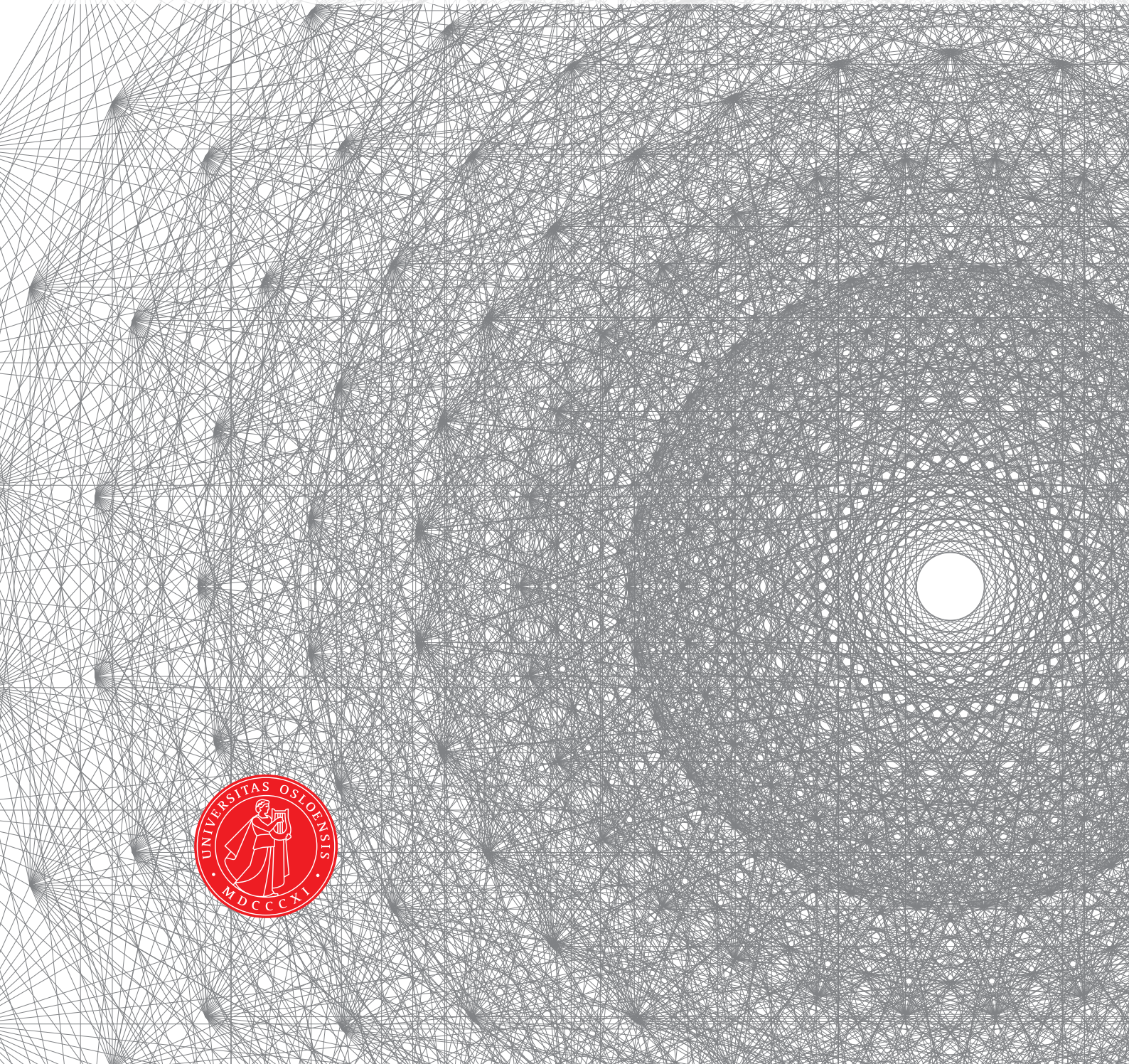
UiO : Department of Mathematics
University of Oslo

Dirichlet's Theorem on primes in arithmetic progressions, and Chebyshev's Bias

Lars Tore Spånberg Slettan

Master's Thesis, Spring 2018



This master's thesis is submitted under the master's programme *Mathematics*, with programme option *Mathematics*, at the Department of Mathematics, University of Oslo. The scope of the thesis is 60 credits.

The front page depicts a section of the root system of the exceptional Lie group E_8, projected into the plane. Lie groups were invented by the Norwegian mathematician Sophus Lie (1842–1899) to express symmetries in differential equations and today they play a central role in various parts of mathematics.

Abstract

In this master thesis we will prove Dirichlet's theorem on primes in arithmetic progressions. Along with the proof, we show a simpler way to derive the theorem for the cases modulo $q = 3, 4, 6$. The reader will be presented by a formula for the number of zeros of the Riemann zeta function $\zeta(s)$. As an extension to Dirichlet's theorem, we will also highlight its connection with Frobenius density theorem and Chebotarëv's density theorem. In addition, we include a discussion on Chebyshev's bias.

Acknowledgements

After two years of working towards a finished master thesis, it is now time for me to raise my head and look back at the people who made this possible. The first name that comes to mind is of course my helpful supervisor Geir Ellingsrud. I am grateful for the level of availability that you have presented me over the last year. Along with correcting my mathematics at multiple occasions, you have also functioned as an english teacher as my thesis was coming to an end.

Secondly, I want to show my appreciation to the people that I shared the study hall B601 alongside with in my first year at UiO. You guys provided great support when dealing with mathematical challenges, as well as providing the occasional laugh.

At last I want to thank three people who has supported me, although not in a mathematical fashion. Firstly, I give a big thanks to my mother Dagny, who has been there for me my whole life. The next person that has encouraged me in my academic pursuits is Simen Nygård Stubbe. I'm grateful of our many talks these two past years. Lastly, I want to thank my cousin Ellen Galaasen and her family for letting me be a part of their home for several months during my first time in Oslo.

Lars Tore Spånberg Slettan
Oslo, May 2018.

Contents

Chapter 1

Introduction

It is a well known fact that there are an infinte number of primes. The first proof of this fact is credited to Euclid, found in his *Elements* around 300 BC. Along with the development of new branches of mathematics has given rise to a multitude of other ways to prove the infinitude of primes. The *prime counting function* $\pi(x)$, defined by

$$\pi(x) := \big|\{p \leq x : p \text{ prime}\}\big|,$$

is a natural object to study. The *prime number theorem* asserts $\pi(x) \sim x/log\ x$, conjectured by Legendre and Gauss, and ultimately proved independently a century later by Hadamard and de la Vallée Poussin in 1896.

In 1837 Dirichlet proved that there are an infinite number of primes in every primitive residue class a modulo q, for any natural number $q \geq 2$. A residue class $a(mod\ q)$ is called primitive if $gcd(a, q) = 1$. It is this theorem of Dirichlet that we will spend most of our attention on. Actually, the reader will be presented with two different theorems of Dirichlet. The first, *Dirichlet's theorem*, is the main theorem in the thesis and will be proven here. The other, *Dirichlet's density theorem*, is merely stated and will not be given a thoroughly treatise.

Dirichlet's theorem contains some intuition pointing to the primes being equidistributed into the primitive residue classes modulo q. It is this intuition that Dirichlet's density theorem makes precise. For a given natural number $q \geq 2$, then $\varphi(q)$ is the number of primitive residue classes modulo q, where φ is *Euler's totient function*. Then Dirichlet's density theorem states that the set of primes p with $p \equiv a(mod\ q)$ has density $1/\varphi(q)$, for $gcd(a, q) = 1$.

In connection to Dirichlet's density theorem we have *Frobenius density theorem*. Frobenius looked at polynomials $f(x) \in \mathbb{Z}[x]$, and his theorem concerns how $f(x)$ factors in $(\mathbb{Z}/p\mathbb{Z})[x]$ for various primes p. For certain integers $q \geq 2$, Frobenius density theorem actually implies Dirichlet's density theorem when choosing the right polynomial $f(x)$. Frobenius conjectured in 1880 what is now called *Chebotarëv's density theorem*, which implies Dirichlet's density theorem when applied to the polynomial $f(x) = x^q - 1$, providing an algebraic point of view on the subject.

Even though the primes are equidistributed by Dirichlet's density theorem, the russian Chebyshev found in 1853 that there are more primes congurent to $3(mod\ 4)$ than to $1(mod\ 4)$. This phenomenon, called *Chebyshev's bias*,

favours the quadratic non-residue classes over the quadratic residue classes for small $q \geq 2$, but vanishes when $q \to \infty$. An equivalent statement of the prime number theorem is $\pi(x) \sim li(x)$, where $li(x)$ is the *logarithmic integral function* defined by

$$li(x) := \int_0^x \frac{dt}{log\ t}.$$

Computation of $li(x)$ has given indication towards $li(x) > \pi(x)$ for all x, since no counterexample is known to this date. However, in 1914 Littlewood proved that $li(x) < \pi(x)$ for an infinite number of x. The work on Chebyshev bias shed some light on this matter, providing the estimate 0,00000026... for the logarithmic density of the set $\{x \in \mathbb{R}^+ : li(x) < \pi(x)\}$ in $\mathbb{R}^+$. An upper bound for the first member of this set was first found by Skewes in 1933, and have since been improved to $< 10^{370}$ due to te Riele, assuming the Riemann hypothesis.

Chapter 2 provides the reader with the necessary tools in later chapters. This will involve defining multiplicative functions, big $\mathcal{O}$ notation and the some different notions of the density of a set.

In chapter 3 we see the prime number theorem along with some other important results concerning the primes. Here is also a proof of Dirichlet's theorem for the case of $q = 3, 4, 6$ using cyclotomic ploynomials. In addition, we get into Euler factorization, the Chebychev functions $\theta(x), \psi(s)$ and the von Mangoldt function $\Lambda(x)$.

Chapter 4 contains Dirichlet's theorem along with its proof. The proof is built on the notion of characters of a group, and an adjustment to define the Dirichlet charaters. A portion of the attention will be spent on L-series $L(s, \chi)$, mostly showing that $L(s, \chi) \neq 0$ at $s = 1$ for both real and complex Dirichlet characters χ.

In chapter 5 we find an estimate for the number of non-trivial zeros of the Riemann zeta function $\zeta(s)$ with imaginary part $< T$, and this number is denoted $N(T)$. The approach rests on complex analysis. We also sketch how to find the similar formula for the number $N(T, \chi)$, which count the non-trivial zeros of the L-function $L(s, \chi)$ with imaginary part less than T in absolute value.

Chapter 6 is much alike as the previous chapter, giving a formula for the Chebychev function $\psi(x)$, again with the help of complex analysis. One can also use Dirichlet characters to define a function $\psi(x, \chi)$ and this chapter will end with a formula for this function.

In chapter 7 we give an algebraic point of view of Dirichlet's density theorem, in the form of Frobenius density theorem and Chebotarëv's density theorem. The theorems apply for polynomials with integer coefficients, and uses Galois theory in order to speak about the density of primes in the primitive residue classes. Unfortunately, Frobenius theorem cannot separate all primitive residue classes for certain q. This issue is what Chebotarëv's density theorem fixed, and hence gives Dirichlet's density theorem.

Chapter 8 we discuss the phenomenon named the Chebyshev bias, along with some results concerning the topic, and a proof of one of the statements. We also see some numerical examples, which give the reader an illustration of

the Chebyshev bias.

This master thesis is built primarily on four sources. First is Paul Pollack's book *Not always buried deep: A second course in elementary number theorey.* Most of the built up to Dirichlet's Theorem in chapters 2, 3 and 4 can be found in this book. For chapters 5 and 6, when dealing with complex analysis, I have used Harold Davenport's *Multiplicative number theorey* as a guideline. For the theorems of Frobenius and Chebotarëv I have used the article of P. Stevenhagen and H.W.Lenstra Jr. named *Chebotarëv and his density theorem.* Finally, when dealing with Chebychev's Bias, I utilized the article *Chebyshev's bias* by M.Rubinstein and P.Sarnak.

Chapter 2

Background

The purpose of this chapter is to provide the reader with the necessary means to follow the line of argument later on. First is the introduction of the arithmetic functions. Then we will look at the useful consept of big $\mathcal{O}$ notation, as this allows us to speak more precisely about asymptotic behaviour of functions. I have also put in a passage concerning density of a set.

2.1 Some useful definitions and results

Number theory is the study of the integers $\mathbb{Z}$, and among the integers we find the natural numbers $\mathbb{N}$. We define the natural numbers as the positive integers. That is, $0 \notin \mathbb{N}$. We define an *arithmetic function* f to be any function $f : \mathbb{N} \to \mathbb{C}$. If f is not identically zero and satisfy

$$f(ab) = f(a)f(b) \qquad \textit{if } gcd(a,b) = 1, \tag{2.1}$$

we call f a *multiplicative arithmetic function*. Here $gcd()$ is an abbreviation of the greatest common divisor. The function f is a *completely multiplicative arithmetic function*, or for short just completely multiplicative, if (2.1) holds for all $a \in \mathbb{N}$.

At numerous occasions we will rely on the big $\mathcal{O}$ notation. For that reason we have included the definition and some useful properties. Firstly, for two functions $f(x), g(x)$ we say that $f = \mathcal{O}(g)$ (read "f is big $\mathcal{O}$ of g") if

$$|f| \leq C|g| \quad \textit{for some positive constant } C, \textit{ and } x \textit{ big enough.}$$

This is to say that g is a function of greater or equal magnitude than f. We will also use the notation $f \ll g$, which suggest the same thing. The functions f and g will be defined on $\mathbb{R}$ (or some subset) with target set $\mathbb{R}$.

We will not give the big $\mathcal{O}$ notation a thoroughly attention, but merely state a couple of useful facts concerning the topic. Let f, g, h be functions. Then

- If $f = \mathcal{O}(g)$, then $h \cdot f = \mathcal{O}(h \cdot g)$.
- If $f = \mathcal{O}(g)$ and k a constant, then $k \cdot f = \mathcal{O}(g)$, i.e. constants can be discarded from the calculations.

- If $f = \sum_{i=1}^{n} f_i, \quad f_1 \ll f_2 \ll ... \ll f_n,$ and $f_n = \mathcal{O}(g)$, then $f = \mathcal{O}(g)$. In other words, if f is a sum of functions, then only the function-term with the greatest magnitude will determine $\mathcal{O}(f)$.
- The $\ll$-relation is transetive, i.e. $(f \ll g) \wedge (g \ll h) \implies f \ll h$.

This following chain of magnitudes of some basic functions will be of some assistance. With the facts above and the chain below one can make use of the big $\mathcal{O}$ notation for a large number of different functions.

$$Constants \ll logarithmic \ll polynomials \ll exponentials \ll factorial.$$

And of course the reciprocal functions of the functions above will be chained in the reverse order.

We will also make use of the little o notation, which is stronger than the big $\mathcal{O}$ notation. For functions f, g, we say that $f(x) = o(g(x))$ if

$$\lim_{x \to \infty} \frac{f(x)}{g(x)} = 0.$$

To save some energy on notation, we have adopted the following convention. In the summation $\sum_{p \leq x}$ we sum over all primes p less than x. For the sum $\sum_{n \leq x}$ we sum over all natural numbers n less than x. Further, in the chapter with Dirichlet characters χ, the sum $\sum_{\chi}$ is indexed over all Dirichlet characters of the group $(\mathbb{Z}/q\mathbb{Z})^*$.

