} & \cdots & e^{-(N-1)i\theta_1} \\ e^{i\theta_2} & 1 & e^{-i\theta_2} & \cdots & e^{-(N-2)i\theta_2} \\ e^{2i\theta_3} & e^{i\theta_3} & 1 & \cdots & e^{-(N-3)i\theta_3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ e^{(N-1)i\theta_N} & e^{(N-2)i\theta_N} & e^{(N-3)i\theta_N} & \cdots & 1 \end{vmatrix} d\theta.$$

We can consider this determinant as the sum of all possible products of N coefficients of the matrix, where each row and each column occurs exactly once, (times a sign). Now note that $\int_0^{2\pi} e^{ix\theta} d\theta = 0$, unless x is a multiple of 2π, in which case this integral is equal to 2π. Since we are only working with integer powers of $e^{i\theta_l}$ in the expression above, we conclude that a term in the summation only has non-zero contribution if the integer power of each $e^{i\theta_l}$ is zero. Hence the above integrals will simplify a lot, as the only term in the sum that has non-zero contribution is the term where we choose only 1's. We conclude that

$$\int\int \begin{vmatrix} 1 & e^{-i\theta_1} & e^{-2i\theta_1} & \cdots & e^{-(N-1)i\theta_1} \\ e^{i\theta_2} & 1 & e^{-i\theta_2} & \cdots & e^{-(N-2)i\theta_2} \\ e^{2i\theta_3} & e^{i\theta_3} & 1 & \cdots & e^{-(N-3)i\theta_3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ e^{(N-1)i\theta_N} & e^{(N-2)i\theta_N} & e^{(N-3)i\theta_N} & \cdots & 1 \end{vmatrix} d\theta. = \int\int 1 \; d\theta = (2\pi)^N.$$

Hence $\int\int \prod_{1\leq j<k\leq N} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta = (2\pi)^N N!$ and we find that

$$1 = \frac{c}{N!} \cdot (2\pi)^N N!$$

Therefore

$$c = \frac{1}{(2\pi)^N}.$$

□

3.4 Dyson's Theorem

In this section we will apply the Weyl integration formula to actually calculate the expectation value of a class function. It is very useful to have seen such a calculation. Moreover the theorem we prove is actually necessary to the proof of theorem 1.1. The theorem was first proven by Dyson [7] in 1970.

Theorem 3.7 (Dyson's Theorem)**.**

$$\mathop{\mathbb{E}}_{g\in U(N)} \left[\left| \mathrm{Tr}(g^k) \right|^2 \right] = \begin{cases} N^2 & \text{if } k = 0 \\ |k| & \text{if } |k| \leq N \text{ and } k \neq 0 \\ N & \text{if } |k| \geq N \text{ and } k \neq 0 \end{cases}$$

Proof. As noted in the beginning of this section $\mathrm{Tr}(g^k) = \sum_{j=1}^{N} \left(e^{i\theta_j}\right)^k$, where $e^{i\theta_j}$ are the eigenvalues of g. Hence

$$\left|\mathrm{Tr}(g^k)\right|^2 = \mathrm{Tr}(g^k)\overline{\mathrm{Tr}(g^k)} = \sum_{n=1}^{N} e^{ki\theta_n} \sum_{m=1}^{N} e^{-ki\theta_m} = \sum_{n,m} e^{ki(\theta_n - \theta_m)}.$$

Hence using theorem 3.5 and lemma 3.6, we see that $\mathbb{E}_{g\in U(N)}\left[\left|\mathrm{Tr}(g^k)\right|^2\right]$ is equal to

$$\sum_{n,m} \frac{1}{(2\pi)^N} \iint e^{ki(\theta_n-\theta_m)} \begin{vmatrix} 1 & e^{-i\theta_1} & e^{-2i\theta_1} & \cdots & e^{-(N-1)i\theta_1} \\ e^{i\theta_2} & 1 & e^{-i\theta_2} & \cdots & e^{-(N-2)i\theta_2} \\ e^{2i\theta_3} & e^{i\theta_3} & 1 & \cdots & e^{-(N-3)i\theta_3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ e^{(N-1)i\theta_N} & e^{(N-2)i\theta_N} & e^{(N-3)i\theta_N} & \cdots & 1 \end{vmatrix} d\theta.$$

Again we consider this determinant as a sum of all possible products of N coefficients of the matrix, where each row and column occurs exactly once, (times a sign). Now $\frac{1}{2\pi}\int_0^{2\pi} e^{ix\theta}d\theta = 0$, unless x is a multiple of 2π, in which case the integral is equal to 1. Since we're only working with integer powers of $e^{i\theta_l}$ in the expression above, we conclude that a term in the summation only has non-zero contribution if the integer power of each $e^{i\theta_l}$ is zero. This is the same argument we used earlier, but this time we need to consider different terms in the sum, due to the factor $e^{ki(\theta_n-\theta_m)}$.

First consider the case $k = 0$. Here we can make the same calculation as in the proof of the Weyl Integration Formula, so we see that

$$\frac{1}{(2\pi)^N} \iint e^{0\cdot i(\theta_n-\theta_m)} \begin{vmatrix} 1 & e^{-i\theta_1} & e^{-2i\theta_1} & \cdots & e^{-(N-1)i\theta_1} \\ e^{i\theta_2} & 1 & e^{-i\theta_2} & \cdots & e^{-(N-2)i\theta_2} \\ e^{2i\theta_3} & e^{i\theta_3} & 1 & \cdots & e^{-(N-3)i\theta_3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ e^{(N-1)i\theta_N} & e^{(N-2)i\theta_N} & e^{(N-3)i\theta_N} & \cdots & 1 \end{vmatrix} d\theta = 1$$

for all n, m. Since we're summing over N^2 pairs (n, m) we conclude that

$$\mathop{\mathbb{E}}_{X\in U(N)}\left[\left|\mathrm{Tr}(X^0)\right|^2\right] = N^2.$$

Next suppose $k \neq 0$ and consider the terms $n = m$. Then $e^{ki(\theta_n-\theta_m)} = 1$, so again the only contribution of the sum of the determinant will come from the products with only 1's. Since we have N pairs (n, m) with $n = m$, these terms sum up to N.

Next consider the terms with $n \neq m$. When $|k| \geq N$, then the factor $e^{ki(\theta_n-\theta_m)} = e^{ki\theta_n}e^{-ki\theta_m}$ cannot be cancelled by the factors in the determinant, as these have largest power $N-1$. Hence when $|k| \geq N$ the contributions of the terms $n \neq m$ are all zero and $\mathbb{E}_{g\in U(N)}\left[\left|\mathrm{Tr}(g^k)\right|^2\right] = N$.

Finally when $|k| = N - j$ with $1 \leq j < N$ and $n \neq m$, it is possible for factors in the determinant to cancel $e^{ki(\theta_n-\theta_m)} = e^{ki\theta_n}e^{-ki\theta_m}$. Hence the products that have non-zero contribution are those that cancel $e^{ki\theta_n}e^{-ki\theta_m}$ and have further only 1's in the product. This can only happen when $m - n = k$. Suppose that $m - n \neq k$ and that the discriminant

has some product that cancels $e^{ki\theta_n}e^{-ki\theta_m}$. Then this product contains the matrix elements $(k+n,n)$ and $(m-k,m)$. We must have that $(k+n,n) = (m,m-k)$, (i.e. that these elements are each other transposed), as otherwise the rest of the product cannot consist of all 1's, (remember that every row and column occur exactly once and there are only 1's on the diagonal). Now there are exactly j such pairs (n,m) and all have sign -1. We conclude that $\mathbb{E}_{g\in U(N)}\left[\left|\mathrm{Tr}(g^k)\right|^2\right] = N - j = |k|$. □

This concludes the introduction to random matrix theory. From the theorems in this chapter, we are specifically interested in Dyson's theorem 3.7. We use it in chapter 6. In the next chapter we will see another matrix integral, but the means to calculate it will be entirely different.

Chapter 4

Introduction to the representations of algebras

4.1 Introduction

In this chapter we will look at representation theory and more specifically the representations of associative algebras. Note that it is also possible to study the representation of groups. This is a matter of semantics. In the later sections we will study the representation of groups by studying the representation of their group algebras.

Representation theory is a large and very useful topic in mathematics. We will only be able to look at a few specific results, which we will need for this thesis. In particular we are interested in two identities, given by corollary 4.8 and corollary 4.20. For further reading and a basic, more general introduction, one could look at [8] or [25]. The same books were consulted in writing this chapter.

Note that in this chapter we will assume to work over $\mathbb{C}$. Although algebras and representations are defined over a field k in general, in this thesis we only need to work over the field $\mathbb{C}$.

Definition 4.1. *An* algebra *is a vector space over $\mathbb{C}$, having a multiplication $a, b \mapsto a \cdot b$, such that for all $x, y, \in A$, $\lambda, \mu \in \mathbb{C}$:*

- $(x + y) \cdot z = x \cdot z + y \cdot z$.
- $x \cdot (y + z) = x \cdot y + x \cdot z$.
- $\lambda x \cdot \mu y = (\lambda\mu)(x \cdot y)$.

It has a unit $1 \in A$, s.t. $1 \cdot a = a \cdot 1 = a$ for all $a \in A$.

Definition 4.2. *Let A and B denote two algebras over $\mathbb{C}$. $\rho : A \to B$ is called a* homomorphism of algebras *if for all $\lambda \in \mathbb{C}$, $x, y \in A$:*

- $\rho(\lambda x) = \lambda\rho(x)$.
- $\rho(x + y) = \rho(x) + \rho(y)$.
- $\rho(x \cdot y) = \rho(x) \cdot \rho(y)$.

Definition 4.3. *Let A be an associative algebra. A* representation *of A is a pair (V, ρ), consisting of a vector space V over $\mathbb{C}$ and a homomorphism of algebras $\rho : A \to \mathrm{End}(V)$.*

In our particular case we're interested in the algebra $\mathrm{End}(\mathbb{C}^N) = \mathrm{Mat}_{N\times N}(\mathbb{C})$, consisting of linear maps $\mathbb{C}^N \to \mathbb{C}^N$, or equivalently of $N \times N$-matrices with complex entries. In this case a representation is given by a vector space V over $\mathbb{C}$, together with a homomorphism $\mathrm{End}(\mathbb{C}^N) \to \mathrm{End}(V)$. Let us list a few examples.

Example 4.4.

- *The trivial representation. $V = \mathbb{C}^N$, $\rho : \mathrm{End}(\mathbb{C}^N) \to \mathrm{End}(\mathbb{C}^N)$, $g \mapsto g$.*
- *The direct sum of two representations. Given two vector spaces V_1, V_2 and two homomorphisms ρ_1, ρ_2, we can define a vector space $V = V_1 \oplus V_2$ with a homomorphism $\rho : g \mapsto (\rho_1(g), \rho_2(g))$.*
- *The tensor product of two representations. Given two vector spaces V_1, V_2 and two homomorphisms ρ_1, ρ_2, we can define a vector space $V = V_1 \otimes V_2$ with a homomorphism $\rho : g \mapsto \rho_1(g) \otimes \rho_2(g)$.*
- *The n-th power tensor product. Take $V = (\mathbb{C}^N)^{\otimes n} = \mathbb{C}^N \otimes \ldots \otimes \mathbb{C}^N$. $\rho : g \mapsto (g \otimes \ldots \otimes g)$.*

This n-th power tensor is important to us, so we will calculate a specific example.

Example 4.5. *Take $N = n = 2$. Suppose that a matrix $M \in \mathrm{End}(\mathbb{C}^2)$ is given by $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Now $V = \mathbb{C}^2 \otimes \mathbb{C}^2$ is the vector space spanned by the vectors $\{e_1 \otimes e_1, e_1 \otimes e_2, e_2 \otimes e_1, e_2 \otimes e_2\}$. We calculate the effect of $\rho(M)$ on each of these vectors.*

$$\begin{aligned}\rho(M)(e_1 \otimes e_1) &= (M \otimes M)(e_1 \otimes e_1) = M(e_1) \otimes M(e_1) = (ae_1 + ce_2) \otimes (ae_1 + ce_2) \\ &= a^2(e_1 \otimes e_1) + ac(e_1 \otimes e_2) + ac(e_2 \otimes e_1) + c^2(e_2 \otimes e_2).\end{aligned}$$

Similarly, we find that

$$\begin{aligned}\rho(M)(e_1 \otimes e_2) &= ab(e_1 \otimes e_1) + ad(e_1 \otimes e_2) + bc(e_2 \otimes e_1) + cd(e_2 \otimes e_2), \\ \rho(M)(e_2 \otimes e_1) &= ab(e_1 \otimes e_1) + bc(e_1 \otimes e_2) + ad(e_2 \otimes e_1) + cd(e_2 \otimes e_2), \\ \rho(M)(e_2 \otimes e_2) &= b^2(e_1 \otimes e_1) + bd(e_1 \otimes e_2) + bd(e_2 \otimes e_1) + d^2(e_2 \otimes e_2).\end{aligned}$$

We conclude that $\rho(M) \in \mathrm{End}(\mathbb{C}^2 \otimes \mathbb{C}^2)$ is given by

$$\rho(M) = \begin{pmatrix} a^2 & ab & ab & b^2 \\ ac & ad & bc & bd \\ ac & bc & ad & bd \\ c^2 & cd & cd & d^2 \end{pmatrix}.$$

4.2 The symmetric n-th power

Note the apparent symmetry in that last example concerning the vectors $e_1 \otimes e_2$ and $e_2 \otimes e_1$. This symmetry motivates us to consider the same n-th tensor power vector space with this symmetry divided out. Let $[N]$ denote the set $\{1, 2, \ldots, N\}$. Then a vector $v \in (\mathbb{C}^N)^{\otimes n}$ is written as

$$v = \sum_{j=(j_1,\ldots,j_n)\in[N]^n} a_j(e_{j_1} \otimes \ldots \otimes e_{j_n}).$$

Now let S_n be the permutation group of n elements. Given a permutation $\sigma \in S_n$, define

$$\sigma(v) = \sum_{j=(j_1,\ldots,j_n)\in[N]^n} a_j(e_{j_{\sigma(1)}} \otimes \ldots \otimes e_{j_{\sigma(n)}}).$$

Let W be the subspace of $(\mathbb{C}^N)^{\otimes n}$, spanned by vectors $v-\sigma(v)$ with $v \in (\mathbb{C}^N)^{\otimes n}$, $\sigma \in S_n$. Define

$$\mathrm{Sym}^n(\mathbb{C}^N) = (\mathbb{C}^N)^{\otimes n}/W.$$

We call this vector space the symmetric n-th power. Since it is given by a representation modulo some vector space, it is itself a representation. Let us write down a basis for this vector space. Since the permutation group S_n is generated by transpositions, we know that $\mathrm{Sym}^n(\mathbb{C}^N)$ is spanned by the vectors

$$\{e_{i_1} \otimes \ldots \otimes e_{i_n} \mid 1 \leq i_1 \leq i_2 \leq \ldots \leq i_n \leq N\}.$$

The easiest way to see this, is to look at an example.

Example 4.6. *In the same setting as example 4.5 the above is easily seen. We already know that* $\mathbb{C}^2 \otimes \mathbb{C}^2$ *was spanned by the vectors* $\{e_1 \otimes e_1, e_1 \otimes e_2, e_2 \otimes e_1, e_2 \otimes e_2\}$. *Now* $e_1 \otimes e_2 - e_2 \otimes e_1 = 0$ *in* $\mathrm{Sym}^2(\mathbb{C}^2)$, *so* $e_1 \otimes e_2 = e_2 \otimes e_1$. *Furthermore no permutation can give a relation between* $e_1 \otimes e_1, e_1 \otimes e_2$ *and* $e_2 \otimes e_2$, *so we conclude that*

$$\mathrm{Sym}^2\mathbb{C}^2 = \mathrm{Span}\left(\{e_1 \otimes e_1, e_1 \otimes e_2, e_2 \otimes e_2\}\right).$$

As the symmetric n-th power is a representation, we can again calculate $\rho(M)$ *in the case* $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. *Now*

$$\begin{aligned}
\rho(M)(e_1 \otimes e_1) &= (M \otimes M)(e_1 \otimes e_1) = M(e_1) \otimes M(e_1) = (ae_1 + ce_2) \otimes (ae_1 + ce_2) \\
&= a^2(e_1 \otimes e_1) + 2ac(e_1 \otimes e_2) + c^2(e_2 \otimes e_2). \\
\rho(M)(e_1 \otimes e_2) &= ab(e_1 \otimes e_1) + (ad + bc)(e_1 \otimes e_2) + cd(e_2 \otimes e_2), \\
\rho(M)(e_2 \otimes e_2) &= b^2(e_1 \otimes e_1) + 2bd(e_1 \otimes e_2) + d^2(e_2 \otimes e_2).
\end{aligned}$$

Hence

$$\rho(M) = \begin{pmatrix} a^2 & ab & b^2 \\ 2ac & ad+bc & 2bd \\ c^2 & cd & d^2 \end{pmatrix} \in \mathrm{End}(\mathrm{Sym}^2(\mathbb{C}^2)).$$

For notational convenience we will denote this homomorphism ρ as Sym^n. This will not cause any inconvenience. Hence $\mathrm{Sym}^n(\mathbb{C}^N)$ is the vector space spanned by $\{e_{i_1} \otimes \ldots \otimes e_{i_n} \mid 1 \leq i_1 \leq i_2 \leq \ldots \leq i_n \leq N\}$, while $\mathrm{Sym}^n M$ is a linear transformation on this vector space.

4.3 The trace of the symmetric n-th power

We are now interested in the trace of the matrix $\mathrm{Sym}^n M$. This will give us one of the identities we will use in chapter 7.

Theorem 4.7. *Let* $M \in \mathrm{End}(\mathbb{C}^N)$ *be a matrix with eigenvalues* $\lambda_1, \ldots, \lambda_N$. *Then*

$$\mathrm{Tr}\ \mathrm{Sym}^n M = \sum_{k_1+\cdots+k_N=n} \lambda_1^{k_1} \ldots \lambda_N^{k_N}.$$

Proof. First note that it is sufficient to prove this for the case that M is an upper-triangular matrix. Given a matrix M, there exist a matrix P, s.t. $P^{-1}MP = J$ is a matrix in Jordan normal form, (so in particular it is upper-triangular). Now

$$\begin{aligned}\text{Tr Sym}^n M &= \text{Tr Sym}^n PJP^{-1} = \text{Tr } (\text{Sym}^n P \text{ Sym}^n J \text{ Sym}^n P^{-1}) \\ &= \text{Tr } (\text{Sym}^n J \text{ Sym}^n P^{-1} \text{ Sym}^n P) = \text{Tr } (\text{Sym}^n J \text{ Sym}^n P^{-1} P) = \text{Tr Sym}^n J,\end{aligned}$$

using the fact that the trace function is invariant under cyclic permutations and that Sym^n is a homomorphism of algebras. Since M and J have the same eigenvalues, we conclude that if the theorem holds for all upper-triangular matrices, then it holds for all matrices.

Now suppose that $M = [x_{ij}]_{i,j=1,\dots,N}$ is an upper triangular matrix, (hence $x_{ij} = 0$ if $i < j$). Then $M(e_i) = \sum_j x_{ij} e_j$. Fix $1 \le i_1 \le \cdots \le i_n \le N$ and calculate $\text{Sym}^n M(e_{i_1} \otimes \cdots \otimes e_{i_n})$. This is given by

$$\left(\sum_j x_{i_1 j} e_j\right) \otimes \cdots \otimes \left(\sum_j x_{i_n j} e_j\right) = \sum_{(j_1,\dots,j_n)\in[N]^n} x_{i_1 j_1} \dots x_{i_n j_n} (e_{j_1} \otimes \cdots \otimes e_{j_n})$$

Now we call two n-tuples $(i_1, \dots i_n), (j_1, \dots, j_n) \in [N]^n$ equivalent if there exists a permutation $\sigma \in S_n$, such that $(i_1, \dots, i_n) = (j_{\sigma(1)}, \dots, j_{\sigma(n)})$. In this case we will write $i \sim j$. Given the tuple $(i_1, \dots i_n)$, fixed above, we see that the coefficient of $e_{i_1} \otimes \cdots \otimes e_{i_n}$ in the expansion of $\text{Sym}^n M(e_{i_1} \otimes \cdots \otimes e_{i_n})$ is given by

$$\sum_{\substack{(j_1,\dots,j_n)\in[N]^n \\ j\sim i}} x_{i_1 j_1} \dots x_{i_n j_n}.$$

Now if M is upper triangular, then for $i \sim j$,

$$x_{i_1 j_1} \dots x_{i_n j_n} = \begin{cases} x_{i_1 i_1} \dots x_{i_n i_n} & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

The first equality is trivial. For the second, note that if two tuples are equivalent, but not equal to each other, there always exists a k s.t. $i_k < j_k$. Hence $x_{i_k j_k} = 0$. Furthermore note that if M is upper triangular, it has its eigenvalues on its diagonal. Hence

$$\text{Tr Sym}^n M = \sum_{1\le i_1\le\cdots\le i_n\le N} x_{i_1 i_1} \dots x_{i_n i_n} = \sum_{1\le i_1\le\cdots\le i_n\le N} \lambda_{i_1} \dots \lambda_{i_n} = \sum_{k_1+\cdots+k_N=n} \lambda_1^{k_1} \dots \lambda_N^{k_N}.$$

□

Corollary 4.8. *Let $M \in \text{End}(\mathbb{C}^N)$. Then*

$$\frac{1}{\det(I_N - Mx)} = \sum_{n\ge 0} (\text{Tr Sym}^n M)\ x^n$$

as formal power series.

Proof. Suppose that M has eigenvalues $\lambda_1, \dots, \lambda_N$. Then

$$\begin{aligned}\frac{1}{\det(I_N - Mx)} &= \prod_{k=1}^N \frac{1}{1-\lambda_k x} = \prod_{k=1}^N \left(\sum_{n\ge 0} \lambda_k^n x^n\right) \\ &= \sum_{n\ge 0} \left(\sum_{k_1+\cdots+k_N=n} \lambda_1^{k_1} \dots \lambda_N^{k_N}\right) x^n = \sum_{n\ge 0} (\text{Tr Sym}^n M)\ x^n.\end{aligned}$$

□

4.4 Irreducible representations

Definition 4.9. *Given an associative algebra A and a representation (V, ρ), a* subrepresentation *of this representation V is a subspace $W \subseteq V$, invariant under $\rho(a)$ for all $a \in A$.*

Definition 4.10. *A representation (V, ρ) is called* irreducible *if there are no subrepresentations other than 0 and V.*

Definition 4.11. *A representation (V, ρ) is called* indecomposable *if there exist no non-trivial subrepresentations V_1, V_2 s.t. $V = V_1 \oplus V_2$.*

Clearly an irreducible representation is indecomposable. The converse needs not necessarily hold. However, if V is a finite dimensional vector space, the converse does hold. We will not prove this, as we do not need it. We will prove the following important property for irreducible representations.

Lemma 4.12 (Schur's lemma)**.** *Let A be an algebra and (V, ρ) be an irreducible representation of A. If $\phi : V \to V$ is a linear operator commuting with $\rho(a)$ for each $a \in A$, then ϕ is a scalar.*

Proof. Let λ be some eigenvalue of ϕ. Denote by $E_\lambda = \{v \in V | \phi(v) = \lambda v\}$ its eigenspace of ϕ. Then for each $v \in E_\lambda$, $a \in A$, we have that

$$\phi(\rho(a)(v)) = \rho(a)\left(\phi(v)\right) = \rho(a)\left(\lambda v\right) = \lambda\rho(a)(v).$$

Hence for each $v \in E_\lambda$, $a \in A$, $\rho(a)(v) \in E_\lambda$. By definition E_λ is a non-empty subrepresentation of V, so $E_\lambda = V$, as V is irreducible. Hence for every vector $v \in V$, we have that $\phi(v) = \lambda v$ and we conclude that ϕ is just multiplication by the scalar λ. □

We also prove that the symmetric n-th power defined in section 4.2 is an irreducible representation. We need this for the important corollary 4.20, which is one of the two identities we set out to prove in this chapter.

Theorem 4.13. *For any $n \geq 0$, $\mathrm{Sym}^n(\mathbb{C}^N)$ is an irreducible representation of $\mathrm{End}(\mathbb{C}^N)$.*

Proof. Let $p_1, \ldots, p_N$ denote the first N prime numbers and let $A \in \mathrm{End}(\mathbb{C}^N)$ be the diagonal matrix with these prime numbers on its diagonal. Now $\mathrm{Sym}^n A$ is also a diagonal matrix. Furthermore each number on its diagonal is a combination of n prime numbers. By the fundamental theorem of arithmetic all these diagonal elements are different, so we conclude that the eigenvalues $\lambda_1, \ldots, \lambda_m$ of $\mathrm{Sym}^n A$ are all distinct. As A is diagonal with distinct eigenvalues, the eigenvectors $v_1, \ldots, v_m$ of $\mathrm{Sym}^n A$ are given by the basis of our vector space $\mathrm{Sym}^n(\mathbb{C}^N)$. So each eigenvector v_j is given by $e_{i_1} \otimes \cdots \otimes e_{i_n}$ for some tuple $1 \leq i_1 \leq \cdots \leq i_n \leq N$. Now suppose $W \subseteq \mathrm{Sym}^n(\mathbb{C}^N)$ is a non-empty subspace of $\mathrm{Sym}^n(\mathbb{C}^N)$, invariant under $\mathrm{Sym}^n M$ for all $M \in \mathrm{End}(\mathbb{C}^N)$. We show that there exists a non-empty subset S of eigenvectors of $\mathrm{Sym}^n A$, such that S is a basis of W. Since W is non-empty, and the the eigenvectors $v_1, \ldots, v_m$ of $\mathrm{Sym}^n A$ form an orthogonal basis of $\mathrm{Sym}^n(\mathbb{C}^N)$, there exists some non zero

$$w = \alpha_1 v_1 + \cdots + \alpha_m v_m \in W.$$

As W is invariant under $\mathrm{Sym}^n M$ for all $M \in \mathrm{End}(\mathbb{C}^N)$, we know that

$$\mathrm{Sym}^n A(w) = \alpha_1 \lambda_1 v_1 + \cdots + \alpha_m \lambda_m v_m \in W.$$

Hence

$$\mathrm{Sym}^n A(w) - \lambda_m w = \alpha_1(\lambda_1 - \lambda_m)v_1 + \cdots + \alpha_{m-1}(\lambda_{m-1} - \lambda_m)v_{m-1} \in W.$$

Continuing the process of eliminating eigenvectors, we find that all vectors

$$\alpha_1 v_1, \ldots, \alpha_m v_m \in W.$$

As w was non zero, we know that some of the α_j are non zero, hence these $v_j \in W$. Repeating this to determine which of $v_1, \ldots, v_m$ are in W, we conclude that there exists a non-empty subset S of eigenvectors of $\mathrm{Sym}^n A$, such that S is a basis of W. Finally we show that S is given by all eigenvectors v_j, (hence $W = \mathrm{Sym}^n(\mathbb{C}^N)$). When S does not consist of all eigenvectors, there exists two tuples $i, i' \in [N]^n$, s.t. $i = (i_1, \ldots, i_{n-1}, k)$ and $i' = (i_1, \ldots, i_{n-1}, l)$ with

$$e_{i_1} \otimes \cdots \otimes e_{i_{n-1}} \otimes e_k \in S$$

and

$$e_{i_1} \otimes \cdots \otimes e_{i_{n-1}} \otimes e_l \notin S.$$

Note that we don't demand these sequences to be increasing. Define the $N \times N$ matrix E_{kl} by having zeroes everywhere, except at position lk, where it is one. Then $E_{kl}e_j = \delta_{kj}e_l$. Now consider $(I_N + E_{kl})$. Since W is invariant under $\mathrm{Sym}^n M$ for all $M \in \mathrm{End}(\mathbb{C}^N)$, we find that

$$\mathrm{Sym}^n(I_N + E_{kl})(e_{i_1} \otimes \cdots \otimes e_{i_{n-1}} \otimes e_k) = (e_{i_1} + \delta_{ki_1}e_l) \otimes \cdots \otimes (e_{i_{n-1}} + \delta_{ki_{n-1}}e_l) \otimes (e_k + e_l) \in W.$$

By eliminating the other vectors with the same procedure shown above, we see that

$$e_{i_1} \otimes \cdots \otimes e_{i_{n-1}} \otimes e_l \in S$$

and we have derived a contradiction. We conclude that S consists of all eigenvectors of $\mathrm{Sym}^n A$ and hence that $W = \mathrm{Sym}^n(\mathbb{C}^N)$. It follows that $\mathrm{Sym}^n(\mathbb{C}^N)$ is an irreducible representation of $\mathrm{End}(\mathbb{C}^N)$. □

4.5 Topological groups and characters

A logical next step in any book on representation theory is the study of representations of finite groups. Here, a representation of a group G is a representation of the algebra $k[G]$. It is obvious that a homomorphism $\rho : k[G] \to \mathrm{End}(V)$ is uniquely determined by $\rho|_G : G \mapsto \mathrm{Aut}(V)$ and vice versa. It will then introduce characters of a finite group and show a lot of nice properties about these characters. Only in a later stage, it will consider compact topological groups, (such as $U(N)$), for which much of the same results apply, although these groups are not finite. For this thesis the study of finite group representations, although very interesting, is not necessary. We will hence skip this part and focus immediately on compact topological groups. We will furthermore only prove the results on characters necessary for this thesis.

