Durham University

The Distribution of Prime Numbers

Oliver Welch

Supervisor: Dr. Pankaj Vishe

April 29, 2020

Abstract

We start by motivating the study of prime numbers on a 'global scale', introducing various prime-counting problems. We then begin to tackle these problems with analytic methods, proving Dirichlet's theorem using the basic theory of L-functions and Dirichlet characters. With the duality between the zeros of L-functions in the complex plane and the primes as motivation, we move on to finding the analytic continuation of L-functions. This allows us to study the positioning and frequency of zeros in the 'critical strip'. Using this knowledge, we prove two more landmark results in number theory: the prime number theorem, and the prime number theorem for arithmetic progressions. Finally, we put our results into perspective by investigating primes in arithmetic progressions on a more local level. We highlight peculiar biases that occur, providing intuitive reasons for this where possible.

Contents

Notation

$\mathbb{N}$ The natural numbers, $\{1, 2, 3, \dots\}$

$\mathbb{Z}$ The integers, $\{0, \pm1, \pm2, \dots\}$

$\mathbb{R}$ The set of real numbers

$\mathbb{C}$ The set of complex numbers

$R^{\times}$ The multiplicative group of units of a ring R

$\exists, \forall$ The quantifiers 'there exists' and 'for all'

$\mathfrak{R}(z)$ The real part of a complex number z

$\mathfrak{I}(z)$ The imaginary part of a complex number z

$\overline{z}$ The complex conjugate of a complex number z

$\overline{n}$ The residue class of an integer n to a given modulus

$[x]$ The greatest integer not exceeding a real number x

$a \equiv b \ (q)$ For $a, b, q \in \mathbb{Z}$, a is congruent to b modulo q

$(a, b) = c$ The greatest common divisor of integers a and b is c

$\phi(q)$ Euler's totient function evaluated at q

$f(x) = O\left(g(x)\right)$ if $\exists\ C,\ x_0 > 0$ s.t. $\forall\ x > x_0,\ |f(x)| \leq Cg(x)$

$f(x) \ll g(x)$ Equivalent to $f(x) = O(g(x))$

$f(x) = g(x) + O\left(h(x)\right)$ if $|f(x) - g(x)| = O\left(h(x)\right)$

$f(x) \sim g(x)$ if $f(x)/g(x) \to 1$ as x tends to some limit (∞ unless otherwise stated)

$f(x) = o(g(x))$ if $f(x)/g(x) \to 0$ as $x \to \infty$

Chapter 1

Introduction

Prime numbers are fundamental in the study of number theory; they are the 'atoms' from which all other integers are built. Indeed, any whole number can be uniquely represented (up to multiplication by ± 1 and re-ordering) as a product of prime numbers. This is known as the *fundamental theorem of arithmetic*. The primes have fascinated mathematicians for hundreds of years, mainly due to their highly non-trivial nature. To quote Euler [1]:

Quote 1.1. *"Mathematicians have tried in vain to this day to discover some order in the sequence of prime numbers, and we have reason to believe that it is a mystery into which the human mind will never penetrate."*

In addition to the beauty and mystery of the primes, they have very practical uses in cryptography such as the RSA algorithm [2]. Our aim therefore is to prove the great Euler wrong (to some extent!). The first natural question about the primes is simply 'how many are there?'. This was known by Euclid over 2000 years ago [3].

Proposition 1.2. *(Euclid, 300 BC) There are infinitely many prime numbers.*

Proof. Euclid's classic proof is by contradiction and is particularly elegant. Assume there are finitely many primes, and enumerate them as $p_1, p_2, \ldots, p_n$ in an exhaustive list. Then, set

$$N = p_1 p_2 \ldots p_n + 1.$$

None of the primes p_i divide N, but N is either prime or it has a prime dividing it. Consequently, there exists another prime not in the list $p_1, \ldots p_n$, which is a contradiction. $\square$

Broadly speaking, problems about the primes can be categorised as either *global problems* or *local problems*. The former are concerned with questions such as 'how many', while the latter studies properties of individual primes p_n, such as the size of $p_{n+1} - p_n$. There is a wealth of unsolved problems in both categories, but the main focus of this report is to solve those of the global kind with analytic methods. According to Davenport [4, p. 1], the adoption of such methods to tackle problems about primes

arguably began with Dirichlet, whose 1837 memoir studied the existence of primes in arithmetic progressions. Such progressions are of the form

$$a, \quad a+q, \quad a+2q, \quad a+3q, \quad \ldots$$

where a and q are integers. Note that if a and q share a common factor, then all numbers in the progression (with the possible exception of a) are not prime. It is therefore only interesting to study those progressions when a and q are coprime. It was conjectured long before the work of Dirichlet that such a progression always contains infinitely many primes - Dirichlet gave a proof of this with the (rather large) assumption of his *class number formula* [4, p. 1], the details of which we will not delve into. For certain examples, it is elementary to show that there are infinitely many primes. For example, consider the progression $4n + 3$, for non-negative integers n. Suppose there were finitely many primes of this form, say $p_1, \ldots, p_n$, and consider the number

$$N = 4p_1 \ldots p_n + 3.$$

N can be written as a product of primes, and since N is odd these are either of the form $4n + 1$ or $4n + 3$. Moreover, none of these can be of the form $4n + 3$, as N is not divisible by any of the p_i. Thus, N is the product of primes of the form $4n + 1$, and we can easily check that this implies N is also of the form $4n + 1$ - a contradiction. This method is limited: what if q is not a number such as 4, but is 400, or 4000? One of our first aims is to prove the existence of infinitely many primes in arithmetic progressions for *arbitrary* coprime integers q and a. This is known as Dirichlet's theorem.

Further advances in analytic number theory followed Riemann's memoir of 1859, which led to important results regarding the number of primes less than a given value. To this end, one studies the function $\pi(x)$, defined as the number of primes $p \leq x$. Gauss had conjectured early in his life that

$$\pi(x) \sim \frac{x}{\log x} \text{ as } x \to \infty,$$

where $f(x) \sim g(x)$ if the limit of f/g is 1. He later refined his estimate to

$$\pi(x) \sim \mathrm{Li}(x) = \int_2^x \frac{\mathrm{d}t}{\log t},$$

which actually gives a much sharper approximation (and implies his previous estimate by integration by parts). We note the numerical evidence detailed in Table 1.1.

x	$\pi(x)$	$|\pi(x) - x/\log x|$	$|\pi(x) - \mathrm{Li}(x)|$
10^2	25	3	5
10^4	1229	143	17
10^6	78498	6116	130
10^8	5761455	332774	754
10^{10}	455052511	20758029	3104