We say that a subset S of the primes has the density δ if

$$\Big(\sum_{p \in S} \frac{1}{p}\Big)\Big(\sum_{p \ prime} \frac{1}{p^s}\Big)^{-1} \to \delta \qquad as \ s \downarrow 1.$$

This is called the *analytic* or the *Dirichlet* density. One can also define density δ of a subset S of the primes by

$$\frac{|\{p \leq x : p \in S\}|}{|\{p \leq x : p \ prime\}|} \to \delta \qquad for \ x \to \infty.$$

This is the *natural* density. We follow the convention that $|S|$ is the size of the set S. Dirichlet proved his theorem using the analytic density, although he never made the notion explicit [5]. De la Vallée-Poussin proved that Dirichlet's statement still holds for natural density as well [5].

It can be shown that if a set of primes has a natural density, then it also has an analytic density, and the densities are equal. However, the converse is not true [5]. We will encounter density theorems of Frobenius and Chebotarëv. These theorems were originally proved with the analytic density, but also hold for natural density.

Chapter 3

The Prime Number Theorem

The main destination of this thesis is Dirichlet's theorem on primes in arithmetic progressions, simply refered to as Dirichet's theorem. However, it would be a shame to discuss primes without mentioning the prime number theorem, a true gem in my opinion. The statement tells us that the prime frequency show some predictability and also connects the primes to the logarithmic function, a feature I find both suprising and interesting.

Along with the prime number theorem, we will also briefly discuss the infinitude of the primes and Euler factorization. We will include a proof of Dirichlet's theorem for the cases $q = 3, 4, 6$ in this chapter, which don't rely on the heavy machinery developed later on. In the end of this chapter we will look at the Chebychev functions $\theta(x)$ and $\psi(x)$, and the von Mangoldt function $\Lambda(n)$, a function which will come in good use in the proof of Dirichlet's theorem.

3.1 The infinitude of the primes and the Prime Number Theorem

Firstly, remember that the prime counting function $\pi(x)$ is defined by

$$\pi(x) = \big|\{p \leq x : p \text{ a prime}\}\big|.$$

A natural question concerning $\pi(x)$ is whether $\pi(x)$ is infinite or bounded as x grows. This have been answered in a multitude of ways in a wide range of mathematical branches. The perhaps most famous proof beares the name of Euclid and relies on the fact that from any finite list L of primes one can construct some natural number with a prime factor not in L.

Theorem 3.1 (The infinitude of the primes)**.** *There are infinitely many primes. That is,*

$$\pi(x) \to \infty \text{ when } x \to \infty.$$

One can ask, how common it is for a natural number to also be a prime. The following theorem give some answer to that question.

Theorem 3.2. *The natural density of the primes as a subset of the natural numbers is zero, that is,*

$$\lim_{x\to\infty} \frac{\pi(x)}{x} = 0.$$

Even though they are infinitely many, the primes are quite rare among the natural numbers. The prime number theorem gives us some information of how fast $\pi(x)$ grows.

Theorem 3.3 (Prime number theorem). *As $x \to \infty$, we have that*

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

Actually, a better estimate for $\pi(x)$ is the *logarithmic integral function*

$$Li(x) := \int_2^x \frac{dt}{log\ t}.$$

The fact $\pi(x) \sim Li(x)$ suggests that the probability of a natural number n being a prime, is roughly $\frac{1}{log\ n}$. This indicates that the primes become less dense among the natural numbers as x grows.

3.2 Proof of infinitude of primes modulo $q = 3, 4, 6$

Dirichlet's theorem will solve the issue concerning proving the infinitude of primes in each primitive residue class modulo any positive integer q. However, there are shortcuts for certain moduli q where we also meet interesting objects such as the cyclotomic polynomials.

The proof is divided in two parts. First, we will show the infinitude of primes in the residue classes $a(mod\ q)$, where $a \not\equiv 1(mod\ q)$. This can be shown in an Euclidean manner. Then we introduce the cyclotomic polynomials to prove the same statement for the residue class $1(mod\ q)$. As the observant reader has noticed, for $q = 3, 4, 6$, there are only two primitive residue classes, namely $a \equiv \pm 1(mod\ q)$. This fact is why the argument below works. We start with the case of $a \not\equiv 1(mod\ q)$.

To show the infinitude of primes $p \not\equiv 1(mod\ q)$ is pretty straightforward with Euclid's proof in mind. So assume there is only a finite number of primes $p \not\equiv 1(mod\ q)$, where $q \geq 3$. Let P denote the product of those primes. Consider the number $N = Pq - 1$. We can factorize N by primes p'. That is, $N = \prod p'$. And no p' is part of the original list. So

$$N = \prod p', \qquad \textit{and all prime factors satisfy} \quad p' \equiv 1(mod\ q).$$

But this is impossible, since $N \not\equiv 1(mod\ q)$. Hence, at least one prime factor p' of N satisfy $p' \not\equiv 1(mod\ q)$, which is the contradiction we were looking for.

To help us prove the case of $a \equiv 1(mod\ q)$ we need the following two lemmas.

Lemma 3.4. *For a non-constant polynomial $f(x) \in \mathbb{Z}[x]$, denote by $\mathbb{P}(f)$ the set of primes which divide $f(t) \neq 0$ for some $t \in \mathbb{Z}$. Then $|\mathbb{P}(f)| = \infty$.*

Proof. The proof will be by contraposition. Assume $\mathbb{P}(f)$ finite and make the list

$$\mathbb{P}(f) = \{p_1, \ldots, p_r\}$$

of distinct primes. We have that

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_0, \quad a_i \in \mathbb{Z}, \quad a_n \neq 0.$$

If $a_0 = 0$ we are finished because of the factorization $f(x) = x \cdot g(x)$ with $g(x) \in \mathbb{Z}[x]$. Then any prime p will be a factor of $f(p)$ simply because $f(p) = p \cdot g(p)$, and thus $|\mathbb{P}(f)| = \infty$.

Now consider $a_0 \neq 0$. Let

$$\mathbf{P} = \prod_{i=1}^{r} p_i,$$

that is, $\mathbf{P}$ is the product of the primes in $\mathbb{P}(f)$. This product is non-empty unless $f(x)$ is constantly equal to 1 or -1, since any prime divisor of $f(0) = a_0$ will be in $\mathbb{P}(f)$. Now

$$f(a_0\mathbf{P}) = a_0(a_n a_0^{n-1}\mathbf{P}^n + \ldots + a_1\mathbf{P} + 1).$$

Assume p' is a prime factor of $f(a_0\mathbf{P})$ which divides the second factor $(a_n a_0^{n-1}\mathbf{P}^n + \ldots + a_1\mathbf{P} + 1)$. Then, by definition, $p' \in \mathbb{P}(f)$ so $p'|\mathbf{P}$. But this implies $p'|1$, which contradicts p' being a prime. Hence

$$(a_n a_0^{n-1}\mathbf{P}^n + \ldots + a_1\mathbf{P} + 1) = 1 \vee -1,$$

leading to

$$f(a_0\mathbf{P}) = a_0 \vee -a_0.$$

We can reuse the argumentation above on the evaluations $f(a_0 j\mathbf{P})$, letting j run through $j = 1, 2, \ldots, (2n+1)$. Among these evaluations of f, we can find $n+1$ different evaluations which have the same value (either a_0 or $-a_0$). A polynomial f of degree at most n can only have n points with the same function value, unless f is a constant. Thus we have proved that $\mathbb{P}(f)$ finite implies $f(x)$ constant, and the proof is finished. □

The next lemma involves cyclotomic polynomials.

Definition (Cyclotomic polynomials)**.** *For a positive integer m, the m-th cyclotomic polynomial is defined by*

$$\Phi_m(x) = \prod_{\substack{1 \leq k \leq m \\ gcd(k,m)=1}} (x - e^{\frac{2\pi i k}{m}}).$$

We see that $\Phi_m(x)$ is the monic polynomial whose roots are the primitive m-th roots of unity. We will use the fact that $\Phi_m(x) \in \mathbb{Z}[x]$ and $x^m - 1 = \prod_{d|m} \Phi_d(x)$. (See [3, p. 17-18] for proofs).

Definition (Multiplicative order)**.** *Let t be an integer and n a positive integer with $gcd(t, n) = 1$. Then the multiplicative order of t modulus n is the smallest positive integer r such that*

$$t^r \equiv 1 (mod\ n).$$

We write this as $ord_n\ t = r$.

Lemma 3.5. *Let m be a positive integer and assume $p|\Phi_m(t)$ for some $t \in \mathbb{Z}$. Then either $p|m$, or $ord_p\ t = m$.*

Proof. Let $ord_p\ t = r$. By assumption $p|\Phi_m(t)$, and since $\Phi_m(t)$ is a factor of $t^m - 1$, we get that $p|t^m - 1$. This is equivalent to $t^m \equiv 1(mod\ p)$. We remember from elementary number theory that $ord_p\ t|m$, so $r|m$. If $r = m$, we are in the latter case of the lemma.

Now assume $r \neq m$. By definition $t^r - 1 \equiv 0(mod\ p)$, hence $p|t^r - 1$. By the fact stated above we have that

$$t^r - 1 = \Phi_r(t) \prod_{\substack{b|r \\ b \neq r}} \Phi_b(t). \tag{3.1}$$

Since p divides the left side in (3.1), p must also divide at least one of the factors on the right side. Say $p|\Phi_c(t)$, and assume $c \neq r$. This would imply that $ord_p\ t < r$, which is a contradiction. To see this, exchange r with c in (3.1) and use the fact that p divides the right side. Then p must also divide the left side, which in turn will lead to $ord_p\ t \leq c$, contradicting $ord_p\ t = r$. Hence $p|\Phi_r(t)$.

Using (3.1) and $r \neq m$ gives us

$$t^m - 1 = \Phi_r(t)\Phi_m(t) \prod_{\substack{b|m \\ b \neq r,m}} \Phi_b(t). \tag{3.2}$$

From (3.2), together with $p|\Phi_r(t)$ and $p|\Phi_m(t)$, we see that the polynomial $x^m - 1(mod\ p)$ has a double root at $x = t$. Thus, p must divide $x^m - 1$ and its derivative mx^{m-1} at $x = t$. By the property of p being a prime, either $p|m$ or $p|t^{m-1}$. But $p|t^{m-1} \implies p|1$, which is obviously false. Hence $p|m$, and we have proved the lemma.

□

We are now going to use the two lemmas to prove that there are infinitely many primes $p \equiv 1(mod\ q)$. Since $\Phi_q(x) \in \mathbb{Z}[x]$, lemma 3.4 applies. So $|\mathbb{P}(\Phi_q)| = \infty$. Let $p \in \mathbb{P}(\Phi_q)$, i.e. there exist a $t \in \mathbb{Z}$ such that $p|\Phi_q(t)$. From lemma 3.5 either $p|q$ or $ord_p\ t = q$. But for fixed q there are only a finite number of primes $p \in \mathbb{P}(\Phi_q)$ for which $p|q$.

So there are infinitely many primes such that $ord_p\ t = q$. By Fermat's little theorem, and using $t \not| p$, we get

$$t^{p-1} \equiv 1(mod\ p).$$

Again, relying on elementary number theory, we have that $ord_p\ t|(p-1)$. But $ord_p\ t = q$, so $q|(p-1)$. This is to say that $p \equiv 1(mod\ q)$ and we are done.

3.3 Euler factorization and the von Mangoldt function

The next theorem gives us a great tool when working on the Riemann zeta function $\zeta(s)$, by letting us express the infinite sum as an infinite product indexed by the primes.

Theorem 3.6 (Euler factorization)**.** *Let f be a multiplicative arithmetic function. Then*

$$\sum_{n=1}^{\infty} f(n) = \prod_{p}(1 + f(p) + f(p^2) + \dots), \tag{3.3}$$

provided either of the two expressions below converges

$$\sum_{n=1}^{\infty} |f(n)|, \qquad \prod_{p}(1 + |f(p)| + |f(p^2)| + \dots).$$

If further f is completely multiplicative, then each factor in the product on the right side of (3.3) is a geometric series. Thus

Corollary 3.6.1. *Let f be a completely multiplicative arithmetic function. Then*

$$\sum_{n=1}^{\infty} f(n) = \prod_{p} \frac{1}{1 - f(p)}$$

if the same conditions are fulfilled as in the theorem.

The Riemann zeta function is defined by

$$\zeta(s) := \sum_{n=1}^{\infty} \frac{1}{n^s}$$

for $\Re(s) > 1$. Since $f(n) = \frac{1}{n^s}$ is a completely multiplicative arithmetic function, and one can use the integral test to show convergence of $\sum_{n=1}^{\infty} |\frac{1}{n^s}|$ when $s > 1$, collorary 3.6.1 applies. So

$$\zeta(s) = \prod_{p} \frac{1}{1 - \frac{1}{p^s}}.$$

Another way to utilize the Euler factorization is to prove the divergence of the sum of the resiprocals of the primes.

Theorem 3.7. *The series $\sum_p 1/p$ diverges.*

Proof. Proof by contradiction. Assume that the series converges and let $\sum_p 1/p = C$. We are going to use the assumption to show convergence of the harmonic series, which is a contradiction. For the harmonic series, $f(n) = 1/n$. Our next step is to create a bound for the product

$$\prod_{p \le x} \left(1 + \frac{1}{p} + \frac{1}{p^2} + \dots\right) \tag{3.4}$$

independent of x.

$$\prod_{p \le x} \left(1 + \frac{1}{p} + \frac{1}{p^2} + \dots\right) = \prod_{p \le x} \frac{1}{1 - \frac{1}{p}} = \prod_{p \le x} \left(1 + \frac{1}{p - 1}\right) \le \prod_{p \le x} \left(1 + \frac{2}{p}\right).$$

The inequality $e^t \geq 1+t$ for $t \geq 0$ follows from truncationg the Taylor expansion of e^t at $t = 0$. So

$$\prod_{p \leq x} \left(1 + \frac{2}{p}\right) \leq \prod_{p \leq x} e^{2/p} = e^{\sum_{p \leq x} 2/p} \leq e^{2C}.$$

As the products in (3.4) form a bounded, increasing sequence as x increases, we can conclude that the infinite product is convergent. Then the Euler factorization theorem applies, and we get

$$\sum_{n=1}^{\infty} \frac{1}{n} = \prod_{p} \frac{1}{1 - \frac{1}{p}} \leq e^{2C},$$

showing convergence of the harmonic series, which is a contradiction.

□

Connected with the prime counting function $\pi(x)$ are these two functions, sometimes called Chebychev functions.

$$\theta(x) := \sum_{p \leq x} log\ p,$$
$$\psi(x) := \sum_{n=1}^{\infty} \theta(x^{1/n}).$$

The usefulness of these functions becomes apparent by the next result.