Definition 4.14. *A* topological group *G is a group, on which a topology is defined, such that both the group operation and the inverse function are continuous. It is compact if it is so as a topological space.*

Definition 4.15. *A representation of a topological group G is a pair (V, ρ) consisting of a finite dimensional vector space V over $\mathbb{C}$, on which a topology is defined, and a continuous group homomorphism $\rho : G \to \mathrm{Aut}(V)$. Here the topology on $\mathrm{Aut}(V)$ is inherited from the product topology on $\mathrm{End}(V)$.*

Definition 4.16. *Given a representation (V, ρ) of a group G, the* character $\chi : G \to \mathbb{C}$ *associated to this representation, is given by $\chi(g) = \mathrm{Tr}_V(\rho(g))$.*

We will now focus on the case that $G = U(N)$. Although this is not (yet) necessary, it will make notation and the proofs somewhat easier. Recall definition 3.3 of a class function on $U(N)$, stating that a function $f : U(N) \to \mathbb{C}$ is a class function if $f(g)$ only depends on the eigenvalues of g. The character associated to a representation of $U(N)$ is a class function, as

$$\chi(h^{-1}gh) = \mathrm{Tr}_V\left(\rho(h^{-1}gh)\right) = \mathrm{Tr}_V\left(\rho(h^{-1})\rho(g)\rho(h)\right) = \mathrm{Tr}_V\left(\rho(g)\rho(h)^{-1}\rho(h)\right) = \mathrm{Tr}_V(g) = \chi(g)$$

and as $\chi(g)$ is symmetric in the eigenvalues of g. Now for two class functions ϕ, ψ on $U(N)$, we define an inner product

$$\langle \phi | \psi \rangle = \int_{U(N)} \phi(g) \cdot \overline{\psi(g)} d\mu(g),$$

where $d\mu(g)$ is the normalised Haar measure. It turns out that the characters of the irreducible representations of $U(N)$ form an orthonormal basis of class functions on this group. (This holds for any compact topological group G). We will prove only a tiny part of this statement.

Theorem 4.17. *Let (V, ρ) be an irreducible representation of $U(N)$. Let χ be the character associated to this representation. Then*

$$|\chi|^2 = \langle \chi | \chi \rangle = 1.$$

Proof. Suppose $\phi : V \to V$ is a linear map. Define the linear map $\phi_0 : V \to V$ by

$$\phi_0(v) = \int_{U(N)} (\rho(g) \circ \phi \circ \rho^{-1}(g))(v) d\mu(g).$$

Now ϕ_0 commutes with $\rho(h)$ for any $h \in U(N)$, as

$$\begin{aligned}
\rho(h) \circ \phi_0 \circ \rho^{-1}(h)(v) &= \int_{U(N)} (\rho(h) \circ \rho(g) \circ \phi \circ \rho^{-1}(g) \circ \rho^{-1}(h))(v) d\mu(g) \\
&= \int_{U(N)} (\rho(hg) \circ \phi \circ \rho^{-1}(hg) \circ)(v) d\mu(g) \\
&= \int_{U(N)} (\rho(g) \circ \phi \circ \rho^{-1}(g))(v) d\mu(g) \\
&= \phi_0(v).
\end{aligned}$$

Since (V, ρ) is an irreducible representation, it follows by Schur's lemma 4.12 that ϕ_0 is equal to multiplication by a scalar λ. Obviously taking the trace of ϕ_0 yields

$$\mathrm{Tr}_V(\phi_0) = \dim(V) \cdot \lambda.$$

However, we also know that

$$\begin{aligned}
\mathrm{Tr}_V(\phi_0) &= \int_{U(N)} \mathrm{Tr}_V(\rho(g) \circ \phi \circ \rho^{-1}(g)) d\mu(g) \\
&= \int_{U(N)} \mathrm{Tr}_V(\phi \circ \rho^{-1}(g) \circ \rho(g)) d\mu(g) \\
&= \int_{U(N)} \mathrm{Tr}_V(\phi) d\mu(g) = \mathrm{Tr}_V(\phi),
\end{aligned}$$

again using the fact that the trace is invariant under cyclic permutations. Hence ϕ_0 is equal to multiplication by a scalar $\lambda = \frac{\mathrm{Tr}_V(\phi)}{\dim(V)}$. Let $v_1, \ldots, v_m$ be an orthonormal basis for V. Note that V is finite dimensional, so it is indeed possible to choose such a basis. Suppose $u, w \in V$ are vectors. Write $u = u_1 v_1 + \cdots + u_m v_m$ and $w = w_1 v_1 + \cdots + w_m v_m$. Denote by

$$\langle u|w\rangle = \sum_{i=1}^{m} \overline{u_i} w_i$$

the standard hermitian inner product on V. Using the bra-ket notation, we see that

$$\begin{aligned}
\langle \chi|\chi\rangle &= \int_{U(N)} \overline{\chi(g)} \cdot \chi(g) d\mu(g) \\
&= \int_{U(N)} \chi(\overline{g}) \cdot \chi(g) d\mu(g) \\
&= \int_{U(N)} \chi(g^{-1}) \cdot \chi(g) d\mu(g) \\
&= \int_{U(N)} \mathrm{Tr}_V(\rho(g^{-1})) \cdot \mathrm{Tr}_V(\rho(g)) d\mu(g) \\
&= \int_{U(N)} \sum_{i=1}^{m} \langle v_i|\rho(g^{-1})|v_i\rangle \sum_{j=1}^{m} \langle v_j|\rho(g)|v_j\rangle d\mu(g) \\
&= \sum_{i,j=1}^{m} \int_{U(N)} \langle v_i|\rho(g^{-1})|v_i\rangle \langle v_j|\rho(g)|v_j\rangle d\mu(g).
\end{aligned}$$

Here $|v_i\rangle\langle v_j|$ is the linear function $V \to V$ mapping a vector w onto the vector $\langle v_j|w\rangle v_i$. By the above the function $\rho(g^{-1})|v_i\rangle\langle v_j|\rho(g)$ is given by the scalar $\lambda_{ij} = \frac{\mathrm{Tr}_V(|v_i\rangle\langle v_j|)}{\dim(V)} = \frac{\mathrm{Tr}_V(|v_i\rangle\langle v_j|)}{m}$. Note that $|v_i\rangle\langle v_j|$ is the matrix with all zeroes, except at the position ij, where it is 1. Hence

$$\lambda_{ij} = \begin{cases} \frac{1}{m} & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

Substituting this in the expression above, we see that

$$\langle \chi|\chi\rangle = \sum_{i,j=1}^{m} \int_{U(N)} \langle v_i|\lambda_{ij}|v_j\rangle d\mu(g) = \sum_{i=1}^{m} \lambda_{ii} \langle v_i|v_i\rangle \int_{U(N)} d\mu(g) = \sum_{i=1}^{m} \frac{1}{m} = 1.$$

□

4.6 Weyl's Unitary trick and a final identity

The last theorem that we need again uses a lot of Lie algebra theory. We will not prove it, but just present the idea of the proof.

Definition 4.18. *Let V be a vector space over $\mathbb{C}$ and let (V, ρ) be a representation of a group G. If we have a hermitian inner product on V, denoted $\langle\ |\ \rangle$, then the representation is* unitary *if for all $v, w \in V$, $g \in G$, we have that $\langle v|w\rangle = \langle \rho(g)(v)|\rho(g)(w)\rangle$.*

Equivalently, (V, ρ) is unitary if the image of ρ belongs to

$$U(V) = \{M \in \mathrm{Aut}(V) |\ M \text{ is unitary w.r.t. the hermitian inner product } \langle\ |\ \rangle\}.$$

For example the representation Sym^n on $U(N)$ is unitary: If $M \in U(N)$ is a unitary matrix with eigenvalues $\lambda_1, \ldots, \lambda_N$ and eigenvectors $v_1, \ldots, v_N$, then $\mathrm{Sym}^n M$ has eigenvectors $v_{i_1} \otimes \cdots \otimes v_{i_n}$ for any $1 \leq i_1 \leq \cdots \leq i_n \leq N$. Such an eigenvector has eigenvalue $\lambda_{i_1} \ldots \lambda_{i_n}$, which has absolute value one. Hence Sym^n is a unitary matrix in the vector space $\mathrm{Sym}^n(\mathbb{C}^N)$. This also follows from Weyl's unitary trick.

Theorem 4.19 (Weyl's unitary trick)**.** *Let G be a simple Lie group and K be its maximal compact subgroup. Then for any complex representation (V, ρ) of G, there exists a hermitian inner product on V, s.t. $(V, \rho|_K)$ is a unitary representation of K. Conversely for any unitary representation (V, ρ) of K, there exists a unique extension $\tilde{\rho}$ of ρ, such that $(V, \tilde{\rho})$ is a representation of G.*

Idea of proof for $G = GL_N(\mathbb{C})$, $K = U(N)$. As V is a complex representation of G, we have a standard hermitian inner product $\langle\ |\ \rangle$. Now define a new inner product on V by

$$\langle v|w\rangle' = \int_{U(N)} \langle \rho(g)(v)|\rho(g)(w)\rangle d\mu(g).$$

Now for some $h \in K$, $v, w \in V$

$$\begin{aligned}
\langle \rho(h)(v)|\rho(h)(w)\rangle' &= \int_{U(N)} \langle \rho(g)\rho(h)(v)|\rho(g)\rho(h)(w)\rangle d\mu(g) \\
&= \int_{U(N)} \langle \rho(gh)(v)|\rho(gh)(w)\rangle d\mu(g) \\
&= \int_{U(N)} \langle \rho(g)(v)|\rho(g)(w)\rangle d\mu(g) \\
&= \langle v|w\rangle'.
\end{aligned}$$

The converse is more difficult to prove, but this follows from the fact that $\mathfrak{g} = \mathfrak{K} \oplus i\mathfrak{K}$, where $\mathfrak{g}$ is the Lie algebra of G and $\mathfrak{K}$ is the Lie algebra of K. (This holds for all G, K as in the statement of the theorem). Now a representation of a group corresponds to a representation of its Lie algebra and we can extend a representation of $\mathfrak{K}$ to $\mathfrak{g}$ using the identity $\mathfrak{g} = \mathfrak{K} \oplus i\mathfrak{K}$. For the full proof, see [24, § 6.5-6.7]. □

In our case $GL_N(\mathbb{C})$ is the simple Lie group G and $U(N)$ is its maximal compact subgroup K. It then follows that the restriction of Sym^n to $U(N)$ is unitary. It however also follows that the restriction of Sym^n to $U(N)$ is irreducible. If it was not irreducible, there would exist a non-trivial subrepresentation of Sym^n on $U(N)$. We could then extend this subrepresentation to the whole of $GL_N(\mathbb{C})$, from which it would follow that Sym^n on $GL_N(\mathbb{C})$ has some non-trivial subrepresentation. It would then not be irreducible on $GL_N(\mathbb{C})$ or on $\mathrm{End}(\mathbb{C}^N)$. This contradicts theorem 4.13. We find the following corollary.

Corollary 4.20. *For any $n, N \geq 0$,*

$$\int_{U(N)} |\mathrm{Tr}\ \mathrm{Sym}^n g|^2 \, d\mu(g) = 1.$$

Proof. By the above Sym^n on $U(N)$ is irreducible. Hence the associated character $\chi = \mathrm{Tr}\ \mathrm{Sym}^n$ has absolute value one. That is

$$1 = \langle \chi | \chi \rangle = \int_{U(N)} \chi(g)\overline{\chi(g)} d\mu(g) = \int_{U(N)} |\mathrm{Tr}\ \mathrm{Sym}^n g|^2 \, d\mu(g).$$

□

Chapter 5

Introduction to number theory in $\mathbb{F}_q[T]$

In this section we give all the notions in finite field theory we need for chapter 6 and 7. Moreover, we prove some additional lemmas, following Keating and Rudnick in [20]. For more theorems and background information, we follow Rosen in [23].

5.1 Prime Polynomial Theorem

We start with the Prime Polynomial Theorem, as introduced in chapter 1. It was first proven by Gauss in his (posthomous) manuscript "Die Lehre von den Resten", see [11, p. 589-629].

Theorem 5.1 (Prime Polynomial Theorem (PPT))**.**

$$\sum_{f \in \mathcal{M}_n} \Lambda(f) = q^n.$$

Proof. We first define the zeta function of a finite field by

$$\zeta_q(s) = \sum_{f \text{ monic}} |f|^{-s} = \sum_{f \text{ monic}} q^{-s \deg(f)}.$$

Recall that $\mathcal{M}_n$ consists of all monic polynomials of degree n. Note that $\#\mathcal{M}_n = q^n$. This follows from the fact that we can write a function $f \in \mathcal{M}_n$ as $x^n + a_{n-1}x^n + \cdots + a_0$. We then have q options for every a_i. Now we can rewrite the zeta function to see that

$$\zeta_q(s) = \sum_{f \text{ monic}} q^{-s \deg(f)} = \sum_{n \geq 0} \sum_{f \in \mathcal{M}_n} q^{-ns} = \sum_{n \geq 0} q^n q^{-ns} = \sum_{n \geq 0} \left(q^{1-s}\right)^n = \frac{1}{1 - q^{1-s}}.$$

As we have unique factorization in $\mathbb{F}_q[T]$, we also have an Euler product for our zeta function. This is given by

$$\zeta_q(s) = \prod_{\substack{P \text{ monic} \\ P \text{ irreducible}}} \frac{1}{1 - |P|^{-s}}.$$

Suppose that we have a_d monic irreducible polynomials of degree d. The product $\prod_{P \text{ irreducible}}^{P \text{ monic}} \frac{1}{1-|P|^{-s}}$ then rewrites as $\prod_{d\geq 1}\left(1-q^{-ds}\right)^{-a_d}$. Writing $u=q^{-s}$ and $\mathcal{Z}_q(u)=\zeta_q(s)$, we find the equation

$$\frac{1}{1-qu}=\mathcal{Z}_q(u)=\prod_{d\geq 1}\left(1-u^d\right)^{-a_d}.$$

We take the logarithmic derivative on both sides. On the left hand side we find

$$\frac{d}{du}\log\left(\frac{1}{1-qu}\right)=-\frac{d}{du}\log\left(1-qu\right)=\frac{q}{1-qu}.$$

On the right hand side we have

$$\frac{d}{du}\log\prod_{d\geq 1}\left(1-u^d\right)^{-a_d}=\sum_{d\geq 1}-a_d\frac{d}{du}\log(1-u^d)=\sum_{d\geq 1}\frac{da_du^{d-1}}{1-u^d}.$$

Multiplying both sides with u and expanding them as formal power series, we see that

$$\sum_{n\geq 1}q^nu^n=\frac{qu}{1-qu}=\sum_{d\geq 1}\frac{da_du^d}{1-u^d}=\sum_{d\geq 1}\sum_{n\geq 0}da_du^{dn+d}=\sum_{n\geq 1}\sum_{d\geq 1}da_du^{dn}=\sum_{n\geq 1}\sum_{d|n}da_du^n.$$

This implies that for $n\geq 1$

$$q^n=\sum_{d|n}da_d=\sum_{\substack{P \text{ irreducible monic}\\ \deg(P)|n}}\deg(P)=\sum_{f\in\mathcal{M}_n}\Lambda(f).$$

□

5.2 Dirichlet Characters over $\mathbb{F}_q[T]$

This section assumes some basic knowledge about characters. Just like a Dirichlet character modulo d in $\mathbb{Z}$ is the extension of a character $\chi:(\mathbb{Z}/d\mathbb{Z})^\times\to\mathbb{C}^\times$ to the whole $\mathbb{Z}$, we can define a Dirichlet character modulo some polynomial Q. In this way a Dirichlet character $\chi:\mathbb{F}_q[T]\to\mathbb{C}$ modulo Q is the extension of some character $\chi:(\mathbb{F}_q[T]/(Q))^\times\to\mathbb{C}^\times$ to the whole $\mathbb{F}_q[T]$. Hence $\chi:\mathbb{F}_q[T]\to\mathbb{C}$ is a Dirichlet character modulo Q if and only if:

1. $\chi(fg)=\chi(f)\chi(g)$ for all $f,g\in\mathbb{F}_q[T]$;
2. $\chi(1)=1$;
3. $\chi(f+gQ)=\chi(f)$ for all $f,g\in\mathbb{F}_q[T]$;
4. $\chi(f)=0$ if $\gcd(f,Q)\neq 1$.

We will assume throughout this chapter that $\deg(Q)\geq 2$. Let us denote by $\varphi(Q)=\#(\mathbb{F}_q[T]/(Q))^\times$ the Euler totient function of Q. As there are exactly as many characters on an Abelian group A as there are elements of A, we know that there exist $\varphi(Q)$ Dirichlet characters modulo Q, (since the Dirichlet characters are extensions of characters on $(\mathbb{F}_q[T]/(Q))^\times$).

Also, again using the fact that we have extensions of characters on $(\mathbb{F}_q[T]/(Q))^\times$, we have the following orthogonality relations:

$$\sum_{f \bmod Q} \chi_1(f)\overline{\chi_2}(f) = \begin{cases} \varphi(Q) & \text{if } \chi_1 = \chi_2. \\ 0 & \text{otherwise.} \end{cases} \tag{5.1}$$

$$\sum_{\chi \bmod Q} \chi(f)\overline{\chi}(g) = \begin{cases} \varphi(Q) & \text{if } f \equiv g \bmod Q \text{ and } \gcd(f,Q) = 1. \\ 0 & \text{otherwise.} \end{cases} \tag{5.2}$$

Here the last sum is taken over all Dirichlet characters mod Q. We also have the following analogous definitions to the number field setting.

Definition 5.2. *Let Q' be a proper divisor of Q. A Dirichlet character χ modulo Q is* induced *by another Dirichlet character χ' modulo Q' if $\chi(f) = \chi'(f)$ for all $f \equiv 1 \bmod Q'$ and $\gcd(f,Q) = 1$. The Q' with the smallest degree, s.t. there exists some χ' modulo Q' inducing χ is called the* conductor *of χ. It is unique up to a unit. If χ is not induced by any other Dirichlet character, χ is called* primitive.

Definition 5.3. *Let χ be a character mod Q. Then χ is* even *if $\chi(c) = 1$ for all $c \in \mathbb{F}_q^\times$. If χ is not even, then χ is* odd.

The first definition is clear to be analogous to the number theory setting. For the second, note that characters on $\mathbb{Z}$ are even when $\chi(u) = 1$ for all units u, (namely ± 1). Note that when c generates $\mathbb{F}_q^\times$, then χ is even precisely when $\chi(c) = 1$. Define the evaluation map

$$\begin{aligned} \psi : \{\chi : (\mathbb{F}_q[T]/(Q))^\times \to \mathbb{C}^\times\} &\to \mathbb{F}_q^\times \\ \chi &\mapsto \chi(c). \end{aligned}$$

This map has exactly the even characters as its kernel. Since $\#\mathbb{F}_q^\times = q-1$, there are exactly $\varphi_{\text{ev}}(Q) := \frac{\varphi(Q)}{q-1}$ even characters and $\varphi_{\text{odd}}(Q) := \varphi(Q)\left(1 - \frac{1}{q-1}\right)$ odd characters modulo Q.

Lemma 5.4. *Suppose χ_1, χ_2 are Dirichlet characters modulo $Q = T^m$ with $m \geq 2$, s.t. $\overline{\chi_1}\chi_2$ is even. Then*

$$\sum_{\substack{B \bmod T^m \\ B(0)=1}} \overline{\chi_1}(B)\chi_2(B) = \delta_{\chi_1\chi_2} q^{m-1}.$$

Proof. We know that

$$\sum_{B \bmod T^m} \overline{\chi_1}(B)\chi_2(B) = \delta_{\chi_1\chi_2}\varphi(T^m) = \delta_{\chi_1\chi_2}(q-1)q^{m-1}.$$

Here we used orthogonality relation (5.1) and the fact that $\varphi(T^m) = (q-1)q^{m-1}$. The latter follows from the fact that if a polynomial $\sum_{n=0}^m a_n T^n$ is coprime to T^m, then you have q choices for the a_n with $1 \leq n \leq m$ and $(q-1)$ choices for a_0, (as $a_0 \neq 0$). Now $\overline{\chi_1}(B)\chi_2(B) = 0$ if and only if B and T^m are not coprime, so when $B(0) = 0$. For any B with $B(0) \neq 0$ there exists a unit u and a polynomial B_1 s.t. $B = uB_1$ with $B_1(0) = 1$. Since $\overline{\chi_1}\chi_2$ is even, we find that

$$\begin{aligned} \sum_{B \bmod T^m} \overline{\chi_1}(B)\chi_2(B) &= \sum_{\substack{B \bmod T^m \\ B(0)\neq 0}} \overline{\chi_1}(B)\chi_2(B) = \sum_{\substack{B_1 \bmod T^m \\ B_1(0)=1}} \sum_{u \in \mathbb{F}_q} \overline{\chi_1}(uB_1)\chi_2(uB_1) \\ &= \sum_{\substack{B_1 \bmod T^m \\ B_1(0)=1}} \sum_{u \in \mathbb{F}_q} \overline{\chi_1}(B_1)\chi_2(B_1) = (q-1) \sum_{\substack{B \bmod T^m \\ B(0)=1}} \overline{\chi_1}(B)\chi_2(B). \end{aligned}$$

Hence

$$\sum_{\substack{B \bmod T^m \\ B(0)=1}} \overline{\chi_1}(B)\chi_2(B) = \frac{1}{q-1} \sum_{B \bmod T^m} \overline{\chi_1}(B)\chi_2(B) = \delta_{\chi_1\chi_2} q^{m-1}.$$

□

Recall the convolution product and Möbius inversion from section 2.2. In $\mathbb{F}_q[T]$ an exact analogue holds as the proofs only relied on unique factorization. If we define the Möbius function

$$\mu(f) = \begin{cases} 0 & \text{if } f \text{ is not square free,} \\ (-1)^t & \text{is } f = cP_1 \dots P_t \text{ for } t \text{ distinct irrducible polynomials,} \end{cases}$$

we see that we have the same sort of Möbius inversion. We apply this to estimate the number of primitive characters and the number of even and odd primitive characters. This will help us later, as we will want to estimate some sum over all characters. Using the estimations we derive now, we will see that it is sufficient to consider only primitive even characters. Denote the number of primitive (even) characters mod Q as $\varphi^{\mathrm{pr}}(Q), \varphi^{\mathrm{pr}}_{\mathrm{ev}}(Q)$ respectively. Since every character has a conductor, for which this character is primitive, we know that $\varphi(Q) = \sum_{D|Q} \varphi^{\mathrm{pr}}(D)$. Möbius inversion gives $\varphi^{\mathrm{pr}}(Q) = \sum_{D|Q} \mu(D)\varphi\left(\frac{Q}{D}\right)$. Hence

$$\left|\frac{\varphi^{\mathrm{pr}}(Q)}{\varphi(Q)} - 1\right| = \left|\sum_{\substack{D|Q \\ D\neq 1}} \mu(D)\frac{\varphi\left(\frac{Q}{D}\right)}{\varphi(Q)}\right| \overset{(a)}{\leq} \sum_{\substack{D|Q \\ D\neq 1}} \frac{\varphi\left(\frac{Q}{D}\right)}{\varphi(Q)} \overset{(b)}{\leq} \sum_{\substack{D|Q \\ D\neq 1}} \frac{1}{q} \overset{(c)}{<} \frac{2^{\deg(Q)}}{q}.$$

Here (a) is due to the triangle inequality. (b) follows from the fact that $\varphi(Q)$ is at least q times as large as $\varphi\left(\frac{Q}{D}\right)$ for any proper divisor D, as $\frac{Q}{D}$ has degree strictly smaller than the degree of Q. Finally at (c) we used that there are at most $2^{\deg(Q)}$ divisors of Q. This follows from the fact that we can write $Q = \prod_{i=1}^{\deg(Q)}(T-\alpha_i)$ with the α_i in the algebraic closure of $\mathbb{F}_q$. Any divisor is of the form $\prod_I (T-\alpha_i)$ with $I \subseteq \{1, 2, \dots, \deg(Q)\}$. Since there are $2^{\deg(Q)}$ such subsets I, we conclude that there are at most $2^{\deg(Q)}$ divisors of Q in $\mathbb{F}_q[T]$.
The inequality above shows that

$$\frac{\varphi^{\mathrm{pr}}(Q)}{\varphi(Q)} = 1 + O\left(\frac{1}{q}\right)$$

for a fixed Q. Hence almost all Dirichlet characters are primitive as $q \to \infty$. Furthermore we know that

$$\varphi^{\mathrm{pr}}_{\mathrm{ev}}(Q) = \sum_{D|Q} \mu(D)\varphi_{\mathrm{ev}}\left(\frac{Q}{D}\right) = \frac{1}{q-1}\sum_{D|Q} \mu(D)\varphi\left(\frac{Q}{D}\right) = \frac{\varphi^{\mathrm{pr}}(Q)}{q-1}.$$

5.3 L-functions

Definition 5.5. *Given a Dirichlet charater χ modulo Q, we can define the* L-function *corresponding to χ by*

$$L(s,\chi) = \sum_{f \text{ monic}} \frac{\chi(f)}{|f|^s} = \prod_{\substack{P \text{ monic} \\ P \text{ irreducible}}} \left(1 - \frac{\chi(P)}{|P|^s}\right)^{-1}.$$

Theorem 5.6. *Given a non-trivial Dirichlet charater χ modulo Q, the L-function $L(s,\chi)$ is a polynomial in $u = q^{-s}$ of degree at most* $\deg(Q) - 1$.

Proof. Define $A(n,\chi) = \sum_{f\in\mathcal{M}_n} \chi(f)$. Then

$$L(s,\chi) = \sum_{f \text{ monic}} \frac{\chi(f)}{|f|^s} = \sum_{n\geq 0} \sum_{f\in\mathcal{M}_n} \frac{\chi(f)}{|f|^s} = \sum_{n\geq 0} A(n,\chi)\left(q^{-s}\right)^n.$$

We show that $A(n,\chi) = 0$ for all $n \geq \deg(Q)$. Suppose that we fix such an n. Then f is written in a unique way as $f = gQ + h$ with $\deg(g) = n - \deg(Q)$, $\deg(h) < \deg(Q)$. Note that g has a fixed leading coefficient, given by the multiplicative inverse of the leading coefficient of Q, (as f is monic). If we let g run through all polynomials with degree $n - \deg(Q)$ with this leading coefficient and we let h run through all polynomials of degree strictly smaller than $\deg(Q)$, then we see that f runs exactly through all monic polynomials of degree n. Since there are $q^{n-\deg(Q)}$ such polynomials g, we conclude that

$$A(n,\chi) = \sum_{f\in\mathcal{M}_n} \chi(f) = \sum_g \sum_h \chi(gQ+h) = \sum_g \sum_h \chi(h) = q^{n-\deg(Q)} \sum_{h \bmod Q} \chi(h) = 0,$$

by orthogonality relation (5.1). □

We can hence write an L-function as a product

$$L(s,\chi) = \prod_{j=1}^{\deg(Q)-1} (1 - \alpha_j(\chi)q^{-s}),$$

where the $\alpha_j(\chi)$ are the inverse roots of the L-function and $\alpha_j(\chi) = 0$ for some j if the polynomial has some degree strictly smaller than $\deg(Q) - 1$. It turns out that if χ is primitive, then the polynomial has exactly degree $\deg(Q) - 1$. As we're more interested in the L-function as a function of $u = q^{-s}$, we now define

$$\mathcal{L}(u,\chi) = L(s,\chi) = \prod_{j=1}^{\deg(Q)-1} (1 - \alpha_j(\chi)u).$$

Taking the logarithmic derivative of the equation

$$\mathcal{L}(u,\chi) = \prod_{\substack{P \text{ monic} \\ P \text{ irreducible}}} \left(1 - \chi(P)u^{\deg(P)}\right)^{-1} = \prod_{j=1}^{\deg(Q)-1} (1 - \alpha_j(\chi)u)$$

yields on the right hand side:

$$\begin{aligned}\frac{d}{du}\log\left(\prod_{j=1}^{\deg(Q)-1} (1 - \alpha_j(\chi)u)\right) &= \sum_{j=1}^{\deg(Q)-1} \frac{d}{du}\log\left(1 - \alpha_j(\chi)u\right) = -\sum_{j=1}^{\deg(Q)-1} \frac{\alpha_j(\chi)}{1 - \alpha_j(\chi)u} \\ &= -\sum_{n\geq 0} \sum_{j=1}^{\deg(Q)-1} \alpha_j(\chi)^{n+1}u^n = -\sum_{n\geq 1} \sum_{j=1}^{\deg(Q)-1} \alpha_j(\chi)^n u^{n-1}.\end{aligned}$$

On the left hand side, you find

$$\begin{aligned}\frac{d}{du}\log\left(\prod_P\left(1-\chi(P)u^{\deg(P)}\right)^{-1}\right) &= -\sum_P \frac{d}{du}\log\left(1-\chi(P)u^{\deg(P)}\right)\\ &= \sum_P \frac{\chi(P)\deg(P)u^{\deg(P)-1}}{1-\chi(P)u^{\deg(P)}}\\ &= \sum_P\sum_{n\geq 0}\deg(P)\chi(P)^{n+1}u^{(n+1)\deg(P)-1}\\ &= \sum_P\sum_{n\geq 0}\deg(P)\chi(P^{n+1})u^{\deg(P^{n+1})-1}\\ &= \sum_{f \text{ monic}}\Lambda(f)\chi(f)u^{\deg(f)-1} = \sum_{n\geq 1}\sum_{f\in\mathcal{M}_n}\chi(f)\Lambda(f)u^{n-1}.\end{aligned}$$

Now define

$$\mathcal{M}(n,\chi\Lambda) = \sum_{f\in\mathcal{M}_n}\chi(f)\Lambda(f).$$

By the above it follows that

$$\mathcal{M}(n,\chi\Lambda) = -\sum_{j=1}^{\deg(Q)-1}\alpha_j(\chi)^n.$$

We will use this function to prove theorem 1.1. The Riemann hypothesis in finite fields was proven in 1948 by Andre Weil. A year later he gives a complete proof in [26]. The theorem tells us that for each (nonzero) inverse root $\alpha_j(\chi) = 1$ or $|\alpha_j(\chi)| = \sqrt{q}$. This implies that $\mathcal{M}(n,\chi\Lambda) = O(nq^{n/2})$. We conclude this section by deriving a relation between $\mathcal{M}(n,\chi\Lambda)$ and the unitarized Frobenius matrix Θ_χ of a primitive character χ, defined below. Before we do so, we first need to look at the completed L-function of the primitive character.