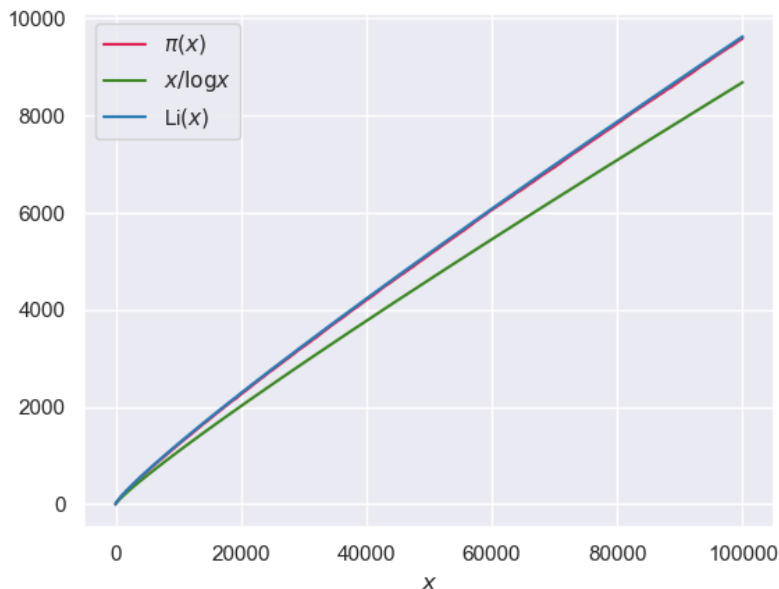Table 1.1: Errors in Gauss's approximations for $\pi(x)$.[1]

Figure 1.1: A comparison of the growth of $\pi(x)$, $\mathrm{Li}(x)$, and $x/\log x$ for values up to $x = 100,000$. This illustrates plainly how well $\mathrm{Li}(x)$ approximates $\pi(x)$.[2]

Note that the number of digits in the right-most column is always roughly half the number of digits in the left-most, suggesting the error in approximating $\pi(x)$ by $\mathrm{Li}(x)$ is approximately $\sqrt{x}$. The problem of quantifying this error term, and proving that $\pi(x) \sim \mathrm{Li}(x)$ is a central problem in this report. Gauss's conjecture was first proved independently by Hadamard and de la Vallée Poussin in 1896, extending the crucial work in Riemann's memoir, with a result known as the *prime number theorem* [5, p. 2].

Results relating to $\pi(x)$ may then be extended using similar methods to results about $\pi(x; q, a)$, defined as the number of primes less than or equal to x in the arithmetic progression $qn + a$. This is the ultimate aim of the report: to say something concrete about how the primes are distributed among each arithmetic progression with respect to a given q, and how many primes there are in each progression less than a given x. This is done through a generalisation of the prime number theorem, namely the *prime number theorem for arithmetic progressions*. We then detail further works relating to the distribution of primes, as well as significant open problems in the area. For now, we begin at the historical starting point of analytic number theory: the basic theory developed by Dirichlet.

---

[1]The On-Line Encyclopedia of Integer Sequences, published electronically at `https://oeis.org`. Sequences A006880, A057835, A057752. [Accessed: 26-04-2020]

[2]Plot made using Python 3.7.4 with the library Seaborn 0.10.0

Chapter 2

Dirichlet L-Functions

2.1 Definition and Connection to Primes

To study the distribution of primes, we first define a very important class of function, known as Dirichlet L-functions. These are defined for a complex variable, whose standard notation is $s = \sigma + it$. Unless otherwise stated, we follow [6] here.

Definition 2.1. For $\sigma > 1$, Dirichlet L-functions are defined as

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

where $\chi(n)$ is a *Dirichlet character* of modulus $q \in \mathbb{N}$.

Dirichlet characters are another special class of function which are interesting in their own right. They are defined as follows.

Definition 2.2. For a given $q \in \mathbb{N}$, a Dirichlet character χ of modulus q is a complex-valued function from the integers satisfying:

- $\chi(n + q) = \chi(n)$ for all $n \in \mathbb{Z}$ (q-periodicity).

- $\chi(mn) = \chi(m)\chi(n)$ (Multiplicativity).

- $\chi(n) \neq 0$ if and only if $(n, q) = 1$.

For the time being, we will not concern ourselves with why these characters are so important, but simply use their properties to exhibit the connection of L-functions to the primes. L-functions are generalisations of the famous *Riemann zeta-function*,

$$\zeta(s) := \sum_{n=1}^{\infty} n^{-s}, \quad (\sigma > 1). \tag{2.1}$$

First note that $\zeta(s)$ is uniformly convergent to an analytic function on $\sigma > 1$. For we can bound the absolute value of each term in the infinite sum by $n^{-1-\varepsilon}$, for some $\varepsilon > 0$. Then since

$$\int_1^\infty x^{-1-\varepsilon}\mathrm{d}x < \infty,$$

the integral test implies convergence of $\sum_n n^{-1-\varepsilon}$. We may therefore bound the absolute value of $\zeta(s)$ on $\sigma > 1$ by this convergent sum independent of s, so by Weierstrass's M-test we have uniform convergence. It is then a consequence of Proposition A3 (see appendix - we will use this result implicitly from now) that $\zeta(s)$ is holomorphic on this region. L-functions are also holomorphic on this region: we may use the same bound as for $\zeta(s)$, following easily from the fact that $|\chi(n)| \leq 1$ for all integers n, which will be shown in the next section. L-functions are intimately connected to the primes via an infinite product representation; we use a theorem from [7, Theorem 1.1] to show this.

Theorem 2.3. *Let $f : \mathbb{N} \to \mathbb{C}$ be a multiplicative function. That is, for every coprime $m, n \in \mathbb{N}$, we have $f(mn) = f(m)f(n)$. Suppose that $\sum_{n=1}^\infty |f(n)| < \infty$. Then*

$$\sum_{n=1}^\infty f(n) = \prod_{p \text{ prime}} \left(\sum_{n=0}^\infty f(p^n) \right),$$

where the infinite product is over every prime number p.

Proof. Using the fact that $\sum_{n=1}^\infty |f(n)| < \infty$, we may expand the brackets of terms in the product over finitely many primes. This gives

$$\prod_{p \leq x} (1 + f(p) + \dots) = \sum_{n \in S_x} f(n).$$

where S_x denotes the subset of natural numbers whose prime factors are all at most x. This follows from the multiplicativity of f and the fundamental theorem of arithmetic, so that each term in the expansion of the finite product is covered once and only once on the right hand side. Then let

$$\sum_{n \in S_x} f(n) = \sum_{n \leq x} f(n) + R(x).$$

Here, $R(x)$ is the sum of $f(n)$ for all the numbers greater than x with prime factors less than or equal to x, so that each term in the left hand side is covered exactly once. Therefore,

$$|R(x)| \leq \sum_{n > x} |f(n)| \to 0 \text{ as } x \to \infty,$$

since the tails of a convergent sum tend to zero. Hence, we let $x \to \infty$, to give

$$\prod_{p \text{ prime}} (1 + f(p) + \dots) = \sum_{n=1}^\infty f(n).$$

$\square$

In the case of L-functions, the hypothesis of the theorem is satisfied: the terms in the series are multiplicative (owing to the properties of Dirichlet characters) and absolutely convergent on $\sigma > 1$. Therefore, we have

$$
\begin{aligned}
L(s, \chi) &= \prod_{p \text{ prime}} \left(\sum_{n=0}^{\infty} \chi(p^n) p^{-ns} \right) \\
&= \prod_{p \text{ prime}} \left(\frac{1}{1 - \chi(p) p^{-s}} \right), \quad (\sigma > 1)
\end{aligned}
\tag{2.2}
$$

where the last step applies the formula for the geometric series, which is justified as the absolute value of each term is less than 1. In the case of the zeta-function, we recover the famous product formula first discovered by Euler:

$$
\zeta(s) = \prod_{p \text{ prime}} \left(\frac{1}{1 - p^{-s}} \right), \quad (\sigma > 1).
\tag{2.3}
$$

These products are the key in answering questions about the distribution of prime numbers. One immediate consequence of (2.3) is another proof that there are infinitely many primes: as $s \to 1$, the zeta-function becomes the harmonic series and diverges. The infinite product therefore never terminates, as otherwise its value would be finite, implying the infinitude of primes. We will see that studying the behaviour of L-functions at $s = 1$ is at the heart of the analogous result for primes in arithmetic progressions - Dirichlet's theorem. Before doing so, we need more knowledge on the properties of Dirichlet characters.