Proposition 3.1. *As $x \to \infty$, we have that*

$$\frac{\theta(x)}{x} = \frac{\pi(x)}{x/log\ x} + o(1),$$
$$\frac{\psi(x)}{x} = \frac{\pi(x)}{x/log\ x} + o(1).$$

So the prime number theorem (theorem 3.3) is equivalent to $\theta(x) \sim x$, or $\psi(x) \sim x$. Knowing this, we can use these functions to estimate the prime counting function $\pi(x)$. In a later chapter we will derive a formula for $\psi(x)$ using zeros of the L-function $L(s, \chi)$.

Definition (The von Mangoldt function)**.** *For $n \in \mathbb{N}$, the von Mangoldt function $\Lambda(n)$ is defined by*

$$\Lambda(n) := \begin{cases} log(p), & \text{if } n = p^k \text{ for some prime } p \text{ and some integer } k, \\ 0 & \text{else.} \end{cases} \tag{3.5}$$

The function $\Lambda(n)$ will come to great use in the proof of Dirichlet's theorem in the next chapter. We will need these two following facts about the von Mangoldt function.

$$log\ n = \sum_{d|n} \Lambda(d), \qquad and \qquad \sum_{d \leq x} \Lambda(d) \ll x.$$

The last thing we shall look at in this section is the sum

$$A(x) := \sum_{p \leq x} \frac{log\ p}{p}. \tag{3.6}$$

This sum will be the object of attention in Dirichlet's theorem, with the modification that the sum only extends over primes in a single residue class a modulo q.

Proposition 3.2.

$$A(x) = log\ x + \mathcal{O}(1) \tag{3.7}$$

Proof. First off, note that

$$\sum_{d \leq x} \frac{\Lambda(d)}{d} = \sum_{p^k \leq x} \frac{log\ p}{p^k} = \sum_{p \leq x} \frac{log\ p}{p} + \sum_{\substack{p^k \leq x \\ k \geq 2}} \frac{log\ p}{p^k}.$$

This last sum on the right is $\mathcal{O}(1)$ [3]. So we will prove the claim for $\sum_{d \leq x} \Lambda(d)/d$. Now

$$\begin{aligned} \sum_{n \leq x} log\ n &= \int_1^x log\ t\ dt + E(x) \\ &= xlog\ x - x + E(x) \\ &= xlog\ x + \mathcal{O}(x), \end{aligned} \tag{3.8}$$

where the error term $E(x) = \mathcal{O}(log\ x)$.

We can also consider this sum by using the von Mangoldt function $\Lambda(n)$, simply because $log\ n = \sum_{d|n} \Lambda(d)$.

$$\begin{aligned} \sum_{n \leq x} log\ n &= \sum_{n \leq x} \sum_{d|n} \Lambda(d) \\ &= \sum_{d \leq x} \sum_{\substack{n \leq x \\ d|n}} \Lambda(d) \\ &= \sum_{d \leq x} \Lambda(d) \lfloor \frac{x}{d} \rfloor \\ &= x \sum_{d \leq x} \frac{\Lambda(d)}{d} + E(x), \end{aligned} \tag{3.9}$$

where the error term $E(x) \ll \sum_{d \leq x} \Lambda(d) \ll x$. Setting the two expressions (3.8) and (3.9) equal eachother and dividing by x gives

$$\sum_{d \leq x} \frac{\Lambda(d)}{d} = log\ x + \mathcal{O}(1). \tag{3.10}$$

We can replace $\sum_{d \leq x} \big(\Lambda(d)/d\big)$ by $\sum_{p \leq x} \big(log\ p/p\big)$ in (3.10) by our initial argumentation, and thus the result follows. □

Chapter 4

Dirichlet's Theorem

This chapter will contain the main theorem in which the thesis is built around, namely a theorem concerning primes in arithmetic progressions by Dirichlet. An immediate consequence is the infinitude of primes in every residual class $a(mod\ q)$ whenever $gcd(a,q) = 1$. A more striking feature of the theorem is how it indicates a good behaviour by the otherwise unruly primes. What we mean by good behaviour is that it seems that the primes divides themselves evenly among the primitive residue classes modulo q for all natural numbers $q \geq 2$.

Theorem 4.1 (Dirichlet's theorem)**.** *Let $x \geq 4$. Then*

$$\sum_{\substack{p \leq x \\ p \equiv a(mod\ q)}} \frac{log\ p}{p} = \frac{1}{\varphi(q)} log\ x + \mathcal{O}(1). \tag{4.1}$$

Remember from (3.7) that

$$\sum_{p \leq x} \frac{log\ p}{p} = log\ x + \mathcal{O}(1).$$

We see that the two sums only differ in that the index of (4.1) goes through the primes p less than x AND $p \equiv a(mod\ q)$, while (3.7) is not constrained to the second condition. All but a finite number of the primes $p \leq x$ must be in one of the $\varphi(q)$ primitive residue classes. Let $a_1, a_2, \ldots, a_{\varphi(q)}$ denote these residue classes. If we disregard the primes dividing q, we can split up the sum in (3.7) as

$$\sum_{p \leq x} \frac{log\ p}{p} = \sum_{\substack{p \leq x \\ p \equiv a_1 (mod\ q)}} \frac{log\ p}{p} + \sum_{\substack{p \leq x \\ p \equiv a_2 (mod\ q)}} \frac{log\ p}{p} + \cdots + \sum_{\substack{p \leq x \\ p \equiv a_{\varphi(q)} (mod\ q)}} \frac{log\ p}{p}. \tag{4.2}$$

Using Dirichlet's theorem and (3.7) on (4.2) produces

$$log\ x = \frac{1}{\varphi(q)} log\ x + \frac{1}{\varphi(q)} log\ x + \cdots + \frac{1}{\varphi(q)} log\ x + \mathcal{O}(1). \tag{4.3}$$

It is by (4.3) that we can get the indication that the primes divide themselves equally into the primitive residue classes.

4.1 Outline of the proof of Dirichlet's theorem

The proof of theorem 4.1 starts with showing that

$$\sum_{\substack{p\leq x\\ p\equiv a(mod\ q)}} \frac{log\ p}{p} = \sum_{\substack{n\leq x\\ n\equiv a(mod\ q)}} \frac{\Lambda(n)}{n} + \mathcal{O}(1). \tag{4.4}$$

By (4.4) we can direct our attention to the sum $\sum \Lambda(n)/n$. We know from before that

$$\sum_{n\leq x} \frac{\Lambda(n)}{n} = log\ x + \mathcal{O}(1). \tag{3.10}$$

The next step is to introduce Dirichlet charachters $\chi(n)$ with the purpose of killing the contribution of all $n \not\equiv a$ terms in the sum in (3.10). Any Dirichlet character is assosiated with an L-series, which is a generalization of the Riemann zeta function. The last part of the proof will involve some effort in showing that the L-series of a non-trivial Dirichlet charachter at the point $s = 1$ is non-zero.

4.2 Dirichlet characters

First off, we will show (4.4). We have that

$$\sum_{\substack{n\leq x\\ n\equiv a(mod\ q)}} \frac{\Lambda(n)}{n} = \sum_{\substack{p^k\leq x\\ p^k\equiv a(mod\ q)}} \frac{log\ p}{p^k} = \sum_{\substack{p\leq x\\ p\equiv a(mod\ q)}} \frac{log\ p}{p} + \sum_{k\geq 2} \sum_{\substack{p\leq x^{1/k}\\ p^k\equiv a(mod\ q)}} \frac{log\ p}{p^k}.$$

The double sum above is convergent, since

$$\sum_{k\geq 2} \sum_{\substack{p\leq x^{1/k}\\ p^k\equiv a(mod\ q)}} \frac{log\ p}{p^k} \leq \sum_{k\geq 2}\sum_{n\geq 2} \frac{log\ n}{n^k} = \sum_{n\geq 2} \frac{log\ n}{n(n-1)} \leq \sum_{n\geq 1} \frac{log\ (n+1)}{n^2},$$

and if we combine the integral test with

$$\int \frac{log(n+1)}{n^2}\ dn = \frac{log(n+1)}{n} + log(n) - log(n+1) + \ constant$$

we have the convergence. Thus, (4.4) follows.

In order to fulfill our plan to prove theorem 4.1, we need to find some way to kill the contribution of all $n \not\equiv a \pmod q$ in the summation $\sum_{n\leq x} \frac{\Lambda(n)}{n}$. This will be done by the help of Dirichlet characters χ. Firstly, let's define the characters of a finite abelian group G.

Definition (Character of a finite abelian group)**.** *A character χ of a finite abelian group G, is a homomorphism*

$$\chi : G \to \mathbb{C}^*,$$

where $\mathbb{C}^*$ denotes the set of non-zero complex numbers.

Then any element in $Hom(G,\ \mathbb{C}^*)$ is a character of G. For our purpose the group G is the multiplicative group of integers modulus q, denoted by $\left(\mathbb{Z}/q\mathbb{Z}\right)^*$. In the language of number theory, the characters are completely multiplicative. We call the character which sends any element of G to 1, for the principal character χ_0. A property of the characters is that the image of χ is contained in the set of the n-th roots of unity, where n is the order of G. To see this let $g \in G$. Then

$$\chi(g)^n = \chi(g^n) = \chi(1) = 1.$$

The complex conjugate $\bar{\chi}$ of a dirichlet character χ is defined by

$$\bar{\chi}(g) := \overline{\chi(g)}.$$

We are now in a position to define the Dirichlet characters, which have been our objective for this paragraph.

Definition (Dirichlet characters)**.** *Let q be a positve integer, and $G = \left(\mathbb{Z}/q\mathbb{Z}\right)^*$. That is, G is the group of multiplicative inverses modulo q. Then for a character χ of G, we can define a Dirichlet character $\tilde{\chi} : \mathbb{Z} \to \mathbb{C}^*$ modulo q by*

$$\tilde{\chi}(a) = \begin{cases} \chi(a \ (mod\ q)) & \text{if } gcd(a,q) = 1, \\ 0 & \text{if } gcd(a,q) > 1. \end{cases}$$

We will from now on only consider Dirichlet characters, and so for simplicity we drop the tilde. Hence, χ from now on will be a Dirichlet character.

The following theorem displays the orthogonality relation for Dirichlet characters. It is by this theorem's help, that we are able to kill the contribution of $n \not\equiv a$ in (4.4).

Theorem 4.2 (Orthagonality relations)**.** *Let q be a positive integer and a, b two integers where $gcd(a,q) = 1$. Then*

$$\sum_{\chi} \bar{\chi}(a)\chi(b) = \begin{cases} \phi(q) & \text{if } a \equiv b \ (mod\ q), \\ 0 & \text{otherwise.} \end{cases}$$

Proof. See [3, p. 79] □

4.3 L-series for complex and real characters

We are going to wrap up the proof of Dirichlet's theorem in this section by the help of L-series. The work will be to show that these L-series $L(s,\chi)$ assosiated with a non-principal Dirichlet character χ, is non-zero at $s = 1$. First the definition.

Definition (L-series of χ)**.** *For a Dirichlet character χ mod q, we define the L-series by*

$$L(s,\chi) := \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s},$$

where $s \in \mathbb{R}$.

A fact about an L-series that we will not prove, is that for $s > 0$ and $\chi \neq \chi_0$ the L-series is convergent. Actually, the only point of interest for us will be at $s = 1$. The reason being, for $s = 1$, we can apply the next lemma.

Lemma 4.3. *Let χ be a non-principal Dirichlet character mod q. Then for $x \geq 4$,*

$$\sum_{n \leq x} \frac{\chi(n)\Lambda(n)}{n} = \begin{cases} \mathcal{O}(1) & \text{if } L(1,\chi) \neq 0, \\ -log\ x + \mathcal{O}(1) & \text{otherwise.} \end{cases} \tag{4.5}$$

Proof. See [3, p. 83] □

In the passage below it will be clear that $L(1, \chi) \neq 0$ for every non-principal χ. So what about the principal Dirichlet character χ_0?

Lemma 4.4. *Let χ_0 be the principal Dirichlet character mod q. Then for $x \geq 4$*

$$\sum_{n \leq x} \frac{\chi_0(n)\Lambda(n)}{n} = log\ x + \mathcal{O}(1).$$

Proof. We have that $\sum_{n \leq x} \frac{\chi_0(n)\Lambda(n)}{n} - \sum_{n \leq x} \frac{\Lambda(n)}{n} = \mathcal{O}(1)$ by

$$\sum_{n \leq x} \frac{\Lambda(n)}{n} - \sum_{n \leq x} \frac{\chi_0(n)\Lambda(n)}{n} = \sum_{p|q} \sum_{p^k \leq x} \frac{log\ p}{p^k} \ll \sum_{p|q} \frac{log\ p}{p-1} \ll 1.$$

And $\sum_{n \leq x} \frac{\chi_0(n)\Lambda(n)}{n}$ is $log\ x + \mathcal{O}(1)$ by (3.10). Thus, we get the result. □

We remember that the image of a Dirichlet character χ modulo q is contained in the set of the $\phi(q)$-th roots of unity. If the image is real for χ, that is, $\chi(\mathbb{Z}) \subseteq \{1, 0, -1\}$, we say that χ is a real Dirichlet charater. Otherwise, we call χ a complex character. To be able to prove Dirichlets theorem we need to show that for any non-principal χ, $L(1, \chi) \neq 0$ holds. Let's begin with the complex characters.