Theorem 5.7. *If χ is an even character, then $u = 1$ is a zero of $\mathcal{L}(u,\chi)$.*

Proof. If χ is even, then $\chi(c) = 1$ for all $c \in \mathbb{F}_q^\times$. Now

$$\begin{aligned}\mathcal{L}(1,\chi) = \sum_{n=0}^{\deg(Q)-1}A(n,\chi) &= \sum_{n=0}^{\deg(Q)-1}\sum_{f\in\mathcal{M}_n}\chi(f) = \frac{1}{q-1}\sum_{c\in\mathbb{F}_q^\times}\chi(c)\sum_{n=0}^{\deg(Q)-1}\sum_{f\in\mathcal{M}_n}\chi(f)\\ &= \frac{1}{q-1}\sum_{c\in\mathbb{F}_q^\times}\sum_{n=0}^{\deg(Q)-1}\sum_{f\in\mathcal{M}_n}\chi(cf) = \frac{1}{q-1}\sum_{g \bmod Q}\chi(g) = 0,\end{aligned}$$

again using orthogonality relation (5.1). □

It turns out that the converse also holds. Odd characters don't have a zero at $u = 1$. We won't specifically prove this, as we're only interested in even characters, but the idea of the proof is the following. First it is possible to prove that $\prod_{\chi \bmod Q} L(s,\chi)$ doesn't have a zero at $s = 1$. As the trivial character χ_0 is the only character where $L(s,\chi_0)$ has a pole at $s = 1$, it is not possible that there are two characters modulo Q with $L(s,\chi)$ having a zero at $s = 1$. Now if χ is odd, then $L(s,\chi)$ cannot have a zero at $s = 1$, as otherwise you indeed have two such characters,

namely χ and $\overline{\chi}$. Hence for odd χ $L(s,\chi)$ does not have a zero at $s=1$, [23, Thm. 4.5]. Now we have the functional equation relating $L(s,\chi)$ and $L(1-s,\overline{\chi})$, [23, Thm. 9.24A]. From this it follows that if χ is odd, then $L(s,\overline{\chi})$ does not have a zero at $s=0$, so $L(s,\chi)$ does not have a zero at $s=0$. Therefore $\mathcal{L}(u,\chi)$ does not have a zero at $u=1$.

Definition 5.8. *Let χ be a primitive character modulo Q. Define*

$$\lambda_\chi = \begin{cases} 1 & \text{if } \chi \text{ is even,} \\ 0 & \text{if } \chi \text{ is odd.} \end{cases}$$

Now define the completed L-function *of χ by*

$$\mathcal{L}^*(u,\chi) = (1-\lambda_\chi u)^{-1}\mathcal{L}(u,\chi).$$

By the above it follows that for primitive χ the completed L-function is a polynomial of degree $N = \deg(Q) - 1 - \lambda_\chi$. By the generalized Riemann hypothesis we can write

$$\mathcal{L}^*(u,\chi) = \prod_{j=1}^{N}(1-\alpha_j(\chi)u) \quad \text{with } |\alpha_j(\chi)| = \sqrt{q}.$$

Now write

$$\alpha_j(\chi) = \sqrt{q}e^{i\theta_j}$$

with each θ_j some angle in $[0, 2\pi[$.

Definition 5.9. *Given a primitive character χ, define the* unitarized Frobenius matrix Θ_χ *as*

$$\Theta_\chi = \begin{pmatrix} e^{i\theta_1} & 0 & \dots & 0 \\ 0 & e^{i\theta_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e^{i\theta_N} \end{pmatrix} \in U(N).$$

The conjugacy class of unitary matrices of Θ_χ in $U(N)$ is called the unitarized Frobenius class *of χ.*

With this definition, it holds that

$$\mathcal{L}^*(u,\chi) = \det(I_n - \sqrt{q}u\Theta_\chi). \tag{5.3}$$

Moreover

$$\mathcal{M}(n,\chi\Lambda) = -\sum_{j=1}^{\deg(Q)-1}\alpha_j(\chi)^n = -q^{\frac{n}{2}}\operatorname{Tr}(\Theta_\chi^n) - \lambda_\chi. \tag{5.4}$$

Both relations are important for our later work in calculating the variance of functions in short intervals.

5.4 A theorem by Katz

In this section we introduce an important theorem, proven by Katz in [16]. This theorem is actually the link between the random matrix theory and the number theory in $\mathbb{F}_q[T]$ in the proof of theorem 1.1 and 1.2. We will not prove this theorem, as this is extremely difficult.

Theorem 5.10. *Fix $n \geq 4$. Working in $\mathbb{F}_q[T]$ and given a primitive character χ modulo T^{n+1}, denote its unitarized Frobenius conjugacy class by $\Theta_{\chi,q}$. In any sequence of finite fields $\mathbb{F}_{q_i}[T]$, not necessarily of the same characteristic, with q_i increasing to ∞, the collection of conjugacy classes*

$$\{\Theta_{\chi,q}\}_{\chi \text{ even, primitive}}$$

becomes equidistributed in $PU(n-1)$.
For $n = 3$ the same result holds, as long as no $\mathbb{F}_{q_i}[T]$ has characteristic 2 or 5.

'Equidistributed' means that the proportion of the collection of conjugacy classes lying in some subset in $PU(n-1)$ is proportional to the measure of that subset. Hence what this theorem tells us, is that if we have a function $f : PU(n-1) \to \mathbb{C}$, then

$$\lim_{q\to\infty} \mathop{\mathbb{E}}_{\substack{\chi \bmod T^{n+1} \\ \chi \text{ even, primitive}}} [f(\Theta_{\chi,q})] = \mathop{\mathbb{E}}_{g\in PU(n-1)} [f(g)] .$$

Here $PU(N)$ is the projective unitary group. It is given by the quotient of the unitary group $U(N)$ by right multiplication of it center. Since the elements of the center of $U(N)$ are just equal to $e^{i\theta} I_N$, we see that elements in $PU(N)$ correspond to the equivalence classes of unitary matrices modulo multiplication by some phase $e^{i\theta}$. Now the Haar measure of $PU(N)$ is the same as the Haar measure on $U(N)$, as $PU(N)$ is defined as $U(N)$, modulo the multiplication by some rotation. If the function f does not change by this phase transition, then we know that

$$\lim_{q\to\infty} \mathop{\mathbb{E}}_{\substack{\chi \bmod T^{n+1} \\ \chi \text{ even, primitive}}} [f(\Theta_{\chi,q})] = \mathop{\mathbb{E}}_{g\in U(n-1)} [f(g)],$$

for which we can apply the Weyl Integration formula 3.5 (if f is also a class function). For example, the function $f(g) = |(\mathrm{Tr}(g)|^2$ does not change by the phase transition, for if a matrix g has eigenvalues $\lambda_1, \ldots, \lambda_N$, then

$$\left|\mathrm{Tr}\left((e^{i\theta} g)^k\right)\right|^2 = \left|\sum_{j=1}^{N} (e^{i\theta}\lambda_j)^k\right|^2 = \left|e^{ik\theta}\right|^2 \left|\sum_{j=1}^{N} \lambda_j^k\right|^2 = \left|\mathrm{Tr}(g^k)\right|^2 .$$

Hence

Corollary 5.11. *Let $k \geq 1$, $n \geq 3$. Then*

$$\lim_{q\to\infty} \mathop{\mathbb{E}}_{\substack{\chi \bmod T^{n+1} \\ \chi \text{ even, primitive}}} \left[\left|(\mathrm{Tr}(\Theta_{\chi,q}^k)\right|^2\right] = \min(n-1, k).$$

Proof. Applying theorem 5.10 to $f(g) = |\mathrm{Tr}(g)|^2$, we see that

$$\lim_{q\to\infty} \mathop{\mathbb{E}}_{\substack{\chi \bmod T^{n+1} \\ \chi \text{ even, primitive}}} \left[\left|(\mathrm{Tr}(\Theta_{\chi,q}^k)\right|^2\right] = \mathop{\mathbb{E}}_{g\in U(n-1)} \left[\left|(\mathrm{Tr}(g^k)\right|^2\right] .$$

Using that $k \geq 1$, we see by theorem 3.7 that this equals $\min(n-1, k)$. □

Chapter 6

The variance of functions in short intervals in $\mathbb{F}_q[T]$

In the first two sections of this chapter we will follow the definitions and proofs of Keating and Rudnick in [19], where they set up a general theory in calculating the variance. The theorems we prove here are also needed in chapter 7, where we study the variance of the Euler totient function in $\mathbb{F}_q[T]$. In the rest of this chapter we will apply these theorems to calculate the variance of the Von Mangoldt function Λ, as Keating and Rudnick did in their (earlier) paper [20].

6.1 Two transformations of functions

Definition 6.1. *Define the following transformation of functions*

$$(.)^* : \mathbb{F}_q[T] \to \mathbb{F}_q[T],$$
$$f^* = T^{\deg(f)} f\left(\frac{1}{T}\right).$$

That is, if $f(T) = \sum_{k=0}^{n} a_k T^k$ with $a_n \neq 0$, so $\deg(f) = n$, then $f^*(T) = \sum_{k=0}^{n} a_{n-k} T^k$. Later on in the proof we will work with Dirichlet characters modulo T^m with $m \geq 1$. Note that

$$\gcd(f, T^m) = 1 \;\Leftrightarrow\; a_0 \neq 0 \;\Leftrightarrow\; f^{**} = f.$$

Furthermore we have some useful statements concerning this transformation.

Lemma 6.2. *If $f, g \in \mathbb{F}_q[T]$, then $(fg)^* = f^* g^*$.*

Proof. Write $f = \sum_{k=0}^{n} a_k T^k$ with $n = \deg(f)$ and $g = \sum_{l=0}^{m} b_l T^l$ with $m = \deg(g)$. Then

$$fg = \sum_{j=0}^{n+m} \sum_{k+l=j} a_k b_l T^j,$$

so

$$(fg)^* = \sum_{j=0}^{n+m} \sum_{k+l=m+n-j} a_k b_l T^j.$$

Now $f^* = \sum_{k=0}^{n} a_{n-k}T^k$ and $g^* = \sum_{l=0}^{m} b_{m-l}T^l$, so

$$f^*g^* = \sum_{j=0}^{n+m} \sum_{k+l=j} a_{n-k}b_{m-l}T^j = \sum_{j=0}^{n+m} \sum_{k'+l'=n+m-j} a_{k'}b_{l'}T^j.$$

□

Corollary 6.3. *Suppose that $f \in \mathbb{F}_q[T]$ is a polynomial with $f(0) \neq 0$. Then f^* is irreducible if and only if f is irreducible.*

Proof. Let f be a polynomial with $f(0) \neq 0$. Since $f^{**} = f$, it is sufficient to prove only one implication. We prove that irreducibility of f^* implies irreducibility of f. Suppose that f^* is irreducible, while f is not. Then we can write $f = gh$ with $\deg(g), \deg(h) > 0$. Since $f(0) \neq 0$, also $g(0) \neq 0$ and $h(0) \neq 0$, so $\deg(g^*), \deg(h^*) > 0$. Now $f^* = g^*h^*$ by lemma 6.2, so f^* is not irreducible. We conclude that the corollary holds. □

We however need another transformation of functions. The fact that whether $(.)^*$ is an involution or not depends on $f(0)$, is not desirable.

Definition 6.4. *For each $n \geq 1$, define a transformation of functions of degree less than or equal to n, by:*

$$\theta_n : \mathcal{P}_{\leq n} \to \mathcal{P}_{\leq n},$$
$$\theta_n(f) = T^n f\left(\frac{1}{T}\right)$$

So for any function $f = a_nT^n + \cdots + a_1T + a_0 \in \mathcal{P}_{\leq n}$, (where a_n is allowed to be zero), we have that $\theta_n(f) = a_0T^n + \cdots + a_{n-1}T + a_n$. Note that θ_n is indeed an involution of functions for each n, (that is $\theta_n(\theta_n(f)) = f$). It is hence a bijection. The following lemma is one of the crucial lemmas in our calculation of the variance, as it allows us to transform a short interval into an arithmetic progression, (for which we can use characters). It is proven by Keating and Rudnick in [19]. Recall that $I(A;h) = \{f \in \mathbb{F}_q[T] : |f - A| \leq q^h\}$.

Lemma 6.5. *Let $h \leq n-2$ and $B \in \mathcal{P}_{n-h-1}$. Then θ_n maps $I(T^{h+1}B;h)$ bijectively to the set $\{g \in \mathcal{P}_{\leq n} : g \equiv \theta_{n-h-1}(B) \bmod T^{n-h}\}$. Furthermore the only functions in $I(T^{h+1}B;h)$ that are mapped to functions of degree n by θ_n are the functions f with $f(0) \neq 0$.*

Proof. The last statement of this lemma follows trivially from the definition of θ_n. Now fix $h \leq n-2$ and let $B \in \mathcal{M}_{n-h-1}$. Write

$$B = b_{n-h-1}T^{n-h-1} + b_{n-h-2}T^{n-h-2} + \cdots + b_1T + b_0.$$

Then we can write for each $f \in I(T^{h+1}B;h)$:

$$f = b_{n-h-1}T^n + b_{n-h-2}T^{n-1} + \cdots + b_1T^{h+2} + b_0T^{h+1} + a_hT^h + a_{h-1}T^{h-1} + \cdots + a_1T + a_0.$$

By definition of θ_n, we know that

$$\theta_n(f) = a_0T^n + a_1T^{n-1} + \cdots + a_hT^{n-h} + b_0T^{n-h-1} + b_1T^{n-h-2} + \cdots + b_{n-h-2}T + b_{n-h-1}.$$

We see that $\theta_n(f)$ lies in the set $\{g \in \mathcal{P}_{\leq n} : g \equiv \theta_{n-h-1}(B) \bmod T^{n-h}\}$. Now $I(T^{h+1}B;h)$ consists of q^{h+1} elements, as we have q choices for each of the coefficients of T^i for all $0 \leq i \leq h$. The set $\{g \in \mathcal{P}_{\leq n} : g \equiv \theta_{n-h-1}(B) \bmod T^{n-h}\}$ also contains q^{h+1} elements, as we have q choices for each of the coefficients of T^i for all $n-h \leq i \leq n$. Since θ_n is a bijection on $\mathbb{P}_{\leq n}$, we conclude that it maps $I(T^{h+1}B;h)$ bijectively to the set $\{g \in \mathcal{P}_{\leq n} : g \equiv \theta_{n-h-1}(B) \bmod T^{n-h}\}$. □

6.2 A general theorem on the variance in short intervals

We are now ready to look at short intervals. In this section we prove a general theorem for the variance in short intervals for a function $\sum_{f\in I(A;h)}\alpha(f)$, where $\alpha:\mathbb{F}_q[T]\to\mathbb{C}$ is function holding some desired properties. Recall the following notation.

Notation 6.6.

- *The average of α over monic polynomials of degree n:*

$$\langle\alpha\rangle_n=\frac{1}{q^n}\sum_{f\in\mathcal{M}_n}\alpha(f).$$

- *The sum of α over a short interval $I(A;h)$, where $A\in\mathcal{M}_n$, $0\le h\le n-2$:*

$$\mathcal{N}_\alpha(A;h)=\sum_{f\in I(A;h)}\alpha(f).$$

- *The average of $\mathcal{N}_\alpha(A;h)$ over all monic polynomials A of degree n:*

$$\langle\mathcal{N}_\alpha(\bullet;h)\rangle_n=\frac{1}{q^n}\sum_{A\in\mathcal{M}_n}\mathcal{N}_\alpha(A;h).$$

- *The variance of $\mathcal{N}_\alpha(A;h)$ over all monic polynomials A of degree n:*

$$\mathrm{Var}_n\mathcal{N}_\alpha(\bullet;h)=\frac{1}{q^n}\sum_{A\in\mathcal{M}_n}\left|\mathcal{N}_\alpha(A;h)-\langle\mathcal{N}_\alpha(\bullet;h)\rangle_n\right|^2.$$

Lemma 6.7.

$$\langle\mathcal{N}_\alpha(\bullet;h)\rangle_n=q^{h+1}\langle\alpha\rangle_n.$$

Proof. Note that we have the partition $\mathcal{M}_n=\cup_{B\in\mathcal{M}_{n-h-1}}I(T^{h+1}B,h)$ for each $0\le h\le n-2$. This follows from the fact that for each $f\in\mathcal{M}_n$, there exists exactly one $B\in\mathcal{M}_{n-h-1}$, such that $|f-T^{n-h}B|\le q^h$, (namely the B s.t. the first $n-h$ coefficients of $T^{h+1}B$ and f coincide). Furthermore if B lies in $I(A,h)$, then $I(A,h)=I(B,h)$. Hence

$$\begin{aligned}\langle\mathcal{N}_\alpha(\bullet;h)\rangle_n&=\frac{1}{q^n}\sum_{A\in\mathcal{M}_n}\sum_{f\in I(A;h)}\alpha(f)=\frac{1}{q^n}\sum_{B\in\mathcal{M}_{n-h-1}}\sum_{A\in I(T^{h+1}B,h)}\sum_{f\in I(A;h)}\alpha(f)\\&=\frac{1}{q^n}\sum_{B\in\mathcal{M}_{n-h-1}}\sum_{A\in I(T^{h+1}B,h)}\sum_{f\in I(T^{h+1}B,h)}\alpha(f)\\&=\frac{q^{h+1}}{q^n}\sum_{B\in\mathcal{M}_{n-h-1}}\sum_{f\in I(T^{h+1}B,h)}\alpha(f)\\&=\frac{q^{h+1}}{q^n}\sum_{f\in\mathcal{M}_n}\alpha(f)=q^{h+1}\langle\alpha\rangle_n.\end{aligned}$$

□

Notation 6.8. *We say that α has property $(*)$, if*

- *α is even. This is the case when $\alpha(f) = \alpha(cf)$ for all $c \in \mathbb{F}_q^\times$ and $f \in \mathbb{F}_q[T]$.*
- *α is multiplicative: for all coprime $f, g \in \mathbb{F}_q[T]$, we have $\alpha(fg) = \alpha(f)\alpha(g)$.*
- *$\alpha(f) = \alpha(f^*)$ for all f with $f(0) \neq 0$.*

Lemma 6.9. *Let $\alpha : \mathbb{F}_q[T] \to \mathbb{C}$ be a function with property $(*)$. Then for all $0 \leq h \leq n-2$, $B \in \mathcal{M}_{n-h-1}$, we have*

$$\mathcal{N}_\alpha(T^{h+1}B; h) = \langle \mathcal{N}_\alpha(\bullet; h)\rangle_n + \frac{1}{\varphi_{\mathrm{ev}}(T^{n-h})} \sum_{m=0}^{n} \alpha(T^{n-m}) \sum_{\substack{\chi \bmod T^{n-h} \\ \chi \neq \chi_0 \\ \chi \text{ even}}} \overline{\chi}(\theta_{n-h-1}(B))\mathcal{M}(m; \alpha\chi),$$

where

$$\mathcal{M}(n; \alpha\chi) = \sum_{f \in \mathcal{M}_n} \alpha(f)\chi(f).$$

Proof. We can write each $f \in \mathcal{M}_n$ uniquely as $f = T^{n-m}f_1$, with $0 \leq m \leq n$ and $f_1 \in \mathcal{M}_m$ s.t. $f_1(0) \neq 0$. Now by construction $\theta_n(f) = \theta_m(f_1) = f_1^*$. Hence

$$\begin{aligned}
\mathcal{N}_\alpha(T^{h+1}B; h) &= \sum_{f \in I(T^{h+1}B; h)} \alpha(f) \\
&= \sum_{m=0}^{n} \sum_{\substack{f_1 \in \mathcal{M}_m \\ f_1(0) \neq 0 \\ T^{n-m}f_1 \in I(T^{h+1}B; h)}} \alpha(T^{n-m}f_1) \\
&\overset{(a)}{=} \sum_{m=0}^{n} \alpha(T^{n-m}) \sum_{\substack{f_1 \in \mathcal{M}_m \\ f_1(0) \neq 0 \\ T^{n-m}f_1 \in I(T^{h+1}B; h)}} \alpha(f_1).
\end{aligned}$$

At (a) we used the fact that α is multiplicative. Now if we fix m and let f_1 run over $\mathcal{M}_m$ s.t. $T^{n-m}f_1 \in I(T^{h+1}B; h)$, then we know by lemma 6.5 that $\theta_n(f) = \theta_m(f_1)$ runs exactly over all polynomials $g \in \mathcal{P}_m$ s.t. $g \equiv \theta_{n-h-1}(B) \bmod T^{n-h}$. Since $\alpha(f_1) = \alpha(f_1^*) = \alpha(\theta_m(f_1))$, we know that

$$\mathcal{N}_\alpha(T^{h+1}B; h) = \sum_{m=0}^{n} \alpha(T^{n-m}) \sum_{\substack{g \in \mathcal{P}_m \\ g \equiv \theta_{n-h-1}(B) \bmod T^{n-h}}} \alpha(g).$$

Now by orthogonality relation (5.1), we know that

$$\sum_{\substack{g \in \mathcal{P}_m \\ g \equiv \theta_{n-h-1}(B) \bmod T^{n-h}}} \alpha(g) = \frac{1}{\varphi(T^{n-h})} \sum_{\chi \bmod T^{n-h}} \overline{\chi}(\theta_{n-h-1}(B)) \sum_{g \in \mathcal{P}_m} \chi(g)\alpha(g).$$

This last sum is now given by

$$\begin{aligned}
\sum_{g \in \mathcal{P}_m} \chi(g)\alpha(g) &= \sum_{f \in \mathcal{M}_m} \sum_{c \in \mathbb{F}_q^\times} \chi(cf)\alpha(cf) \overset{(b)}{=} \sum_{f \in \mathcal{M}_m} \chi(f)\alpha(f) \sum_{c \in \mathbb{F}_q^\times} \chi(c) \\
&= \begin{cases} (q-1)\mathcal{M}(m; \alpha\chi) & \text{if } \chi \text{ is even,} \\ 0 & \text{if } \chi \text{ is odd.} \end{cases}
\end{aligned}$$

At (b) we used that α is even. As $\varphi_{\mathrm{ev}}(T^{n-h}) = \frac{\varphi(T^{n-h})}{q-1}$, we find that

$$\sum_{\substack{g\in\mathcal{P}_m \\ g\equiv\theta_{n-h-1}(B) \bmod T^{n-h}}} \alpha(g) = \frac{1}{\varphi_{\mathrm{ev}}(T^{n-h})} \sum_{\substack{\chi \bmod T^{n-h} \\ \chi \text{ even}}} \overline{\chi}(\theta_{n-h-1}(B))\mathcal{M}(m;\alpha\chi).$$

Hence

$$\mathcal{N}_\alpha(T^{h+1}B;h) = \frac{1}{\varphi_{\mathrm{ev}}(T^{n-h})} \sum_{m=0}^{n} \alpha(T^{n-m}) \sum_{\substack{\chi \bmod T^{n-h} \\ \chi \text{ even}}} \overline{\chi}(\theta_{n-h-1}(B))\mathcal{M}(m;\alpha\chi).$$

Finally we calculate the contribution of the trivial character, which we will show to be $\langle \mathcal{N}_\alpha(\bullet;h)\rangle_n$. Since $\frac{1}{\varphi_{\mathrm{ev}}(T^{n-h})} = \frac{q-1}{\varphi(T^{n-h})} = \frac{q^{h+1}}{q^n}$, the trivial character χ_0 contributes:

$$\frac{1}{\varphi_{\mathrm{ev}}(T^{n-h})} \sum_{m=0}^{n} \alpha(T^{n-m})\overline{\chi_0}(\theta_{n-h-1}(B)) \sum_{f\in\mathcal{M}_m} \alpha(f)\chi_0(f)$$

$$= \frac{q^{h+1}}{q^n} \sum_{m=0}^{n} \sum_{\substack{f\in\mathcal{M}_m \\ f(0)\neq 0}} \alpha(T^{n-m})\alpha(f) = \frac{q^{h+1}}{q^n} \sum_{f\in\mathcal{M}_n} \alpha(f) \overset{\text{(lemma 6.7)}}{=} \langle \mathcal{N}_\alpha(\bullet;h)\rangle_n .$$

We conclude that

$$\mathcal{N}_\alpha(T^{h+1}B;h) = \langle \mathcal{N}_\alpha(\bullet;h)\rangle_n + \frac{1}{\varphi_{\mathrm{ev}}(T^{n-h})} \sum_{m=0}^{n} \alpha(T^{n-m}) \sum_{\substack{\chi \bmod T^{n-h} \\ \chi\neq\chi_0 \text{ even}}} \overline{\chi}(\theta_{n-h-1}(B))\mathcal{M}(m;\alpha\chi).$$

□

This gives rise to the following theorem:

Theorem 6.10. *Let $\alpha : \mathbb{F}_q[T] \to \mathbb{C}$ be a function with property $(*)$. For $0 \le h \le n-2$, we have*

$$\mathrm{Var}_n \mathcal{N}_\alpha(\bullet;h) = \frac{1}{\varphi_{\mathrm{ev}}(T^{n-h})^2} \sum_{\substack{\chi \bmod T^{n-h} \\ \chi\neq\chi_0 \text{ even}}} \sum_{m_1,m_2=0}^{n} \alpha(T^{n-m_1})\overline{\alpha(T^{n-m_2})}\mathcal{M}(m_1;\alpha\chi)\overline{\mathcal{M}(m_2;\alpha\chi)}.$$

Proof. By the same reasoning as in the proof of lemma 6.7, we see that

$$\begin{aligned}\mathrm{Var}_n \mathcal{N}_\alpha(\bullet;h) &= \frac{1}{q^n} \sum_{A\in\mathcal{M}_n} |\mathcal{N}_\alpha(A;h) - \langle \mathcal{N}_\alpha(\bullet;h)\rangle_n|^2 \\ &= \frac{q^{h+1}}{q^n} \sum_{B\in\mathcal{M}_{n-h-1}} \left|\mathcal{N}_\alpha(T^{h+1}B;h) - \langle \mathcal{N}_\alpha(\bullet;h)\rangle_n\right|^2 .\end{aligned}$$

Now we apply lemma 6.9 to see that this equals

$$\frac{1}{\varphi_{\mathrm{ev}}(T^{n-h})^2} \sum_{\substack{\chi_1,\chi_2 \bmod T^{n-h} \\ \chi_1,\chi_2\neq\chi_0 \text{ even}}} \sum_{m_1,m_2=0}^{n} \alpha(T^{n-m_1})\overline{\alpha(T^{n-m_2})}\mathcal{M}(m_1;\alpha\chi_1)\overline{\mathcal{M}(m_2;\alpha\chi_2)}$$

$$\cdot \frac{1}{q^{n-h-1}} \sum_{B \in \mathcal{M}_{n-h-1}} \overline{\chi_1}(\theta_{n-h-1}(B))\chi_2(\theta_{n-h-1}(B)).$$

Now as B ranges over $\mathcal{M}_{n-h-1}$, $\theta_{n-h-1}(B)$ ranges over all polynomials $\{g \in \mathcal{P}_{\leq n-h-1} \mid g(0) = 1\}$. Hence for this last factor, we see that

$$\begin{aligned}\sum_{B \in \mathcal{M}_{n-h-1}} \overline{\chi_1}(\theta_{n-h-1}(B))\chi_2(\theta_{n-h-1}(B)) &= \sum_{\substack{C \in \mathcal{P}_{\leq n-h-1} \\ C(0)=1}} \overline{\chi_1}(C)\chi_2(C) \\ &= \sum_{\substack{C \bmod T^{n-h} \\ C(0)=1}} \overline{\chi_1}(C)\chi_2(C) \\ &\overset{\text{(lemma 5.4)}}{=} \delta_{\chi_1\chi_2} q^{n-h-1}.\end{aligned}$$

Substituting this in the expression above, we find the statement of the theorem. □

6.3 Calculating the variance of $\sum \Lambda(f)$

In this section we look at the von Mangoldt function, as Keating and Rudnick do in [20]. Recall that it is given by

$$\begin{aligned}&\Lambda : \mathbb{F}_q[T] \to \mathbb{C}, \\ &\quad \Lambda(f) = \begin{cases} \deg(P) & \text{if } f = cP^k, \\ 0 & \text{otherwise.} \end{cases}\end{aligned}$$

Lemma 6.11. *Let $f \in \mathbb{F}_q[T]$, coprime to T^m. Then $\Lambda(f) = \Lambda(f^*)$.*