2.2 Properties of Dirichlet Characters

The easiest way to construct the Dirichlet characters modulo q is algebraically. Consider $G = (\mathbb{Z}/q\mathbb{Z})^{\times}$. Then, if χ' is a group homomorphism $G \to \mathbb{C}^{\times}$, we may construct a Dirichlet character χ by setting $\chi(n) = 0$ if $(n, q) > 1$, and $\chi(n) = \chi'(\overline{n})$ otherwise, where $\overline{n}$ is the residue class of n modulo q. Clearly all three properties of a Dirichlet character hold for χ, with multiplicativity coming from the homomorphism property of χ'.

Conversely, suppose χ is a Dirichlet character. From the third property in Definition 2.2, we must have that χ is only non-zero at integers coprime to q: precisely those whose residue classes are in G. Furthermore, q-periodicity means that χ must be single-valued on each residue class modulo q, while multiplicativity guarantees that χ restricted to G is a homomorphism $G \to \mathbb{C}^{\times}$. Therefore, the Dirichlet characters modulo q are precisely those following this construction.

First, note that the function $\chi_0(n) = 1$ for all n coprime to q, $\chi_0(n) = 0$ otherwise is always a character - we call this the *trivial character*. Furthermore, since group homomorphisms always map identity to identity, we have $\chi(1) = 1$ for all characters χ. Consequently, all non-zero values of χ must be $\phi(q)$-th roots of unity, where ϕ denotes Euler's totient function. Indeed, since for all n coprime to q, $n^{\phi(q)} \equiv 1$ (q), we have $1 = \chi(1) = \chi(n^{\phi(q)}) = \chi(n)^{\phi(q)}$, using the periodic and multiplicative properties. It follows that characters always have absolute value 1 on values n coprime to q, so that $1 = |\chi(n)|^2 = \chi(n)\overline{\chi(n)}$. Thus, the inverse to each character is its complex conjugate. Note that the complex conjugate also defines a character, denoted $\overline{\chi}$, since for m, n coprime to q,

$$\overline{\chi}(mn) = \overline{\chi(mn)} = \overline{\chi(m)}\,\overline{\chi(n)} = \overline{\chi}(m)\overline{\chi}(n).$$

We now bring in some facts from representation theory, exposed in [8]. Since the Dirichlet characters restricted to G are homomorphisms to $\mathbb{C}^\times$, they are one-dimensional (and by default irreducible) representations of the group G. Conversely, since G is abelian, all such irreducible representations are one-dimensional, hence giving Dirichlet characters. Thus, the group characters (the trace of the irreducible representations) coincide with the Dirichlet characters restricted to G. Therefore, facts about group characters of G from representation theory also hold for Dirichlet characters.

Proposition 2.4. *There are precisely $\phi(q)$ Dirichlet characters modulo q.*

This follows from the fact that the number of group characters is the number of conjugacy classes of the group [8, p. 19, Theorem 7]. Since G is abelian, this number is the size of the group, $\phi(q)$. It is also a fact that group characters are "orthogonal" [8, p. 15, Theorem 3], which gives the following.

Proposition 2.5. *If χ and ψ are Dirichlet characters modulo q, and $a, b \in \mathbb{Z}$, then*

1. $\sum_{n=0}^{q-1} \chi(n)\overline{\psi(n)} = \left\{ \begin{array}{ll} \phi(q) & \text{if } \chi = \psi, \\ 0 & \text{if } \chi \neq \psi. \end{array} \right\}$

2. $\sum_\chi \chi(a)\overline{\chi(b)} = \left\{ \begin{array}{ll} \phi(q) & \text{if } a \equiv b \ (q) \ \& \ (a, q) = 1, \\ 0 & \text{otherwise.} \end{array} \right\}$

where the sum is over all Dirichlet characters χ modulo q.

We also note the important concept of *induced* and *primitive* Dirichlet characters, from [5, Chapter 11]. A character ψ to modulus q is said to be *induced* by another character χ of modulus $m < q$ if $\chi(n) = \psi(n)$ for all n coprime to q. For example, consider the character of modulus 4, given as

n	1	2	3	4
$\chi(n)$	1	0	-1	0

We may then produce an induced character ψ of modulus 8, given as:

n	1	2	3	4	5	6	7	8
$\psi(n)$	1	0	-1	0	1	0	-1	0

If a character is not induced by any other character of lower modulus, then it is said to be *primitive*. Thus, in this example χ is in fact primitive, while ψ is induced. L-functions of an induced character are related to the L-function of the corresponding primitive character as follows. Suppose χ of modulus q is induced by the primitive character χ_1. Then for $\sigma > 1$, by the product formula,

$$L(s,\chi) = \prod_{p \text{ prime}} (1 - \chi(p)p^{-s})^{-1} = L(s,\chi_1)\prod_{p|q}(1 - \chi_1(p)p^{-s}). \qquad (2.4)$$

In particular, notice that $\zeta(s)$ is the special case when $q = 1$ and the character is trivial. All subsequent trivial characters are induced by this character, so if χ_0 is the trivial character modulo q,

$$L(s,\chi_0) = \zeta(s)\prod_{p|q}(1 - p^{-s}) \quad (\sigma > 1). \qquad (2.5)$$

Now, using the properties of Dirichlet characters, we may sum L-functions over all characters modulo q, which will yield interesting results concerning primes in arithmetic progressions. The first of which is *Dirichlet's Theorem*, which we study next.

2.3 Dirichlet's Theorem

This section aims to outline the proof of Dirichlet's theorem as in [6] - the fact that there are infinitely many primes congruent to each a coprime to any modulus q, with "equal density" among each of these residue classes. In order to state this idea rigorously, we must first make an important definition.

Definition 2.6. (Dirichlet Density) Given a set $\mathcal{P}$ of positive prime numbers, we say that $\mathcal{P}$ has *Dirichlet density* if

$$\lim_{s \to 1} \frac{\sum_{p \in \mathcal{P}} p^{-s}}{\log(s-1)^{-1}}.$$

exists. We denote the value of the limit by $d(\mathcal{P})$, named the Dirichlet density of $\mathcal{P}$.