Theorem 4.5. *Let χ be a complex Dirichlet character mod q. Then $L(1, \chi) \neq 0$.*

Proof. Let N denote the number of non-principal Dirichlet characters χ such that $L(1, \chi) = 0$. Then by lemma 4.3 and lemma 4.4, we get

$$\sum_{\chi} \sum_{n \leq x} \frac{\chi(n)\Lambda(n)}{n} = (1 - N)log\ x + \mathcal{O}(1).$$

Another way to evaluate this double sum is by using the orthogonality relations in theorem 4.2, taking $a = 1$. Then

$$\sum_{\chi} \sum_{n \leq x} \frac{\chi(n)\Lambda(n)}{n} = \phi(q) \sum_{\substack{n \leq x \\ n \equiv 1 \pmod q}} \frac{\Lambda(n)}{n}$$

and this sum is obviously non-negative. The only way these to representations can be equal, is if $N \leq 1$. Assume that there exists a complex character χ such that $L(1,\chi) = 0$. Then

$$L(1,\bar{\chi}) = \sum_{n=1}^{\infty} \frac{\bar{\chi}(n)}{n} = \overline{\sum_{n=1}^{\infty} \frac{\chi(n)}{n}} = \overline{L(1,\chi)} = 0,$$

hence $\bar{\chi}$ is also a Dirichlet character for which $L(1,\bar{\chi}) = 0$, and since χ is assumed to be complex, we must have that $\chi \neq \bar{\chi}$, and so $N > 1$. We have now reached a contradiction and the result follows. □

Theorem 4.6. *Let χ be a real Dirichlet character mod q and $\chi \neq \chi_0$. Then $L(1,\chi) \neq 0$.*

Proof. Sketch of proof: For the full details see [3, p. 86]. First define the auxiliary function for $x \in (0,1)$

$$f(x) := \sum_{n=1}^{\infty} \chi(n) \frac{x^n}{1-x^n}.$$

The function $f(x)$ is an absolute convergent series, and $f(x) \to \infty$ when $x \uparrow 1$. By using this facts, we are going to prove the statement by contradiction. So let's assume $L(1,\chi) = 0$. Then for $x \in (0,1)$

$$\begin{aligned} -f(x) &= \frac{L(1,\chi)}{1-x} - f(x) \\ &= \sum_{n=1}^{\infty} \chi(n)\Big(\frac{1}{n(1-x)} - \frac{x^n}{1-x^n}\Big) \\ &= \sum_{n=1}^{\infty} \chi(n) b_n(x), \qquad b_n(x) = \Big(\frac{1}{n(1-x)} - \frac{x^n}{1-x^n}\Big). \end{aligned}$$

It can be shown that the b_i-s satisfy

$$b_1(x) \geq b_2(x) \geq b_3 \geq ... \geq 0. \tag{4.6}$$

Define $S(x) := \sum_{n \leq x} \chi(n)$. We have the bound $|S(x)| \leq q$ from [3, p. 82]. Consider

$$\begin{aligned} \sum_{n=1}^{M} \chi(n) b_n(x) &= \sum_{n=1}^{M} \Big(S(n) - S(n-1)\Big) b_n(x) \\ &= S(M) b_M(x) + \sum_{n=1}^{M-1} S(n)\Big(b_n(x) - b_{n+1}(x)\Big). \end{aligned}$$

Using (4.6) and the triangle inequality provides the following bound.

$$\begin{aligned} \Big|\sum_{n=1}^{M} \chi(n) b_n(x)\Big| &\leq q b_M(x) + q \sum_{n=1}^{M-1} \Big(b_n(x) - b_{n+1}(x)\Big) \\ &= q b_M(x) + q(b_1(x) - b_M(x)) = q b_1 = q. \end{aligned} \tag{4.7}$$

So $\left|\sum_{n=1}^{M} \chi(n)b_n(x)\right| \leq q$ for any natural number M and $x \in (0,1)$. If M goes to infinity, this tells us that $|f(x)| \leq q$ for $x \in (0,1)$, which contradicts $f(x) \to \infty$ as $x \uparrow 1$. Thus our assumption of $L(1,\chi) = 0$ is false.

□

4.4 Proof of Dirichlet's Theorem

We are now going to put everything together to prove Dirichlet's theorem. Let q and a be two positive integers with $gcd(a,q) = 1$. From the lemmas 4.3 and (4.4) we know that

$$\sum_{n \leq x} \frac{\chi(n)\Lambda(n)}{n} = \begin{cases} \mathcal{O}(1) & \text{if } \chi \neq \chi_0, \\ log\ x + \mathcal{O}(1) & \text{if } \chi = \chi_0. \end{cases}$$

Using the orthogonality relations (4.2)

$$\begin{aligned} \sum_{\substack{n \leq x \\ n \equiv a (mod\ q)}} \frac{\Lambda(n)}{n} &= \sum_{n \leq x} \frac{\Lambda(n)}{n} \Big[\frac{1}{\phi(q)} \sum_{\chi} \bar{\chi}(a)\chi(n) \Big] \\ &= \frac{1}{\phi(q)} \sum_{\chi} \bar{\chi}(a) \sum_{n \leq x} \frac{\chi(n)\Lambda(n)}{n} \\ &= \frac{1}{\phi(q)} \bar{\chi_0}(a) log\ x + \mathcal{O}(1) \\ &= \frac{1}{\phi(q)} log\ x + \mathcal{O}(1). \end{aligned}$$

As we showed in the start of this section

$$\sum_{\substack{p \leq x \\ p \equiv a (mod\ q)}} \frac{log\ p}{p} = \sum_{\substack{n \leq x \\ n \equiv a (mod\ q)}} \frac{\Lambda(n)}{n} + \mathcal{O}(1),$$

hence

$$\sum_{\substack{p \leq x \\ p \equiv a (mod\ q)}} \frac{log\ p}{p} = \frac{1}{\phi(q)} log\ x + \mathcal{O}(1).$$

and Dirichlet's theorem is proven.

Chapter 5

The number of zeros of the Riemann zeta function $\zeta(s)$

The Riemann zeta function $\zeta(s)$ is an infinite sum, which converges for $\Re(s) > 1$. By analytic continuation, one can extend $\zeta(s)$ to be an analytic function on the whole complex plane, with the exception for the pole at $s = 1$. The zeros of $\zeta(s)$ are located at $s = -2, -4, -6, \ldots$, called the trivial zeros, and inside the area between the lines $\Re(s) = 0$ and $\Re(s) = 1$, called the non-trivial zeros. This last area is refered to as the critical stip. The Riemann hypothesis conjecture that all non-trivial zeros of $\zeta(s)$ has real part equal $\frac{1}{2}$.

We are interested in estimating how $N(T)$ grows as T grows, where $N(T)$ denotes the number of zeros of $\zeta(s)$ with an imaginary part smaller than T and a real part in $(0, 1)$. In other words, $N(T)$ is the number of zeros of $\zeta(s)$ in the part of the critical strip above the real axis and below the line $\Im(s) = T$. Our approach will be by the means of complex analysis.

For an L-series $L(s, \chi)$, one can also do analytic continuation to define the L-function $L(s, \chi)$. The number $N(T, \chi)$ counts similarly as $N(T)$ the zeros of the Dirichlet L-function in the rectancle $0 < \sigma < 1$, $|t| <$ T. This chapter rests heavily on the chapters 15 and 16 in [1], with additional results found in the same book. The last chapter in this thesis will concern Chebyshev's bias. The technique of calculating these biases makes use of zeros of L-functions, and one need to know how $N(T, \chi)$ grows when T grows.

5.1 Replacing $\zeta(s)$ by $\xi(s)$.

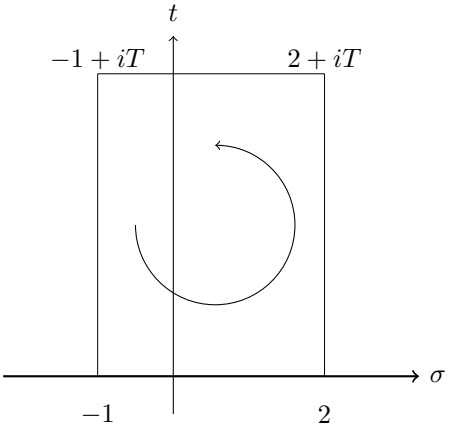
The first order of business is to replace $\zeta(s)$ by the function $\xi(s)$ for practical purposes. By convention, denote the complex variable s by $s = \sigma + it$. Now

$$\xi(s) = \frac{1}{2}s(s-1)\Pi^{-\frac{1}{2}}\Gamma(\frac{s}{2})\zeta(s). \tag{5.1}$$

By inspecting $\xi(s)$, we see that $\xi(s) = 0$ if and only if $\zeta(s) = 0$ in the rectangle $0 < \sigma < 1$, $0 < t < T$. So $N(T)$ is also the number of zeros of $\xi(s)$ in the same rectangle.

The next step is to apply the argument principle to the meromorphic function $\xi(s)$. We remember that a function is meromorphic on an open subset $D \subseteq \mathbb{C}$ if it is holomorphic in D, with the exception of isolated points

in which the function has poles. Then $\xi(s)$ satisfies these conditions since it is a product of meromorphic and holomorphic functions. We choose the boundary of the rectangle R as the path of integration, where R has vertices $\{2, 2+iT, -1+iT, -1\}$. Assume for simplicity that T is not an ordinate of a zero of $\xi(s)$, so our path of integration does not hit a pole (See figure below).



Path of integration

Theorem 5.1 (The Argument Principle)**.** *Let $f(s)$ be a meromorphic function inside and on the contour C, and assume f has no poles or zeros on C, Then*

$$\oint_C \frac{f'(s)}{f(s)} ds = 2\pi i(N - P),$$

where N denote the number of zeros of f inside C, and P the number of poles of f inside C.

The assumption of $\xi(s)$ being without zeros and poles on the contour we have chosen seems at first to be violated when looking at (5.1). However, the zero at $s = 0$ is cancelled by the simple pole of $\Gamma(s)$, and the zero at $s = 1$ is cancelled by the simple pole of $\zeta(s)$. With the argument principle in mind, $\xi(s)$ has no poles inside R, so we are left with counting zeros of $\xi(s)$, or equvivalently, counting zeros of $\zeta(s)$ when evaluating the integral

$$\oint_R \frac{\xi'(s)}{\xi(s)} ds.$$

From the argument principle we have that

$$\oint_R \frac{\xi'(s)}{\xi(s)} ds = 2\pi i \ N(T),$$

looking at the imaginary part gives

$$\Delta_R \ arg \ \xi(s) = 2\pi N(T),$$

where $\Delta_R \ arg \ \xi(s)$ is understood as the change of the argument of $\xi(s)$ as we integrate along the boundary of R in a positive sence. Now we can evaluate

the change in the argument of $\xi(s)$ by adding together each contribution from the factors found in (5.1).

Our goal is to evaluate $\Delta_R \; arg \; \xi(s)$. Observe that for s traversing through the real part from -1 to 2, there is no change in $arg \; \xi(s)$ because the function value is real and nowhere zero on the described path. Hence, this path gives no contribution to $\Delta_R \; arg \; \xi(s)$. Separate the remainding path of integration into the paths L_1 and L_2, where

$$L_1 : \textit{The union of the lines } \{2, 2+iT\}, \textit{ and } \{2+iT, \frac{1}{2}+iT\},$$
$$L_2 : \textit{The union of the lines } \{\frac{1}{2}+iT, -1+iT\}, \textit{ and } \{-1+iT, -1\}.$$

From the function relation

$$\xi(s) = \xi(1-s) = \overline{\xi(\overline{1-s})} \tag{5.2}$$

we get that

$$\Delta_{L_1} \; arg \; \xi(s) = \Delta_{L_2} \; arg \; \xi(s). \tag{5.3}$$

To see this fact from $\xi(s) = \xi(1-s)$, let $s_1, s_2 \in \mathbb{C}$ which are symmetric around $s = \frac{1}{2}$. This is to say that the points s_1, s_2 satisfy

$$s_1 + \frac{s_2 - s_1}{2} = \frac{1}{2}.$$

Solving for s_2 gives $s_2 = 1 - s_1$, and consequently, $\xi(s_1) = \xi(s_2)$.

If the auxiliary paths $\overline{L}_1$ and $\overline{L}_2$ denotes the complex conjugate paths of L_1 and L_2 respectively, then by the functional relation $\xi(1-s) = \overline{\xi(\overline{1-s})}$ we have that

$$\Delta_{L_1} \; arg \; \xi(s) = -\Delta_{\overline{L}_1} \; arg \; \xi(s),$$

and $\xi(s) = \xi(1-s)$ tells us that

$$\Delta_{L_2} \; arg \; \xi(s) = -\Delta_{\overline{L}_1} \; arg \; \xi(s).$$

So the contribution from the paths L_1 and L_2 to $\Delta_R \; arg \; \xi(s)$ are equal, which unburden us from calculating both paths. We will continue with the L_1-path, and from now on just call it L.

By using a property of the Γ-function, we can rewrite (5.1) as

$$\xi(s) = (s-1)\Pi^{-\frac{s}{2}}\Gamma(\frac{s}{2}+1)\zeta(s).$$

We are now in a position to start calculating $\Delta_R \; arg \; \xi(s)$, since

$$\Delta_R \; arg \; \xi(s) = 2\Delta_L \; arg \; \xi(s) =$$
$$2\Delta_L \Big[\; arg \; (s-1) + arg \; \Pi^{-\frac{s}{2}} + arg \; \Gamma(\frac{s}{2}+1) + arg \; \zeta(s) \Big].$$

To get to the estimate of $N(T)$ we need to find

$$\Delta_L \Big[\; arg \; (s-1) + arg \; \Pi^{-\frac{s}{2}} + arg \; \Gamma(\frac{s}{2}+1) + arg \; \zeta(s) \Big]. \tag{5.4}$$

5.2 Calculating the argument of the factors of $\xi(s)$

From (5.4) it is clear that we can find estimates of each of the four terms individually. So we start with $\Delta_L\ arg\ (s-1)$. Here

$$\Delta_L\ arg\ (s-1) = arg(-\frac{1}{2}+iT) = \frac{\pi}{2}+\mathcal{O}(T^{-1}). \tag{5.5}$$

The error term comes from the Taylor expansion of $arctan(T)$ around ∞.

Next up is $\Delta_L\ arg(\Pi^{-\frac{s}{2}})$. Now we get

$$\begin{aligned}\Delta_L\ arg\ \Pi^{-\frac{s}{2}} &= e^{-\frac{s}{2}log\ \pi} = \Delta_L\ arg\ (e^{-\frac{\sigma+it}{2}log\ \pi}) \\ &= \Delta_L\ arg(e^{-\frac{\sigma\ log\ \pi}{2}}e^{i(-\frac{t}{2}log\ \pi)}) = \Delta_L(-\frac{t}{2}log\ \pi) \\ &= -\frac{1}{2}T\ log\ \pi. \end{aligned}\tag{5.6}$$

As we turn to $\Delta_L\ arg\ (\Gamma(\frac{s}{2}+1))$, we make use of Stirling's asymptotic formula for the gamma-function, which can be stated

$$log\ \Gamma(s) = (s-\frac{1}{2})log(s) - s + \frac{1}{2}log(2\pi) + \mathcal{O}(|s|^{-1}). \tag{5.7}$$

With this estimation in mind

$$\begin{aligned}\Delta_L\ arg\ \Gamma(\frac{s}{2}+1) &= \Im(log\ \Gamma(\frac{iT}{2}+\frac{5}{4})) \\ &= \Im\Big[(\frac{iT}{2}+\frac{3}{4})log(\frac{iT}{2}+\frac{5}{4}) - \frac{iT}{2} - \frac{5}{4} + \frac{1}{2}log(2\pi) + \mathcal{O}(T^{-1})\Big].\end{aligned}$$

Remember that $log\ z = log\ |z| + i\arg z$ for $z \in \mathbb{C}$. So

$$\Im(\frac{iT}{2}log(\frac{iT}{2}+\frac{5}{4})) = \frac{T}{2}log\sqrt{(\frac{T}{2})^2+(\frac{5}{4})^2} = \frac{T}{2}log\frac{T}{2} + \mathcal{O}(T^{-1}),$$

and

$$\Im(\frac{3}{4}log(\frac{iT}{2}+\frac{5}{4})) = \frac{3}{4}arg(\frac{iT}{2}+\frac{5}{4}) = \frac{3}{4}arg(\frac{iT}{2}+\frac{5}{4}),$$

and $arg(\frac{iT}{2}+\frac{5}{4}) = \frac{\pi}{2}+\mathcal{O}(T^{-1})$. This gives

$$\Delta_L\ arg\ \Gamma(\frac{s}{2}+1) = \frac{1}{2}Tlog(\frac{T}{2}) - \frac{T}{2} + \frac{3}{8}\pi + \mathcal{O}(T^{-1}). \tag{5.8}$$

Calculating $\Delta_L\ arg\ \zeta(s)$ will demand some work. The claim is that $\Delta_L\ arg\ \zeta(s)$ is $\mathcal{O}(log\ T)$. In order to do this we need the following lemma.