Proof. First suppose that $f = cP^k$ with $c \in \mathbb{F}_q^\times$ and P irreducible. Then $P(0) \neq 0$, as $f(0) \neq 0$, so P^* is irreducible by corollary 6.3. As $f^* = c(P^*)^k$ by lemma 6.2, we conclude that $\Lambda(f) = \Lambda(f^*)$. Furthermore if $f = gh$ with g, h coprime and $\deg(g), \deg(h) > 0$, then $g(0) \neq 0$, $h(0) \neq 0$, so $f^* = g^*h^*$ with g^*, h^* coprime and $\deg(g^*), \deg(h^*) > 0$. In that case $\Lambda(f^*) = 0 = \Lambda(f)$. □

If you want to calculate the variance of $\sum \alpha$ for α hoving property $(*)$, such as the Möbius function μ, the divisor function σ_0 or the Euler totient function φ, this is just a matter of applying theorem 6.10 and estimating $\mathcal{M}(m; \alpha\chi)$. This last bit can be quite challenging, as you can see in the next chapter. The von Mangoldt function Λ however does not have property $(*)$. It is even, but it is not multiplicative. We first need to work around this problem. Fortunately, an even stronger property holds. If f, g are coprime, then $\Lambda(fg) = 0$. In the proof of lemma 6.9, we used the multiplicativity at two points. First to split the sum over f into a sum over T^{n-m} and over a sum f_1 with $f_1(0) \neq 0$. Later we used multiplicativity to put this sum back together. With the van Mangoldt function, we don't need to split these sums, as $f = T^{n-m}f_1$ with $f_1(0) \neq 0$, $m < n$, implies that $\Lambda(f) = 0$, except when $f = T^n$. This case is however negligible. Note that by theorem 5.1 and lemma 6.7 we have

$$\langle \mathcal{N}_\Lambda(\bullet; h) \rangle_n = q^{h+1}.$$

Now $\Lambda(T^n) = 1$ has no influence on the variance, as $q \to \infty$. It is easier for us to discard it right away and consider

$$\mathcal{N}^*_\Lambda(A; h) = \sum_{\substack{f \in I(A;h) \\ f(0) \neq 0}} \Lambda(f).$$

This enables us to reason in exactly the same way as in lemma 6.9 and theorem 6.10 to see that

$$\mathcal{N}^*_\Lambda(T^{h+1}B;h) = \sum_{\substack{f_1\in\mathcal{M}_n\\ f_1(0)\neq 0\\ f_1\in I(T^{h+1}B;h)}} \Lambda(f_1)$$

for any $B\in\mathcal{M}_{n-h-1}$, $0\leq h\leq n-2$. We also find

$$\mathrm{Var}_n\mathcal{N}^*_\Lambda(\bullet;h) = \frac{1}{\varphi_{\mathrm{ev}}(T^{n-h})^2}\sum_{\substack{\chi \bmod T^{n-h}\\ \chi\neq\chi_0 \text{ even}}} \left|\mathcal{M}(n;\Lambda\chi)\right|^2.$$

We now see that it is useful to have some estimations on $\mathcal{M}(n;\Lambda\chi)$. In general after applying theorem 6.10 for some α, you want some estimates on $\mathcal{M}(m;\alpha\chi)$ for all $0\leq m\leq n$. You get these estimations by expressing $\mathcal{M}(m;\alpha\chi)$ as a function of the zeroes of $\mathcal{L}(u,\chi)$ and applying the Riemann hypothesis for finite fields (from which it follows that the zeroes have absolute value at most $q^{\frac{1}{2}}$). In this case we have already done that, by deriving relation (5.4). There are q^{n-h-1} even characters modulo T^{n-h}. Of these q^{n-h-1} even characters only $O(q^{n-h-2})$ are non-primitive (as $q\to\infty$). By the Riemann Hypothesis $\mathcal{M}(n;\Lambda\chi)=O(nq^{n/2})$, so the non-primitive terms only contribute a factor $\frac{1}{q^{2(n-h-1)}}O(q^{n-h-2})O(n^2q^n)=O(n^2q^h)$ to the equation above. Furthermore for primitive, even characters we derived that $\mathcal{M}(n;\Lambda\chi)=-q^{\frac{n}{2}}\mathrm{Tr}(\Theta^n_\chi)-\lambda_\chi$. Hence for these characters

$$\left|\mathcal{M}(n;\Lambda\chi)\right|^2 = q^n\left|\mathrm{Tr}(\Theta^n_\chi)\right|^2 + O\left((n-h)q^{\frac{n}{2}}\right).$$

It then follows that

$$\mathrm{Var}_n\mathcal{N}^*_\Lambda(\bullet;h) = q^{h+1}\left(\frac{1}{q^{n-h-1}}\sum_{\substack{\chi \bmod T^{n-h}\\ \chi \text{ even, primitive}}}\left|\mathrm{Tr}(\Theta^n_\chi)\right|^2 + O\left(\frac{n^2}{q^{\frac{1}{2}}}+\frac{n^2}{q}+\frac{n-h}{q^{n/2}}\right)\right),$$

where the implied constant depends only on n.
Finally proving theorem 1.1 is now rather straight forward.

Theorem 6.12 (Restatement of theorem 1.1)**.**
Let $0<h<n-3$. As $q\to\infty$, we have that

$$\mathrm{Var}_n\mathcal{N}_\Lambda(\bullet;h)\sim q^{h+1}(n-h-2).$$

Proof. There are $q^{n-h-1}(1-\frac{1}{q})$ even primitive characters modulo T^{n-h}, so the expectation value of $\left|\mathrm{Tr}(\Theta^n_\chi)\right|^2$ is given by

$$\mathop{\mathbb{E}}_{\substack{\chi \bmod T^{n-h}\\ \chi \text{ even, primitive}}}\left[\left|\mathrm{Tr}(\Theta^n_\chi)\right|^2\right] = \frac{1}{q^{n-h-1}(1-\frac{1}{q})}\sum_{\substack{\chi \bmod T^{n-h}\\ \chi \text{ even, primitive}}}\left|\mathrm{Tr}(\Theta^n_\chi)\right|^2.$$

By corollary 5.11, we know that

$$\mathop{\mathbb{E}}_{\substack{\chi \bmod T^{n-h}\\ \chi \text{ even, primitive}}}\left[\left|\mathrm{Tr}(\Theta^n_\chi)\right|^2\right] = \min(n,n-h-2) = n-h-2.$$

We conclude that as $q \to \infty$

$$\begin{aligned}\mathrm{Var}_n \mathcal{N}_\Lambda^*(\bullet; h) &= q^{h+1}\left(\frac{1}{q^{n-h-1}} \sum_{\substack{\chi \bmod T^{n-h} \\ \chi \text{ even, primitive}}} \left|\mathrm{Tr}(\Theta_\chi^n)\right|^2 + O\left(\frac{n^2}{q^{\frac{1}{2}}} + \frac{n^2}{q} + \frac{n-h}{q^{n/2}}\right)\right) \\ &= q^{h+1}\left(\left(1-\frac{1}{q}\right) \mathop{\mathbb{E}}_{\substack{\chi \bmod T^{n-h} \\ \chi \text{ even, primitive}}} \left[\left|\mathrm{Tr}(\Theta_\chi^n)\right|^2\right] + O\left(\frac{n^2}{q^{\frac{1}{2}}} + \frac{n^2}{q} + \frac{n-h}{q^{n/2}}\right)\right) \\ &\sim q^{h+1}(n-h-2)\end{aligned}$$

As the term $f = T^N$ in $\sum_{f \in I(A;h)} \Lambda(f)$ was negligible, we also conclude that

$$\mathrm{Var}_n \mathcal{N}_\Lambda(\bullet; h) \sim q^{h+1}(n-h-2).$$

□

Chapter 7

The Euler totient function over $\mathbb{F}_q[T]$

7.1 The average of $\frac{\varphi(f)}{|f|}$

Now we will look at the Euler function over $\mathbb{F}_q[T]$. In the first section we will look at the avarage of $\frac{\varphi(f)}{|f|}$. You would expect to find analogous results to those in section 2.6. It turns out this is indeed the case. In the second section of this chapter we calculate the variance of the Euler totient function in short intervals, applying theorem 6.10 to $\alpha = \varphi$.

Definition 7.1. *Define*

$$\begin{aligned}\varphi : \mathbb{F}_q[T] &\to \mathbb{Z}, \\ f &\mapsto \#\left(\mathbb{F}_q[T]/(f)\right)^\times.\end{aligned}$$

That is, $\varphi(f)$ is the number of polynomials with degree strictly smaller than the degree of f, coprime to f. We first note a few properties of φ, which are actually the exact analogues of the properties in lemma 2.33.

Lemma 7.2. *For all $f \in \mathbb{F}_q[T]$, $f \neq 0$, the following statements hold:*

1. $|f| = \sum_{\substack{g \mid f, \\ g \text{ monic}}} \varphi(g)$.
2. $\varphi = I_1 * \mu$, *where I_1 is the arithmetic function given by $I_1(f) = |f|$.*
3. $$\varphi(f) = |f| \prod_{\substack{P \mid f \\ \text{monic,} \\ \text{irreducible}}} \left(1 - \frac{1}{|P|}\right).$$

Proof. The proof is exactly the same as that of lemma 2.33.

1. Let g be a monic divisor of f. Denote

$$V_g = \{h \in \mathbb{F}_q[T] |\ \deg(h) < \deg(f),\ \gcd(f,h) = g\}.$$

By dividing every element of V_g by g, you see that $\#V_g = \varphi\left(\frac{f}{g}\right)$. Now for every h in $(\mathbb{F}_q[T]/(f))$, there is exactly one monic g, such that $h \in V_g$, so $(\mathbb{F}_q[T]/(f)) = \cup_{\substack{g|f \\ g \text{ monic}}} V_g$, (as sets). Hence

$$|f| = \#\left(\mathbb{F}_q[T]/(f)\right) = \sum_{\substack{g|f \\ g \text{ monic}}} \#V_g = \sum_{\substack{g|f \\ g \text{ monic}}} \varphi\left(\frac{f}{g}\right) = \sum_{\substack{g|f, \\ g \text{ monic}}} \varphi(g).$$

2. We know that $I_1 = \varphi * E$, so using Möbius inversion for $\mathbb{F}_q[T]$, we see that $\varphi = I_1 * \mu$.

3. By 2. φ is a multiplicative function, so it is sufficient to calculate $\varphi(P^k)$ for all monic irreducibles P. Now

$$\varphi(P^k) = \sum_{l=0}^{k} \mu\left(P^l\right) I_1\left(P^{k-l}\right) = \left|P^k\right| - \left|P^{k-1}\right| = |P|^k - |P|^{k-1} = \left|P^k\right|\left(1 - \frac{1}{|P|}\right).$$

The statement follows.

□

Note that $\varphi(f) = |f| \prod_{\substack{P|f \\ \text{monic}, \\ \text{irreducible}}} \left(1 - \frac{1}{|P|}\right) = q^n \prod_{\substack{P|f \\ \text{monic}, \\ \text{irreducible}}} \left(1 - \frac{1}{|P|}\right)$, if f has degree n. We are of course interested in the average of $\frac{\varphi(f)}{q^n}$. Analogous to theorem 2.35, one might expect that this average will be $\frac{1}{\zeta_q(2)}$ in the large n-limit. It turns out an even stronger result holds.

Theorem 7.3. *For all $n \geq 1$, we have*

$$\frac{1}{q^n} \sum_{f \in \mathcal{M}_n} \frac{\varphi(f)}{q^n} = \frac{1}{\zeta_q(2)}.$$

Proof. We compute the generating function of $\sum_{f \in \mathcal{M}_n} \varphi(f)$:

$$\begin{aligned}
\sum_{n \geq 0} \left(\sum_{f \in \mathcal{M}_n} \varphi(f) \right) u^n &= \sum_{f \text{ monic}} \varphi(f) u^{\deg(f)} \\
&\overset{\text{(lemma 7.2.2)}}{=} \sum_{f \text{ monic}} \sum_{\substack{g|f, \\ g \text{ monic}}} |g| \, \mu\left(\frac{f}{g}\right) u^{\deg(f)} \\
&= \sum_{g \text{ monic}} \sum_{h \text{ monic}} q^{\deg(g)} \mu(h) u^{\deg(g)+\deg(h)} \\
&= \left(\sum_{g \text{ monic}} (qu)^{\deg(g)} \right) \left(\sum_{h \text{ monic}} \mu(h) u^{\deg(h)} \right) \\
&= \frac{\mathcal{Z}_q(qu)}{\mathcal{Z}_q(u)} \overset{(a)}{=} \frac{1 - qu}{1 - q^2 u} = (1 - qu) \sum_{n \geq 0} (q^2 u)^n \\
&= \sum_{n \geq 0} q^{2n} u^n - \sum_{n \geq 0} q^{2n+1} u^{n+1} \\
&= 1 + \sum_{n \geq 1} q^{2n} \left(1 - \frac{1}{q}\right) u^n.
\end{aligned}$$

Here we applied that $\mathcal{Z}(u) = \frac{1}{1-qu}$, which we have seen in the proof of theorem 5.1. We conclude that for $n \geq 1$, we have

$$\frac{1}{q^n} \sum_{f \in \mathcal{M}_n} \frac{\varphi(f)}{q^n} = \left(1 - \frac{1}{q}\right) = \frac{1}{\zeta_q(2)},$$

as $\zeta_q(s) = \frac{1}{1-q^{1-s}}$. □

Corollary 7.4. *Let $n \geq h \geq 1$. Then*

$$\langle \mathcal{N}_{\frac{\varphi(f)}{|f|}}(\bullet; h) \rangle_n = \frac{q^{h+1}}{\zeta_q(2)}.$$

Proof. This follows directly from theorem 7.3 and lemma 6.7. □

Note that this corollary is the direct analogue of theorem 2.36

7.2 The variance of $\sum \frac{\varphi(f)}{|f|}$ in short intervals

In this section we calculate the variance of $\sum \frac{\varphi(f)}{|f|}$ in short intervals. First note that it is sufficient to calculate the variance of $\mathcal{N}_\varphi$ in short intervals, as we have that

$$\begin{aligned}
\mathrm{Var}_n \mathcal{N}_{\frac{\varphi(f)}{|f|}}(\bullet; h) &= \frac{1}{q^n} \sum_{A \in \mathcal{M}_n} \left| \mathcal{N}_{\frac{\varphi(f)}{|f|}}(A; h) - \langle \mathcal{N}_{\frac{\varphi(f)}{|f|}}(\bullet; h) \rangle_n \right|^2 \\
&= \frac{1}{q^n} \sum_{A \in \mathcal{M}_n} \left| \sum_{f \in I(A;h)} \frac{\varphi(f)}{q^n} - \frac{1}{q^n} \sum_{B \in \mathcal{M}_n} \sum_{f \in I(B;h)} \frac{\varphi(f)}{q^n} \right|^2 \\
&= \frac{1}{q^{3n}} \sum_{A \in \mathcal{M}_n} \left| \sum_{f \in I(A;h)} \varphi(f) - \frac{1}{q^n} \sum_{B \in \mathcal{M}_n} \sum_{f \in I(B;h)} \varphi(f) \right|^2 \\
&= \frac{1}{q^{2n}} \mathrm{Var}_n \mathcal{N}_\varphi(\bullet; h)
\end{aligned}$$

We will hence calculate $\mathrm{Var}_n \mathcal{N}_{\frac{\varphi(f)}{|f|}}(\bullet; h)$ by calculating $\mathrm{Var}_n \mathcal{N}_\varphi(\bullet; h)$. We do this by applying theorem 6.10 to $\alpha = \varphi$. We first prove that φ has property $(*)$.

Lemma 7.5. *Let $f \in \mathbb{F}_q[T]$, s.t. $f(0) \neq 0$. Then $\varphi(f) = \varphi(f^*)$.*

Proof. Fix $f \in \mathbb{F}_q[T]$ with $f(0) \neq 0$ and suppose that f factors into irreducible polynomials as $f = P_1^{k_1} \dots P_n^{k_n}$. By lemma 6.2 $f^* = (P_1^*)^{k_1} \dots (P_n^*)^{k_n}$. Since $f(0) \neq 0$, we know that $P_i(0) \neq 0$ for all i. Hence by corollary 6.3 we know that the P_i^* are again irreducible. Furthermore we know that $\deg(f) = \deg(f^*)$, as $f(0) \neq 0$. Applying lemma 7.2.3, we conclude that

$$\varphi(f^*) = |f^*| \prod_{\substack{P | f^* \\ \text{monic,} \\ \text{irreducible}}} \left(1 - \frac{1}{|P|}\right) = |f^*| \prod_{i=1}^{n} \left(1 - \frac{1}{|P_i^*|}\right) = |f| \prod_{i=1}^{n} \left(1 - \frac{1}{|P_i|}\right) = \varphi(f).$$

□

In the proof of lemma 7.2 we saw that φ is multiplicative and from the definition it follows that φ is even, so we conclude that φ has property $(*)$. Applying theorem 6.10 and using that $\varphi(T^m) = (q-1)q^{m-1}$, we see that for $0 \leq h \leq n-2$

$$\mathrm{Var}_n \mathcal{N}_\varphi(\bullet; h) = \frac{1}{\varphi_{\mathrm{ev}}(T^{n-h})^2} \sum_{\substack{\chi \bmod T^{n-h} \\ \chi \neq \chi_0 \\ \text{even}}} \sum_{m_1, m_2 = 0}^{n} (q-1)^2 q^{2n-m_1-m_2-2} \mathcal{M}(m_1; \varphi\chi) \overline{\mathcal{M}(m_2; \varphi\chi)}. \tag{7.1}$$

We need to find a relation between $\mathcal{M}(m; \varphi\chi) = \sum_{f \in \mathcal{M}_m} \varphi(f)\chi(f)$ and the zeroes of $\mathcal{L}(u, \chi)$, before we can estimate the terms in the sum. Let χ be an even character. We will first assume that χ is also primitive and later adapt this proof in the case that χ is not primitive. Recall from section 5.3 that we have defined the L-function associated to a character χ by $\mathcal{L}(u, \chi) = \sum_{f \text{ monic}} \chi(f) u^{\deg(f)}$. Furthermore we have seen that for even characters $\mathcal{L}(u, \chi) = (1-u)\mathcal{L}^*(u, \chi)$. Here $\mathcal{L}^*(u, \chi)$ is the completed L-function of χ. It is a finite polynomial in u of degree at most $N = \deg Q - 2 = n - h - 2$. Moreover if χ is primitive it has exactly degree N. Hence we can write $\mathcal{L}^*(u, \chi) = \prod_{j=1}^{N} (1 - \alpha_j(\chi)u)$. By the Riemann Hypothesis, we know that $|\alpha_j(\chi)| = \sqrt{q}$ for each j. Therefore we have the (final) expression $\mathcal{L}^*(u, \chi) = \det(I_N - \sqrt{q}u\Theta_\chi)$, where $\Theta_\chi \in U(N)$ is the unitarized Frobenius matrix of χ. We have the following lemma.

Lemma 7.6. *Let χ be a primitive even character $\bmod T^{n-h}$. $\mathcal{M}(m; \varphi\chi)$ is a polynomial in $q^{\frac{1}{2}}$ of degree $3m$ if $m \leq N$. If $m \geq N+1$, then $\mathcal{M}(m; \varphi\chi)$ is a polynomial in $q^{\frac{1}{2}}$ of degree $m + 2N + 1$.*

Proof. We first compute the generating function of $\mathcal{M}(m; \varphi\chi)$:

$$\begin{aligned}
\sum_{m \geq 0} \mathcal{M}(m; \varphi\chi) u^m &= \sum_{m \geq 0} \left(\sum_{f \in \mathcal{M}_m} \varphi(f)\chi(f) \right) u^m \\
&= \sum_{f \text{ monic}} \varphi(f)\chi(f) u^{\deg(f)} \\
&\overset{\text{(lemma 7.2.2)}}{=} \sum_{f \text{ monic}} \sum_{\substack{g \mid f, \\ g \text{ monic}}} |g| \, \mu\left(\frac{f}{g}\right) \chi(g \cdot \frac{f}{g}) u^{\deg(f)} \\
&= \sum_{g \text{ monic}} \sum_{h \text{ monic}} q^{\deg(g)} \mu(h)\chi(g)\chi(h) u^{\deg(g)+\deg(h)} \\
&= \left(\sum_{g \text{ monic}} \chi(g)(qu)^{\deg(g)} \right) \left(\sum_{h \text{ monic}} \chi(h)\mu(h) u^{\deg(h)} \right) = \frac{\mathcal{L}(qu, \chi)}{\mathcal{L}(u, \chi)}.
\end{aligned}$$

Denote by $\lambda_j(M)$ the coefficient of x^j in the expression $\det(I_N - xM)$. In particular $\lambda_0(M) = 1$,

$\lambda_1(M) = -\text{Tr } M$ and $\lambda_N(M) = (-1)^N \det(M)$. We see that

$$\begin{aligned}\mathcal{L}(qu,\chi) &= (1-qu)\det(I_N - uq^{\frac{3}{2}}\Theta_\chi) = (1-qu)\sum_{j=0}^{N}\lambda_j(\Theta_\chi)q^{\frac{3j}{2}}u^j \\ &= \sum_{j=0}^{N}\lambda_j(\Theta_\chi)q^{\frac{3j}{2}}u^j - \sum_{j=0}^{N}\lambda_j(\Theta_\chi)q^{\frac{3j+2}{2}}u^{j+1} \\ &= \sum_{j=0}^{N}\lambda_j(\Theta_\chi)q^{\frac{3j}{2}}u^j - \sum_{j=1}^{N+1}\lambda_{j-1}(\Theta_\chi)q^{\frac{3j-1}{2}}u^j \\ &= \sum_{j=0}^{N+1}\left(\lambda_j(\Theta_\chi) - \lambda_{j-1}(\Theta_\chi)q^{-\frac{1}{2}}\right)q^{\frac{3j}{2}}u^j,\end{aligned}$$

where we have defined $\lambda_{-1}(\Theta_\chi) = 0 = \lambda_{N+1}(\Theta_\chi)$. To express $\frac{1}{\mathcal{L}(u,\chi)}$, we use corollary 4.8, which says that

$$\frac{1}{\det(I_N - xM)} = \sum_{k\geq 0}\text{Tr Sym}^k M\ x^k.$$

Then

$$\frac{1}{\mathcal{L}(u,\chi)} = \frac{1}{(1-u)\det(I_N - uq^{\frac{1}{2}}\Theta_\chi)} = \sum_{l\geq 0}u^l\sum_{k\geq 0}q^{\frac{k}{2}}\text{Tr Sym}^k\Theta_\chi\ u^k.$$

We hence find that

$$\begin{aligned}\sum_{m\geq 0}\mathcal{M}(m;\varphi\chi)u^m &= \sum_{l\geq 0}\sum_{j=0}^{N}\sum_{k\geq 0}\left(\lambda_j(\Theta_\chi) - \lambda_{j-1}(\Theta_\chi)q^{-\frac{1}{2}}\right)\text{Tr Sym}^k\Theta_\chi q^{\frac{3j+k}{2}}u^{j+k+l} \\ &= \sum_{m\geq 0}\sum_{\substack{j+k\leq m\\ 0\leq j\leq N+1\\ k\geq 0}}\left(\lambda_j(\Theta_\chi) - \lambda_{j-1}(\Theta_\chi)q^{-\frac{1}{2}}\right)\text{Tr Sym}^k\Theta_\chi q^{\frac{3j+k}{2}}u^m.\end{aligned}$$

Defining

$$A_{j,k}(\chi) := \lambda_j(\Theta_\chi)\text{Tr Sym}^k\Theta_\chi,$$

we conclude that

$$\mathcal{M}(m;\varphi\chi) = \sum_{\substack{j+k\leq m\\ 0\leq j\leq N+1\\ k\geq 0}} A_{j,k}(\chi)q^{\frac{3j+k}{2}} - A_{j-1,k}(\chi)q^{\frac{3j+k-1}{2}}. \tag{7.2}$$

To estimate this, we want to know the maximum q-dependency. That is, we want to know the term with the largest exponent, such that the corresponding coefficient is non-zero. Looking at the exponent $\frac{3j+k}{2}$, it is clear that this is maximal when you first choose j maximal, and then k maximal. When $m \leq N$, this means that $j = m$ and $k = 0$. Note that in this case $A_{j,k}(\chi) = A_{m,0}(\chi) = \lambda_m(\Theta_\chi)$. In general this is non-zero. When $m \geq N+1$, this means that $j = N+1$ and $k = m-(N+1)$. This however gives us a problem, as in this case $A_{j,k}(\chi) = A_{N+1,m-(N+1)}(\chi) = 0$, as $\lambda_{N+1}(\Theta_\chi) = 0$. The next possibility for a maximum q dependency is given by the same choice of j and k, when we look at the other term in the sum,

which is $A_{j-1,k}(\chi)q^{\frac{3j+k-1}{2}}$. For this term $A_{j-1,k}(\chi) = A_{N,m-(N+1)}(\chi)$, which is not zero in general. Now $3j+k-1 = 3(N+1)+m-(N+1)-1 = m+2N+1$. We hence find that $\mathcal{M}(m;\varphi\chi)$ is a polynomial in $q^{\frac{1}{2}}$ of degree $3m$ if $m \leq N$. If $m \geq N+1$, then $\mathcal{M}(m;\varphi\chi)$ is a polynomial in $q^{\frac{1}{2}}$ of degree $m+2N+1$. □

We now want substitute this into relation (7.1), to estimate the maximum q-dependency of the variance, but this is not as easy as it looks, as it turns out there is a lot of cancellation of terms. Let us first consider $\mathcal{M}(m,\varphi\chi)$. By lemma 7.6, we know that we can write this as follows.

Definition 7.7. *For $m \leq N$, define the coefficients $x_l(m)$ with $0 \leq l \leq 3m$, s.t.*

$$\mathcal{M}(m,\varphi\chi) = \sum_{0\leq l\leq 3m} x_l(m)q^{\frac{3m-l}{2}} = x_0(m)q^{\frac{3m}{2}} + x_1(m)q^{\frac{3m-1}{2}} + \cdots + x_{3m}(m)$$

For $m \geq N+1$, define the coefficients $y_l(m)$ with $0 \leq l \leq m+2N+1$, s.t.

$$\mathcal{M}(m,\varphi\chi) = \sum_{0\leq l\leq m+2N+1} y_l(m)q^{\frac{m+2N+1-l}{2}} = y_0(m)q^{\frac{m+2N+1}{2}} + y_1(m)q^{\frac{m+2N}{2}} + \cdots + y_{m+2N+1}(m)$$

Furthermore we generalize our definition of $A_{j,k}(\chi)$, for some notational convenience.

Definition 7.8. *Define*

$$A_{j,k}(\chi) = \begin{cases} \lambda_j(\Theta_\chi)\text{Tr Sym}^k\Theta_\chi & \text{if } 0 \leq j \leq N, 0 \leq k \\ 0 & \text{otherwise} \end{cases}$$

We will often abbreviate this to $A_{j,k}$. Just remember that $A_{j,k}$ still depends on χ. From relation (7.2) it follows that for $m \leq N$

$$x_l(m) = \sum_{\substack{3j+k=3m-l \\ j+k\leq m}} A_{j,k} - \sum_{\substack{3j+k-1=3m-l \\ j+k\leq m}} A_{j-1,k} = \sum_{\substack{3j+k=3m-l \\ j+k\leq m}} A_{j,k} - \sum_{\substack{3j+k=3m-l-2 \\ j+k\leq m-1}} A_{j,k}. \quad (7.3)$$

Using the same reasoning for $m \geq N+1$, we also see that

$$y_l(m) = \sum_{\substack{3j+k=m+2N+1-l \\ j+k\leq m}} A_{j,k} - \sum_{\substack{3j+k=m+2N-1-l \\ j+k\leq m-1}} A_{j,k}. \quad (7.4)$$

Notice that these are finite sums by our definition of $A_{j,k}$, so we don't need to worry about convergence. We now prove some useful lemmas about the coefficients x_l and y_l.

Lemma 7.9. *Fix $m \leq N$ and $l \leq 3m$, and suppose $L \leq \min(m,l)$. Then*

$$\sum_{i=0}^{L} x_{l-i}(m-i) = \sum_{\substack{3j+k=3m-l \\ j+k\leq m}} A_{j,k} - \sum_{\substack{3j+k=3m-l-2-2L \\ j+k\leq m-1-L}} A_{j,k}.$$

Furthermore if $L = \min(l,m)$, then this last term is zero, so

$$\sum_{i=0}^{L} x_{l-i}(m-i) = \sum_{\substack{3j+k=3m-l \\ j+k\leq m}} A_{j,k}.$$

Proof. The first equality follows directly from (7.3).

$$\begin{aligned}
\sum_{i=0}^{L} x_{l-i}(m-i) &= \sum_{i=0}^{L} \left(\sum_{\substack{3j+k=3m-l-2i \\ j+k\le m-i}} A_{j,k} - \sum_{\substack{3j+k=3m-l-2-2i \\ j+k\le m-1-i}} A_{j,k} \right) \\
&= \sum_{i=0}^{L} \sum_{\substack{3j+k=3m-l-2i \\ j+k\le m-i}} A_{j,k} - \sum_{i=0}^{L} \sum_{\substack{3j+k=3m-l-2-2i \\ j+k\le m-1-i}} A_{j,k} \\
&= \sum_{i=0}^{L} \sum_{\substack{3j+k=3m-l-2i \\ j+k\le m-i}} A_{j,k} - \sum_{i=1}^{L+1} \sum_{\substack{3j+k=3m-l-2i \\ j+k\le m-i}} A_{j,k} \\
&= \sum_{\substack{3j+k=3m-l \\ j+k\le m}} A_{j,k} - \sum_{\substack{3j+k=3m-l-2-2L \\ j+k\le m-1-L}} A_{j,k}.
\end{aligned}$$

Furthermore if $L = \min(m, l)$, then either $m = L$ or $l = L$. In the first case $m - L - 1 = -1$, so if $j + k \le m - L - 1$, then one of j, k is negative and $A_{j,k} = 0$. In the last case the last summation becomes

$$\sum_{\substack{3j+k=3m-3l-2 \\ j+k\le m-l-1}} A_{j,k}.$$

Now $j + k \le m - l - 1$ implies that j is at most $m - l - 1$ or k would be negative. Now the largest $3j + k$ can become for non-negative k and j is $3m - 3l - 3$. Hence this summation is also zero. □

For y we prove almost the same cancellation.