It is clear from the definition that if a set of primes has non-zero Dirichlet density, then the set contains infinitely many primes. Moreover, if $\mathcal{P}$ can be written as the disjoint union of two sets, then the Dirichlet density of $\mathcal{P}$ will be equal to the sum of the Dirichlet densities of the disjoint components. It is a fact that if $\mathcal{P}$ contains all but finitely many positive primes, then its Dirichlet density is 1. Indeed, from (2.3), we may write the logarithm of $\zeta(s)$ (for real values $s > 1$) as

$$\log \zeta(s) = \sum_p -\log(1 - p^{-s})$$

$$= \sum_p \sum_{m=1}^{\infty} m^{-1}p^{-ms}$$

$$= \sum_p p^{-s} + \sum_p \sum_{m=2}^{\infty} m^{-1}p^{-ms}, \qquad (2.6)$$

where the second step comes from the Taylor expansion of $\log(1-x)$. Furthermore, we have that

$$\sum_p \sum_{m=2}^{\infty} m^{-1} p^{-ms} < \sum_p \sum_{m=2}^{\infty} p^{-ms} = \sum_p p^{-2s}(1-p^{-s})^{-1}$$
$$\leq (1-2^{-s})^{-1} \sum_p p^{-2s} < 2\zeta(2),$$

so $\log \zeta(s) = \sum_p p^{-s} + R(s)$ where R is a function which remains bounded as $s \to 1$. We will also see that $\zeta(s)$ has residue 1 at $s = 1$, so in particular $\lim_{s \to 1+}(s-1)\zeta(s) = 1$. Letting $\rho(s) = (s-1)\zeta(s)$, we have

$$\frac{\log \zeta(s)}{\log(s-1)^{-1}} = 1 + \frac{\log \rho(s)}{\log(s-1)^{-1}}.$$

Since $\rho(s) \to 1$ as $s \to 1$, its logarithm goes to zero. In particular

$$\lim_{s \to 1+} \frac{\sum_p p^{-s}}{\log(s-1)^{-1}} = \lim_{s \to 1+} \frac{\log \zeta(s)}{\log(s-1)^{-1}} = 1. \tag{2.7}$$

Thus, the Dirichlet density of all the primes with the exception of finitely many is 1. In summary, the Dirichlet density is a number between 0 and 1 which measures (in a sense) the "proportion" of primes which lie in a set. We may now use this to state Dirichlet's theorem.

Theorem 2.7. *(Dirichlet's Theorem) Suppose a and q are coprime integers. Let $\mathcal{P}(a;q)$ denote the set of all primes p such that $p \equiv a \ (q)$. Then*

$$d\left(\mathcal{P}(a;q)\right) = 1/\phi(q).$$

In particular, the primes have equal Dirichlet density among residue classes coprime to q, and there are infinitely many primes in each such residue class.

This may be proved in a similar manner to proving the Dirichlet density of the set of all primes is 1. Rather than taking the logarithm of $\zeta(s)$, we wish to take the logarithm of $L(s, \chi)$. However, in this case there is the technicality of branch cuts to consider: even when restricting to real $s > 1$, the L-function may still take on complex values. Therefore, we must proceed in a more subtle manner. Consider

$$G(s, \chi) = \sum_p \sum_{k=1}^{\infty} k^{-1} \chi(p^k) p^{-ks}.$$

Notice that this is similar to the infinite series defining $\log \zeta(s)$. Also, since the absolute value of each term of G is bounded by p^{-ks}, G is bounded by $\zeta(s)$ which is uniformly

convergent for real $s > 1$, so the same is true for G. Moreover, taking the exponential map, we see that

$$\exp(G(s,\chi)) = \prod_p \exp\left(\sum_{k=1}^{\infty} k^{-1}(\chi(p)p^{-s})^k\right)$$
$$= \prod_p (1 - \chi(p)p^{-s})^{-1} = L(s,\chi),$$

where the penultimate step comes from the Taylor series of $\log(1 - \chi(p)p^{-s})$, justified since $|\chi(p)p^{-s}| < 1$. Thus, the infinite series G gives a definition of $\log L(s,\chi)$ without the need to worry about branch cuts until later. We therefore proceed working directly with G. In a similar manner to before, we have that

$$G(s,\chi) = \sum_p \chi(p)p^{-s} + \sum_p \sum_{k=2}^{\infty} k^{-1}\chi(p^k)p^{-ks}.$$

Since $\left|\chi(p^k)\right| \le 1$, we may bound the second term in the same way as for $\log\zeta(s)$ by a function R_χ which remains bounded as $s \to 1$. Also note that $\chi(p) = 0$ when p is a prime dividing q. Therefore,

$$G(s,\chi) = \sum_{p \nmid q} \chi(p)p^{-s} + R_\chi(s). \tag{2.8}$$

Now, let a be coprime to q, multiply both sides of (2.8) by $\overline{\chi(a)}$, and sum over all Dirichlet characters modulo q. This gives

$$\sum_\chi \overline{\chi(a)}G(s,\chi) = \sum_{p \nmid q} p^{-s} \sum_\chi \chi(p)\overline{\chi(a)} + \sum_\chi \overline{\chi(a)}R_\chi. \tag{2.9}$$

Now, we may apply the second orthogonality relation in Proposition 2.5: the sum over characters in the first term in the right hand side of (2.9) gives $\phi(q)$ whenever $p \equiv a(q)$, and zero otherwise. Thus,

$$\sum_\chi \overline{\chi(a)}G(s,\chi) = \phi(q) \sum_{p \equiv a(q)} p^{-s} + R_{\chi,a}(s), \tag{2.10}$$

where $R_{\chi,a}$ is again a function which is bounded as $s \to 1^+$. We are now very close to proving the theorem: the following proposition will allow us to complete the proof.

Proposition 2.8. *If* χ_0 *denotes the trivial character, then*

$$\lim_{s \to 1^+} \frac{G(s,\chi)}{\log(s-1)^{-1}} = \left\{ \begin{array}{ll} 1 & \text{if } \chi = \chi_0, \\ 0 & \text{if } \chi \ne \chi_0. \end{array} \right\}$$

13

The first case in this proposition is clear from (2.5) - $\zeta(s)$ has a pole of identical residue at $s = 1$ to $L(s, \chi_0)$, so this is equivalent to the Dirichlet density of all primes being 1. The latter statement is not so clear, and is equivalent to the fact that $L(1, \chi)$ is non-zero and analytic for non-trivial characters χ - let us assume this for now, as we will come back to this in more detail later.

If $L(1, \chi) \neq 0$ for non-trivial characters χ, we may define $\log L(s, \chi)$ on a small interval $(1, 1 + \delta)$. Note that $G(s, \chi)$ and $\log L(s, \chi)$ are both mapped by exp to $L(s, \chi)$, so that $G(s, \chi) = \log L(s, \chi) + 2k\pi i$. Since $\log L(s, \chi)$ tends to a limit as $s \to 1^+$, $G(s, \chi)$ must remain bounded. Hence, the second part of the proposition is proved. Thus, we divide both sides of (2.10) by $\log(s - 1)^{-1}$ and take the limit as $s \to 1^+$. The limit on the left hand side is 1 (the only term in the sum without limit zero is the one for the trivial character). On the right hand side, the first term approaches $\phi(q)d\left(\mathcal{P}(a; q)\right)$, while the second term goes to zero since the function $R_{\chi,a}$ remains bounded as $s \to 1^+$. We therefore divide both sides by $\phi(q)$, which gives Dirichlet's Theorem:

$$d\left(\mathcal{P}(a; q)\right) = 1/\phi(q).$$

We have therefore proved a deep theorem about primes using the properties of L-functions and their Dirichlet characters, assuming the key step that L-functions of non-trivial character are finite and non-zero at 1. The study of where the zeros of L-functions occur and do not occur is key to the study of the distribution of primes (and is still subject to research today) - this theorem provides a taste of this.