Lemma 5.2. *Let $\rho = \beta + i\gamma$ be the notation of a non-trivial zero of $\zeta(s)$, then for large T*

$$\sum_{\rho}\frac{1}{1+(T-\gamma)^2} = \mathcal{O}(log\ T),$$

where the index goes over the non-trivial zeros ρ.

Proof. From [1] we have

$$-\Re(\frac{\zeta'(s)}{\zeta(s)}) < A\ log\ t - \sum_{\rho} \Re(\frac{1}{\rho} + \frac{1}{s-\rho})$$

for $1 \leq \sigma \leq 2$ and $t \geq 2$ and A is a positive constant. If we insert $s = 2 + iT$ in the inequality and use the fact that $|\frac{\zeta'(s)}{\zeta(s)}|$ is bounded for such s, we get

$$\sum_{\rho} \Re(\frac{1}{\rho} + \frac{1}{s-\rho}) < A\ log\ T \tag{5.9}$$

by just adjusting the constant A. Since $s = 2 + iT$ and $\beta \in (0, 1)$, we get

$$\Re(\frac{1}{s-\rho}) = \frac{2-\beta}{(2-\beta)^2 + (T-\gamma)^2} \geq \frac{1}{4 + (T-\gamma)^2}.$$

Hence

$$\sum_{\rho} \frac{1}{4 + (T-\gamma)^2} \leq \sum_{\rho} \Re(\frac{1}{s-\rho}) \leq A\ log\ T$$

and we can conclude that the series $\sum_{\rho} \frac{1}{4+(T-\gamma)^2} \frac{1}{log\ T}$ converges. By applying the limit comparison test to this series and the series $\sum_{\rho} \frac{1}{1+(T-\gamma)^2} \frac{1}{log\ T}$, we find that the latter series is convergent, and hence the result of the lemma follows. □

From lemma 5.2 we deduce two corollaries and use them to show that $\Delta_L\ arg\ \zeta(s) = \mathcal{O}(log\ T)$. Still let $\rho = \beta + i\gamma$ denote a zero of $\zeta(s)$ and let $G_T = \{\rho \mid T - 1 < \gamma < T + 1\}$. Then

Corollary 5.2.1. $|G_T| = \mathcal{O}(log\ T)$.

Corollary 5.2.2. $\sum_{\rho \notin G_T} \frac{1}{(T-\gamma)^2} = \mathcal{O}(log\ T)$.

The only thing that is left is to show how lemma 5.2 and its corollaries imply $\Delta_L\ arg\ \zeta(s) = \mathcal{O}(log\ T)$. By [1, p. 99]

$$\frac{\zeta'(s)}{\zeta(s)} = \mathcal{O}(1) + \sum_{\rho} (\frac{1}{s-\rho} - \frac{1}{2 + iT - \rho}),$$

so

$$\frac{\zeta'(s)}{\zeta(s)} = \mathcal{O}(1) + \sum_{\rho \notin G_T} (\frac{1}{s-\rho} - \frac{1}{2+iT-\rho}) + \sum_{\rho \in G_T} (\frac{1}{s-\rho}) - \sum_{\rho \in G_T} (\frac{1}{2+iT-\rho}). \tag{5.10}$$

Each term in the first sum on the right-hand side of (5.10) can be shown to be $\leq \frac{3}{|\gamma - T|^2}$ from the fact that $\rho \notin G_T$. Corollary 5.2.2 states that there are $\mathcal{O}(log\ T)$ such terms, so the first sum gives the contribution $\mathcal{O}(log\ T)$. By similarly reasoning, each term in the third sum of (5.10) is ≤ 1 by the fact that every zero ρ of $\zeta(s)$ satisfies $\Re(\rho) \leq 1$. By corollary 5.2.1 the third sum gives the contribution $\mathcal{O}(log\ T)$. Hence, we can rewrite (5.10) as

$$\frac{\zeta'(s)}{\zeta(s)} = \sum_{\rho \in G_T} \frac{1}{s-\rho} + \mathcal{O}(log\ T). \tag{5.11}$$

We apply (5.11) when calculating $\Delta_L \; arg \; \zeta(s)$.

$$\begin{aligned}
\Delta_L \; arg \; \zeta(s) &= \int_L \Im(\frac{\zeta'(s)}{\zeta(s)})ds \\
&= \int_2^{2+iT} \Im(\frac{\zeta'(s)}{\zeta(s)})ds - \int_{\frac{1}{2}+iT}^{2+iT} \Im(\frac{\zeta'(s)}{\zeta(s)})ds \\
&= \mathcal{O}(1) - \int_{\frac{1}{2}+iT}^{2+iT} \Im(\frac{\zeta'(s)}{\zeta(s)})ds \\
&= \mathcal{O}(1) - \sum_{\rho \in G_T} \Big[\int_{\frac{1}{2}+iT}^{2+iT} \Im(\frac{1}{s-\rho}) \Big] - \mathcal{O}(log \; T)ds \\
&= \mathcal{O}(1) - \sum_{\rho \in G_T} \Big[\int_{\frac{1}{2}+iT}^{2+iT} \Im(\frac{1}{s-\rho})ds \Big] + \mathcal{O}(log \; T) \\
&= \mathcal{O}(log \; T), \qquad (5.12)
\end{aligned}$$

because $\int_{\frac{1}{2}+iT}^{2+iT} \Im(\frac{1}{s-\rho})ds \leq \pi$, and hence each term is $\mathcal{O}(1)$, and from corollary 5.2.1 we know there are $\mathcal{O}(log \; T)$ such terms.

By putting (5.5), (5.6), (5.8) and (5.12) together, we end up with the following estimate for the number of non-trivial zeros of $\zeta(s)$ with positive imaginary part less than T:

$$N(T) = \frac{T}{2\pi} log \; \frac{T}{2\pi} - \frac{T}{2\pi} + \frac{7}{8} + \mathcal{O}(log \; T) \qquad (5.13)$$

5.3 Finding the number $N(T, \chi)$

We remind ourself that an L-series $L(s, \chi)$ is defined by

$$L(s, \chi) := \sum_{n=1}^{\infty} \frac{\chi(s)}{n^s},$$

where $\chi(s)$ is a Dirichlet charachter. We notice that $\zeta(s)$ is just an L-series assosiated with the principal Dirichlet charachter χ_0. After an analytic continuation of the L-series $L(s, \chi)$, we now refeer to $L(s, \chi)$ as an L-function. As in the previous subsection we want to find an estimate for the number $N(T, \chi)$ of zeros of $L(s, \chi)$ in the rectangle $0 < \sigma < 1$, $|t| < T$. Notice here that because the zeros of $L(s, \chi)$ are not symmetrical around the real axis, as they were for $\zeta(s)$, we want to include the zeros with negative imaginary part as well.

Because $\zeta(s)$ is an L-function, the derivation of $N(T, \chi)$ will follow the same path of the derivation of $N(T)$. Thus the treatise will only highligth the main lines.

First, let's assume χ is a primitive Dirichlet character modulo q. That is, χ is not induced by a smaller modulo. An example of a non-primitive Dirichlet

character is the Dirichlet character χ_x modululo 8 defined by

$$\chi_x(a) = \begin{cases} 1 & \text{if } a = 1, 5, \\ -1 & \text{if } a = 3, 7, \\ 0 & \text{else.} \end{cases}$$

The Dirichlet charater χ_x is not primitive because it is identical with χ_y modulo 4, where $\chi_y(1) = 1$ and $\chi_y(3) = -1$.

As with $\zeta(s)$, we define

$$\xi(s, \chi) = (\frac{q}{\pi})^{\frac{s}{2}+\frac{\alpha}{2}}\Gamma(\frac{s}{2} + \frac{\alpha}{2})L(s, \chi),$$

where

$$\alpha = \begin{cases} 0 & \text{if } \chi(-1) = 1, \\ 1 & \text{if } \chi(-1) = -1. \end{cases}$$

We want to calculate $\Delta_R \; arg \; \xi(s)$, R being the rectangle with vertices

$$\{\frac{5}{2} - iT, \frac{5}{2} + iT, -\frac{3}{2} + iT, -\frac{3}{2} - iT\}.$$

Inside R is all non-trivial zeros of $L(s, \chi)$ with $|t| < T$, but also one trivial zero at either $s = 0$ or $s = -1$ depending on α. This zero is being cancelled by the pole of $\Gamma(s)$. So

$$\Delta_R \; arg \; \xi(s) = 2\pi N(T, \chi).$$

We have the following relation

$$arg \; \xi(s, \chi) = arg \; \overline{\xi(\overline{1 - s}, \chi)} + c,$$

c being a constant independent of s. This relation gives us the opportunity to only consider the path of integration L, consisting of the three lines

$$\tfrac{1}{2} - iT \text{ to } \tfrac{5}{2} - iT, \quad \tfrac{5}{2} - iT \text{ to } \tfrac{5}{2} + iT, \quad \text{and } \tfrac{5}{2} + iT \text{ to } \tfrac{1}{2} + iT,$$

instead of integrating along the whole of the boundary of R. After some calculations we end up with the following estimate for $N(T, \chi)$, which will be useful to us when considering Chebyshev's bias.

$$N(T, \chi) = \frac{T}{\pi} log(\frac{qT}{2\pi}) - \frac{T}{\pi} + \mathcal{O}(log \; T + log \; q). \tag{5.14}$$

Chapter 6

Formula for $\psi(x)$

The objective of this chapter is to find a formula for the Chebychev function $\psi(x)$. This will be accomplished by complex integration. As for the previous chapter, all results are due to Davenport from [1], with my contribution being to merely fill in the intermediate calculations. We have that

$$\psi(x) = \sum_{n \leq x} \Lambda(n) = \sum_{p^k \leq x} log\ p,$$

where $\Lambda(n)$ is the von Mangoldt function

$$\Lambda(n) = \begin{cases} log\ p & \text{if } n = p^k \text{ for some prime power } p^k, \\ 0 & \text{else.} \end{cases}$$

We observe that $\psi(x)$ makes a jump for every prime power p^k. For technical reasons, it is more useful to work with the function

$$\psi_0(x) = \begin{cases} \psi(x) & \text{if } x \text{ is not a prime power,} \\ \psi(x) + \frac{1}{2} log\ p & \text{if } x = p^k \text{ for some prime power.} \end{cases}$$

That is, $\psi_0(x)$ and $\psi(x)$ are equal for all x, except when x is some prime power p^k. In that case, $\psi_0(p^k)$ is sort of the average of $\psi(x)$ in the interval around the point $x = p^k$.

Then for $x > 1$, we have the following formula for $\psi(x)$, which we will justify in this chapter.

$$\psi_0(x) = x - \sum_{\rho} \frac{x^\rho}{\rho} - \frac{\zeta'(0)}{\zeta(0)} - \frac{1}{2} log(1 - x^{-2}), \tag{6.1}$$

ρ being the non-trivial zeros of $\zeta(s)$ and the sum goes over all ρ.

6.1 A useful integral

As mentioned above, we are going to prove (6.1) with the help of complex integration. For the variable $s = \sigma + it$, we have that

$$\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s} = -\frac{\zeta'(s)}{\zeta(s)} \tag{6.2}$$

whenever $\sigma > 1$. To see this

$$\begin{aligned}\frac{\zeta'(s)}{\zeta(s)} &= \frac{d}{ds} log\ \zeta(s) = \frac{d}{ds}\big[-\sum_p log(1-p^{-s})\big] \\ &= -\sum_p \frac{p^{-s} log\ p}{1-p^{-s}} = -\sum_p p^{-s} log\ p \sum_{k\geq 0} p^{-ks} \\ &= -\sum_{p,k\geq 0} p^{-(k+1)s} log\ p = -\sum_{n=1}^{\infty} n^{-s}\Lambda(n),\end{aligned}$$

and (6.2) follows.

We will use the following integral to find a formula for $\psi_0(x)$.

$$\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} y^s \frac{ds}{s} = \begin{cases} 0 & \text{if } 0 < y < 1, \\ \frac{1}{2} & \text{if } y = 1, \\ 1 & \text{if } y > 1, \end{cases} \tag{6.3}$$

where $c > 0$. The next computation will demonstrate how we can find $\psi_0(x)$ with the help of complex integration.

$$\begin{aligned}\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\Big[-\frac{\zeta'(s)}{\zeta(s)}\Big]\frac{x^s}{s}ds &= \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\sum_{n=1}^{\infty}\Lambda(n)n^{-s}\frac{x^s}{s}ds \\ &= \sum_{n=1}^{\infty}\Lambda(n)\big[\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}(\frac{x}{n})^s\frac{ds}{s}\big] \\ &= \sum_{n\leq x}\Lambda(n)(+\frac{1}{2}\Lambda(x) \quad \text{if } x \text{ is power of a prime }) \\ &= \psi_0(x).\end{aligned}$$

Lemma 6.1. *Let $\delta(y)$ be the function of y displayed in* (6.3)*, and*

$$I(y,T) = \frac{1}{2\pi i}\int_{c-iT}^{c+iT}\frac{y^s}{s}ds.$$

Then for $y > 0, c > 0, T > 0$, the following is true

$$|I(y,T) - \delta(y)| < \begin{cases} y^c\ min\{1, T^{-1}|log\ y|^{-1}\} & \text{if } y \neq 1 \\ cT^{-1} & \text{if } y = 1 \end{cases} \tag{6.4}$$

Notice that as $T \to \infty$, $I(y,T) \to \delta(y)$.

Proof. First let's assume $0 < y < 1$. Then $\delta(y) = 0$. Since $y^s/s \to 0$ as $\sigma \to \infty$, and this convergence is uniform with respect to t, we can change path of integration to the two horizontal integrals

$$I(y,T) = -\frac{1}{2\pi i}\int_{c+i\infty}^{\infty+iT}\frac{y^s}{s}ds + \frac{1}{2\pi i}\int_{c-iT}^{\infty-iT}\frac{y^s}{s}ds. \tag{6.5}$$

Let's look at the first integral. Now $\frac{1}{|s|} \leq \frac{1}{T}$ for all s in the path, and $y^\sigma \to 0$ when $\sigma \to \infty$, since $0 < y < 1$. Let $\gamma(\sigma) = \sigma + iT$ be the parametrization of the path, and notice that $\gamma'(\sigma) = 1$. Then

$$\Big|\frac{1}{2\pi i}\int_{c+iT}^{\infty+iT} \frac{y^s}{s} ds\Big| = \Big|\frac{1}{2\pi i}\int_c^\infty \frac{y^{\sigma+iT}}{\sigma+iT} d\sigma\Big| \leq \frac{1}{2\pi T}\int_c^\infty y^\sigma d\sigma \leq \frac{1}{2T}\frac{y^c}{|log\ y|}$$

The exact same bound can be found for the other integral in (6.5), and by adding the two integrals and their bounds we achieve one of the bounds in the *min*-parenthesis in (6.4).