Lemma 7.10. *Fix $m \ge N + 1$ and $l \le m + 2N + 1$, and suppose $L \le l$. Then*

$$\sum_{i=0}^{L} y_{l-i}(m+i) = - \sum_{\substack{3j+k=m+2N-1-l \\ j+k\le m-1}} A_{j,k} + \sum_{\substack{3j+k=m+2N+1-l+2L \\ j+k\le m+L}} A_{j,k}.$$

Furthermore if $L = l$, then this last term is zero, so

$$\sum_{i=0}^{l} y_{l-i}(m+i) = - \sum_{\substack{3j+k=m+2N-1-l \\ j+k\le m-1}} A_{j,k}.$$

Proof. The proof is the same as in 7.9. By (7.4), we have

$$\begin{aligned}
\sum_{i=0}^{L} y_{l-i}(m+i) &= \sum_{i=0}^{L} \left(\sum_{\substack{3j+k=m+2N+1-l+2i \\ j+k\leq m+i}} A_{j,k} - \sum_{\substack{3j+k=m+2N-1-l+2i \\ j+k\leq m+i-1}} A_{j,k} \right) \\
&= \sum_{i=0}^{L} \sum_{\substack{3j+k=m+2N+1-l+2i \\ j+k\leq m+i}} A_{j,k} - \sum_{i=0}^{L} \sum_{\substack{3j+k=m+2N-1-l+2i \\ j+k\leq m+i-1}} A_{j,k} \\
&= \sum_{i=1}^{L+1} \sum_{\substack{3j+k=m+2N-1-l+2i \\ j+k\leq m+i-1}} A_{j,k} - \sum_{i=0}^{L} \sum_{\substack{3j+k=m+2N-1-l+2i \\ j+k\leq m+i-1}} A_{j,k} \\
&= - \sum_{\substack{3j+k=m+2N-1-l \\ j+k\leq m-1}} A_{j,k} + \sum_{\substack{3j+k=m+2N+1-l+2L \\ j+k\leq m+L}} A_{j,k}.
\end{aligned}$$

Furthermore if $L = l$, then the last summation becomes

$$\sum_{\substack{3j+k=m+2N+1+l \\ j+k\leq m+l}} A_{j,k}.$$

Since $j \leq N$, (or else $A_{j,k} = 0$), the largest $3j + k$ occurs when $j = N$ and $k = m + l - N$. In this case $3j + k = 3N + m + l - N = 2N + m + l \leq m + 2N + 1 + l$. Hence this summation is zero. □

Now we can start to estimate the terms in relation (7.1).

Theorem 7.11. *Fix n and $0 \leq h \leq n-2$. For any primitive even character χ modulo T^{n-h}*

$$\sum_{m=0}^{N} q^{-m} \mathcal{M}(m, \varphi\chi) = \sum_{\alpha=-2N}^{N} \left(\sum_{\substack{3j+k=2N+\alpha \\ j+k\leq N}} A_{j,k} \right) q^{\frac{\alpha}{2}}.$$

Proof. As we can use lemma 7.9, we only need to rewrite the sums. Note that m ranges from 0

to N, so we know that

$$\begin{aligned}
\sum_{m=0}^{N} q^{-m}\mathcal{M}(m,\varphi\chi) &= \sum_{m=0}^{N}\sum_{l=0}^{3m} x_l(m) q^{\frac{m-l}{2}} \\
&\overset{(a)}{=} \sum_{m=0}^{N}\sum_{\alpha=-2m}^{m} x_{m-\alpha}(m) q^{\frac{\alpha}{2}} \\
&\overset{(b)}{=} \sum_{\alpha=0}^{N}\sum_{m=\alpha}^{N} x_{m-\alpha}(m) q^{\frac{\alpha}{2}} + \sum_{\alpha=-2N}^{-1}\sum_{m=-\frac{\alpha}{2}}^{N} x_{m-\alpha}(m) q^{\frac{\alpha}{2}} \\
&\overset{(c)}{=} \sum_{\alpha=0}^{N}\sum_{i=0}^{N-\alpha} x_{N-\alpha-i}(N-i) q^{\frac{\alpha}{2}} + \sum_{\alpha=-2N}^{-1}\sum_{i=0}^{N+\frac{\alpha}{2}} x_{N-\alpha-i}(N-i) q^{\frac{\alpha}{2}} \\
&\overset{(d)}{=} \sum_{\alpha=0}^{N}\left(\sum_{\substack{3j+k=2N+\alpha \\ j+k\le N}} A_{j,k}\right) q^{\frac{\alpha}{2}} + \sum_{\alpha=-2N}^{-1}\left(\sum_{\substack{3j+k=2N+\alpha \\ j+k\le N}} A_{j,k} - \sum_{\substack{3j+k=-2 \\ j+k\le \frac{\alpha}{2}-1}} A_{j,k}\right) q^{\frac{\alpha}{2}} \\
&\overset{(e)}{=} \sum_{\alpha=-2N}^{N}\left(\sum_{\substack{3j+k=2N+\alpha \\ j+k\le N}} A_{j,k}\right) q^{\frac{\alpha}{2}}.
\end{aligned}$$

Here at (a) we used the transformation $\alpha = m - l$. Next at (b) we interchanged the order of summation and at (c) we used the transformation $i = N - m$. Finally at (d) we applied lemma 7.9 with $l = N - \alpha$, $m = N$ and $L = N - \alpha$, $L = N + \frac{\alpha}{2}$ respectively. At (e), we used that the sum $\sum_{\substack{3j+k=-2 \\ j+k\le \frac{\alpha}{2}-1}} A_{j,k}$ is trivially zero. □

Theorem 7.12. *Fix n and $0 \le h \le n-2$. For any primitive even character χ modulo T^{n-h}*

$$\sum_{m=N+1}^{n} q^{-m}\mathcal{M}(m,\varphi\chi) = -\sum_{\alpha=-2N}^{N}\left(\sum_{\substack{3j+k=2N+\alpha \\ j+k\le N}} A_{j,k}\right) q^{\frac{\alpha}{2}} + \sum_{\alpha=-2n}^{2N-n}\left(\sum_{\substack{3j+k=2n+\alpha \\ j+k\le n}} A_{j,k}\right) q^{\frac{\alpha}{2}}.$$

Proof. This time we apply lemma 7.10. so we know that

$$\begin{aligned}
\sum_{m=N+1}^{n} q^{-m}\mathcal{M}(m,\varphi\chi) &= \sum_{m=N+1}^{n}\sum_{l=0}^{m+2N+1} y_l(m)q^{\frac{-m+2N+1-l}{2}} \\
&\overset{(a)}{=} \sum_{m=N+1}^{n}\sum_{\alpha=-2m}^{2N+1-m} y_{2N+1-m-\alpha}(m)q^{\frac{\alpha}{2}} \\
&\overset{(b)}{=} \sum_{\alpha=-2n}^{N}\sum_{m=\max(N+1,-\frac{\alpha}{2})}^{\min(n,2N+1-\alpha)} y_{2N+1-m-\alpha}(m)q^{\frac{\alpha}{2}} \\
&= \sum_{\alpha=2N+1-n}^{N}\sum_{m=N+1}^{2N+1-\alpha} y_{2N+1-m-\alpha}(m)q^{\frac{\alpha}{2}} \\
&\quad + \sum_{\alpha=-2N-2}^{2N-n}\sum_{m=N+1}^{n} y_{2N+1-m-\alpha}(m)q^{\frac{\alpha}{2}} \\
&\quad + \sum_{\alpha=-2n}^{-2N-3}\sum_{m=\frac{\alpha}{2}}^{n} y_{2N+1-m-\alpha}(m)q^{\frac{\alpha}{2}} \\
&\overset{(c)}{=} \sum_{\alpha=2N+1-n}^{N}\sum_{i=0}^{N-\alpha} y_{N-\alpha-i}(N+1+i)q^{\frac{\alpha}{2}} \\
&\quad + \sum_{\alpha=-2N-2}^{2N-n}\sum_{i=0}^{n-(N+1)} y_{N-\alpha-i}(N+1+i)q^{\frac{\alpha}{2}} \\
&\quad + \sum_{\alpha=-2n}^{-2N-3}\sum_{i=0}^{n+\frac{\alpha}{2}} y_{2N+1-\frac{\alpha}{2}-i}(-\frac{\alpha}{2}-i)q^{\frac{\alpha}{2}} \\
&\overset{(d)}{=} \sum_{\alpha=2N+1-n}^{N}\left(-\sum_{\substack{3j+k=2N+\alpha\\ j+k\le N}} A_{j,k}\right)q^{\frac{\alpha}{2}} \\
&\quad + \sum_{\alpha=-2N-2}^{2N-n}\left(-\sum_{\substack{3j+k=2N+\alpha\\ j+k\le N}} A_{j,k} + \sum_{\substack{3j+k=2n+\alpha\\ j+k\le n}} A_{j,k}\right)q^{\frac{\alpha}{2}} \\
&\quad + \sum_{\alpha=-2n}^{-2N-3}\left(-\sum_{\substack{3j+k=-2\\ j+k\le -\frac{\alpha}{2}-1}} A_{j,k} + \sum_{\substack{3j+k=2n+\alpha\\ j+k\le n}} A_{j,k}\right)q^{\frac{\alpha}{2}} \\
&\overset{(e)}{=} -\sum_{\alpha=-2N}^{N}\left(\sum_{\substack{3j+k=2N+\alpha\\ j+k\le N}} A_{j,k}\right)q^{\frac{\alpha}{2}} + \sum_{\alpha=-2n}^{2N-n}\left(\sum_{\substack{3j+k=2n+\alpha\\ j+k\le n}} A_{j,k}\right)q^{\frac{\alpha}{2}}.
\end{aligned}$$

Here at (a) we used the transformation $\alpha = 2N+1-m-l$. Next at (b) we interchanged the order of summation and at (c) we used the transformation $i = m-(N+1)$ at the first two sums

and the transformation $i = m + \frac{\alpha}{2}$ at the last sum. Next at (d) we applied lemma 7.10 three times. On the first sum with $l = N - \alpha$, $m = N + 1$ and $L = N - \alpha$. On the second sum with $l = N-\alpha$, $m = N+1$ and $L = n-(N+1)$. On the third sum with $l = 2N+1-\frac{\alpha}{2}$, $m = -\frac{\alpha}{2}$ and $L = n + \frac{\alpha}{2}$. Finally at (e) we rearranged the terms and used that $\sum_{\substack{3j+k=-2 \\ j+k\leq -\frac{\alpha}{2}-1}} A_{j,k}$ is trivially zero. Also at (e) we applied the fact that $\sum_{\substack{3j+k=2N+\alpha \\ j+k\leq N}} A_{j,k}$ is trivially zero for $\alpha < -2N$. □

Adding theorem 7.11 and theorem 7.12 together, we find that for primitive even characters χ

$$\sum_{m=0}^{n} q^{-m}\mathcal{M}(m,\varphi\chi) = \sum_{\alpha=-2n}^{2N-n} \left(\sum_{\substack{3j+k=2n+\alpha \\ j+k\leq n}} A_{j,k} \right) q^{\frac{\alpha}{2}}.$$

Looking at the term $\alpha = 2N - n = n - 2h - 4$, we see that this has coefficient $\sum_{\substack{3j+k=n+2N \\ j+k\leq n}} A_{j,k}$. The only $0 \leq j \leq N$, $k \geq 0$ that solve this are $j = N, k = n - N = h + 2$. Hence we conclude that

Theorem 7.13. *Fix n and $0 \leq h \leq n-2$. For any primitive even character χ modulo T^{n-h}*

$$\sum_{m=0}^{n} q^{-m}\mathcal{M}(m,\varphi\chi) = A_{N,h+2}(\chi) q^{\frac{n-2h-4}{2}} + O\left(q^{\frac{n-2h-5}{2}}\right).$$

Before we substitute this into relation (7.1), we turn our attention to non-primitive characters. We claim that

Theorem 7.14. *Fix n and let $0 \leq h \leq n-2$. Suppose χ is a non-primitive even character modulo T^{n-h}. Then*

$$\sum_{m=0}^{n} q^{-m}\mathcal{M}(m,\varphi\chi) = A'_{N,h+2}(\chi) q^{\frac{n-2h-4}{2}} + O\left(q^{\frac{n-2h-5}{2}}\right)$$

for some $A'_{j,k}(\chi)$, defined below.

Proof. Note that to show the cancellation that occurred for primitive characters, we didn't actually use the definition of $A_{j,k}$, except that it is zero when $j, k < 0$ or $j > N$. Now we will show that for a non-primitive character χ, we also have a formula like relation (7.2), but for some different function $A'_{j,k}$. Since this function $A'_{j,k}$ will also have the property that it is zero when $j, k < 0$ or $j > N$, we can go through the exact same arguments, only with $A'_{j,k}$ instead of $A_{j,k}$, to see that for non-primitive characters this theorem holds. For non-primitive even characters χ modulo T^{n-h}, we know that we can write $\mathcal{L}(u,\chi) = (1-u)\prod_{j=1}^{N}(1-\alpha_j(\chi)u)$. Here $\alpha_j(\chi)$ are the inverse roots of χ. By the Riemann hypothesis, we know that either $\alpha_j(\chi) = 0$ or $|\alpha_j(\chi)| = q^{\frac{1}{2}}$. Now write $\alpha_j(\chi) = \beta_j(\chi)q^{\frac{1}{2}}$, (hence $\beta_j(\chi) = 0$ or $|\beta_j(\chi)| = 1$). Define $\lambda'_j(\chi)$ to be the coefficients of the polynomial of degree at most N in u, given by $\prod_{j=1}^{N}(1-\beta_j(\chi)u)$. Then

$\prod_{j=1}^{N}(1-\beta_j(\chi)u) = \sum_{j=0}^{N}\lambda'_j(\chi)u^j$. The proof of $\mathcal{M}(m,\varphi\chi) = \frac{\mathcal{L}(qu,\chi)}{\mathcal{L}(u,\chi)}$ still holds, so we find that

$$\begin{aligned}
\mathcal{M}(m,\varphi\chi) &= \frac{\mathcal{L}(qu,\chi)}{\mathcal{L}(u,\chi)} = \frac{(1-qu)\prod_{j=1}^{N}(1-\beta_j(\chi)q^{\frac{3}{2}}u)}{(1-u)\prod_{j=1}^{N}(1-\beta_j(\chi)q^{\frac{1}{2}}u)} \\
&= (1-qu)\left(\sum_{j=0}^{N}\lambda'_j(\chi)q^{\frac{3j}{2}}u^j\right)\left(\sum_{k_0\geq 0}u^{k_0}\right)\prod_{j=1}^{N}\left(\sum_{k_i\geq 0}\beta_j(\chi)^{k_i}q^{\frac{k_i}{2}}u^{k_i}\right) \\
&= \left(\sum_{j=0}^{N+1}\left(\lambda'_j(\chi)-\lambda'_{j-1}q^{-\frac{1}{2}}\right)q^{\frac{3j}{2}}u^j\right)\left(\sum_{k_0\geq 0}u^{k_0}\right)\left(\sum_{k\geq 0}\left(\sum_{k_1+\cdots+k_N=k}\prod_{i=1}^{N}\beta_i(\chi)^{k_i}\right)q^{\frac{k}{2}}u^k\right) \\
&= \left(\sum_{m\geq 0}\sum_{\substack{j+k\leq m\\ 0\leq j\leq N+1\\ k\geq 0}}\lambda'_j(\chi)\left(\sum_{k_1+\cdots+k_N=k}\prod_{i=1}^{N}\beta_i(\chi)^{k_i}\right)q^{\frac{3j+k}{2}}u^m\right) \\
&\quad -\left(\sum_{m\geq 0}\sum_{\substack{j+k\leq m\\ 0\leq j\leq N+1\\ k\geq 0}}\lambda'_{j-1}(\chi)\left(\sum_{k_1+\cdots+k_N=k}\prod_{i=1}^{N}\beta_i(\chi)^{k_i}\right)q^{\frac{3j+k-1}{2}}u^m\right).
\end{aligned}$$

Hence if we define

$$A'_{j,k}(\chi) = \begin{cases} \lambda'_j(\chi)\left(\sum_{k_1+\cdots+k_N=k}\prod_{i=1}^{N}\beta_i(\chi)^{k_i}\right) & \text{if } 0\leq j\leq N, 0\leq k, \\ 0 & \text{otherwise,}\end{cases}$$

we see that for non-primitive χ

$$\mathcal{M}(m;\varphi\chi) = \sum_{\substack{j+k\leq m\\ 0\leq j\leq N+1\\ k\geq 0}} A'_{j,k}(\chi)q^{\frac{3j+k}{2}} - A'_{j-1,k}(\chi)q^{\frac{3j+k-1}{2}}. \tag{7.5}$$

Comparing relation (7.5) to (7.2), we see that they are the same up to the coefficients $A_{j,k}$ and $A'_{j,k}$. One can now make exactly the same arguments we made to prove theorem 7.13, to prove theorem 7.14. □

Finally we substitute our results of theorems 7.13 and 7.14 into relation (7.1). This relation consists of a sum over characters, which we can split into a sum over primitive and a sum over non-primitive characters. By our work in section 5.2, we know that there are $O\left(\frac{\varphi(T^{n-h})}{q}\right) = O(q^{n-h-2})$ non-primitive characters modulo T^{n-h}. It follows that the sum over all non-primitive characters is estimated by

$$\frac{1}{\varphi_{\text{ev}}(T^{n-h})^2}O(q^{n-h-2})(q-1)^2q^{2n-2}O(q^{n-2h-4}) = O(q^{2n-h-4}).$$

Here we have also used that $\varphi_{\text{ev}}(T^{n-h}) = q^{n-h-1}$. We then find

$$\begin{aligned}\mathrm{Var}_n\mathcal{N}_\varphi(\bullet;h) &= \frac{1}{\varphi_{\mathrm{ev}}(T^{n-h})^2}\sum_{\substack{\chi \bmod T^{n-h}\\ \chi\neq\chi_0 \text{ even}}}\sum_{m_1,m_2=0}^{n}(q-1)^2q^{2n-m_1-m_2-2}\mathcal{M}(m_1;\varphi\chi)\overline{\mathcal{M}(m_2;\varphi\chi)}\\ &= \frac{1}{\varphi_{\mathrm{ev}}(T^{n-h})^2}\sum_{\substack{\chi \bmod T^{n-h}\\ \chi\neq\chi_0 \text{ even, primitive}}}\left(\left|A_{N,h+2}(\chi)\right|^2q^{3n-2h-4}+O\left(q^{3n-2h-\frac{9}{2}}\right)\right)+O(q^{2n-h-4})\\ &= \frac{\sum_{\substack{\chi \bmod T^{n-h}\\ \chi\neq\chi_0 \text{ even, primitive}}}\left|A_{N,h+2}(\chi)\right|^2}{q^{n-h-1}\left(1-\frac{1}{q}\right)}q^{2n-h-3}+O\left(q^{2n-h-\frac{7}{2}}\right)+O\left(q^{2n-h-4}\right)\\ &= \mathop{\mathbb{E}}_{\substack{\chi \bmod T^{n-h}\\ \chi \text{ even, primitive}}}\left[\left|A_{N,h+2}(\chi)\right|^2\right]q^{2n-h-3}+O\left(q^{2n-h-\frac{7}{2}}\right).\end{aligned}$$

Now let us consider $|A_{N,h+2}(\chi)|^2$. By definition it is given by

$$\left|A_{N,h+2}(\chi)\right|^2=\left|\lambda_N(\Theta_\chi)\right|^2\left|\mathrm{Tr}\ \mathrm{Sym}^{h+2}\Theta_\chi\right|^2=\left|(-1)^N\det(\Theta_\chi)\right|^2\left|\mathrm{Tr}\ \mathrm{Sym}^{h+2}\Theta_\chi\right|^2=\left|\mathrm{Tr}\ \mathrm{Sym}^{h+2}\Theta_\chi\right|^2.$$

By theorem 5.10, as $q\to\infty$,

$$\mathop{\mathbb{E}}_{\substack{\chi \bmod T^{n-h}\\ \chi \text{ even, primitive}}}\left[\left|\mathrm{Tr}\ \mathrm{Sym}^{h+2}\Theta_\chi\right|^2\right]=\int_{PU(N)}\left|\mathrm{Tr}\ \mathrm{Sym}^{h+2}U\right|^2dU=\int_{U(N)}\left|\mathrm{Tr}\ \mathrm{Sym}^{h+2}U\right|^2dU.$$

Finally by corollary 4.20, this last integral is equal to 1. We conclude that

Theorem 7.15. *Fix $0<h<n-3$. As $q\to\infty$,*

$$\mathrm{Var}_n\mathcal{N}_\varphi(\bullet;h)\sim q^{2n-h-3}.$$

Furthermore we conclude

Theorem 7.16 (Restatement of theorem 1.2)**.**
Fix $0<h<n-3$. As $q\to\infty$,

$$\mathrm{Var}_n\mathcal{N}_{\frac{\varphi(f)}{|f|}}(\bullet;h)\sim q^{-h-3}.$$

Chapter 8

The variance of the Euler totient function over $\mathbb{Z}$

8.1 Introduction

In this chapter we will revisit the Euler totient function over $\mathbb{Z}$ and look more closely at the variance in short intervals. One might expect, in analogue of theorem 1.2, that the averaged variance of $\sum_{n=x+1}^{x+H} \frac{\varphi(n)}{n}$ would be inversely proportional with the size of the interval H. It turns out this is not the case. Numerical simulations suggest that the averaged variance is not inversely proportional, but constant. See for example figure 8.1, which takes $H = x$. That is

$$\frac{1}{X}\sum_{x=1}^{X}\left[\left(\sum_{x<n<x+H}\frac{\varphi(n)}{n} - \frac{H}{\zeta(2)}\right)^2\right] \xrightarrow{X\to\infty} C, \tag{8.1}$$

for some constant C. Note that, even though it is not inversely proportional, relation (8.1) is still quite a strong result. We know by theorem 2.36 that $\sum_{n=x+1}^{x+H} \frac{\varphi(n)}{n}$ gets infinitely large, roughly of size $\frac{H}{\zeta(2)}$. Furthermore the error term we got was very bad, as it also becomes infinitely large, (just slower than $\frac{H}{\zeta(2)}$). Result (8.1) would suggest that this error term is actually finite, so that $\frac{H}{\zeta(2)}$ would be an almost perfect approximation of $\sum_{n=x+1}^{x+H} \frac{\varphi(n)}{n}$. Unfortunately, we cannot prove relation (8.1). We can get very close and even make the prediction that $C = \frac{1}{6\zeta(2)} - \frac{1}{6\zeta(2)^2}$, but we can only give a definitive proof under certain assumptions.

In section 8.4 we prove this for $H = x$. In section 8.6 we also look at short intervals, even though this is much more difficult. We can however make another assumption to make it easier to calculate the variance. It turns out that in this case the variance in this case also converges to the constant $\frac{1}{6\zeta(2)} - \frac{1}{6\zeta(2)^2}$.

Theorem 8.1 (Restatement of theorem 1.3)**.**
Let $H = \Theta(X^\delta)$, $0 < \delta \leq 1$. Assuming 8.21 and 8.24, we find

$$\frac{1}{X}\sum_{x=1}^{X}\left[\left(\sum_{x<n<x+H}\frac{\varphi(n)}{n} - \frac{H}{\zeta(2)}\right)^2\right] \xrightarrow{X\to\infty} \frac{1}{6\zeta(2)} - \frac{1}{6\zeta(2)^2}.$$

Later in this chapter we see what assumptions we actually need.

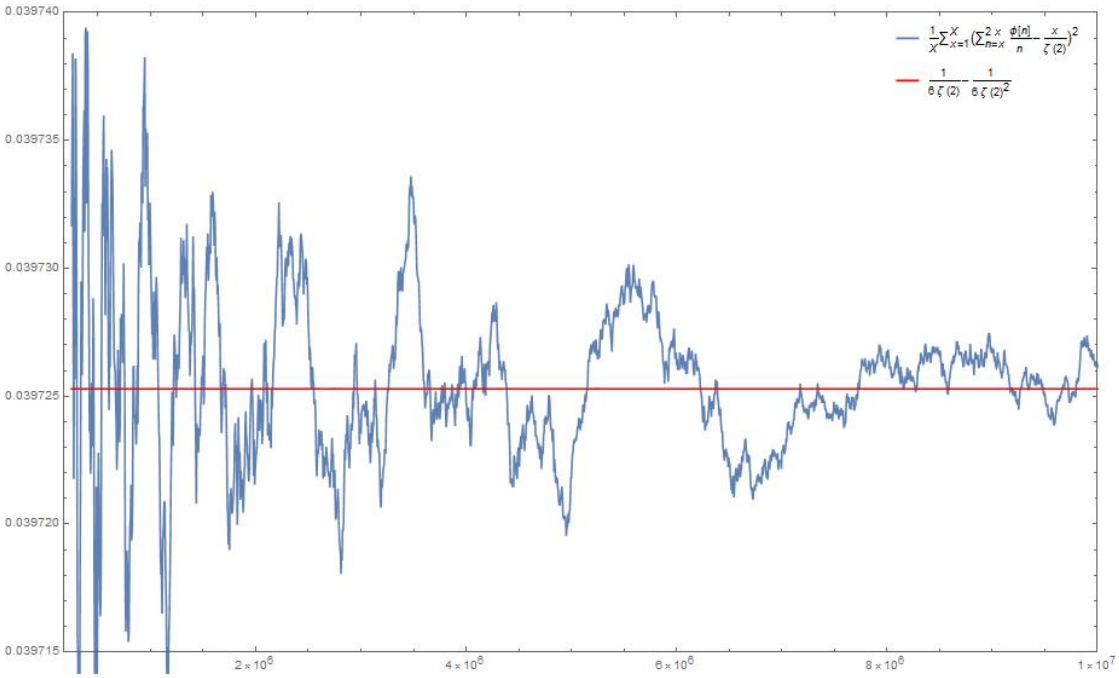

Figure 8.1: $\frac{1}{X}\sum_{x=1}^{X}\left(\sum_{n=x+1}^{2x}\frac{\varphi(n)}{n}-\frac{x}{\zeta(2)}\right)^2$ for $2.5\cdot 10^5 \le X \le 10^7$.

8.2 A new function $G(y)$

Recall the fractional part function, defined in the proof of corollary 2.17 by $\{x\} = x - [x]$. Define $G : \mathbb{R}_{\ge 1} \to \mathbb{R}$ by

$$G(y) := \sum_{n=1}^{\infty} \frac{\mu(n)}{n}\left\{\frac{y}{n}\right\}. \tag{8.2}$$

Note that this infinite sum converges for every $y \in \mathbb{R}_{\ge 1}$. This follows from the fact that $n > y$ implies $\left\{\frac{y}{n}\right\} = \frac{y}{n}$. Hence for some fixed y the sequence $\left(\sum_{n=1}^{N}\frac{\mu(n)}{n}\left\{\frac{y}{n}\right\}\right)$ is a Cauchy sequence, as for $n > y$ the terms become $\frac{y\mu(n)}{n^2}$. We prove two statements about $G(y)$.