2.4   Introduction to the Prime Number Theorem

We now turn our attention to the *growth* of the prime numbers, and in particular the prime counting function, defined as

$$\pi(x) := \#\{p \leq x \ : \ p \text{ prime}\}.$$

The aim of the Prime Number Theorem is to find the *asymptotic distribution* of $\pi(x)$ - meaning a function which essentially approximates $\pi(x)$ well. Or, more precisely, one whose ratio with $\pi(x)$ tends to 1 as x becomes large. It turns out that it is more natural to study the function $\psi(x)$, defined as

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

where $\Lambda(n)$ is the von Mangoldt function, defined as $\log p$ whenever n is a power of a prime p, and zero otherwise. These two functions are related by the following lemma [5, Lemma 3].

Lemma 2.9. *For $x \geq 2$,*

$$\pi(x) = \frac{\psi(x)}{\log x} + \int_2^x \frac{\psi(t)}{t \log^2 t} \mathrm{d}t + O(x^{1/2}). \tag{2.11}$$

Proof. First, set

$$\theta(x) = \sum_{p \leq x} \log p.$$

Then we have

$$
\begin{aligned}
\int_2^x \frac{\theta(t)}{t \log^2 t} \mathrm{d}t &= \int_2^x \sum_{p \leq t} \frac{\log p}{t \log^2 t} \mathrm{d}t \\
&= \sum_{p \leq x} \int_p^x \frac{\log p}{t \log^2 t} \mathrm{d}t \\
&= \sum_{p \leq x} \left[-\frac{\log p}{\log t} \right]_p^x \\
&= \pi(x) - \frac{\theta(x)}{\log x}.
\end{aligned}
$$

Thus,

$$\pi(x) = \frac{\theta(x)}{\log x} + \int_2^x \frac{\theta(t)}{t \log^2 t} \mathrm{d}t. \tag{2.12}$$

Now consider the relationship between θ and ψ. We have that

$$\psi(x) = \theta(x) + \sum_{k \geq 2} \sum_{p^k \leq x} \log p.$$

In the double sum, the primes are at most $x^{1/2}$ since the exponent k is at least 2, so there are at most $x^{1/2}$ such terms. Furthermore, having $p^k \leq x$ implies $k \leq \frac{\log x}{\log p}$. Therefore,

$$\psi(x) \leq \theta(x) + \frac{\log x}{\log p} \sum_{p^k \leq x} \log p \leq \theta(x) + x^{1/2} \log x.$$

So,

$$\psi(x) = \theta(x) + O\left(x^{1/2} \log x\right).$$

Inserting this into (2.12), we obtain the required result. □

For a coprime to q, there is an analogous prime counting function for the residue class a modulo q, defined as

$$\pi(x; q, a) := \#\{p \leq x \ : \ p \equiv a \ (q), \ p \text{ prime}\},$$

as well as a corresponding ψ function,

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

Lemma 2.9 also holds for $\pi(x; q, a)$ and $\psi(x; q, a)$, simply by restricting the sums to the relevant residue class. Remarkably, $\psi(x)$ and $\psi(x; q, a)$ are intimately connected with the Riemann zeta-function and Dirichlet L-functions respectively - in particular their zeros in the complex plane. Instead of studying these functions on a restricted region where they converge, we will use analytic continuation to study them on the whole of $\mathbb{C}$. In the case of $\psi(x)$, it turns out that it can be written explicitly for any $T \geq x \geq 1$ as

$$\psi(x) = x - \sum_{\rho : |\gamma| < T} \frac{x^\rho}{\rho} + O(\frac{x \log^2 x}{T}), \tag{2.13}$$

where the ρ represent the (infinitely many) so-called non-trivial zeros of the analytically continued zeta-function [5, Theorem 18]. Standard notation for the ρ is $\beta + i\gamma$, which we will use throughout. We shall see in the next section that these all lie in the *critical strip*, $0 \leq \sigma \leq 1$. The problem of estimating $\psi(x)$ and hence $\pi(x)$ is then essentially a question of estimating the contribution to this formula of the zeros as well as possible. A similar formula holds for the function

$$\psi(x, \chi) := \sum_{n \leq x} \chi(n) \Lambda(n), \tag{2.14}$$

but this time over the non-trivial zeros of L-functions, which also lie in the critical strip. For a coprime to q we have

$$\frac{1}{\phi(q)} \sum_\chi \overline{\chi}(a) \psi(x, \chi) = \frac{1}{\phi(q)} \sum_{n \leq x} \Lambda(n) \left(\sum_\chi \overline{\chi}(a) \chi(n) \right) = \sum_{\substack{n \leq x \\ n \equiv a(q)}} \Lambda(n) = \psi(x; q, a).$$

$$\tag{2.15}$$

So, the analogous explicit formula for $\psi(x, \chi)$ reveals the relationship of the zeros of L-functions to $\psi(x; q, a)$, and hence $\pi(x; q, a)$ - the growth of primes in different residue classes. The purpose of mentioning these formulas is merely demonstrative: to show how the zeros of L-functions influence the distribution of primes. We shall not prove them in their entirety, and instead use a more refined method. However, ultimately we still require essentially the same knowledge about the non-trivial zeros, so this should serve as the motivation behind the following two chapters.

In order to say anything about the non-trivial zeros of L-functions and in particular $\zeta(s)$, we must study their analytic continuation to the whole complex plane. The analytic continuation of $\zeta(s)$ was Riemann's key insight in his 1859 memoir, and provides much of the foundation for big results in analytic number theory. We therefore study this in the next chapter, before studying the zeros of the analytically continued functions and ultimately proving very important results about the distribution of primes.

Chapter 3

Analytic Continuation and the Functional Equations

3.1 Continuation to the Right Half-Plane

Our general strategy in analytic continuation of L-functions, and in particular $\zeta(s)$, is to extend them to a slightly larger region first, before proving a powerful functional equation, which allows us to extend each function to $\mathbb{C}$. We first consider the case of L-functions of non-trivial character (which are to some extent easier to extend than $\zeta(s)$ owing to their lack of a singularity at $s = 1$) and refer to [6]. First, define

$$S(x) = \sum_{n \leq x} \chi(n) \tag{3.1}$$

for a non-trivial character χ to modulus q. For $\sigma > 1$,

$$\sum_{n=1}^{N} S(n) \left(n^{-s} - (n+1)^{-s} \right) = \sum_{n \leq N} \chi(n) n^{-s} - S(N)(N+1)^{-s} \to L(s, \chi) \text{ as } N \to \infty,$$

Note that the secondary term goes to zero since $|S(N)|$ is certainly less than N. Therefore:

$$\begin{aligned} L(s, \chi) &= \sum_{n=1}^{\infty} S(n) \left(n^{-s} - (n+1)^{-s} \right) \\ &= s \sum_{n=1}^{\infty} S(n) \int_{n}^{n+1} x^{-s-1} \mathrm{d}x \\ &= s \int_{1}^{\infty} S(x) x^{-s-1} \mathrm{d}x \quad (\sigma > 1), \end{aligned} \tag{3.2}$$

where taking $S(x)$ inside the integral is valid as it only changes at discrete values. We now need the following Lemma.