To find the other bound in the *min*-parenthesis, consider changing the integral path of $I(y,T)$ from the vertical path with real part c, to the new path starting at the same points but now is a circular path with center at the origin. This is justified since y^c/s is meromorphic, and we stay clear of the pole at $s = 0$. Let R denote the distance between the origin and the point $s = (c - iT)$. We use that $|y^s| \leq y^c$ and $|s| = R$ for this path. If we apply the ML-inequality, we get

$$|I(y,T)| \leq \frac{1}{2}\int \frac{|y^s|}{|s|} ds \leq \frac{1}{2\pi}\pi R\frac{y^c}{R} < y^c.$$

This finish the proof of the lemma in the case of $0 < y < 1$.

Next, we will assume $y > 1$. Now $\delta(y) = 1$. Let V denote the vertical integration path of $I(y,T)$, and A be the path starting at $s = (c + iT)$ going as an arc with positive orientation ending up at $s = (c - iT)$ and with center at the origin and with radius R. Denote $-A$ as the path going in the opposite direction of A. Observe the closed path $V + A$ contains the pole at $s = 0$ for y^s/s in which the residue is 1. Using Cauchy's residue theorem gives

$$\begin{aligned}
|I(y,T) - \delta(y)| &= \Big|\frac{1}{2\pi i}\int_V \frac{y^s}{s} ds - 1\Big| \\
&= \Big|\frac{1}{2\pi i}\Big(\int_V + \int_A + \int_{-A}\Big)\frac{y^s}{s} ds - 1\Big| \\
&= \Big|\frac{1}{2\pi i}\int_{-A} \frac{y^s}{s} ds + \frac{1}{2\pi i}\oint_{V+A} \frac{y^s}{s} ds - 1\Big| \\
&= \Big|\frac{1}{2\pi i}\int_{-A} \frac{y^s}{s} ds\Big|.
\end{aligned}$$

On $-A$ we have $|y^s| \leq y^c$ and $|s| = R$. So

$$\Big|\frac{1}{2\pi i}\int_{-A} \frac{y^s}{s} ds\Big| \leq \frac{1}{2\pi}\int_{-A} |\frac{y^s}{s}| ds < \frac{1}{2\pi}\pi R\frac{y^c}{R} < y^c,$$

thus, proving the claim for $y \neq 1$.

Lastly, assume $y = 1$. In this case we can simply compute the integral. Perform the substitution $s = c + it$. Then $ds = i\ dt$. Then

$$I(1,T) = \frac{1}{2\pi}\Big[\int_{-T}^0 \frac{dt}{c+it} + \int_0^T \frac{dt}{c+it}\Big]. \tag{6.6}$$

For the first integral in (6.6) we have that

$$\int_{-T}^{0} \frac{dt}{c+it} = -\int_{0}^{-T} \frac{dt}{c+it} = \int_{0}^{T} \frac{dt}{c-it} = \int_{0}^{T} \frac{c+it}{c^2+t^2} dt,$$

and for the second integral in (6.6)

$$\int_{0}^{T} \frac{c-it}{c^2+t^2} = \int_{T}^{0} \frac{c-it}{c^2+t^2} dt.$$

Putting this together gives

$$\begin{aligned} I(1,T) &= \frac{1}{\pi}\int_{0}^{T} \frac{c}{c^2+t^2} = \frac{1}{\pi}\int_{0}^{\frac{T}{c}} \frac{du}{1+u^2} = \frac{1}{\pi}\int_{0}^{\infty} \frac{du}{1+u^2} - \frac{1}{\pi}\int_{\frac{T}{c}}^{\infty} \frac{du}{1+u^2} \\ &< \frac{1}{2} - \frac{c}{T}. \end{aligned}$$

The last inequality comes from the fact that $\frac{d}{du} arctan(u) = \frac{1}{1+u^2}$, and by using the Taylor expansion of $arctan(u)$ at $u = \infty$. Since $\delta(1) = \frac{1}{2}$, we are finished proving the lemma. □

We apply lemma 6.1 on $\psi_0(x)$. Let

$$J(x,T) = \frac{1}{2\pi i}\int_{c-iT}^{c+iT} \Big[-\frac{\zeta'(s)}{\zeta(s)} \Big] \frac{x^s}{s} ds. \tag{6.7}$$

Then

$$\begin{aligned} |\psi_0(x) - J(x,T)| &= \sum_{n=1}^{\infty} \Lambda(n) \Big| \delta(\frac{x}{n}) - I(\frac{x}{n}, T) \Big| \\ &< \sum_{\substack{n=1 \\ n \neq x}}^{\infty} \Big[\Lambda(n) (\frac{x}{n})^c min\{1, T^{-1} |log \frac{x}{n}|^{-1}\} \Big] + cT^{-1}\Lambda(x), \end{aligned}$$

where $c > 1$. It can been shown [1] that

$$|\psi_0(x) - J(x,T)| = \mathcal{O}\Big(\frac{x(log\ x)^2}{T} + (log\ x)\ min\{1, x(T < x >)^{-1}\} \Big), \tag{6.8}$$

where $< x >$ denote the closest prime power to x. As we see from (6.8), for x fixed, $J(x,T) \to \psi_0(x)$ as $T \to \infty$. Our investigation will therefore continue with evaluating $J(x,T)$, and let T go to ∞. This will produce an estimate for $\psi_0(x)$ along with an error term.

6.2 Evaluation of the integral $J(x,T)$

The path of integration of $J(x,T)$, as defined by (6.7), is the vertical line from $s = c - iT$ to $s = c + iT$. The way of our approach to evaluate $J(x,T)$, will be by integrating the integrand of $J(x,T)$ along a closed path, where the vertical line mentioned is part of the closed path. What is left is to subtract the contribution of the path of integration which is not the original vertical line.

The closed path will be the boundary of a rectangle with vertices

$$c - iT, \qquad c + iT, \qquad -U + iT, \qquad -U - iT.$$

Here U is a large odd integer and the purpose is to avoid any trivial zeros of $\zeta(s)$. Denote the closed path of integration by R. By the residue theorem

$$\frac{1}{2\pi i}\int_R \Big[-\frac{\zeta'(s)}{\zeta(s)}\Big]\frac{x^s}{s}ds = x - \sum_{|\gamma|<T}\frac{x^\rho}{\rho} - \frac{\zeta'(0)}{\zeta(0)} + \sum_{\substack{m\in\mathbb{N}\\ 0<2m<U}}\frac{x^{-2m}}{2m}. \tag{6.9}$$

The first sub-path of R we treat is the horizontal paths where $s = \sigma \pm iT$ and $-1 \leq \sigma \leq 2$. Denote the union of these paths by H. From (5.11) in the previous chapter, the following formula is valid for $s \in H$.

$$\frac{\zeta'(s)}{\zeta(s)} = \sum_{\rho\in G_T}\frac{1}{s-\rho} + \mathcal{O}(log\ T).$$

Let

$$max\ G_T = max_{\rho\in G_T}\{|\gamma - T|\}.$$

We remember that G_T is the set of zeros ρ of $\zeta(s)$ with an imaginary γ such that $|T-\gamma| < 1$. By corollary 5.2.1, $|G_T| = \mathcal{O}(log\ T)$. Together with $\frac{1}{s-\rho} = \mathcal{O}(log\ T)$ for $s \in H$, we can conclude that $\frac{\zeta'(s)}{\zeta(s)} = \mathcal{O}(log^2\ T)$ for all $s \in H$ (See [1, p. 108] for details). From this we can see that

$$\frac{1}{2\pi i}\int_H \Big[-\frac{\zeta'(s)}{\zeta(s)}\Big]\frac{x^s}{s}ds \ll log^2\ T\int_{-1}^{c}\Big|\frac{x^s}{s}\Big|d\sigma \ll \frac{log^2\ T}{T}\int_{-\infty}^{c}x^\sigma d\sigma \ll \frac{xlog^2\ T}{T\ log\ x} \tag{6.10}$$

The remainding horizontal path of $R\backslash H$ gives the following contribution to (6.9)

$$\ll \frac{log\ T}{Tx\ log\ x}, \tag{6.11}$$

and the left vertical path with real part $-U$ contributes

$$\ll \frac{T\ log\ U}{Ux^U}. \tag{6.12}$$

Now we see that (6.11) is smaller than (6.10) for sufficiently large x and T, and can therefor be incorporated in the latter. For the contribution (6.12), we see that it disappears when $U \to \infty$. Which in turn leads to the following formula for $J(x,T)$ with an error term $E(x,T)$.

$$J(x,T) = x - \sum_{|\gamma|<T}\frac{x^\rho}{\rho} - \frac{\zeta'(0)}{\zeta(0)} - \frac{1}{2}log(1-x^{-2}) + E(x,T).$$

Putting this together with (6.8) ends our quest for a formula of $\psi_0(x)$.

$$\psi_0(x) = x - \sum_{|\gamma|<T}\frac{x^\rho}{\rho} - \frac{\zeta'(0)}{\zeta(0)} - \frac{1}{2}log(1-x^{-2}) + E(x,T).$$

Here the error term $E(x,T)$ is $\mathcal{O}\Big(\frac{x\log^2\ (xT)}{T} + (log\ x)min\{1, \frac{x}{T<x>}\}\Big)$, and will vanish when $T \to \infty$.

6.3 Formula for $\psi(x, \chi)$

We used the Dirichlet characters χ to generalize the Riemann zeta function $\zeta(s)$ to L-series $L(s, \chi)$. In the same fashion, define $\psi(x, \chi)$ by

$$\psi(x, \chi) := \sum_{n \leq x} \chi(n)\Lambda(n).$$

That is, the Chebychev function $\psi(s)$ can be rewritten as $\psi(s) = \psi(s, \chi_0)$, where χ_0 is the principal character. As Davenport shows in [1], a formula can be obtained for $\psi(x, \chi)$ using similar approach as we showed for $\psi(x)$.

$$\psi(x, \chi) = -\sum_{|\lambda| \leq T} \frac{x^\rho}{\rho} + \mathcal{O}\Big(\frac{xlog^2(xT)}{T} + log\ x\Big). \tag{6.13}$$

The formula (6.13) will be used in the next chapter when we encounter Chebyshev's bias. Notice from (6.13) that, in order to calculate $\psi(x, T)$, we need to know the zeros $\rho = \beta + i\gamma$ of the L-function $L(s\chi)$ with $|\gamma| \leq T$. By choosing x and T large enough, we can get the estimate of $\psi(s, \chi)$ as accurate as we want.

Chapter 7

Frobenius and Chebotarëv's Density Theorems

In this chapter we will highlight the connection between what we have deciced to call Dirichlet's density theorem, and two density theorems credited to Frobenius and Chebotarëv. The reader should be aware that here is a difference between Dirichlet's density theorem, and Dirichlet's theorem, although the theorems are linked. The content in this chapter relies on an article of P. Stevenhagen and H.W.Lenstra, Jr.[5].

First we will state the Dirichlet's density theorem 7.1. Then we need to define decomposition type of a polynomial f, and cycle pattern of a permutation in order to formulate Frobenius density theorem (theorem 7.2). To see the connection between the two mentioned theorems, we will use an example from [5], where I have filled in the details. These calculations will involve Galois theory, and a couple of results concerning the Legendre symbol.

As we will see, Frobenius density theorem will imply Dirichlet's density theorem for certain moduli q and polynomials f. However, there is gap for which the implication fails. This gap is covered by Chebotarëv's density theorem (theorem 7.4), letting us prove Dirchlet's density theorem with Chebotarë's density theorem.

7.1 Frobenius density theorem

To start, consider Dirichlet's density theorem.

Theorem 7.1 (Dirichlet's density theorem)**.** *Let $q \in \mathbb{N}$. Then for each integer a with $gcd(a, q) = 1$, the set of primes p with $p \equiv a (mod\ q)$ has density $1/\varphi(q)$.*

This is a precise statement of how the primes divide themselves evenly into the primitive residue classes. Dirichet's density theorem was first proved with respect to analytic density, but is also true for natural density. In order to state Frobenius density theorem, we need the notion of a decomposition types of a polynomial f.

Definition (Decomposition type)**.** *Let $f(x) \in \mathbb{Z}[x]$ be a polynomial and p be a prime. Further,*

$$f = g_1, g_2, \ldots, g_r \ (mod\ p),$$

where the g_i-s are irreducible polynomials over $\mathbb{Z}/p\mathbb{Z}$, and $i < j \implies deg(g_i) \leq deg(g_j)$. Then the decomposition type of f modulo p is defined by

$$dec_p(f) := \{deg(g_1), deg(g_2), \ldots, deg(g_r)\}.$$

The next example will show how one can use $dec_p(f)$ for different primes to determine whether $f(x)$ is irreducible in $\mathbb{Z}[x]$.

Let $f(x)$ be a monic polynomial in $\mathbb{Z}[x]$ of degree 4, and let p, q be two primes. Assume $dec_p(f) = \{1, 3\}$ and $dec_q(f) = \{2, 2\}$. Consider the map sending $f \in \mathbb{Z}[x]$ to $f(mod\ p) \in (\mathbb{Z}/p\mathbb{Z})[x]$, and use the fact that there is a one-to-one correspondence between a factor of f, and a factor g_i of $f(mod\ p)$, or possibly several factors g_i-s.

In the example above, $dec_p(f) = \{1, 3\}$ tells us that f has no factor of degree 2, and $dec_q(f) = \{2, 2\}$ makes it impossible for f to have a factor of degree 1 or 3. Hence, the fourth degree polynomial f can only have factors of degree 4, which makes it irreducible.

One can determine the irreducibility of $f(x) \in \mathbb{Z}[x]$ by a single prime p if $dec_p(f) = \{deg(f)\}$.

This next paragraph will be a setup to the theorem of Frobenius. Let $f(x) \in \mathbb{Z}[x]$ be a monic polynomial of degree n with non-zero discriminant $\Delta(f)$. Then f has n distinct zeros $\alpha_1, \alpha_2, \ldots, \alpha_n$ in an extension field $K = \mathbb{Q}(\alpha_1, \ldots, \alpha_n)$.

Denote the Galois group of f by G. Then from Galois theory, we know that an element $\sigma \in G$ permutes the zeros of f. Any $\sigma \in G$ can be written as a product of disjoint cycles(possibly of length 1). Now list the length of these cycles and sort them in an increasing order. This sorted list $n_1, n_2, \ldots, n_t$ is a partition of n, and is called the *cycle pattern* of σ.

Observe that we produce a partition of n in two different ways. Either by the cycle pattern of a given $\sigma \in G$. Or we could choose a prime p, and find $dec_p(f)$. These partitions are connected by Frobenius density theorem.

Theorem 7.2 (Frobenius density theorem)**.** *The density of the set of primes p for which $dec_p(f) = \{n_1, n_2, \ldots, n_t\}$ exists, and is equal to*

$$\frac{|A|}{|G|},$$

where

$$A = \{\sigma \in G | \textit{cycle pattern of } \sigma \textit{ is } n_1, n_2, \ldots, n_t\}.$$

We are going to apply Frobenius density theorem to the polynomial $f(x) = x^{12} - 1$, in order to prove Dirichlet's density theorem in the case of $q = 12$.