Lemma 8.2. *Let m be a positive integer and let $0 < \epsilon < 1$. Then*

$$G(m+\epsilon) - G(m) = \frac{\epsilon}{\zeta(2)}.$$

Proof. Fix $m \in \mathbb{N}$ and $0 < \epsilon < 1$. For each n, we have that $\left[\frac{m}{n}\right] = \left[\frac{m+\epsilon}{n}\right]$, as otherwise there would be an integer between $\frac{m}{n}$ and $\frac{m+\epsilon}{n}$, implying that there would be a multiple of n between m and $m+\epsilon$. Hence

$$\left\{\frac{m+\epsilon}{n}\right\} - \left\{\frac{m}{n}\right\} = \frac{m+\epsilon}{n} - \left[\frac{m+\epsilon}{n}\right] - \frac{m}{n} + \left[\frac{m}{n}\right] = \frac{\epsilon}{n}.$$

It follows that

$$\begin{aligned} G(m+\epsilon) - G(m) &= \sum_{n=1}^{\infty} \frac{\mu(n)}{n} \left\{ \frac{m+\epsilon}{n} \right\} - \sum_{n=1}^{\infty} \frac{\mu(d)}{n} \left\{ \frac{m}{n} \right\} \\ &= \sum_{n=1}^{\infty} \frac{\mu(n)}{n} \left(\left\{ \frac{m+\epsilon}{n} \right\} - \left\{ \frac{m}{n} \right\} \right) \\ &= \epsilon \sum_{n=1}^{\infty} \frac{\mu(n)}{n^2} = \frac{\epsilon}{\zeta(2)}, \end{aligned}$$

applying corollary 2.14. □

Theorem 8.3. *Let $x, H \in \mathbb{R}_{\geq 1}$. Then*

$$\sum_{x < n \leq x+H} \frac{\varphi(n)}{n} - \frac{H}{\zeta(2)} = G(x) - G(x+H).$$

Proof. Fix $n, m \in \mathbb{Z}_{\geq 1}$. Note that

$$\left\{ \frac{m+1}{n} \right\} - \left\{ \frac{m}{n} \right\} = \frac{m+1}{n} - \left[\frac{m+1}{n} \right] - \frac{m}{n} + \left[\frac{m}{n} \right] = \frac{1}{n} + \left[\frac{m}{n} \right] - \left[\frac{m+1}{n} \right].$$

Now $\left[\frac{m}{n}\right] = \left[\frac{m+1}{n}\right]$, except when there is an integer k, such that $\frac{m}{n} < k \leq \frac{m+1}{n}$. This happens when $m < kn \leq (m+1)$, which is exactly the case when $n | m+1$. In this last case $\left[\frac{m}{d}\right] = \left[\frac{m+1}{d}\right] - 1$. We see that

$$\begin{aligned} G(m+1) - G(m) &= \sum_{n=1}^{\infty} \frac{\mu(n)}{n} \left\{ \frac{m+1}{n} \right\} - \sum_{n=1}^{\infty} \frac{\mu(n)}{n} \left\{ \frac{m}{n} \right\} \\ &= \sum_{n=1}^{\infty} \frac{\mu(n)}{n} \left(\left\{ \frac{m+1}{n} \right\} - \left\{ \frac{m}{n} \right\} \right) \\ &= \sum_{n=1}^{\infty} \frac{\mu(n)}{n^2} - \sum_{n | m+1} \frac{\mu(n)}{n} \\ &= \frac{1}{\zeta(2)} - \frac{\varphi(m+1)}{m+1}. \end{aligned}$$

For the last equation we applied corollary 2.14 and corollary 2.34. Now for $x, H \in \mathbb{R}_{\geq 1}$, we see that

$$\begin{aligned} G(x+H) - G(x) &= G(x+H) - G([x+H]) + \sum_{m=[x]}^{[x+H]-1} (G(m+1) - G(m)) - (G(x) - G([x])) \\ &\overset{\text{(lemma 8.2)}}{=} \frac{\{x+H\}}{\zeta(2)} + \left(\frac{[x+H] - [x]}{\zeta(2)} - \sum_{n=x+1}^{x+[H]} \frac{\varphi(n)}{n} \right) - \frac{\{x\}}{\zeta(2)} \\ &= \frac{H}{\zeta(2)} - \sum_{x < n \leq x+H} \frac{\varphi(n)}{n}. \end{aligned}$$

□

8.3 The expected value of $G(y)^2$

As we are interested in the variance of $\sum_{x<n\leq x+H} \frac{\varphi(n)}{n}$, by the previous theorem we are interested in $(G(x+H)-G(x))^2$. More specifically we are interested in

$$\mathop{\mathbb{E}}_{x\in\mathbb{Z}_{\geq 1}}\left[\left(\sum_{n=1}^{\infty}\frac{\mu(n)}{n}\left\{\frac{x+H}{n}\right\}-\sum_{n=1}^{\infty}\frac{\mu(n)}{n}\left\{\frac{x}{n}\right\}\right)^2\right],$$

where we introduced the symbol $\mathbb{E}_{x\in\mathbb{Z}_{\geq 1}}$ for $\lim_{X\to\infty}\frac{1}{X}\sum_{x=1}^{X}$. (This is definitely an abuse of notation, as we don't have a probability measure on $\mathbb{Z}_{\geq 1}$. It however does make our intentions clear.) We are interested in the expectation value, when we take x uniformly from $\{1,...,X\}$ and $X\to\infty$. This expectation value turns out to be very difficult to calculate. What we can however calculate is the same expression with the sum and expectation value interchanged. For example, we might be interested in $\mathbb{E}_{x\in\mathbb{Z}_{\geq 1}}\left[\sum_{m,n=1}^{\infty}\frac{\mu(m)\mu(n)}{mn}\left\{\frac{x}{n}\right\}\left\{\frac{x}{m}\right\}\right]$, which is one of the terms if you expand the square. We cannot calculate this, but we can calculate

$$\sum_{m,n=1}^{\infty}\frac{\mu(m)\mu(n)}{mn}\mathop{\mathbb{E}}_{x\in\mathbb{Z}_{\geq 1}}\left[\left\{\frac{x}{n}\right\}\left\{\frac{x}{m}\right\}\right]=\frac{1}{6\zeta(2)^2}+\frac{1}{12\zeta(2)}. \tag{8.3}$$

The question is whether these two values are the same. Numerical simulations suggest that the expectation value we are interested in indeed coincides with this value, as you can see in figure 8.2. Whether we are allowed to interchange the sum and expectation value is a question we will come back to in section 8.5. Let us first show how to prove relation (8.3).

Lemma 8.4. *Let m,n be positive integers. Fix $x\in\mathbb{Z}_{\geq 1}$. Then*

$$\sum_{k=1}^{mn}\left\{\frac{x+k}{m}\right\}\left\{\frac{x+k}{n}\right\}=\frac{(m-1)(n-1)}{4}+\frac{\gcd(m,n)^2-1}{12}.$$

Proof. Note that if $x\equiv r_m \bmod m$ and $x\equiv r_n \bmod n$ with $0\leq r_m<m$, $0\leq r_n<n$, then $\left\{\frac{x}{m}\right\}\left\{\frac{x}{n}\right\}=\frac{r_m}{m}\frac{r_n}{n}$. It follows that $\left\{\frac{x+mn}{m}\right\}\left\{\frac{x+mn}{n}\right\}=\left\{\frac{x}{m}\right\}\left\{\frac{x}{n}\right\}$. We say that $\left\{\frac{x}{m}\right\}\left\{\frac{x}{n}\right\}$ is periodic with period mn. It is hence sufficient to prove that

$$\sum_{k=1}^{mn}\left\{\frac{k}{m}\right\}\left\{\frac{k}{n}\right\}=\frac{(m-1)(n-1)}{4}+\frac{\gcd(m,n)^2-1}{12}.$$

If m and n are coprime, then when k runs from 1 to mn, the pair $(k \bmod m, k \bmod n)$ runs over all pairs (r_m,r_n) exactly once. Hence for coprime m,n the lemma is trivial as

$$\sum_{k=1}^{mn}\left\{\frac{k}{m}\right\}\left\{\frac{k}{n}\right\}=\left(\sum_{r_m=0}^{m-1}\frac{r_m}{m}\right)\left(\sum_{r_n=0}^{n-1}\frac{r_n}{n}\right)=\left(\frac{m-1}{2}\right)\left(\frac{n-1}{2}\right).$$

Let us turn our attention to the general case. Write $d=\gcd(m,n)$. First note that $\left\{\frac{x}{m}\right\}\left\{\frac{x}{n}\right\}$ is not only periodic with period mn, but also with period $\operatorname{lcm}(m,n)=\frac{mn}{d}$. Hence

$$\sum_{k=1}^{mn}\left\{\frac{k}{m}\right\}\left\{\frac{k}{n}\right\}=d\sum_{k=1}^{\frac{mn}{d}}\left\{\frac{k}{m}\right\}\left\{\frac{k}{n}\right\}.$$

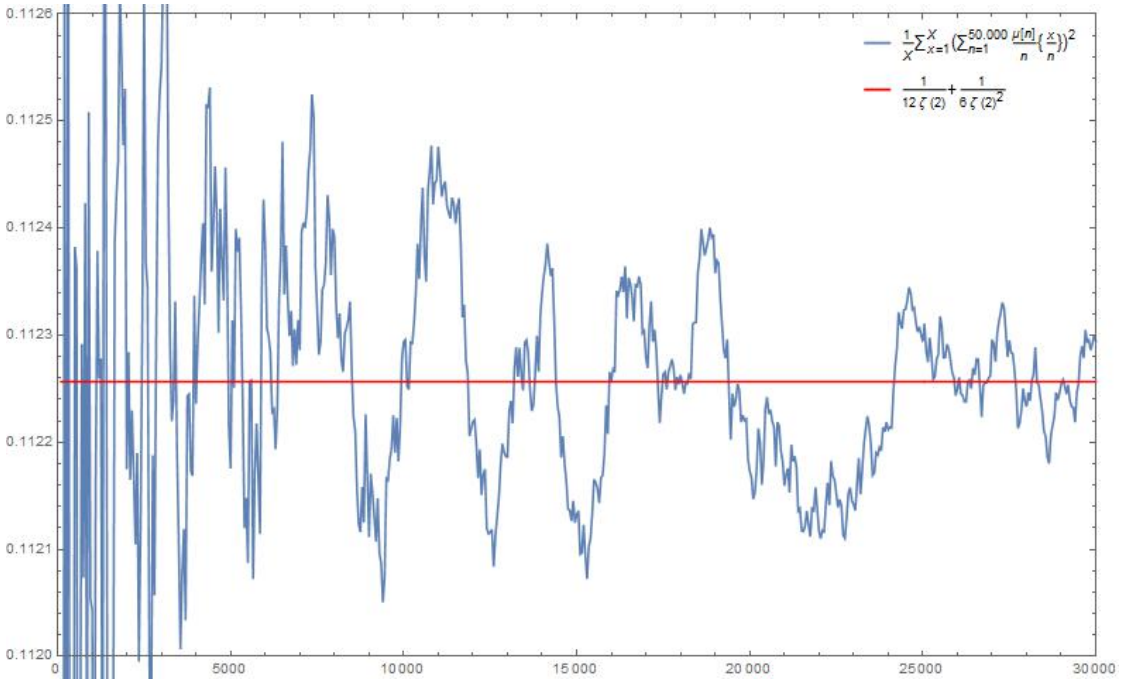

Figure 8.2: $\frac{1}{X}\sum_{x=1}^{X}\left(\sum_{n=1}^{50000}\frac{\varphi(n)}{n}\left\{\frac{x}{n}\right\}\right)^2$ for $10^2 \leq X \leq 3\cdot 10^4$.

The next question is which pairs ($k \bmod m, k \bmod n$) are attained when k runs over the integers from 1 to $\frac{mn}{d}$, as it clearly cannot reach all pairs. We claim that k reaches exactly all pairs (r_m, r_n) such that $r_m \equiv r_n \bmod d$. Suppose that for some k we find that ($k \bmod m, k \bmod n$) $= (r_m, r_n)$. Then $k = r_m + l_m m = r_n + l_n n$, implying that $r_m - r_n = l_n n - l_m m$, so d has to divide $r_m - r_n$. Furthermore there are $\frac{mn}{d}$ pairs (r_m, r_n) with $0 \leq r_m < m$, $0 \leq r_n < n$ s.t. $r_m \equiv r_n \bmod d$: if you fix r_n, then there are $\frac{m}{d}$ choices for $r_m \equiv r_n \bmod d$. If there is a pair that ($k \bmod m, k \bmod n$) does not attain, then by the pigeon hole principle there exists a pair that ($k \bmod m, k \bmod n$) attains twice, (when k runs from 1 to $\frac{mn}{d}$). Hence there exist $0 \leq k < k' < \frac{mn}{d}$ s.t. $k \equiv k' \bmod m$ and $k \equiv k' \bmod n$. By the first statement the difference between k and k' is a multiple of m, by the second it is a multiple of n. Hence the difference between k and k' is a multiple of $\operatorname{lcm}(m,n) = \frac{mn}{d}$. This is not possible. We conclude that the claim is true. Now we use the fact that ($k \bmod m, k \bmod n$) attains exactly the pairs (r_m, r_n) with $r_m \equiv r_n \bmod d$ in our

calculation:

$$\begin{aligned}
\sum_{k=1}^{mn}\left\{\frac{k}{m}\right\}\left\{\frac{k}{n}\right\} &= d\sum_{k=1}^{\frac{mn}{d}}\left\{\frac{k}{m}\right\}\left\{\frac{k}{n}\right\} \\
&= d\sum_{l=0}^{d-1}\left(\sum_{\substack{0\le r_m<m\\ r_m\equiv l \bmod d}}\frac{r_m}{m}\right)\left(\sum_{\substack{0\le r_n<n\\ r_n\equiv l \bmod d}}\frac{r_n}{n}\right) \\
&= d\sum_{l=0}^{d-1}\left(\sum_{t_m=0}^{\frac{m}{d}-1}\frac{l+dt_m}{m}\right)\left(\sum_{t_n=0}^{\frac{n}{d}-1}\frac{l+dt_n}{n}\right) \\
&\overset{(a)}{=} d\sum_{l=0}^{d-1}\left(\frac{l}{d}+\frac{m-d}{2d}\right)\left(\frac{l}{d}+\frac{n-d}{2d}\right) \\
&= \frac{1}{d}\sum_{l=0}^{d-1}\left(l^2+\frac{m+n-2d}{2}l+\frac{(m-d)(n-d)}{4}\right) \\
&\overset{(a,b)}{=} \frac{(d-1)(2d-1)}{6}+\frac{(m+n-2d)(d-1)}{4}+\frac{(m-d)(n-d)}{4} \\
&= \frac{(m-1)(n-1)}{4}+\frac{d^2-1}{12}.
\end{aligned}$$

Here at (a) we used the general identity $\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$, while at (b) we used that $\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$. □

Corollary 8.5. *Let m, n be positive integers. Then*

$$\mathop{\mathbb{E}}_{x\in\mathbb{Z}_{\ge 1}}\left[\left\{\frac{x}{m}\right\}\left\{\frac{x}{n}\right\}\right] = \frac{1}{mn}\left(\frac{(m-1)(n-1)}{4}+\frac{\gcd(m,n)^2-1}{12}\right).$$

Proof. By applying lemma 8.4 to the first $mn\left[\frac{X}{mn}\right]$ integers and using $0 \le \left\{\frac{x}{m}\right\}\left\{\frac{x}{n}\right\} < 1$ for the remaining $mn\left\{\frac{X}{mn}\right\}$ integers, it follows that

$$\frac{1}{X}\sum_{x=1}^{X}\left\{\frac{x}{m}\right\}\left\{\frac{x}{n}\right\} \ge \frac{1}{X}\left[\frac{X}{mn}\right]\left(\frac{(m-1)(n-1)}{4}+\frac{\gcd(m,n)^2-1}{12}\right)$$

and

$$\frac{1}{X}\sum_{x=1}^{X}\left\{\frac{x}{m}\right\}\left\{\frac{x}{n}\right\} \le \frac{1}{X}\left[\frac{X}{mn}\right]\left(\frac{(m-1)(n-1)}{4}+\frac{\gcd(m,n)^2-1}{12}\right)+\frac{mn}{X}\left\{\frac{X}{mn}\right\}.$$

Now take the limit $X \to \infty$ to see that the required statement holds. □

Lemma 8.6.

$$\sum_{m,n=1}^{\infty}\frac{\mu(m)\mu(n)}{m^2n^2}\cdot\gcd(m,n)^2 = \frac{1}{\zeta(2)}.$$

Proof. This proof requires a couple of variable transformations. Start by writing $d = \gcd(m, n)$ and $m = dm'$, $n = dn'$. Then we see that

$$\sum_{m,n=1}^{\infty} \frac{\mu(m)\mu(n)}{m^2n^2} \cdot \gcd(m,n)^2 = \sum_{d=1}^{\infty} \sum_{\substack{m',n'=1 \\ \gcd(m',n')=1}}^{\infty} \frac{\mu(dm')\mu(dn')}{(dm')^2(dn')^2} \cdot d^2 = \sum_{d=1}^{\infty} \sum_{\substack{m',n'=1 \\ \gcd(m',n')=1}}^{\infty} \frac{\mu(dm')\mu(dn')}{(dm'n')^2}.$$

We now claim that

$$\mu(dm')\mu(dn') = \mu(d)\mu(dm'n')$$

for coprime m', n'. Note if d and m' or d and n' are not coprime or if one of the d, m', n' contains a square, then both sides are zero. If d, m' and n' are the product of respectively t_d, t_m and t_n distinct primes, then

$$\mu(dm')\mu(dn') = (-1)^{t_d+t_m}(-1)^{t_d+t_n} = (-1)^{t_m+t_n} = (-1)^{t_d}(-1)^{t_d+t_m+t_n} = \mu(d)\mu(dm'n').$$

We see that

$$\sum_{d=1}^{\infty} \sum_{\substack{m',n'=1 \\ \gcd(m',n')=1}}^{\infty} \frac{\mu(dm')\mu(dn')}{(dm'n')^2} = \sum_{d=1}^{\infty} \sum_{\substack{m',n'=1 \\ \gcd(m',n')=1}}^{\infty} \frac{\mu(d)\mu(dm'n')}{(dm'n')^2} = \sum_{d=1}^{\infty} \sum_{m',n'=1}^{\infty} \frac{\mu(d)\mu(dm'n')}{(dm'n')^2}.$$

In the last equality we dropped the condition in the sum that m', n' are coprime. If they are not, then $\mu(dm'n') = 0$. Now write $k = m'n'$ and change the sum over m', n' into a sum over k and a sum over $n'|k$. Then

$$\sum_{d=1}^{\infty} \sum_{m',n'=1}^{\infty} \frac{\mu(d)\mu(dm'n')}{(dm'n')^2} = \sum_{d=1}^{\infty} \sum_{k=1}^{\infty} \sum_{n'|k} \frac{\mu(d)\mu(dk)}{(dk)^2} = \sum_{d=1}^{\infty} \sum_{k=1}^{\infty} \frac{\mu(d)\mu(dk)\sigma_0(k)}{(dk)^2},$$

where σ_0 denotes the divisor function. Finally we write $l = dk$ and transform the sum over d and k into a sum over l and $k|l$. Then

$$\sum_{d=1}^{\infty} \sum_{k=1}^{\infty} \frac{\mu(d)\mu(dk)\sigma_0(k)}{(dk)^2} = \sum_{l=1}^{\infty} \frac{\mu(l)}{l^2} \sum_{k|l} \sigma_0(k)\mu\left(\frac{l}{k}\right).$$

In example 2.11, we have seen that $E = \sigma_0 * \mu$. Applying corollary 2.14 for the last equality, we see that

$$\sum_{m,n=1}^{\infty} \frac{\mu(m)\mu(n)}{m^2n^2} \cdot \gcd(m,n)^2 = \sum_{l=1}^{\infty} \frac{\mu(l)}{l^2} \sum_{k|l} \sigma_0(k)\mu\left(\frac{l}{k}\right) = \sum_{l=1}^{\infty} \frac{\mu(l)E(l)}{l^2} = \sum_{l=1}^{\infty} \frac{\mu(l)}{l^2} = \frac{1}{\zeta(2)}.$$

□

Theorem 8.7.

$$\sum_{m,n=1}^{\infty} \frac{\mu(m)\mu(n)}{mn} \mathop{\mathbb{E}}_{x \in \mathbb{Z}_{\geq 1}} \left[\left\{\frac{x}{n}\right\}\left\{\frac{x}{m}\right\}\right] = \frac{1}{6\zeta(2)^2} + \frac{1}{12\zeta(2)}.$$

Proof.

$$\sum_{m,n=1}^{\infty} \frac{\mu(m)\mu(n)}{mn} \mathop{\mathbb{E}}_{x \in \mathbb{Z}_{\geq 1}} \left[\left\{\frac{x}{n}\right\}\left\{\frac{x}{m}\right\}\right] \overset{\text{(corollary 8.5)}}{=} \sum_{m,n=1}^{\infty} \frac{\mu(m)\mu(n)}{m^2n^2} \left(\frac{(m-1)(n-1)}{4} + \frac{\gcd(m,n)^2 - 1}{12}\right)$$

$$= \frac{1}{6}\sum_{m,n}^{\infty} \frac{\mu(m)\mu(n)}{m^2n^2} - \frac{1}{4}\sum_{m,n}^{\infty} \frac{\mu(m)\mu(n)}{mn^2} - \frac{1}{4}\sum_{m,n}^{\infty} \frac{\mu(m)\mu(n)}{m^2n} + \frac{1}{4}\sum_{m,n}^{\infty} \frac{\mu(m)\mu(n)}{mn} + \frac{1}{12}\sum_{m,n}^{\infty} \frac{\mu(m)\mu(n)}{m^2n^2} \cdot \gcd(m,n)^2.$$

Now we can split each of the first four sums into a product of a sum over m and a sum over n. We then use theorem 2.31 and corollary 2.14, stating that $\sum_{k=1}^{\infty} \frac{\mu(k)}{k} = 0$ and $\sum_{k=1}^{\infty} \frac{\mu(k)}{k^2} = \frac{1}{\zeta(2)}$, to see that the first sum equals $\frac{1}{\zeta(2)^2}$ and the other three sums are zero. For the last sum we apply lemma 8.6. We conclude that the theorem holds. □

8.4 The variance of $\sum \frac{\varphi(n)}{n}$ in the interval $[x, 2x]$

Recall that we are interested in $(G(x+H) - G(x))^2$. Unfortunately the technique described in the previous section does not help us when $H = x^{\delta}$, because knowing x mod n, you do not have a clue about $[x^{\delta}]$ mod n. It can quite literally be every value between 0 and n. Hence the combinatorial arguments in the proof of lemma 8.4 cannot be replicated, while these arguments were actually the key to dropping the difficult fractional part function. In section 8.6 we make an additional assumption to be able to replicate some of the arguments. This assumption is however not necessary when we look at the interval $[x, 2x]$. Knowing x mod n, we clearly know $2x$ mod n. We hence turn our attention here to calculate the sum

$$\sum_{m,n=1}^{\infty} \frac{\mu(m)\mu(n)}{mn} \left(\mathop{\mathbb{E}}_{x \in \mathbb{Z}_{\geq 1}} \left[\left\{ \frac{x}{n} \right\} \left\{ \frac{x}{m} \right\} \right] - \mathop{\mathbb{E}}_{x \in \mathbb{Z}_{\geq 1}} \left[\left\{ \frac{2x}{n} \right\} \left\{ \frac{x}{m} \right\} \right] - \mathop{\mathbb{E}}_{x \in \mathbb{Z}_{\geq 1}} \left[\left\{ \frac{x}{n} \right\} \left\{ \frac{2x}{m} \right\} \right] + \mathop{\mathbb{E}}_{x \in \mathbb{Z}_{\geq 1}} \left[\left\{ \frac{2x}{n} \right\} \left\{ \frac{2x}{m} \right\} \right] \right).$$

This calculation will turn out to be more tedious as we need to consider the parity of m and n. We hence need some additional lemmas before we can prove theorem 8.13, stating that the sum above equals

$$\frac{1}{6\zeta(2)} - \frac{1}{6\zeta(2)^2}.$$

Lemma 8.8. *Let m, n be positive integers and suppose that m is odd. Fix $x \in \mathbb{Z}_{\geq 1}$. Then*

$$\sum_{k=1}^{mn} \left\{ \frac{2(x+k)}{m} \right\} \left\{ \frac{x+k}{n} \right\} = \frac{(m-1)(n-1)}{4} + \frac{\gcd(m,n)^2 - 1}{24}.$$

Proof. As in the proof of lemma 8.4, it is sufficient to consider the case $x = 0$, as $\left\{ \frac{2x}{m} \right\} \left\{ \frac{x}{n} \right\}$ is periodic with period mn. Since m is odd, this period is the same as that of $\left\{ \frac{x}{m} \right\} \left\{ \frac{x}{n} \right\}$. Defining $d = \gcd(m, n)$ we can argue analogously as in lemma 8.4, to see that when k runs over the integers 1 to $\frac{mn}{d}$, the pair $(2k \bmod m, k \bmod n)$ runs over the pairs (r_m, r_n) with $0 \leq r_m < m$,

$0 \leq r_n < n$, and $r_m \equiv 2r_n \bmod d$. Note that d is odd, because m is odd. We find that

$$\begin{aligned}
\sum_{k=1}^{mn} \left\{\frac{2k}{m}\right\}\left\{\frac{k}{n}\right\} &= d\sum_{k=1}^{\frac{mn}{d}} \left\{\frac{2k}{m}\right\}\left\{\frac{k}{n}\right\} \\
&= d\sum_{l=0}^{d-1} \left(\sum_{\substack{0\leq r_m<m \\ r_m\equiv 2l \bmod d}} \frac{r_m}{m}\right)\left(\sum_{\substack{0\leq r_n<n \\ r_n\equiv l \bmod d}} \frac{r_n}{n}\right) \\
&= d\sum_{l=0}^{\frac{d-1}{2}} \left(\sum_{t_m=0}^{\frac{m}{d}-1} \frac{2l+dt_m}{m}\right)\left(\sum_{t_n=0}^{\frac{n}{d}-1} \frac{l+dt_n}{n}\right) \\
&\qquad + d\sum_{l=\frac{d+1}{2}}^{d-1} \left(\sum_{t_m=-1}^{\frac{m}{d}-2} \frac{2l+dt_m}{m}\right)\left(\sum_{t_n=0}^{\frac{n}{d}-1} \frac{l+dt_n}{n}\right) \\
&= d\sum_{l=0}^{d-1} \left(\sum_{t_m=0}^{\frac{m}{d}-2} \frac{2l+dt_m}{m}\right)\left(\sum_{t_n=0}^{\frac{n}{d}-1} \frac{l+dt_n}{n}\right) \\
&\qquad + d\sum_{l=0}^{\frac{d-1}{2}} \left(\frac{2l+m-d}{m}\right)\left(\sum_{t_n=0}^{\frac{n}{d}-1} \frac{l+dt_n}{n}\right) \\
&\qquad + d\sum_{l=\frac{d+1}{2}}^{d-1} \left(\frac{2l-d}{m}\right)\left(\sum_{t_n=0}^{\frac{n}{d}-1} \frac{l+dt_n}{n}\right) \\
&= d\sum_{l=0}^{d-1} \left(\sum_{t_m=-1}^{\frac{m}{d}-2} \frac{2l+dt_m}{m}\right)\left(\sum_{t_n=0}^{\frac{n}{d}-1} \frac{l+dt_n}{n}\right) + d\sum_{l=0}^{\frac{d-1}{2}} \left(\sum_{t_n=0}^{\frac{n}{d}-1} \frac{l+dt_n}{n}\right) \\
&= d\sum_{l=0}^{d-1} \left(\frac{2l}{m}\frac{m}{d} + \frac{d}{m}\sum_{t_m=-1}^{\frac{m}{d}-2} t_m\right)\left(\frac{l}{n}\frac{n}{d} + \frac{d}{n}\sum_{t_n=0}^{\frac{n}{d}-1} t_n\right) \\
&\qquad + d\sum_{l=0}^{\frac{d-1}{2}} \left(\frac{l}{n}\frac{n}{d} + \frac{d}{n}\sum_{t_n=0}^{\frac{n}{d}-1} t_n\right) \\
&\overset{(a)}{=} d\sum_{l=0}^{d-1} \left(\frac{2l}{d} + \frac{m-3d}{2d}\right)\left(\frac{l}{d} + \frac{n-d}{2d}\right) + d\sum_{l=0}^{\frac{d-1}{2}} \left(\frac{l}{d} + \frac{n-d}{2d}\right) \\
&= \frac{1}{d}\sum_{l=0}^{d-1} \left(2l^2 + \frac{m+2n-5d}{2}l + \frac{(m-3d)(n-d)}{4}\right) + \sum_{l=0}^{\frac{d-1}{2}} \left(l + \frac{n-d}{2}\right) \\
&\overset{(a,b)}{=} \frac{(d-1)(2d-1)}{3} + \frac{(m+2n-5d)(d-1)}{4} + \frac{(m-3d)(n-d)}{4} \\
&\qquad + \frac{(d-1)(d+1)}{8} + \frac{(d+1)(n-d)}{4} \\
&= \frac{(m-1)(n-1)}{4} + \frac{d^2-1}{24}.
\end{aligned}$$

Again at (a) we used $\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$, while at (b) we used that $\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$. □

Corollary 8.9. *Let m, n be positive integers. Then*

$$\mathop{\mathbb{E}}_{x \in \mathbb{Z}_{\geq 1}} \left[\left\{ \frac{2x}{m} \right\} \left\{ \frac{x}{n} \right\} \right] = \begin{cases} \frac{1}{mn} \left(\frac{(m-2)(n-1)}{4} + \frac{\gcd(\frac{m}{2}, n)^2 - 1}{6} \right) & \text{if } m \text{ is even,} \\ \frac{1}{mn} \left(\frac{(m-1)(n-1)}{4} + \frac{\gcd(m,n)^2 - 1}{24} \right) & \text{if } m \text{ is odd.} \end{cases}$$

Proof. If m is even, we can define $m' = \frac{m}{2}$. Applying corollary 8.5, we see that

$$\begin{aligned} \mathop{\mathbb{E}}_{x \in \mathbb{Z}_{\geq 1}} \left[\left\{ \frac{2x}{m} \right\} \left\{ \frac{x}{n} \right\} \right] &= \mathop{\mathbb{E}}_{x \in \mathbb{Z}_{\geq 1}} \left[\left\{ \frac{x}{m'} \right\} \left\{ \frac{x}{n} \right\} \right] \\ &= \frac{1}{m'n} \left(\frac{(m'-1)(n-1)}{4} + \frac{\gcd(m', n)^2 - 1}{12} \right) \\ &= \frac{2}{mn} \left(\frac{(\frac{m}{2}-1)(n-1)}{4} + \frac{\gcd(\frac{m}{2}, n)^2 - 1}{12} \right) \\ &= \frac{1}{mn} \left(\frac{(m-2)(n-1)}{4} + \frac{\gcd(\frac{m}{2}, n)^2 - 1}{6} \right). \end{aligned}$$