Lemma 3.1. *For a non-trivial character χ to modulus q,*

$$|S(x)| \leq \phi(q) \text{ for all } x > 0.$$

Assuming this gives

$$\left| s \int_1^\infty S(x) x^{-s-1} \mathrm{d}x \right| \leq |s| \int_1^\infty |S(x)| x^{-\sigma-1} \mathrm{d}x$$

$$\leq \phi(q)|s| \int_1^\infty x^{-\sigma-1} \mathrm{d}x < \infty, \quad (\sigma > 0).$$

The lemma therefore implies the following.

Proposition 3.2. *For χ a non-trivial character to modulus q, and $S(x)$ defined as above,*

$$L(s, \chi) = s \int_1^\infty S(x) x^{-s-1} \mathrm{d}x,$$

which defines an analytic function on $\sigma > 0$.

To prove Lemma 3.1, for a given $x > 0$, write $x = mq + r$, where $0 \leq r < q$. By the q-periodicity of χ,

$$\sum_{n \leq x} \chi(n) = m \left(\sum_{n=0}^{q-1} \chi(n) \right) + \sum_{n \leq r} \chi(n).$$

By the first orthogonality relation in Proposition 2.5, the first term is 0. Hence,

$$|S(x)| \leq \sum_{n \leq r} |\chi(n)| \leq \phi(q),$$

where the last inequality comes from the fact that there are at most $\phi(q)$ non-zero terms, all of modulus 1.

Having found an analytic continuation of L-functions of non-trivial character to $\sigma > 0$, we turn to the case of $\zeta(s)$, and by (2.5) the case of L-functions of trivial character. Here, we have the slight issue of the pole at $s = 1$ corresponding to the harmonic series, so we intend to essentially remove it. In this case we have that $S(x) = [x]$, since $q = 1$ so $\chi(n) = 1$ for all n. So, for $\sigma > 1$

$$\zeta(s) = s \int_1^\infty [x] x^{-s-1} \mathrm{d}x,$$

as in (3.2). Therefore, using the trick of adding and subtracting the same term as in [7, p. 3],

$$\zeta(s) = s \int_1^\infty x^{-s} \mathrm{d}x - s \int_1^\infty (x - [x]) x^{-s-1} \mathrm{d}x$$

$$= \frac{s}{s-1} - s \int_1^\infty (x - [x]) x^{-s-1} \mathrm{d}x \quad (\sigma > 1). \tag{3.3}$$

18

Since $|x - [x]| \leq 1$ for all x, we have for $\sigma > 0$

$$\left| \int_1^\infty (x - [x]) x^{-s-1} \mathrm{d}x \right| \leq \int_1^\infty x^{-\sigma-1} \mathrm{d}x < \infty,$$

so it follows that (3.3) gives an analytic continuation of $\zeta(s)$ to $\sigma > 0$. Moreover, it shows plainly the value of the residue is 1 at $s = 1$, which we used repeatedly throughout the previous chapter.

3.2 The Key Step in Dirichlet's Theorem

Having continued L-functions of trivial and non-trivial characters to the right half plane, we now have the tools required to complete the proof of Dirichlet's theorem. We may now study with greater ease the behaviour of L-functions of non-trivial character near $s = 1$, safe in the knowledge that they are analytic here; it therefore suffices to show that they are non-zero at $s = 1$. We follow the argument in [6]. Consider the function

$$F(s) = \prod_\chi L(s, \chi),$$

where the product is over all Dirichlet characters modulo q. We claim that $F(s) \geq 1$ for all real $s > 1$. Recall that

$$G(s, \chi) = \sum_p \sum_{k=1}^\infty k^{-1} \chi(p^k) p^{-ks}, \quad (\sigma > 1).$$

It follows from the orthogonality relations of characters that

$$\sum_\chi G(s, \chi) = \phi(q) \sum_{\substack{k \geq 1 \\ p^k \equiv 1 \ (q)}} k^{-1} p^{-ks} \quad (\sigma > 1),$$

so that $\sum_\chi G(s, \chi) \geq 0$ for real $s > 1$ since each term is positive. By taking the exponential map, we have the claim that $F(s) \geq 1$ for all real $s > 1$. Now, consider the behaviour of $F(s)$ as $s \to 1$. If an L-function of complex character had $L(1, \chi) = 0$, it would follow that $L(1, \overline{\chi}) = \overline{L(1, \chi)} = 0$, so two terms of $F(s)$ are zero. Then, since the only term in the product with a pole at $s = 1$ is the trivial character, and this pole is simple, it follows that $F(1) = 0$. This clearly contradicts $F(s) \geq 1$. Thus, $L(1, \chi) \neq 0$ for all complex characters χ. We must use a different approach for real characters as they coincide with their conjugate characters. We assume for a contradiction that χ is a real character modulo q with $L(1, \chi) = 0$, and consider

$$\tilde{F}(s) = \frac{L(s, \chi) L(s, \chi_0)}{L(2s, \chi_0)},$$

where χ_0 is the trivial character modulo q as usual. Since the zero of $L(s, \chi)$ cancels the pole of $L(s, \chi_0)$ at $s = 1$, the numerator is analytic on $\sigma > 0$, while the denominator is

analytic when $\sigma > 1/2$. It follows that $F(s)$ is analytic on $\sigma > 1/2$, and $F(s) \to 0$ as $s \to 1$, by virtue of the pole of the denominator. By the Euler products,

$$\tilde{F}(s) = \prod_p \frac{(1 - \chi_0(p)p^{-2s})}{(1 - \chi(p)p^{-s})(1 - \chi_0(p)p^{-s})}.$$

We have $\chi(p) = \pm 1$ if $(p, q) = 1$ since χ is real, and if $\chi(p) = -1$, the numerator will coincide with the denominator. Moreover, if $\chi(p) = 0$, then p and q are not coprime and each term is 1. Thus,

$$\tilde{F}(s) = \prod_{\chi(p)=1} \frac{1 + p^{-s}}{1 - p^{-s}}, \quad (\sigma > 1).$$

Since

$$\frac{1 + p^{-s}}{1 - p^{-s}} = (1 + p^{-s})\left(\sum_{k=0}^{\infty} p^{-ks}\right) = 1 + 2p^{-s} + 2p^{-2s} + \dots,$$

we may write $\tilde{F}(s)$ as a series of the form $\sum_{n \geq 1} a_n n^{-s}$, where $a_n \geq 0$, convergent for $\sigma > 1$. Note that $a_1 = 1$. Since $\tilde{F}(s)$ is analytic for $\sigma > 1/2$, we may expand it as a power series around $s = 2$ with radius of convergence at least $3/2$, say $\tilde{F}(s) = \sum_{m=1}^{\infty} b_m(s-2)^m$. Each b_m is given by $b_m = \tilde{F}^{(m)}/m!$, and by its representation as an infinite series we have $\tilde{F}^{(m)}(2) = \sum_{n=1}^{\infty} a_n(-\log n)^m n^{-2} = (-1)^m c_m$, with $c_m \geq 0$. Therefore, we have $\tilde{F}(s) = \sum_{m=0}^{\infty} d_m(2 - s)^m$ with $d_m \geq 0$ and $d_0 = \tilde{F}(2) = \sum_{n \geq 1} a_n n^{-2} \geq a_1 = 1$. It follows that for real values $1/2 < s < 2$, we have $\tilde{F}(s) \geq d_0 = 1$, which contradicts $\tilde{F}(s) \to 0$ as $s \to 1/2$. In conclusion, L-functions of any character are non-zero at $s = 1$, which completes the proof of Dirichlet's theorem.