7.2 Connection between the density theorems of Frobenius and Dirichlet

The following example is found in [5], but is presented here with more details. The idea is to choose a polynomial f in such a way that the decomposition type of a given prime p, is only dependent on which residue class p belongs to modulo q.

So let $f(x) = x^{12} - 1$. Then we can initially factorize f in $\mathbb{Z}[x]$ in this way.

$$x^{12} - 1 = (x-1)(x+1)(x^2+1)(x^2-x+1)(x^2+x+1)\Phi_{12}(x), \qquad (7.1)$$

where $\Phi_{12}(x)$ is the 12-th cyclotomic polynomial, and is of degree $\varphi(12) = 4$. It turns out that the reducibility of the factors in (7.1)(in $\left(\mathbb{Z}/p\mathbb{Z}\right)[x]$) is totally determined by which residue class p belong to modulo 12. For the remainder of this example, all polynomials will be thought of to be in the polynomial ring $\left(\mathbb{Z}/p\mathbb{Z}\right)[x]$.

To help us decide how $\Phi_{12}(x)$ factors, we have this part of theorem 2.47 found in [2].

Theorem 7.3. *Let $gcd(a,q) = 1$. Then $\Phi_q(x) \in \mathbb{Z}[x]$ factors into $\varphi(q)/d$ distinct monic irreducible polynomials in $\left(\mathbb{Z}/p\mathbb{Z}\right)[x]$ of the same degree d, where d is the least positive integer such that*

$$a^d \equiv 1 (mod\ q). \qquad (7.2)$$

Applying theorem 7.3 on $\Phi_{12}(x)$, we see that $\Phi_{12}(x)$ factors into four linear factors if $p \equiv 1(mod\ 12)$, and factors into two quadratic polynomials if $p = 5, 7, 11(mod\ 12)$. The only thing left is to determine whether each of the quadratic polynomials in (7.1) is irreducible. This will be accomplished by the Legendre symbol.

From the law of quadratic reciprocity we have the following fact for the Legendre symbol.

$$\left(\frac{-1}{p}\right) = \begin{cases} 1 & \text{if } p \equiv 1(mod\ 4), \\ -1 & \text{if } p \equiv 3(mod\ 4). \end{cases}$$

Hence $x^2 + 1$ is reducible when $p \equiv 1, 5(mod\ 12)$, and irreducible for $p \equiv 7, 11(mod\ 12)$.

Now $x^2 - x + 1$ and $x^2 + x + 1$ are reducible if and only if their discriminant is a quadratic residue modulo p. Notice that the two polynomials have the same discriminant. So we evaluate the Legendre symbol $\left(\frac{-3}{p}\right)$. Putting the fact together that the Legendre symbol is multiplicative,

$$\left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}}, \qquad \text{and(for } p > 3) \qquad \left(\frac{3}{p}\right) = \begin{cases} 1 & \text{if } p \equiv \pm 1(mod\ 12), \\ -1 & \text{if } p \equiv \pm 5(mod\ 12), \end{cases}$$

we get that $x^2 - x + 1$ and $x^2 + x + 1$ are irreducible for $p \equiv 5, 11(mod\ 12)$, and reducible for $p \equiv 1, 7(mod\ 12)$.

Under is a summarization of what we have found for $f(x) = x^{12} - 1$.

$$\begin{aligned} dec_{p\equiv 1}(f) &= \{1,1,1,1,1,1,1,1,1,1,1,1\} \\ dec_{p\equiv 5}(f) &= \{1,1,1,1,2,2,2,2\} \\ dec_{p\equiv 7}(f) &= \{1,1,1,1,1,1,2,2,2\} \\ dec_{p\equiv 11}(f) &= \{1,1,2,2,2,2,2\} \end{aligned} \qquad (7.3)$$

By the above, We have found the decomposition type for a given prime p modulo 12. Now we want to know the cycle pattern of an element in the Galois group of $f(x) = x^{12} - 1$. The Galois group consists of field automophisms of the splitting field of $x^{12} - 1$, which permutes the zeros of f. Notice that the zeros of f are the 12-th unit roots, and their structure as a group is naturally isomorphic to $\mathbb{Z}/12\mathbb{Z}$. An automorphism of $\mathbb{Z}/12\mathbb{Z}$ must send a generator, say 1, to another generator, and this totally determine the automorphism.

Under is all four automorhisms of $\mathbb{Z}/12\mathbb{Z}$ along with each cycle pattern.

automorphism	cycle pattern
$1 \mapsto 1$	1,1,1,1,1,1,1,1,1,1,1,1
$1 \mapsto 5$	1,1,1,1,2,2,2,2
$1 \mapsto 7$	1,1,1,1,1,1,2,2,2
$1 \mapsto 11$	1,1,2,2,2,2,2

(7.4)

By comparing (7.3) and (7.4), and use Frobenius density theorem, we have now proved Dirichlet's density theorem for the case of $q = 12$. Unfortunately this procedure does not work for all q using the polynomial $f(x) = x^q - 1$. Take $q = 10$. Then $dec_{p\equiv 3}(f) = dec_{p\equiv 7}(f)$, which makes us unable to separate primes $p \equiv 3(mod\ 10)$ from $p \equiv 7(mod\ 10)$.

Here, Chebotarëv enters with his density theorem. He managed to prove a conjecture of Frobenius, which solves the issue at hand. Namely, to associate each prime p, which does not divide $\Delta(f)$, to an element $\sigma \in G$ with the cycle pattern of σ being equal to $dec_p(f)$. The next section will concern this connection.

7.3 Chebotarëv's density theorem

This section will solve the following problem: How to construct a one-to-one association between a prime $p(mod\ q)$, and an element σ_p in the Galois group of $f(x)$. In the case of Frobenius, we matched primes $p(mod\ q)$ with an automorphism σ_p of G, if $dec_p(f)$ was the same as the cycle pattern of σ_p. But some primes could have the same decomposition type, which stopped us from proving Dirichlet's density theorem using Frobenius density theorem. So we are going to look at what is called Frobenius substitution of p, denoted σ_p, for a given prime p.

Firstly, for a given prime p, we have the Frobenius map $Frob : \overline{\mathbb{F}_p} \to \overline{\mathbb{F}_p}$, defined by

$$Frob(\alpha) = \alpha^p.$$

Here $\overline{\mathbb{F}_p}$ denotes the algebraic closure of $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$. The Frobenius map is actually a field automorphism. What is left now, is to relate the automorphism $Frob$ of the field $\overline{\mathbb{F}_p}$, to an automorphism σ_p of the field $\mathbb{K} = \mathbb{Q}(\alpha_1, \alpha_2, \ldots, \alpha_n)$. The α-s are the zeros of f.

Definition. *(place) A place of* $\mathbb{K}$ *over* p *is a map* $\psi : \mathbb{K} \to \overline{\mathbb{F}_p} \cup \{\infty\}$ *with the following properties.*

- $\psi^{-1}(\overline{\mathbb{F}_p})$ *is a subring of* $\mathbb{K}$*, and* $\psi : \phi^{-1}(\overline{\mathbb{F}_p}) \to \overline{\mathbb{F}_p}$ *is a ring homomorphism.*
- $\psi(x) = \infty$ *if and only if* $\psi(x^{-1}) = 0$*, for any non-zero* $x \in \mathbb{K}$*.*

If p is a prime not dividing $\Delta(f)$ then there exist a place ψ of $\mathbb{K}$ over p. Unfortunately, the choice of place is not unique. Luckily, a different choice of place would give us a different Frobenius substitution σ_p', which belongs to the same conjugacy class in G as the original Frobenius substitution σ_p. And since conjugate permutations have the same cycle pattern, we are okay. Here, we are looking at the automorphisms in G as permutations of the zeros of f.

Now, for the composition $Frob \circ \psi$, which is also a place of $\mathbb{K}$ over p, defining $Frob(\infty) = \infty$, we can define the Frobenius substitution.

Definition. *(Frobenius substitution) There is a unique element* $Frob_\psi \in G$ *such that.*

$$\psi \circ Frob_\psi = Frob \circ \psi. \tag{7.5}$$

This element is called the Frobenius substitution of p*.*

The notation of the Frobenius substitution may cause some confusion to the reader since we have used both $Frob_\psi$ and σ_p in G to denote the Frobenius substitution. Keep in mind that $Frob_\psi$ is unique for a given place ψ over p, while the notation σ_p suggesting that we think of the conjugacy class, which σ_p belongs to in G.

The Frobenius substitution $Frob_\psi$ satisfies

$$\psi(Frob_\psi(x)) = Frob(\psi(x)) \quad \forall x \in \mathbb{K}. \tag{7.6}$$

We can read from (7.6) that $Frob_\psi$ permutes the zeros $\alpha_1, \ldots, \alpha_n$ of f in the same way as $Frob$ permutes the zeros $\psi(\alpha_1), dots, \psi(\alpha_n)$ of $f(mod\ p)$. Thus, the cycle pattern of $Frob_\psi$ is the same as the decomposition type of $f(mod\ p)$. This finishes our setup, and we are ready to formulate Chebotarëv's density theorem.

Theorem 7.4 (Chebotarëv's density theorem)**.** *Let* $f(x) \in \mathbb{Z}[x]$ *be a monic polynomial, and assume that the discriminant* $\Delta(f)$ *is non-zero. Let* C *be a conjugacy class of the Galois group* G *of* f*. Then the set of primes* p *not dividing* $\Delta(f)$*, for which* σ_p *belongs to* C*, has a density, and this density equals* $|C|/|G|$*.*

In the case $f(x) = x^q - 1$, the Galois group is naturally isomorphic to the multiplicative group of units $\left(\mathbb{Z}/q\mathbb{Z}\right)^*$, which is abelian. So every conjugucy class C of G is of size 1, and $|G| = \varphi(q)$, leading to Dirichlet's density theorem. We end this section by something that Stevenhagen and Lenstra writtes in their article [5, p. 31].

"Chebotarëv's density theorem may be regarded as the least common generalization of Dirichlet's theorem on primes in arithmetic progressions(1837) and a theorem of Frobenius."

Chapter 8

Chebyshev's Bias

Dirichlet's density theorem states that the primes divide themselves equally among the primitive residue classes modulo any given positive integer q. However, empirical data would suggest some primitive residue classes is prefered over others for certain moduli. Chebyshev noticed that there are more primes $3(mod\ 4)$ than $1(mod\ 4)$ in 1853. In fact, there is a bias towards quadratic non-residue classes. Remember that a residue class $a(mod\ q)$ is called quadratic non-residue if there is no $b \in \mathbb{Z}/q\mathbb{Z}$ such that $b^2 \equiv a(mod\ q)$.

In this chapter we will develop the necessary theory to show Chebyshev's bias. five theorems will be stated along with the proof of one of the theorems. All results in this part originates from an article of M. Rubinstein and P. Sarnak [4].

8.1 Some theorems concerning Chebyshev's bias

Let's start with some definitions. For $a, q \in \mathbb{N}, x \in \mathbb{R}$, let

$$\pi(x, q, a) = \big|\{p \text{ prime } : p \leq x \text{ and } p \equiv a(mod\ q)\}\big|.$$

That is, $\pi(x, q, a)$ is the prime counter function for primes $p \equiv a(mod\ q)$. Further, let $a_1, a_2, \ldots, a_r \in \big(\mathbb{Z}/q\mathbb{Z}\big)^*$. Then we can define the following set, which is what most of our discussion will revolve around.

$$P_{q;a_1,a_2,\ldots,a_r} = \big\{x \geq 2 : \pi(x, q, a_1) > \pi(x, q, a_2) > \cdots > \pi(x, q, a_r)\big\}.$$

Definition. *(Logarithmic density) The logarithmic density $\delta(P)$ of the set P is equal $\delta(P) = \underline{\delta}(P) = \overline{\delta}(P)$, where*

$$\underline{\delta}(P) = \liminf_{x\to\infty} \frac{1}{x} \int_{t\in P\cap[2,x]} \frac{dt}{t},$$
$$\overline{\delta}(P) = \limsup_{x\to\infty} \frac{1}{x} \int_{t\in P\cap[2,x]} \frac{dt}{t},$$

if the two limits are equal.

Because of the non-existence of a natural density for the set $P_{q;a_1,a_2,\ldots,a_r}$, we need to use logaritmic density [4, p. 174]. To help us find the densities and biases of the sets $P_{q;a_1,\ldots,a_r}$ we introduce the vector-valued function $E_{q;a_1,\ldots,a_r}(x)$,

for $x \geq 2$.

$$E_{q;a_1,\ldots,a_r}(x) = \frac{log\ x}{\sqrt{x}}(\varphi(q)\pi(x,q,a_1) - \pi(x), \ldots, \varphi(q)\pi(x,q,a_r) - \pi(x)). \quad (8.1)$$

To see why (8.1) is a tool for exposing biases, consider the i-th entry in (8.1) if there where no bias towards certain residue classes. Then we would have $\varphi(q)\pi(x,q,a_i) \approx \pi(x)$, which in turn would make us expect that each entry of $E_{q;a_1,\ldots,a_r}(x)$ is close to 0 for most $x \in \mathbb{R}$.

We will need to assume the Generalized Riemann Hypothesis(GRH) in our results below, that is, GRH assumes that the real part of every non-trivial zero of a Dirichlet L-function $L(s,\chi)$ is $\frac{1}{2}$.

Theorem 8.1. *Assume GRH. Then $E_{q;a_1,\ldots,a_r}$ has a limiting distribution $\mu_{q;a_1,\ldots,a_r}$ on $\mathbb{R}^r$, that is,*

$$\lim_{X\to\infty} \frac{1}{log\ X} \int_2^X f(E_{q;a_1,\ldots,a_r}(x))\frac{dx}{x} = \int_{\mathbb{R}^r} f(x) d\mu_{q;a_1,\ldots,a_r}(x)$$

for all bounded continuous functions f on $\mathbb{R}^r$.

The messure $\mu_{q;a_1,\ldots,a_r}$ will contain the information regarding densities and biases that we looking for. If $\mu_{q;a_1,\ldots,a_r}$ is absolutely continuous we can use the messure $\mu_{q;a_1,\ldots,a_r}$ to find the logarithmic density of $P_{q;a_1,\ldots,a_r}$ by

$$\delta(P_{q;a_1,\ldots,a_r}) = \mu_{q;a_1,\ldots,a_r}(\{x \in \mathbb{R}^r : x_1 > x_2 > \cdots > x_r\}).$$

For the next theorem, define the sets

$$\begin{aligned} B'_R &= \{x \in \mathbb{R}^r :\ |x| \geq R\}, \\ B^+_R &= \{x \in \mathbb{R}^r :\ \epsilon(a_j)x_j > 0\}, \\ B^-_R &= -B^+_R, \end{aligned}$$

where

$$\epsilon(a_j) = \begin{cases} 1 & \text{if } a \equiv 1 (mod\ q), \\ -1 & \text{otherwise.} \end{cases}$$

Theorem 8.2. *Assume GRH. Then there exist $c_1, c_2, c_3, c_4 \in \mathbb{R}^+$, only depending on q such that*

$$\begin{aligned} \mu_{q;a_1,\ldots,a_r}(B'_R) &\leq c_1 e^{(-c_2\sqrt{R})}, \\ \mu_{q;a_1,\ldots,a_r}(B^\pm_R) &\geq c_3 e^{(-e^{c_4 R})}. \end{aligned}$$

The theorem says that the tails of the distributions are small. At first glance the set $B^\pm_R$ seems like a strange set to consider. The reason being that we are able to produce a non-zero lower bounds for this bound.