If m is odd, the statement follows directly from lemma 8.8, in exactly the same way as corollary 8.5 follows from lemma 8.4. □

Lemma 8.10. *Let m, n be positive integers. Then*

$$\mathop{\mathbb{E}}_{x \in \mathbb{Z}_{\geq 1}} \left[\left\{ \frac{2x}{m} \right\} \left\{ \frac{2x}{n} \right\} \right] = \begin{cases} \frac{1}{mn} \left(\frac{(m-2)(n-2)}{4} + \frac{\gcd(m,n)^2 - 4}{12} \right) & \text{if } m \text{ is even, } n \text{ is even,} \\ \frac{1}{mn} \left(\frac{(m-2)(n-1)}{4} + \frac{\gcd(m,n)^2 - 1}{12} \right) & \text{if } m \text{ is even, } n \text{ is odd,} \\ \frac{1}{mn} \left(\frac{(m-1)(n-2)}{4} + \frac{\gcd(m,n)^2 - 1}{12} \right) & \text{if } m \text{ is odd, } n \text{ is even,} \\ \frac{1}{mn} \left(\frac{(m-1)(n-1)}{4} + \frac{\gcd(m,n)^2 - 1}{12} \right) & \text{if } m \text{ is odd, } n \text{ is odd.} \end{cases}$$

Proof. If both m and n are even, then we define $m' = \frac{m}{2}$, $n' = \frac{n}{2}$ and apply corollary 8.5 to see that

$$\begin{aligned} \mathop{\mathbb{E}}_{x \in \mathbb{Z}_{\geq 1}} \left[\left\{ \frac{2x}{m} \right\} \left\{ \frac{2x}{n} \right\} \right] &= \mathop{\mathbb{E}}_{x \in \mathbb{Z}_{\geq 1}} \left[\left\{ \frac{x}{m'} \right\} \left\{ \frac{x}{n'} \right\} \right] \\ &= \frac{1}{m'n'} \left(\frac{(m'-1)(n'-1)}{4} + \frac{\gcd(m', n')^2 - 1}{12} \right) \\ &= \frac{4}{mn} \left(\frac{(\frac{m}{2}-1)(\frac{n}{2}-1)}{4} + \frac{\gcd(\frac{m}{2}, \frac{n}{2})^2 - 1}{12} \right) \\ &= \frac{1}{mn} \left(\frac{(m-2)(n-2)}{4} + \frac{\gcd(m, n)^2 - 4}{12} \right). \end{aligned}$$

At the final step we used that $\gcd(\frac{m}{2}, \frac{n}{2}) = \frac{\gcd(m,n)}{2}$. Next if m is odd, n is even, we define

$n = 2n'$ and apply corollary 8.9:

$$\begin{aligned}\mathop{\mathbb{E}}_{x\in\mathbb{Z}_{\geq 1}}\left[\left\{\frac{2x}{m}\right\}\left\{\frac{2x}{n}\right\}\right] &= \mathop{\mathbb{E}}_{x\in\mathbb{Z}_{\geq 1}}\left[\left\{\frac{2x}{m}\right\}\left\{\frac{x}{n'}\right\}\right] \\ &= \frac{1}{mn'}\left(\frac{(m-1)(n'-1)}{4} + \frac{\gcd(m,n')^2-1}{24}\right) \\ &= \frac{2}{mn}\left(\frac{(m-1)(\frac{n}{2}-1)}{4} + \frac{\gcd(m,\frac{n}{2})^2-1}{24}\right) \\ &= \frac{1}{mn}\left(\frac{(m-1)(n-2)}{4} + \frac{\gcd(m,n)^2-1}{12}\right).\end{aligned}$$

Here we applied at the final step the fact that $\gcd(m, \frac{n}{2}) = \gcd(m, n)$, which follows from the fact that m is odd. The proof for m even and n odd is of course analogous, as the expression is symmetric in m and n. Finally if both m and n are odd, then $(2x \bmod m, 2x \bmod n)$ runs over the same pairs as $(x \bmod m, x \bmod n)$ as x runs from 1 to $\frac{mn}{\gcd(m,n)}$, only in a different order. Hence

$$\begin{aligned}\mathop{\mathbb{E}}_{x\in\mathbb{Z}_{\geq 1}}\left[\left\{\frac{2x}{m}\right\}\left\{\frac{2x}{n}\right\}\right] &= \mathop{\mathbb{E}}_{x\in\mathbb{Z}_{\geq 1}}\left[\left\{\frac{x}{m}\right\}\left\{\frac{x}{n}\right\}\right] \\ &= \frac{1}{mn}\left(\frac{(m-1)(n-1)}{4} + \frac{\gcd(m,n)^2-1}{12}\right).\end{aligned}$$

□

Finally we need two more lemmas, stating the values of some infinite sums over even or odd integers.

Lemma 8.11.

$$\sum_{n \text{ even}} \frac{\mu(n)}{n^2} = -\frac{1}{3\zeta(2)} \text{ and } \sum_{n \text{ odd}} \frac{\mu(n)}{n^2} = \frac{4}{3\zeta(2)}.$$

Proof. We know that

$$\sum_{n \text{ even}} \frac{\mu(n)}{n^2} = \sum_{n} \frac{\mu(2n)}{(2n)^2} \overset{(a)}{=} \sum_{n \text{ odd}} \frac{\mu(2n)}{(2n)^2} \overset{(b)}{=} -\frac{1}{4}\sum_{n \text{ odd}} \frac{\mu(n)}{n^2}.$$

At (a) we used the fact that if n is even, then $2n$ is not square free, so $\mu(2n) = 0$. At (b) we used the fact that if n is odd, then $\mu(2n) = -\mu(n)$. Now apply corollary 2.14 to see that

$$\frac{1}{\zeta(2)} = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^2} = \sum_{n \text{ even}} \frac{\mu(n)}{n^2} + \sum_{n \text{ odd}} \frac{\mu(n)}{n^2} = \frac{3}{4}\sum_{n \text{ odd}} \frac{\mu(n)}{n^2}.$$

This immediately implies both statements from the lemma. □

Lemma 8.12.

1. $$\sum_{\substack{m \text{ odd},\\ n \text{ odd}}} \frac{\mu(m)\mu(n)}{m^2n^2} \gcd(m,n)^2 = \frac{4}{3\zeta(2)}.$$

2.

$$\sum_{\substack{m \text{ even},\\ n \text{ odd}}} \frac{\mu(m)\mu(n)}{m^2n^2} \gcd(m,n)^2 = \sum_{\substack{m \text{ odd},\\ n \text{ even}}} \frac{\mu(m)\mu(n)}{m^2n^2} \gcd(m,n)^2 = -\frac{1}{3\zeta(2)}.$$

3.

$$\sum_{\substack{m \text{ even},\\ n \text{ even}}} \frac{\mu(m)\mu(n)}{m^2n^2} \gcd(m,n)^2 = \frac{1}{3\zeta(2)}.$$

Proof. Using the same reasoning as in the proof of lemma 8.11 and using that $\gcd(2m, 2n) = 2\gcd(m,n)$, we see that

$$\sum_{\substack{m \text{ even},\\ n \text{ even}}} \frac{\mu(m)\mu(n)}{m^2n^2} \gcd(m,n)^2 = \sum_{\substack{m \text{ odd},\\ n \text{ odd}}} \frac{\mu(2m)\mu(2n)}{(2m)^2(2n)^2} \gcd(2m,2n)^2 = \frac{1}{4} \sum_{\substack{m \text{ odd},\\ n \text{ odd}}} \frac{\mu(m)\mu(n)}{m^2n^2} \gcd(m,n)^2$$

and

$$\sum_{\substack{m \text{ even},\\ n \text{ odd}}} \frac{\mu(m)\mu(n)}{m^2n^2} \gcd(m,n)^2 = \sum_{\substack{m \text{ odd},\\ n \text{ odd}}} \frac{\mu(2m)\mu(n)}{(2m)^2n^2} \gcd(2m,n)^2 = -\frac{1}{4} \sum_{\substack{m \text{ odd},\\ n \text{ odd}}} \frac{\mu(m)\mu(n)}{m^2n^2} \gcd(m,n)^2.$$

This last statement also holds for the sum with m odd and n even. Now by lemma 8.6 we know that

$$\frac{1}{\zeta(2)} = \sum_{m,n=1}^{\infty} \frac{\mu(m)\mu(n)}{m^2n^2} \gcd(m,n)^2 = \frac{3}{4} \sum_{\substack{m \text{ odd},\\ n \text{ odd}}} \frac{\mu(m)\mu(n)}{m^2n^2} \gcd(m,n)^2.$$

All the statements from the lemma immediately follow. □

Finally we apply all these lemmas to derive the following theorem.

Theorem 8.13.

$$\sum_{m,n=1}^{\infty} \frac{\mu(m)\mu(n)}{mn} \left(\mathop{\mathbb{E}}_{x \in \mathbb{Z}_{\geq 1}} \left[\left\{ \frac{x}{n} \right\} \left\{ \frac{x}{m} \right\} \right] - \mathop{\mathbb{E}}_{x \in \mathbb{Z}_{\geq 1}} \left[\left\{ \frac{2x}{n} \right\} \left\{ \frac{x}{m} \right\} \right] - \mathop{\mathbb{E}}_{x \in \mathbb{Z}_{\geq 1}} \left[\left\{ \frac{x}{n} \right\} \left\{ \frac{2x}{m} \right\} \right] + \mathop{\mathbb{E}}_{x \in \mathbb{Z}_{\geq 1}} \left[\left\{ \frac{2x}{n} \right\} \left\{ \frac{2x}{m} \right\} \right] \right) = \frac{1}{6\zeta(2)} - \frac{1}{6\zeta(2)^2}.$$

Proof. Applying corollary 8.5, corollary 8.9 and lemma 8.10, we see that

$$\begin{aligned}
&\sum_{m,n=1}^{\infty} \frac{\mu(m)\mu(n)}{mn} \left(\mathop{\mathbb{E}}_{x\in\mathbb{Z}_{\geq 1}} \left[\left\{ \frac{x}{n} \right\} \left\{ \frac{x}{m} \right\} \right] - \mathop{\mathbb{E}}_{x\in\mathbb{Z}_{\geq 1}} \left[\left\{ \frac{2x}{n} \right\} \left\{ \frac{x}{m} \right\} \right] - \mathop{\mathbb{E}}_{x\in\mathbb{Z}_{\geq 1}} \left[\left\{ \frac{x}{n} \right\} \left\{ \frac{2x}{m} \right\} \right] \right. \\
&\qquad \left. + \mathop{\mathbb{E}}_{x\in\mathbb{Z}_{\geq 1}} \left[\left\{ \frac{2x}{n} \right\} \left\{ \frac{2x}{m} \right\} \right] \right) \\
&= \sum_{\substack{m \text{ odd},\\ n \text{ odd}}} \frac{\mu(m)\mu(n)}{m^2n^2} \left(\frac{\gcd(m,n)^2-1}{12} \right) \\
&\qquad + \sum_{\substack{m \text{ even},\\ n \text{ odd}}} \frac{\mu(m)\mu(n)}{m^2n^2} \left(\frac{5(\gcd(m,n)^2-1)}{24} - \frac{\gcd(\frac{m}{2},n)^2-1}{6} \right) \\
&\qquad + \sum_{\substack{m \text{ odd},\\ n \text{ even}}} \frac{\mu(m)\mu(n)}{m^2n^2} \left(\frac{5(\gcd(m,n)^2-1)}{24} - \frac{\gcd(m,\frac{n}{2})^2-1}{6} \right) \\
&\qquad + \sum_{\substack{m \text{ odd},\\ n \text{ odd}}} \frac{\mu(m)\mu(n)}{m^2n^2} \left(\frac{\gcd(m,n)^2}{6} - \frac{\gcd(\frac{m}{2},n)^2}{6} - \frac{\gcd(m,\frac{n}{2})^2}{6} + \frac{1}{6} \right) \\
&\overset{(a)}{=} \sum_{\substack{m \text{ odd},\\ n \text{ odd}}} \frac{\mu(m)\mu(n)}{m^2n^2} \left(\frac{\gcd(m,n)^2-1}{12} \right) - \sum_{\substack{m \text{ even},\\ n \text{ odd}}} \frac{\mu(m)\mu(n)}{m^2n^2} \left(\frac{\gcd(m,n)^2-1}{24} \right) \\
&\qquad - \sum_{\substack{m \text{ odd},\\ n \text{ even}}} \frac{\mu(m)\mu(n)}{m^2n^2} \left(\frac{\gcd(m,n)^2-1}{24} \right) + \sum_{\substack{m \text{ odd},\\ n \text{ odd}}} \frac{\mu(m)\mu(n)}{m^2n^2} \left(\frac{\gcd(m,n)^2+2}{12} \right).
\end{aligned}$$

At (a) we applied that $\gcd(\frac{m}{2},n) = \gcd(m,n)$ is m is even and n is odd, and $\gcd(m,\frac{n}{2}) = \gcd(m,n)$ is m is odd and n is even. We also used that $\gcd(\frac{m}{2},n) = \gcd(m,\frac{n}{2}) = \frac{\gcd(m,n)}{2}$ if both m and n are even and square free. We now apply lemma 8.12 for the odd/even sums with a factor $\gcd(m,n)$. We split the sums with constant factor into odd/even sums over m or n and apply lemma 8.11. We see that the expression above equals

$$\frac{1}{12}\cdot\frac{4}{3\zeta(2)} - \frac{1}{12}\cdot\frac{16}{9\zeta(2)^2} + \frac{1}{24}\cdot\frac{1}{3\zeta(2)} - \frac{1}{24}\cdot\frac{4}{9\zeta(2)^2} + \frac{1}{24}\cdot\frac{1}{3\zeta(2)} - \frac{1}{24}\cdot\frac{4}{9\zeta(2)^2} + \frac{1}{12}\cdot\frac{1}{3\zeta(2)} + \frac{2}{12}\cdot\frac{1}{9\zeta(2)^2}$$

$$\begin{aligned}
&= \left(\frac{8}{72} + \frac{1}{72} + \frac{1}{72} + \frac{2}{72} \right) \frac{1}{\zeta(2)} + \left(-\frac{32}{216} - \frac{4}{216} - \frac{4}{216} + \frac{4}{216} \right) \frac{1}{\zeta(2)^2} \\
&= \frac{1}{6\zeta(2)} - \frac{1}{6\zeta(2)^2}.
\end{aligned}$$

□

Again the question is whether interchanging the sum and the expectation value is allowed. That is, the question is if

$$\lim_{X\to\infty} \frac{1}{X} \sum_{x=1}^{X} \left[\left(\sum_{n=x}^{2x} \frac{\varphi(n)}{n} - \frac{x}{\zeta(2)} \right)^2 \right] = \lim_{X\to\infty} \frac{1}{X} \sum_{x=1}^{X} \left[\left(\sum_{n=1}^{\infty} \frac{\mu(n)}{n} \left\{ \frac{x+H}{n} \right\} - \sum_{n=1}^{\infty} \frac{\mu(n)}{n} \left\{ \frac{x}{n} \right\} \right)^2 \right]$$

coincides with the value we found above. In the introduction we have seen figure 8.1, where this value is shown for X upto 10^7. It clearly looks to coincide, but this is of course not a proof. In the next section we research the interchanging of sum and expectation value.

8.5 On convergence and interchanging the average and the infinite sum

The question we address in this section is whether we are allowed to interchange the two limits in the expression

$$\lim_{M,N\to\infty}\sum_{m=1}^{M}\sum_{n=1}^{N}\lim_{X\to\infty}\frac{1}{X}\sum_{x=1}^{X}f(x,m,n) \tag{8.4}$$

for different f, without changing its value. Of course for us the most interesting case is the case where f is given by

$$f(x,m,n)=\frac{\mu(m)\mu(n)}{mn}\left(\left\{\frac{x}{n}\right\}\left\{\frac{x}{m}\right\}-\left\{\frac{2x}{n}\right\}\left\{\frac{x}{m}\right\}-\left\{\frac{x}{n}\right\}\left\{\frac{2x}{m}\right\}+\left\{\frac{2x}{n}\right\}\left\{\frac{2x}{m}\right\}\right),$$

as this case would give us a result for the variance of $\frac{\varphi(n)}{n}$. The short answer is that we do not know if we are allowed to interchange the two limits. There are not many theorems on interchanging an infinite sum and an 'infinite average'. The only result the author could find was applying Lebesgue's dominated convergence theorem for summations, see for example [4]. This gives us the following theorem.

Theorem 8.14. *Suppose we have a function $f:(\mathbb{Z}_{\geq 1})^3\to\mathbb{R}$ and a function $g:(\mathbb{Z}_{\geq 1})^2\to\mathbb{R}$, such that*

$$|f(x,m,n)|\leq g(m,n) \text{ for all } x\in\mathbb{Z}_{\geq 1}.$$

If

$$\lim_{M,N\to\infty}\sum_{m=1}^{M}\sum_{n=1}^{N}g(m,n)<\infty,$$

then

$$\lim_{M,N\to\infty}\sum_{m=1}^{M}\sum_{n=1}^{N}\lim_{X\to\infty}\frac{1}{X}\sum_{x=1}^{X}f(x,m,n)=\lim_{X\to\infty}\frac{1}{X}\sum_{x=1}^{X}\lim_{M,N\to\infty}\sum_{m=1}^{M}\sum_{n=1}^{N}f(x,m,n).$$

These requirements are however far stronger than we have. The only thing we do know is that for the interesting f, stated above, we have

$$\lim_{M,N\to\infty}\sum_{m=1}^{M}\sum_{n=1}^{N}\left|\lim_{X\to\infty}\frac{1}{X}\sum_{x=1}^{X}f(x,m,n)\right|<\infty \tag{8.5}$$

This is of course a much weaker statement than the requirement in theorem 8.14. As weak as relation (8.5) may seem, it turns out that it is not even true for

$$f(x,m,n)=\frac{\mu(m)\mu(n)}{mn}\left\{\frac{x}{n}\right\}\left\{\frac{x}{m}\right\}.$$

This is the case we studied in section 8.3. Even though figure 8.2 seems to imply that the two values studied in this section do coincide, we have no justification of this, as the expression (8.4) does not converge absolutely. Even though we cannot prove it, we are convinced that the calculated value coincides with the variance of $\frac{\varphi(n)}{n}$ in short intervals, due to the numerical simulations. We hence make the following assumption.

Assumption 8.15. *Changing the two limits in the expression*

$$\lim_{M,N\to\infty} \sum_{m=1}^{M} \sum_{n=1}^{N} \lim_{X\to\infty} \frac{1}{X} \sum_{x=1}^{X} f(x,m,n)$$

does not change its value, for

$$f(x,m,n) = \frac{\mu(m)\mu(n)}{mn} \left(\left\{ \frac{x}{n} \right\} \left\{ \frac{x}{m} \right\} - \left\{ \frac{2x}{n} \right\} \left\{ \frac{x}{m} \right\} - \left\{ \frac{x}{n} \right\} \left\{ \frac{2x}{m} \right\} + \left\{ \frac{2x}{n} \right\} \left\{ \frac{2x}{m} \right\} \right).$$

Corollary 8.16 (Corollary of theorem 8.13)**.** *Assuming 8.15, we find that*

$$\lim_{X\to\infty} \frac{1}{X} \sum_{x=1}^{X} \left[\left(\sum_{n=x}^{2x} \frac{\varphi(n)}{n} - \frac{x}{\zeta(2)} \right)^2 \right] = \frac{1}{6\zeta(2)} - \frac{1}{6\zeta(2)^2}.$$

We conclude this section by proving and disproving relation (8.5) for the different f mentioned above.

Lemma 8.17. *The expressions*

$$\lim_{M,N\to\infty} \sum_{m=1}^{M} \sum_{n=1}^{N} \frac{\mu(m)\mu(n)}{m^2n^2}$$

and

$$\lim_{M,N\to\infty} \sum_{m=1}^{M} \sum_{n=1}^{N} \frac{\mu(m)\mu(n)}{m^2n^2} \cdot \gcd(m,n)^2$$

converge absolutely.

Proof. The first is clear as

$$\sum_{m,n=1}^{\infty} \frac{|\mu(m)\mu(n)|}{(mn)^2} \leq \sum_{m,n=1}^{\infty} \frac{1}{(mn)^2} = \left(\sum_{m=1}^{\infty} \frac{1}{m^2} \right) \left(\sum_{n=1}^{\infty} \frac{1}{n^2} \right) = \zeta(2)^2.$$

For the second expression, we reason as in lemma 8.6 and see that

$$\begin{aligned} \sum_{m,n=1}^{\infty} \frac{|\mu(m)\mu(n)|}{(mn)^2} \cdot \gcd(m,n)^2 &= \sum_{d=1}^{\infty} \sum_{\substack{m',n'=1 \\ \gcd(m',n')=1}}^{\infty} \frac{|\mu(dm')\mu(dn')|}{(dmn)^2} \\ &= \sum_{d=1}^{\infty} \sum_{m',n'=1}^{\infty} \frac{|\mu(d)\mu(dmn)|}{(dmn)^2} \\ &\leq \left(\sum_{d=1}^{\infty} \frac{1}{d^2} \right) \left(\sum_{m=1}^{\infty} \frac{1}{m^2} \right) \left(\sum_{n=1}^{\infty} \frac{1}{n^2} \right) = \zeta(2)^3. \end{aligned}$$

□

Corollary 8.18. *For*

$$f(x,m,n)=\frac{\mu(m)\mu(n)}{mn}\left(\left\{\frac{x}{n}\right\}\left\{\frac{x}{m}\right\}-\left\{\frac{2x}{n}\right\}\left\{\frac{x}{m}\right\}-\left\{\frac{x}{n}\right\}\left\{\frac{2x}{m}\right\}+\left\{\frac{2x}{n}\right\}\left\{\frac{2x}{m}\right\}\right),$$

we have

$$\lim_{M,N\to\infty}\sum_{m=1}^{M}\sum_{n=1}^{N}\left|\lim_{X\to\infty}\frac{1}{X}\sum_{x=1}^{X}f(x,m,n)\right|<\infty.$$

Proof. Following the exact reasoning as in the proof of theorem 8.13, we see by applying corollary 8.5, corollary 8.9 and lemma 8.10 that

$$\begin{aligned}\lim_{M,N\to\infty}&\sum_{m=1}^{M}\sum_{n=1}^{N}\left|\lim_{X\to\infty}\frac{1}{X}\sum_{x=1}^{X}f(x,m,n)\right|\\&=\sum_{\substack{m\text{ odd},\\ n\text{ odd}}}\frac{|\mu(m)\mu(n)|}{m^2n^2}\left(\frac{\gcd(m,n)^2-1}{12}\right)-\sum_{\substack{m\text{ even},\\ n\text{ odd}}}\frac{|\mu(m)\mu(n)|}{m^2n^2}\left(\frac{\gcd(m,n)^2-1}{24}\right)\\&\quad-\sum_{\substack{m\text{ odd},\\ n\text{ even}}}\frac{|\mu(m)\mu(n)|}{m^2n^2}\left(\frac{\gcd(m,n)^2-1}{24}\right)+\sum_{\substack{m\text{ odd},\\ n\text{ odd}}}\frac{|\mu(m)\mu(n)|}{m^2n^2}\left(\frac{\gcd(m,n)^2+2}{12}\right).\end{aligned}$$

Now the respective summations converge absolutely by lemma 8.17. □

Lemma 8.19. *The expression*

$$\lim_{N\to\infty}\sum_{n=1}^{N}\frac{\mu(n)}{n}$$

does not converge absolutely.

Proof. Note that

$$\sum_{n=1}^{N}\frac{|\mu(n)|}{n}\geq\sum_{p_i\leq N}\frac{1}{p_i},$$

where the last sum is taken over all primes smaller or equal to N. This inequality follows from the fact that for every prime p_i, $|\mu(p_i)|=1$. It is well known that the latter series diverges as $N\to\infty$. We give the proof as is written in [1]. If the series does converge, then there exist a M s.t.

$$\sum_{p_i>M}\frac{1}{p_i}<\frac{1}{2}.$$

Define $Q=\prod_{p_i\leq M}p_i$. If a prime p divides the integer $1+nQ$ for some n, then by definition of Q it follows that $p>M$. Hence

$$\sum_{n=1}^{N}\frac{1}{1+nQ}\leq\sum_{t=1}^{\infty}\left(\sum_{p_i>M}\frac{1}{p_i}\right)^t<\sum_{t=1}^{\infty}\left(\frac{1}{2}\right)^t=1.$$

However

$$\sum_{n=1}^{N}\frac{1}{1+nQ}\geq\int_1^N\frac{dx}{1+xQ}=\frac{\log(1+xQ)}{Q}\bigg|_{x=1}^{x=N}\geq\frac{\log(1+NQ)}{Q}.$$

This diverges as $N\to\infty$, so we have derived a contracdiction. □

Corollary 8.20. *For*

$$f(x,m,n) = \frac{\mu(m)\mu(n)}{mn}\left\{\frac{x}{n}\right\}\left\{\frac{x}{m}\right\},$$

the expression

$$\lim_{M,N\to\infty}\sum_{m=1}^{M}\sum_{n=1}^{N}\left|\lim_{X\to\infty}\frac{1}{X}\sum_{x=1}^{X}f(x,m,n)\right|$$

does not converge.

Proof. Following the proof of theorem 8.7, we see that

$$\begin{aligned}
&\lim_{M,N\to\infty}\sum_{m=1}^{M}\sum_{n=1}^{N}\left|\lim_{X\to\infty}\frac{1}{X}\sum_{x=1}^{X}f(x,m,n)\right| \\
&\qquad= \sum_{m,n=1}^{\infty}\frac{|\mu(m)\mu(n)|}{m^2n^2}\left(\frac{(m-1)(n-1)}{4}+\frac{\gcd(m,n)^2-1}{12}\right) \\
&\qquad\geq \sum_{m,n=1}^{\infty}\frac{|\mu(m)\mu(n)|}{m^2n^2}\left(\frac{mn}{4}+\frac{\gcd(m,n)^2}{12}\right) \\
&\qquad= \frac{1}{12}\sum_{m,n}^{\infty}\frac{|\mu(m)\mu(n)|}{m^2n^2}\cdot\gcd(m,n)^2+\frac{1}{4}\sum_{m,n}^{\infty}\frac{|\mu(m)\mu(n)|}{mn}.
\end{aligned}$$

Now the last sum does not converge by lemma 8.19. □

8.6 The variance of $\sum\frac{\varphi(n)}{n}$ in short intervals

The last section of this chapter is maybe even less definitive than the other sections. We calculate the variance of $\sum\frac{\varphi(n)}{n}$ in the interval $[x,x+H]$ where $H=\Theta(x^\delta)$. Again we look at $\mathbb{E}_{x\in\mathbb{Z}_{\geq 1}}\left[\left\{\frac{x+H}{n}\right\}\left\{\frac{x}{m}\right\}\right]$, but in this section we see this as an entirely stochastic process, instead of a deterministic sum. (So we calculate the chance that $\left\{\frac{x+H}{n}\right\}\left\{\frac{x}{m}\right\}$ is some value and then use this to calculate the 'expectation value'). This enables us to actually calculate it, which we would not be able to do otherwise.

As noted at the start of section 8.4, we do not know anything about $[x^\delta] \bmod m$ if we know $x \bmod n$ for some m,n. This is easily seen. For example if $\delta=\frac{1}{2}$, then $y=[x^\delta]$ implies $y^2\leq x<(y+1)^2$. Since the gaps between y^2 and $(y+1)^2$ get larger and larger, at some point the gaps will be much bigger than n. Hence if we fix some large $y\in\mathbb{Z}$ and let $x\in\mathbb{Z}$ range in between y^2 and $(y+1)^2$, we will find every possible value for $x \bmod n$ and with about the same probability, (because the gap is so large). This qualitative argument persuades us to make the following assumption.

Assumption 8.21. *Fix $m,n\in\mathbb{Z}_{\geq 1}$. For any $0<\delta<1$, there exists no correlation between $x \bmod n$ and $[x^\delta] \bmod m$.*

We will need this assumption in this entire section. It will enable us to make predictions about the variance of $\sum\frac{\varphi(n)}{n}$ in the interval $[x,x+H]$ for $H=\Theta(x^\delta)$. We will apply the same techniques

as in section 8.4, so again we will only be ably to calculate

$$\sum_{m,n=1}^{\infty} \frac{\mu(m)\mu(n)}{mn} \left(\mathop{\mathbb{E}}_{x\in\mathbb{Z}_{\geq 1}} \left[\left\{ \frac{x}{n} \right\} \left\{ \frac{x}{m} \right\} \right] - \mathop{\mathbb{E}}_{x\in\mathbb{Z}_{\geq 1}} \left[\left\{ \frac{x+H}{n} \right\} \left\{ \frac{x}{m} \right\} \right] \right.$$
$$\left. - \mathop{\mathbb{E}}_{x\in\mathbb{Z}_{\geq 1}} \left[\left\{ \frac{x}{n} \right\} \left\{ \frac{x+H}{m} \right\} \right] + \mathop{\mathbb{E}}_{x\in\mathbb{Z}_{\geq 1}} \left[\left\{ \frac{x+H}{n} \right\} \left\{ \frac{x+H}{m} \right\} \right] \right)$$

and show that the numerical simulations suggest $\frac{1}{X} \sum_{x=1}^{X} \left(\sum_{n=x+1}^{x+H} \frac{\varphi(n)}{n} - \frac{x}{\zeta(2)} \right)^2$ to converge to the same value. We first start by taking $H = [x^\delta]$ and show that our assumption predicts the same value $\frac{1}{6\zeta(2)} - \frac{1}{\zeta(2)^2}$ as derived in section 8.4 for the variance in the interval $[x, 2x]$. Note that by assumption 8.21 the specific value of δ does not matter.