Having studied analytic continuation of L-functions to the right half plane and putting this into use, we now turn to analytic continuation on the whole of $\mathbb{C}$, which is a result of relating $L(s, \chi)$ to $L(1 - s, \overline{\chi})$. The idea is that since $L(s, \chi)$ is analytic on the right half-plane, this relation gives the value of $L(s, \overline{\chi})$ on the left half-plane. Such a relation is known as the functional equation, which we now derive.

3.3 The Functional Equations for $\zeta(s)$ and $L(s, \chi)$

This section follows the derivation in [4, Chapter 9]. Throughout, we assume that χ is a primitive character of modulus q. Euler's Gamma function is defined for $\sigma > 0$ as

$$\Gamma(s) := \int_0^{\infty} x^{s-1} e^{-x} \mathrm{d}x.$$

For properties of $\Gamma(s)$, the reader is referred to the appendix. Also contained there are properties of Gauss sums, defined as

$$\tau(\chi) := \sum_{m=1}^{q} \chi(m) e^{2\pi i m/q},$$

and a particular application of the Poisson summation formula, which we shall call upon throughout. Now, for $\sigma > 1$,

$$
\pi^{-s/2}q^{s/2}\Gamma(s/2)n^{-s} = \left(\frac{q}{n^2\pi}\right)^{s/2}\int_0^\infty y^{s/2-1}e^{-y}\mathrm{d}y
$$
$$
= \int_0^\infty x^{s/2-1}e^{-n^2\pi x/q}\mathrm{d}x \quad (\text{set } y = n^2\pi x/q). \qquad (3.4)
$$

There are two main cases to consider depending on the value of $\chi(-1)$: note that since $\chi(-1)^2 = \chi(1) = 1$, we have $\chi(-1) = \pm 1$. First, suppose $\chi(-1) = 1$. Multiplying both sides by $\chi(n)$ and summing over all positive integers n gives

$$
\pi^{-s/2}q^{s/2}\Gamma(s/2)L(s,\chi) = \int_0^\infty x^{s/2-1}\left(\sum_{n=1}^\infty \chi(n)e^{-n^2\pi x/q}\right)\mathrm{d}x
$$
$$
= \frac{1}{2}\int_0^\infty x^{s/2-1}\psi(x,\chi)\mathrm{d}x,
$$

where the interchange of summation and integration is justified by the absolute convergence of the left hand side on $\sigma > 1$, and the function ψ is defined as

$$
\psi(x,\chi) := \sum_{n=-\infty}^\infty \chi(n)e^{-n^2\pi x/q} = 2\sum_{n=1}^\infty \chi(n)e^{-n^2\pi x/q}.
$$

Note that the last equality here relies on the fact that $\chi(n) = \chi(-1)\chi(n) = \chi(-n)$ and $\chi(0) = 0$. We now appeal to Lemma A6 of Gauss sums and the relation (A7). Indeed,

$$
\tau(\overline{\chi})\psi(x,\chi) = \sum_{n=-\infty}^\infty \tau(\overline{\chi})\chi(n)e^{-n^2\pi x/q}
$$
$$
= \sum_{m=1}^q \overline{\chi}(m)\sum_{n=-\infty}^\infty e^{-n^2\pi x/q + 2\pi imn/q} \quad (\text{by Lemma A6}).
$$

Replacing x by x/q and α by m/q in (A7), we therefore have

$$
\tau(\overline{\chi})\psi(x,\chi) = \sum_{m=1}^q \overline{\chi}(m)(q/x)^{1/2}\sum_{n=-\infty}^\infty e^{-(n+m/q)^2\pi qx^{-1}}
$$
$$
= \left(\frac{q}{x}\right)^{1/2}\sum_{m=1}^q \overline{\chi}(m)\sum_{n=-\infty}^\infty e^{-(nq+m)^2\pi x^{-1}/q}
$$
$$
= \left(\frac{q}{x}\right)^{1/2}\sum_{k=-\infty}^\infty \overline{\chi}(k)e^{-k^2\pi x^{-1}/q}
$$
$$
= \left(\frac{q}{x}\right)^{1/2}\psi(x^{-1},\overline{\chi}). \qquad (3.5)
$$

Now, applying this relation gives

$$\pi^{-s/2} q^{s/2} \Gamma(s/2) L(s, \chi) = \frac{1}{2} \int_0^\infty x^{s/2-1} \psi(x, \chi) \mathrm{d}x$$

$$= \frac{1}{2} \int_1^\infty x^{s/2-1} \psi(x, \chi) \mathrm{d}x + \frac{1}{2} \int_0^1 x^{s/2-1} \psi(x, \chi) \mathrm{d}x$$

$$= \frac{1}{2} \int_1^\infty x^{s/2-1} \psi(x, \chi) \mathrm{d}x + \frac{1}{2} \int_1^\infty x^{-s/2-1} \psi(x^{-1}, \chi) \mathrm{d}x$$

$$= \frac{1}{2} \int_1^\infty x^{s/2-1} \psi(x, \chi) \mathrm{d}x + \frac{1}{2} \frac{q^{1/2}}{\tau(\overline{\chi})} \int_1^\infty x^{(1-s)/2-1} \psi(x, \overline{\chi}) \mathrm{d}x.$$

$$(3.6)$$

Here, the penultimate step changes variables from x to x^{-1}, while the last step appeals to (3.5). We now claim that the right hand side defines an analytic function on $\mathbb{C}$. Indeed,

$$\left| \int_1^\infty \psi(x, \chi) x^{s/2-1} \mathrm{d}x \right| \ll \int_1^\infty \sum_{n=1}^\infty e^{-\pi n^2 x} x^{\sigma/2-1} \mathrm{d}x$$

$$\ll \int_1^\infty e^{-\frac{\pi x}{2}} x^{\sigma/2-1} \mathrm{d}x.$$

There are two cases. If $\sigma \leq 2$, we have

$$\int_1^\infty e^{-\frac{\pi x}{2}} x^{\sigma/2-1} \mathrm{d}x \ll \int_1^\infty e^{-\frac{\pi x}{2}} \mathrm{d}x < \infty,$$

and if $\sigma > 2$,

$$\int_1^\infty e^{-\frac{\pi x}{2}} x^{\sigma/2-1} \mathrm{d}x \ll \Gamma(\frac{\sigma}{2} - 1) < \infty$$

by a change of variables. Thus, the first integral in (3.6) is analytic for all $s \in \mathbb{C}$, and so is the second similarly. Moreover, switching s to $1 - s$, and χ to $\overline{\chi}$, we obtain

$$\frac{1}{2} \frac{q^{1/2}}{\tau(\chi)} \int_1^\infty x^{s/2-1} \psi(x, \chi) \mathrm{d}x + \frac{1}{2} \int_1^\infty x^{(1-s)/2-1} \psi(x, \overline{\chi}) \mathrm{d}x,$$

which we claim is the expression in (3.6) multiplied by $q^{1/2}/\tau(\chi)$. Indeed, since $\chi(n) = \chi(-n)$,

$$\overline{\tau(\chi)} = \sum_1^q \overline{\chi}(m) e^{-2\pi i m/q}$$

$$= \sum_1^q \overline{\chi}(m) e^{2\pi i m/q} \quad (\text{using } \chi(m) = \chi(-m))$$