For historical reasons it is worth looking at the density of

$$P_1 = \{x \geq 2 : \pi(x) > li(x)\},$$

where $li(x) = \int_0^x dt/(log\ t)dt$ is the logarithmic integral function. Empirical evidence would imply that $li(x) > \pi(x)$ for all $x \geq 2$. But in 1914 Littlewood

showed that $li(x) - \pi(x)$ actually change sign infinitely many times as $x \to \infty$. Even though P_1 is proven to be an infinitely set, not a single member of P_1 is found. A number of people have worked on finding an upper bound for the first member of P_1. This first member is called Skewes's number, named after Stanley Skewes. Skewes found an upper bound, which have later been improved by te Riele to $< 10^{370}$. By calculation, $\delta(P_1) = 0,00000026\ldots$, an indication of how rare the $x \geq 2$ are, for which $\pi(x) > li(x)$.

This next example deals with what happens density-wise if we group quadratic residue classes and quadratic non-residue classes together. We have that $P_{3;2,1} = 0,9990\ldots$ and $P_{4;3,1} = 0,9959\ldots$, where we notice that $2(mod\ 3)$ and $3(mod\ 4)$ are quadratic non-residue classes. For small moduli q, there are a favouritism for primes to be in quadratic non-residue classes over quadratic residue classes. To make this precise, let

$$P_{q;N,R} = \{x \geq 2 : \pi_N(x,q) > \pi_R(x,q)\},$$
$$P_{q;R,N} = \{x \geq 2 : \pi_R(x,q) > \pi_N(x,q)\},$$

where $\pi_R(x,q)$ is the set of primes $\leq x$ which are quadratic residues, while $\pi_N(x,q)$ is the set of primes $\leq x$ which are quadratic non-residues. This discussion would be trivial for q with more quadratic non-residue classes than quadratic residue classes. So we will assume that q is on the form $q = 4, p^k, 2p^k$ for a prime $p \geq 3$ and $k \in \mathbb{N}$. For such q, the multiplicative group of integers $\left(\mathbb{Z}/q\mathbb{Z}\right)^*$ is cyclic, and this leads to the same amount of quadratic residue classes compared to quadratic non-residue classes modulo q. Below is some numbers collected from [4, p. 188] which shows the bias towards quadratic non-residue classes for certain small q.

$$\delta(P_{3;N,R}) = 0,9990\ldots,$$
$$\delta(P_{4;N,R}) = 0,9959\ldots,$$
$$\delta(P_{5;N,R}) = 0,9954\ldots,$$
$$\delta(P_{7;N,R}) = 0,9782\ldots,$$
$$\delta(P_{11;N,R}) = 0,9167\ldots,$$
$$\delta(P_{13;N,R}) = 0,9443\ldots.$$

For the next two theorems we will also need another assumption, namely the Grand Simplicity Hypothesis(GSH). GSH states that , for a given primitive Dirichlet character χ, the set of $\gamma \geq 0$ such that $L(\frac{1}{2} + i\gamma, \chi) = 0$ is linearly independent over $\mathbb{Q}$.

The next theorem will give us some conditions to when we can expect a bias to apper. First, we say that $(q; a_1, \ldots, a_r)$ is unbiased if the density function of $\mu_{q;a_1,\ldots,a_r}$ is invariant under permutation of $(x_1, x_2, \ldots, x_r)$.

Theorem 8.3. *Under GRH and GSH, $(q; a_1, \ldots, a_r)$ is unbiased if and only if either of the two conditions are fulfilled.*

- $r = 2$ *and* $c(q, a_1) = c(q, a_2)$.

- $r = 3$ *and therer exists* $\rho \neq 1$ *such that*

$$\begin{aligned} \rho^3 &\equiv 1 (mod\ q), \\ a_2 &\equiv a_1 \rho (mod\ q), \\ a_3 &\equiv a_1 \rho (mod\ q). \end{aligned}$$

Here,

$$c(q,a) = -1 + \sum_{\substack{b^2 \equiv a (mod\ q) \\ 0 \leq b \leq q-1}} 1. \tag{8.2}$$

That is, $c(q,a)$ is the number of square roots of a (modulo q) minus 1. The number $c(q,a)$ will be useful when proving theorem 8.1.

For an unbiased $(q; a_1, \ldots, a_r)$, the logarithmic density will be $\delta(P_{q;a_1,\ldots,a_r}) = (r!)^{-1}$. An interesting fact worth mentioning is that as q grows to infinity, the bias will actually vanish. Chebychev's bias is a phenomenom for small q.

Theorem 8.4. *Assume GRH and GSH. Then for fixed* r*,*

$$\max_{a_1,\ldots,a_r \in (\mathbb{Z}/q\mathbb{Z})^*} \left| \delta(P_{a_1,\ldots,a_r}) - \frac{1}{r!} \right| \to 0$$

as $q \to \infty$*.*

To finish, We mention that one can generalize this discussion on Chebychev's bias to apply to the realm of the theory of number fields and hyperbolic geometry.

8.2 Proof of theorem 8.1

In this section we will prove theorem 8.1 along the lines of [4, p. 178-181]. The idea of the proof is to show that we can estimate each entry of $E_{q;a_1,\ldots,a_r}$ by computing an approximation using a finite number of zeros ρ of the L-function $L(s,\chi)$, and this approximation will have an error which we can controll. Then we apply lemma 8.6, which ensures the existence of a limiting distribution of the approximation of an entry of $E_{q;a_1,\ldots,a_r}$.

Let $E(x,q,a_i)$ denote the i-th entry of $E_{q;a_1,\ldots,a_r}$. That is,

$$E(x,q,a) := (\varphi(q)\pi(x,q,a) - \pi(x)) \frac{log\ x}{\sqrt{x}}.$$

We also have that

$$\psi(x,\chi) = \sum_{n \leq x} \chi(n)\Lambda(n),$$

for which we found that

$$\psi(x,\chi) = -\sum_{|\gamma| \leq T} \frac{x^\rho}{\rho} + \mathcal{O}\Big(\frac{x log^2(xT)}{T} + log\ x\Big) \tag{6.13}$$

in a previous section. The summation is of all non-trivial zeros $\rho = \beta + i\gamma$ of the associated $L(s,\chi)$ with imaginary part γ smaller than T in absolute value.

Since we assume the Generalized Riemann Hypothesis, assuming $\beta = \frac{1}{2}$ for all zeros ρ, we get

$$\psi(x,\chi) = -\sqrt{x}\sum_{|\gamma|\leq T}\frac{x^{i\gamma}}{\frac{1}{2}+i\gamma} + \mathcal{O}\Big(\frac{xlog^2(xT)}{T} + log\ x\Big). \tag{8.3}$$

This formula will be used in the first of a total of three lemmas.

Lemma 8.5. *As $x \to \infty$, we have that*

$$E(x,q,a) = -c(q,a) + \sum_{\chi\neq\chi_0}\bar{\chi}(a)\frac{\psi(x,\chi)}{\sqrt{x}} + \mathcal{O}(\frac{1}{log\ x}).$$

The constant $-c(q,a)$ is defined by (8.2), and is the source for the bias towards quadratic non-residues.

Proof. First, let

$$\theta(x,q,a) := \sum_{\substack{p\leq x\\ p\equiv a(mod\ p)}} log\ p,$$

and

$$\psi(x,q,a) := \sum_{\substack{n\leq x\\ n\equiv a(mod\ q)}} \Lambda(n).$$

We can rewrite $\psi(x,q,a)$ as

$$\begin{aligned}\psi(x,q,a) &= \sum_{k\geq 1}\sum_{\substack{p^k\leq x\\ p^k\equiv a(mod\ q)}} log\ p \\ &= \sum_{\substack{p\leq x\\ p\equiv a(mod\ q)}} log\ p + \sum_{\substack{p^2\leq x\\ p^2\equiv a(mod\ q)}} log\ p + \sum_{k\geq 3}\sum_{\substack{p^k\leq x\\ p^k\equiv a(mod\ q)}} log\ p. \end{aligned}\tag{8.4}$$

The first sum of (8.4) is $\theta(x,q,a)$. We know that $\sum_{p^2\leq x} log\ p = \theta(\sqrt{x})$. Using proposition 3.1 and the prime number theorem (theorem 3.3) gives $\theta(\sqrt{x}) \sim \sqrt{x}$. So we get that

$$\sum_{\substack{p^2\leq x\\ p^2\equiv a(mod\ q)}} log\ p = \Big(\sum_{b^2\equiv a(mod\ q)} 1\Big)\frac{\sqrt{x}}{\varphi(q)} + \mathcal{O}(\frac{\sqrt{x}}{log\ x}),$$

from the fact that the primes divides themselves evenly among the primitive residue classes (Dirichlet's density theorem 4.1), and that we need to count each residue class b which is a solution to $b^2 \equiv a(mod\ q)$. The third sum in (8.4) can be incorporated in the error term $\mathcal{O}(\frac{\sqrt{x}}{log\ x})$. So

$$\psi(x,q,a) = \theta(x,q,a) + \Big(\sum_{b^2\equiv a(mod\ q)} 1\Big)\frac{\sqrt{x}}{\varphi(q)} + \mathcal{O}(\frac{\sqrt{x}}{log\ x}). \tag{8.5}$$

We need these two facts to prove the lemma.

$$\psi(x,q,a) = \frac{1}{\varphi(q)} \sum_{\chi} \bar{\chi}(a)\psi(x,\chi), \tag{8.6}$$

and

$$\pi(x,q,a) = \int_2^x \frac{d\theta(t,q,a)}{log\ t}. \tag{8.7}$$

Solving (8.5) for $\theta(x,q,a)$, and using (8.6) and (8.7) when substituting, we get

$$\pi(x,q,a) = \frac{1}{\varphi(q)} \int_2^x \frac{d\psi(t)}{log\ t} + \frac{1}{\varphi(q)} \sum_{\chi \neq \chi_0} \bar{\chi}(a) \int_2^x \frac{d\psi(t,\chi)}{log\ t} - \frac{1}{\varphi(q)} \Big(\sum_{b^2 \equiv a (mod\ q)} 1 \Big) \frac{\sqrt{x}}{log\ x} + \mathcal{O}\big(\frac{\sqrt{x}}{log^2 x}\big). \tag{8.8}$$

By integration by parts we can simplify (8.8). We will merely highlight the crusial point of this calculation. To begin with, let

$$G(x,\chi) = \int_2^x \psi(t,\chi)\ dt.$$

We know from (8.3), after integration and letting $X \to \infty$, that

$$G(x,\chi) = -\sum_{\gamma} \frac{x^{3/2+i\gamma}}{(\frac{1}{2}+i\gamma)(\frac{3}{2}+i\gamma)} + \mathcal{O}(x\ log\ x).$$

This series is absolute convergent by inspecting the formula for $N(T,\chi)$, which we found in (5.14), saying

$$N(T,\chi) = \frac{T}{\pi} log(\frac{qT}{2\pi}) - \frac{T}{\pi} + \mathcal{O}(log\ T + log\ q). \tag{5.14}$$

It follows that $G(x,\chi) \ll x^{3/2}$, and we find that

$$\pi(x,q,a) - \frac{\pi(x)}{\varphi(q)} = -\frac{c(q,a)\sqrt{x}}{\varphi(q) log\ x} + \frac{1}{\varphi(q) log\ x} \sum_{\chi \neq \chi_0} \bar{\chi}(a)\psi(x,\chi) + \mathcal{O}\big(\frac{\sqrt{x}}{log^2 x}\big),$$

which finishes the proof. □

We would very much like to compute $E(x,q,a)$. Lemma 8.5 together with the estimate (6.13) for $\psi(x,\chi)$ gives

$$E(x,q,a) = -c(q,a) - \sum_{|\gamma| \leq T} \frac{x^{i\gamma}}{\frac{1}{2}+i\gamma} + \epsilon_a(x,T).$$

The error term $\epsilon_a(x,T)$ can be made arbitrarily small for T large enough. (lemma 2.2 [4]) So we can estimate $E(x,q,a)$ by

$$E_j^{(T)}(y) = -c(q,a) - \sum_{|\gamma| \leq T} \frac{x^{i\gamma}}{\frac{1}{2}+i\gamma}$$

using non-trivial zeros $\rho = \beta + i\gamma$ of the L-function $L(s, \chi)$ with $|\gamma| \leq T$. Further, we can estimate $E_{q;a_1,\ldots,a_r}$ by

$$E^{(T)}(y) = (E_1^{(T)}(y),\ E_2^{(T)}(y),\ \ldots,\ E_r^{(T)}(y)).$$

This next lemma ensures a probability measure for $E^{(T)}(y)$, which in turn ensures a probability measure μ for $E_{q;a_1,\ldots,a_r}$ (See [4, p. 180-181]).

Lemma 8.6. *For each T there exists a probability measure ν_T on $\mathbb{R}^r$ such that*

$$\nu_t(f) := \int_{\mathbb{R}^r} f(x)d\nu_T(x) = \lim_{Y\to\infty} \frac{1}{Y}\int_{log\ 2}^{Y} f(E^{(T)}(y))dy$$

for all bounded continuous functions f on $\mathbb{R}^r$. Also, there is a constant k_q dependent on q such that the support of ν_T lies in the ball $B(0, k_q\ log^2\ T)$.

To summarize, the idea have been to estimate $E_{q;a_1,\ldots,a_r}$ by $E^{(T)}(y)$. This give rise to an error term, but the error term can be controlled by letting T be big enough. That is, using a necessary amount of non-trivial zeros of the L-function $L(s, \chi)$ to estimate $\psi(x, \chi)$.

Bibliography

[1] Harold Davenport. *Multiplicative number theory*. Third. Vol. 74. Graduate Texts in Mathematics. Revised and with a preface by Hugh L. Montgomery. Springer-Verlag, New York, 2000, pp. xiv+177. ISBN: 0-387-95097-4.

[2] Rudolf Lidl and Harald Niederreiter. *Introduction to finite fields and their applications*. Cambridge university press, 1986.

[3] Paul Pollack. *Not always buried deep*. A second course in elementary number theory. American Mathematical Society, Providence, RI, 2009, pp. xvi+303. ISBN: 978-0-8218-4880-7. URL: https://doi.org/10.1090/mbk/068.

[4] Michael Rubinstein and Peter Sarnak. "Chebyshev's bias". In: *Experiment. Math.* 3.3 (1994), pp. 173–197. ISSN: 1058-6458. URL: http://projecteuclid.org/euclid.em/1048515870.

[5] P. Stevenhagen and H. W. Lenstra Jr. "Chebotarëv and his density theorem". In: *Math. Intelligencer* 18.2 (1996), pp. 26–37. ISSN: 0343-6993. URL: https://doi.org/10.1007/BF03027290.