Lemma 8.22. *Assuming 8.21, we have for any m, n*

$$\mathop{\mathbb{E}}_{x\in\mathbb{Z}_{\geq 1}} \left[\left\{ \frac{x}{m} \right\} \left\{ \frac{x+[x^\delta]}{n} \right\} \right] = \frac{1}{mn} \left(\frac{(m-1)(n-1)}{4} \right).$$

Proof. By assumption 8.21 there is no correlation between $x \bmod m$ and $[x^\delta] \bmod n$. Hence any pair $(x \bmod m, x + [x^\delta] \bmod n)$ is attained with equal probability. Since there are mn different such pairs, we conclude that

$$\mathop{\mathbb{E}}_{x\in\mathbb{Z}_{\geq 1}} \left[\left\{ \frac{x}{m} \right\} \left\{ \frac{x+[x^\delta]}{n} \right\} \right] = \frac{1}{mn} \left(\sum_{r_m=0}^{m-1} \frac{r_m}{m} \right) \left(\sum_{r_n=0}^{n-1} \frac{r_n}{n} \right) = \frac{1}{mn} \left(\frac{(m-1)(n-1)}{4} \right).$$

□

Theorem 8.23. *Assuming 8.21, we have*

$$\sum_{m,n=1}^{\infty} \frac{\mu(m)\mu(n)}{mn} \left(\mathop{\mathbb{E}}_{x\in\mathbb{Z}_{\geq 1}} \left[\left\{ \frac{x}{n} \right\} \left\{ \frac{x}{m} \right\} \right] - \mathop{\mathbb{E}}_{x\in\mathbb{Z}_{\geq 1}} \left[\left\{ \frac{x+[x^\delta]}{n} \right\} \left\{ \frac{x}{m} \right\} \right] \right.$$
$$\left. - \mathop{\mathbb{E}}_{x\in\mathbb{Z}_{\geq 1}} \left[\left\{ \frac{x}{n} \right\} \left\{ \frac{x+[x^\delta]}{m} \right\} \right] + \mathop{\mathbb{E}}_{x\in\mathbb{Z}_{\geq 1}} \left[\left\{ \frac{x+[x^\delta]}{n} \right\} \left\{ \frac{x+[x^\delta]}{m} \right\} \right] \right)$$
$$= \frac{1}{6\zeta(2)} - \frac{1}{6\zeta(2)^2}.$$

Proof. Note that by assumption 8.21 and corollary 8.5

$$\mathop{\mathbb{E}}_{x\in\mathbb{Z}_{\geq 1}} \left[\left\{ \frac{x+[x^\delta]}{n} \right\} \left\{ \frac{x+[x^\delta]}{m} \right\} \right] = \mathop{\mathbb{E}}_{x\in\mathbb{Z}_{\geq 1}} \left[\left\{ \frac{x}{n} \right\} \left\{ \frac{x}{m} \right\} \right] = \frac{1}{mn} \left(\frac{(m-1)(n-1)}{4} + \frac{\gcd(m,n)^2 - 1}{12} \right).$$

Substituting this, together with lemma 8.22, we see that

$$\sum_{m,n=1}^{\infty} \frac{\mu(m)\mu(n)}{mn} \left(\mathop{\mathbb{E}}_{x\in\mathbb{Z}_{\geq 1}} \left[\left\{ \frac{x}{n} \right\} \left\{ \frac{x}{m} \right\} \right] - \mathop{\mathbb{E}}_{x\in\mathbb{Z}_{\geq 1}} \left[\left\{ \frac{x+[x^\delta]}{n} \right\} \left\{ \frac{x}{m} \right\} \right] \right.$$
$$\left. - \mathop{\mathbb{E}}_{x\in\mathbb{Z}_{\geq 1}} \left[\left\{ \frac{x}{n} \right\} \left\{ \frac{x+[x^\delta]}{m} \right\} \right] + \mathop{\mathbb{E}}_{x\in\mathbb{Z}_{\geq 1}} \left[\left\{ \frac{x+[x^\delta]}{n} \right\} \left\{ \frac{x+[x^\delta]}{m} \right\} \right] \right)$$
$$= \sum_{m,n=1}^{\infty} \frac{\mu(m)\mu(n)}{m^2n^2} \left(\frac{\gcd(m,n)^2 - 1}{6} \right)$$
$$= \frac{1}{6\zeta(2)} - \frac{1}{6\zeta(2)^2}.$$

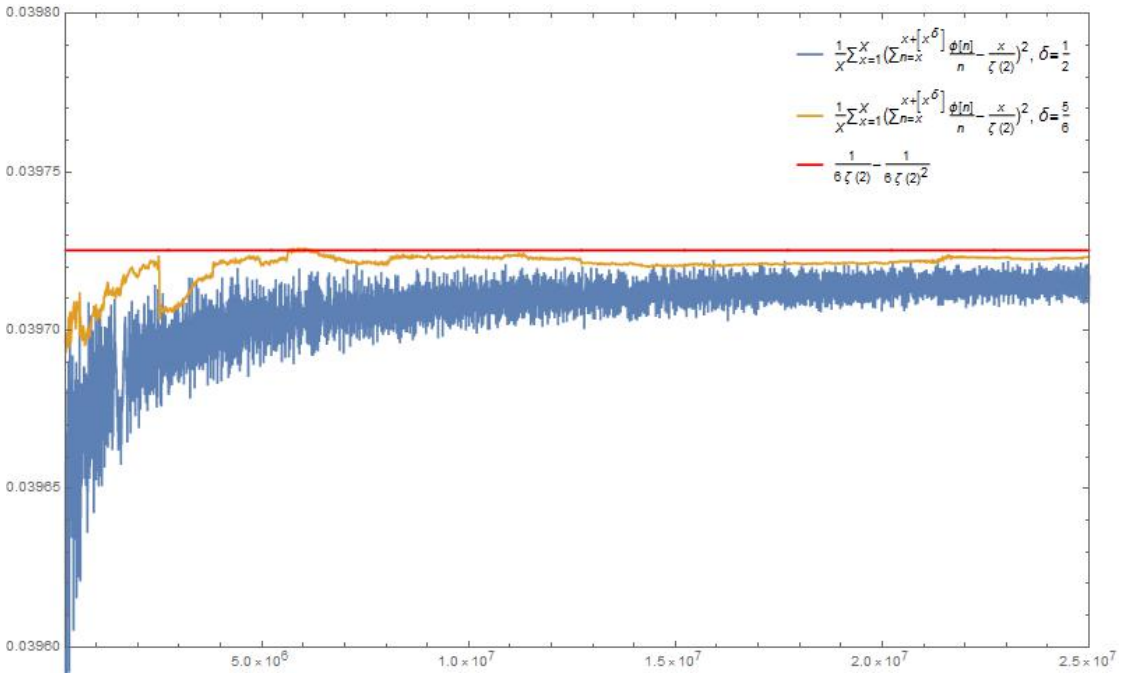

Figure 8.3: $\frac{1}{X}\sum_{x=1}^{X}\left(\sum_{n=x}^{x+2[x^\delta]}\frac{\varphi(n)}{n}-\frac{x}{\zeta(2)}\right)^2$, where X ranges from 0 to $2.5\cdot 10^7$ and $\delta=\frac{1}{2},\frac{5}{6}$; together with the predicted value for the variance.

Here we applied lemma 8.6 and corollary 2.14. □

Again we make another assumption to make this into a final theorem.

Assumption 8.24. *Changing the two limits in the expression*

$$\lim_{M,N\to\infty}\sum_{m=1}^{M}\sum_{n=1}^{N}\lim_{X\to\infty}\frac{1}{X}\sum_{x=1}^{X}f(x,m,n)$$

does not change its value, for $f(x,m,n)$ equal to

$$\frac{\mu(m)\mu(n)}{mn}\left(\left\{\frac{x}{n}\right\}\left\{\frac{x}{m}\right\}-\left\{\frac{x+[x^\delta]}{n}\right\}\left\{\frac{x}{m}\right\}-\left\{\frac{x}{n}\right\}\left\{\frac{x+[x^\delta]}{m}\right\}+\left\{\frac{x+[x^\delta]}{n}\right\}\left\{\frac{x+[x^\delta]}{m}\right\}\right).$$

Theorem 8.25 (Restatement of theorem 1.3)**.**
Let $H=\Theta(x^\delta)$, $0<\delta\leq 1$. Assuming 8.21 and 8.24, we find

$$\frac{1}{X}\sum_{x=1}^{X}\left[\left(\sum_{x<n<x+H}\frac{\varphi(n)}{n}-\frac{H}{\zeta(2)}\right)^2\right]\xrightarrow{X\to\infty}\frac{1}{6\zeta(2)}-\frac{1}{6\zeta(2)^2}.$$

In figure 8.3 you see the numerical simulations for different $0<\delta<1$ for X up to $2.5\cdot 10^7$. Clearly the convergence is much slower than in the $[x,2x]$-case. Even more, for $\delta=\frac{1}{4}$, (not shown in the figure), it is not even clear whether we have convergence at all, although the variance does seem to oscillate around the right value. A reason could be that $x\leq X$ must be very large,

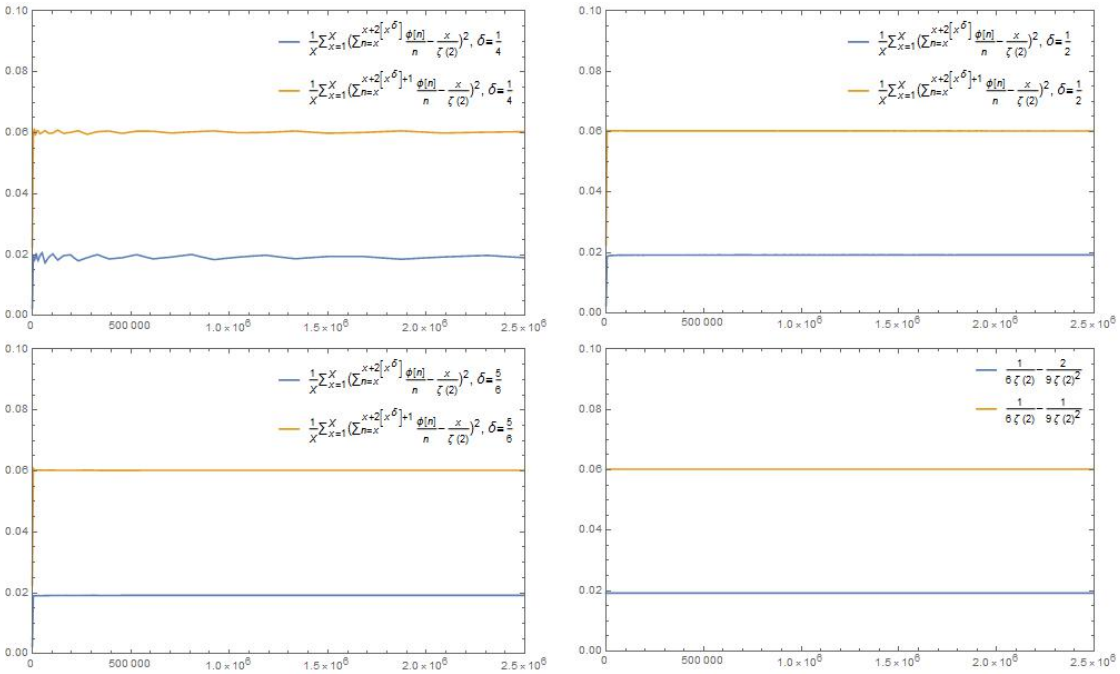

Figure 8.4: $\frac{1}{X}\sum_{x=1}^{X}\left(\sum_{n=x}^{x+2[x^\delta]}\frac{\varphi(n)}{n}-\frac{x}{\zeta(2)}\right)^2$ and $\frac{1}{X}\sum_{x=1}^{X}\left(\sum_{n=x}^{x+2[x^\delta]+1}\frac{\varphi(n)}{n}-\frac{x}{\zeta(2)}\right)^2$, where X ranges from 0 to $2.5\cdot 10^6$ and $\delta=\frac{1}{4},\frac{1}{2},\frac{5}{6}$; together with the predicted values for the respective variance.

before $x \mod n$ and $[x^\delta] \mod m$ become uncorrelated. Another reason might be that it turns out the parity of H strongly influences the variance. This is seen when you take $H=2[x^\delta]$ or $H=2[x^\delta]+1$, as in figure 8.4. For the even H the variance is much lower then for the odd H. It turns out that our model predicts the same results!

Lemma 8.26. *Assuming 8.21, we have for any m,n*

$$\mathop{\mathbb{E}}_{x\in\mathbb{Z}_{\geq 1}}\left[\left\{\frac{x}{m}\right\}\left\{\frac{x+2[x^\delta]}{n}\right\}\right]=\begin{cases}\frac{1}{mn}\left(\frac{(m-1)(n-1)}{4}+\frac{1}{4}\right) & \text{if } m \text{ is even, } n \text{ is even}\\ \frac{1}{mn}\left(\frac{(m-1)(n-1)}{4}\right) & \text{otherwise.}\end{cases}$$

Proof. If n is odd, then the statement is clear by the same arguments as in the proof of lemma 8.22. Suppose that m is odd and n is even. Then if $x \mod m$ is known, there are a $\frac{n}{d}$ possible values for $x \mod n$, where $d=\gcd(m,n)$. Exactly half of these values are odd, and half of these values are even. Since by assumption 8.21, there is no correlation between $x \mod m$ and $[x^\delta] \mod n$, we only know that $2[x^\delta] \mod n$ is even and every even value is equally probable. From this we conclude that if we know $x \mod m$, then every value of $x+2[x^\delta] \mod n$ is equally probable. The statement of the lemma follows in the same way as in the proof of lemma 8.22. Finally if both m and n are even, then $x \mod m$ and $x \mod n$ have the same parity. Hence $x \mod m$ and $x+2[x^\delta] \mod n$ have the same parity. There are $\frac{2}{mn}$ pairs $(r_m \mod m, r_n \mod n)$

with the same parity and each of these pairs is equally probable. We conclude that

$$\begin{aligned}
&\mathop{\mathbb{E}}_{x\in\mathbb{Z}_{\geq 1}}\left[\left\{\frac{x}{m}\right\}\left\{\frac{x+2[x^\delta]}{n}\right\}\right] \\
&\quad= \frac{2}{mn}\sum_{l=0}^{1}\left(\sum_{\substack{0\leq r_m<m\\ r_m\equiv l \bmod 2}}\frac{r_m}{m}\right)\left(\sum_{\substack{0\leq r_n<n\\ r_n\equiv l \bmod 2}}\frac{r_n}{n}\right) \\
&\quad= \frac{2}{mn}\left(\sum_{t_m=0}^{\frac{m}{2}-1}\frac{2t_m}{m}\right)\left(\sum_{t_n=0}^{\frac{n}{2}-1}\frac{2t_n}{n}\right)+\frac{2}{mn}\left(\sum_{t_m=0}^{\frac{m}{2}-1}\frac{1+2t_m}{m}\right)\left(\sum_{t_n=0}^{\frac{n}{2}-1}\frac{1+2t_n}{n}\right) \\
&\quad= \frac{2}{mn}\left(\frac{2}{m}\sum_{t_m=0}^{\frac{m}{2}-1}t_m\right)\left(\frac{2}{n}\sum_{t_n=0}^{\frac{n}{2}-1}t_n\right)+\frac{2}{mn}\left(\frac{1}{2}+\frac{2}{m}\sum_{t_m=0}^{\frac{m}{2}-1}t_m\right)\left(\frac{1}{2}+\frac{2}{n}\sum_{t_n=0}^{\frac{n}{2}-1}t_n\right) \\
&\quad\overset{(a)}{=} \frac{2}{mn}\left(\frac{(m-2)(n-2)}{16}+\left(\frac{1}{2}+\frac{m-2}{4}\right)\left(\frac{1}{2}+\frac{n-2}{4}\right)\right) \\
&\quad= \frac{1}{mn}\left(\frac{1}{2}+\frac{m+n-4}{4}+\frac{(m-2)(n-2)}{4}\right) \\
&\quad= \frac{1}{mn}\left(\frac{(m-1)(n-1)}{4}+\frac{1}{4}\right).
\end{aligned}$$

Again at (a) we applied the identity $\sum_{k=1}^{n}k=\frac{n(n+1)}{2}$ □

Lemma 8.27. *Assuming 8.21, we have for any m,n*

$$\mathop{\mathbb{E}}_{x\in\mathbb{Z}_{\geq 1}}\left[\left\{\frac{x}{m}\right\}\left\{\frac{x+2[x^\delta]+1}{n}\right\}\right]=\begin{cases}\frac{1}{mn}\left(\frac{(m-1)(n-1)}{4}-\frac{1}{4}\right) & \text{if } m \text{ is even, } n \text{ is even}\\ \frac{1}{mn}\left(\frac{(m-1)(n-1)}{4}\right) & \text{otherwise.}\end{cases}$$

Proof. If either m or n is odd, we could use the same arguments as in the proof of lemma 8.26 to show the statement. If both m and n are odd, then $x \bmod m$ and $x+2[x^\delta]+1 \bmod n$ have different parity, so we conclude that

$$\begin{aligned}
&\mathop{\mathbb{E}}_{x\in\mathbb{Z}_{\geq 1}}\left[\left\{\frac{x}{m}\right\}\left\{\frac{x+2[x^\delta]+1}{n}\right\}\right] \\
&\quad= \frac{2}{mn}\sum_{l=0}^{1}\left(\sum_{\substack{0\leq r_m<m\\ r_m\equiv l \bmod 2}}\frac{r_m}{m}\right)\left(\sum_{\substack{0\leq r_n<n\\ r_n\not\equiv l \bmod 2}}\frac{r_n}{n}\right) \\
&\quad= \frac{2}{mn}\left(\left(\sum_{r_m=0}^{m-1}\frac{r_m}{m}\right)\left(\sum_{r_n=0}^{n-1}\frac{r_n}{n}\right)-\left(\sum_{l=0}^{1}\left(\sum_{\substack{0\leq r_m<m\\ r_m\equiv l \bmod 2}}\frac{r_m}{m}\right)\left(\sum_{\substack{0\leq r_n<n\\ r_n\not\equiv l \bmod 2}}\frac{r_n}{n}\right)\right)\right) \\
&\quad= \frac{1}{mn}\left(\frac{(m-1)(n-1)}{4}-\frac{1}{4}\right),
\end{aligned}$$

applying the previous lemma. □

Theorem 8.28. *Assuming 8.21, we have*

$$\sum_{m,n=1}^{\infty} \frac{\mu(m)\mu(n)}{mn} \left(\mathop{\mathbb{E}}_{x\in\mathbb{Z}_{\geq 1}} \left[\left\{ \frac{x}{n} \right\} \left\{ \frac{x}{m} \right\} \right] - \mathop{\mathbb{E}}_{x\in\mathbb{Z}_{\geq 1}} \left[\left\{ \frac{x+2[x^\delta]}{n} \right\} \left\{ \frac{x}{m} \right\} \right] \right.$$
$$\left. - \mathop{\mathbb{E}}_{x\in\mathbb{Z}_{\geq 1}} \left[\left\{ \frac{x}{n} \right\} \left\{ \frac{x+2[x^\delta]}{m} \right\} \right] + \mathop{\mathbb{E}}_{x\in\mathbb{Z}_{\geq 1}} \left[\left\{ \frac{x+2[x^\delta]}{n} \right\} \left\{ \frac{x+2[x^\delta]}{m} \right\} \right] \right)$$
$$= \frac{1}{6\zeta(2)} - \frac{2}{9\zeta(2)^2}$$

Proof. Applying lemma 8.26, we see that the value we want to calculate equals

$$\sum_{m,n=1}^{\infty} \frac{\mu(m)\mu(n)}{m^2n^2} \left(\frac{\gcd(m,n)^2 - 1}{6} \right) - \frac{1}{2} \sum_{\substack{m \text{ even}, \\ n \text{ even}}} \frac{\mu(m)\mu(n)}{m^2n^2} = \frac{1}{6\zeta(2)} - \frac{2}{9\zeta(2)^2}.$$

At the last equality we applied corollary 2.14, lemma 8.6 and lemma 8.12. □

Theorem 8.29. *Assuming 8.21, we have*

$$\sum_{m,n=1}^{\infty} \frac{\mu(m)\mu(n)}{mn} \left(\mathop{\mathbb{E}}_{x\in\mathbb{Z}_{\geq 1}} \left[\left\{ \frac{x}{n} \right\} \left\{ \frac{x}{m} \right\} \right] - \mathop{\mathbb{E}}_{x\in\mathbb{Z}_{\geq 1}} \left[\left\{ \frac{x+2[x^\delta]+1}{n} \right\} \left\{ \frac{x}{m} \right\} \right] \right.$$
$$\left. - \mathop{\mathbb{E}}_{x\in\mathbb{Z}_{\geq 1}} \left[\left\{ \frac{x}{n} \right\} \left\{ \frac{x+2[x^\delta]+1}{m} \right\} \right] + \mathop{\mathbb{E}}_{x\in\mathbb{Z}_{\geq 1}} \left[\left\{ \frac{x+2[x^\delta]+1}{n} \right\} \left\{ \frac{x+2[x^\delta]+1}{m} \right\} \right] \right)$$
$$= \frac{1}{6\zeta(2)} - \frac{1}{9\zeta(2)^2}$$

Proof. Applying lemma 8.27, we see that the value we want to calculate equals

$$\sum_{m,n=1}^{\infty} \frac{\mu(m)\mu(n)}{m^2n^2} \left(\frac{\gcd(m,n)^2 - 1}{6} \right) + \frac{1}{2} \sum_{\substack{m \text{ even}, \\ n \text{ even}}} \frac{\mu(m)\mu(n)}{m^2n^2} = \frac{1}{6\zeta(2)} - \frac{1}{9\zeta(2)^2}.$$

Again we applied corollary 2.14, lemma 8.6 and lemma 8.12. □

Chapter 9

Concluding remarks

9.1 The relation between theorems 1.2 and 1.3

In this thesis we have seen a couple of examples of analogous statements in $\mathbb{Z}$ and $\mathbb{F}_q[T]$, such as the prime number theorem 2.20 and the prime polynomial theorem 5.1. We have also seen relation (1.1) in the introduction of this thesis and its analogue in $\mathbb{F}_q[T]$, theorem 1.1. It might be surprising that the results about the variance of the Euler totient function, in particular theorem 1.2 and the theorem 1.3, are not analogous. This is due to the high amount of cancellation in $\mathbb{F}_q[T]$. Recall lemma 7.6, stating that $\mathcal{M}(m;\varphi\chi)$ is a finite polynomial in $q^{\frac{1}{2}}$ of degree $3m$ if $m \leq N$, and of degree $m+2N+1$ if $m \geq N+1$. In chapter 7 the next step was to focus on the cancellation of the coefficients, before we went on to calculate the variance of φ. If one ignores the cancellation and you would immediately calculate the variance, you would find that

$$\mathrm{Var}_n \mathcal{N}_{\frac{\varphi(f)}{|f|}}(\bullet; h) \sim \frac{1}{q}.$$

In this case you would find an analogue of theorem 1.3, (up to a constant), as $\frac{1}{\zeta_q(2)} - \frac{1}{\zeta_q(2)^2} = \frac{1}{q} - \frac{1}{q^2}$. We conclude that the analogue of this theorem does not hold in $\mathbb{F}_q[T]$ due to the high amount of cancellation.

Note that this is not new. We give one easy example of a similar phenomenon. Recall that the statement $\sum_{n<x} \mu(n) = o(x)$ is equivalent to the prime number theorem. In number theory we also have a conjecture stating that $\sum_{n<x} \mu(n) = O\left(x^{\frac{1}{2}+\epsilon}\right)$ for any $\epsilon > 0$. It turns out that this statement is equivalent to the Riemann hypothesis, see for example [14]. Now in $\mathbb{F}_q[T]$ it turns out that a much stronger statement is true, due to the high amount of cancellation in this ring. It turns out that for any $n \geq 2$

$$\sum_{f \in \mathcal{M}_n} \mu(f) = 0.$$

A proof of this is for example given in [3].

Bibliography

[1] T. M. Apostol. *Introduction to analytic number theory.* Springer-Verlag, New York-Heidelberg, 1976. Undergraduate Texts in Mathematics.

[2] R. Ayoub. *An introduction to the analytic theory of numbers.* Mathematical Surveys, No. 10. American Mathematical Society, Providence, R.I., 1963.

[3] S. Bae, B. Cha, and H. Jung. Möbius function in short intervals for function fields. *Finite Fields Appl.*, 34:235–249, 2015.

[4] A. Browder. *Mathematical analysis.* Undergraduate Texts in Mathematics. Springer-Verlag, New York, 1996. An introduction.

[5] H. Davenport. *Multiplicative number theory*, volume 74 of *Graduate Texts in Mathematics.* Springer-Verlag, New York-Berlin, second edition, 1980. Revised by Hugh L. Montgomery.

[6] C. J. de La Vallée Poussin. Recherches analytiques sur la théorie des nombres; Première partie: La fonction $\zeta(s)$ de Riemann et les nombres premiers en général. *Analles de la Soc. scientifique de Bruxelles*, 20_2:183–256, 1896.

[7] F. J. Dyson. Correlations between eigenvalues of a random matrix. *Comm. Math. Phys.*, 19:235–250, 1970.

[8] P. Etingof, O. Golberg, S. Hensel, T. Liu, A. Schwendner, D. Vaintrob, and E. Yudovina. *Introduction to representation theory*, volume 59 of *Student Mathematical Library.* American Mathematical Society, Providence, RI, 2011.

[9] W. Fulton and J. Harris. *Representation theory*, volume 129 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, 1991. A first course, Readings in Mathematics.

[10] A. Gamburd. Some applications of symmetric functions theory in random matrix theory. In *Ranks of elliptic curves and random matrix theory*, volume 341 of *London Math. Soc. Lecture Note Ser.*, pages 143–169. Cambridge Univ. Press, Cambridge, 2007.

[11] C. F. Gauss. *Untersuchungen über höhere Arithmetik.* Deutsch herausgegeben von H. Maser. Chelsea Publishing Co., New York, 1965.

[12] D. A. Goldston and H. L. Montgomery. Pair correlation of zeros and primes in short intervals. In *Analytic number theory and Diophantine problems (Stillwater, OK, 1984)*, volume 70 of *Progr. Math.*, pages 183–203. Birkhäuser Boston, Boston, MA, 1987.

[13] J. Hadamard. Sur la distribution des zéros de la fonction $\zeta(s)$ et ses conséquences arithmétiques. *Bull. Soc. Math. France*, 24:199–220, 1896.

[14] H. Iwaniec and E. Kowalski. *Analytic number theory*, volume 53 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2004.

[15] N. M. Katz. On a question of Keating and Rudnick about primitive Dirichlet characters with squarefree conductor. *Int. Math. Res. Not. IMRN*, (14):3221–3249, 2013.

[16] N. M. Katz. Witt vectors and a question of Keating and Rudnick. *Int. Math. Res. Not. IMRN*, (16):3613–3638, 2013.

[17] J. Keating. L-functions and random matrix theory. International Conference in Number Theory and Physics, 2015. `http://video.impa.br/index.php?page=international-conference-in-number-theory-and-physics`.

[18] J. Keating, B. Rodgers, E. Roditty-Gershon, and Z. Rudnick. Sums of divisor functions in $F_q[t]$ and matrix integrals. `https://arxiv.org/abs/1504.07804`.

[19] J. Keating and Z. Rudnick. Squarefree polynomials and Möbius values in short intervals and arithmetic progressions. *Algebra Number Theory*, 10(2):375–420, 2016.

[20] J. P. Keating and Z. Rudnick. The variance of the number of prime polynomials in short intervals and in residue classes. *Int. Math. Res. Not. IMRN*, (1):259–288, 2014.

[21] S. J. Miller and R. Takloo-Bighash. *An invitation to modern number theory*. Princeton University Press, Princeton, NJ, 2006.

[22] H. L. Montgomery. The pair correlation of zeros of the zeta function. pages 181–193, 1973.

[23] M. Rosen. *Number theory in function fields*, volume 210 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2002.

[24] W. Rossmann. *Lie groups*, volume 5 of *Oxford Graduate Texts in Mathematics*. Oxford University Press, Oxford, 2002. An introduction through linear groups.

[25] C. Teleman. Representation theory. `https://math.berkeley.edu/~teleman/math/RepThry.pdf`, 2005.

[26] A. Weil. Numbers of solutions of equations in finite fields. *Bull. Amer. Math. Soc.*, 55:497–508, 1949.

[27] H. Weyl. Theorie der Darstellung kontinuierlicher halb-einfacher Gruppen durch lineare Transformationen. I. *Math. Z.*, 23(1):271–309, 1925.

[28] H. Weyl. Theorie der Darstellung kontinuierlicher halb-einfacher Gruppen durch lineare Transformationen. II. *Math. Z.*, 24(1):328–376, 1926.

[29] H. Weyl. Theorie der Darstellung kontinuierlicher halb-einfacher Gruppen durch lineare Transformationen. III. *Math. Z.*, 24(1):377–395, 1926.