$$= \tau(\overline{\chi}),$$

22

so that $q = |\tau(\chi)|^2 = \tau(\chi)\tau(\overline{\chi})$ by (A6), and the claim follows. Thus, we have obtained a functional equation in the case of a primitive character χ with $\chi(-1) = 1$. Defining

$$\xi_1(s, \chi) = \pi^{-s/2} q^{s/2} \Gamma(s/2) L(s, \chi),$$

we obtain the relation

$$\xi_1(1-s, \overline{\chi}) = \frac{q^{1/2}}{\tau(\chi)} \xi_1(s, \chi).$$

This gives a relation between the (known) value of $L(s, \chi)$ on the right half plane, and the unknown value of $L(1-s, \overline{\chi})$ on the left half-plane. In the case where $\chi(-1) = -1$, we must proceed differently, as in this case the function $\psi(x, \chi)$ is zero! Shifting s to $s + 1$, and multiplying by another factor of n gives

$$\xi_2(s, \chi) := \pi^{-(s+1)/2} q^{(s+1)/2} \Gamma\left(\frac{1}{2}(s+1)\right) L(s, \chi)$$

$$= \frac{1}{2} \int_0^\infty \psi_1(x, \chi) x^{\frac{s}{2} - \frac{1}{2}} \mathrm{d}x, \tag{3.7}$$

where we define

$$\psi_1(x, \chi) = \sum_{n=-\infty}^\infty n\chi(n) e^{-n^2 \pi x/q}.$$

Using the same bounds in proving that (3.6) was analytic, both sides of (A7) are uniformly convergent, allowing termwise differentiation with respect to α. This gives

$$\sum_{n=-\infty}^\infty n e^{-n^2 \pi x + 2\pi i n \alpha} = i x^{-3/2} \sum_{n=-\infty}^\infty (n + \alpha) e^{-(n+\alpha)^2 \pi/x}.$$

Setting x to x/q and α to m/q as before, we obtain

$$\sum_{n=-\infty}^\infty n e^{\frac{-n^2 \pi x}{q} + \frac{2\pi i n m}{q}} = i \left(\frac{q}{x}\right)^{3/2} \sum_{n=-\infty}^\infty (n + m/q) e^{-(n+m/q)^2 \pi x^{-1} q}.$$

Therefore, following the exact same procedure as in the case of $\psi(x, \chi)$ gives

$$\tau(\overline{\chi}) \psi_1(x, \chi) = i q^{1/2} x^{-3/2} \psi_1(x^{-1}, \overline{\chi}).$$

We then apply a similar splitting of the integral in (3.7) and use this relation to obtain

$$\xi_2(s, \chi) = \frac{1}{2} \int_1^\infty \psi_1(x, \chi) x^{-(1-s)/2} \mathrm{d}x + \frac{1}{2} \frac{i q^{1/2}}{\tau(\overline{\chi})} \int_1^\infty \psi_1(x, \overline{\chi}) x^{-s/2} \mathrm{d}x. \tag{3.8}$$

This is again analytic on $\mathbb{C}$. By an analogous argument to the previous case, we have $\overline{\tau(\chi)} = -\tau(\overline{\chi})$, so that $\tau(\chi)\tau(\overline{\chi}) = -q$. Thus, replacing s by $1-s$ and χ by $\overline{\chi}$ in (3.8) yields

$$\frac{1}{2} \frac{i q^{1/2}}{\tau(\chi)} \int_1^\infty \psi_1(s, \chi) x^{-(1-s)/2} \mathrm{d}x + \frac{1}{2} \int_1^\infty \psi_1(x, \overline{\chi}) x^{-s/2} \mathrm{d}x,$$

which is precisely (3.8) multiplied by $iq^{1/2}/\tau(\chi)$. We conclude that

$$\xi_2(1-s, \overline{\chi}) = \frac{iq^{1/2}}{\tau(\chi)} \xi_2(s, \chi).$$

The two functional equations for ξ_1 and ξ_2 may be combined as follows. Define

$$a(\chi) = \left\{ \begin{array}{ll} 0 & \text{if } \chi(-1) = 1 \\ 1 & \text{if } \chi(-1) = -1 \end{array} \right\}.$$

Then define

$$\xi(s, \chi) = (q/\pi)^{\frac{1}{2}(s+a)} \Gamma\left(\frac{1}{2}(s+a)\right) L(s, \chi),$$

so that ξ coincides with both ξ_1 and ξ_2, depending on the value of a. The two previous relations are then written compactly as

$$\xi(1-s, \overline{\chi}) = \frac{i^a q^{1/2}}{\tau(\chi)} \xi(s, \chi).$$

This completes the analytic continuation of L-functions of primitive characters to $\mathbb{C}$. Indeed, given a value s on the left half plane, we have that $1 - \sigma \geq 1$, on which region $L(1-s, \overline{\chi})$ is analytic. Since everything is therefore analytic, it is possible to rearrange to write $L(s, \chi)$ as a product of analytic functions on the left half plane. Note that dividing by the Gamma-function is valid as it is nowhere zero (see appendix).

This argument is easily adapted to the case of $\zeta(s)$. Setting $q = 1$, we note that the trivial character of this modulus is (trivially) primitive, so defining

$$\xi(s) = \pi^{-s/2} \Gamma(s/2) \zeta(s),$$

and noting that $\tau(\chi) = q = 1$, we have the corresponding functional equation

$$\xi(s) = \xi(1-s).$$

Again, since $\zeta(s)$ is analytic on the right half plane, this relation implies the analytic continuation of $\zeta(s)$ to $\mathbb{C}$. Note however that $\xi(s)$ has poles at $s = 0, 1$, so we often prefix it with $s(s-1)$ so that it is entire (this does not change the functional equation). Now we have the continuations of $\zeta(s)$ and L-functions of primitive characters to the complex plane, we study their zeros. It is immediate from the functional equation that there are "trivial zeros" of L-functions either at the negative even integers (and zero) or at the negative odd integers, depending on the value of a. These correspond to the poles of $\Gamma(s/2 + a/2)$, cancelling them to make an analytic function.

Furthermore, since L-functions are non-zero on $\sigma > 1$ (by the product formula), it follows that they are also non-zero on $\sigma < 0$ (if not coinciding with a trivial zero). Thus, any other zeros must lie in the *critical strip* $0 \leq \sigma \leq 1$, and such zeros are deemed non-trivial. The exact same arguments hold for $\zeta(s)$. Having already mentioned the duality between the non-trivial zeros of L-functions and the primes, we now study both their frequency and positioning in the critical strip.

24

Chapter 4

The Zeros of $\zeta(s)$ and $L(s, \chi)$

4.1 The Hadamard Product Formula

Now, we aim to write $\xi(s)$ and $\xi(s, \chi)$ as *infinite products* in terms of their zeros. Analysing the properties of the ξ functions, we use this to make conclusions with regard to the distribution of these zeros, and ultimately the distribution of primes. We begin by introducing a special class of function (the conditions for which the ξ functions satisfy).

Definition 4.1. An entire function $f(z)$ is of *finite order* if it satisfies

$$f(z) = O(e^{|z|^{\alpha}}) \text{ as } |z| \to \infty \tag{4.1}$$

for some number α.

Note that $\alpha > 0$ for all non-constant functions, so we can define the order of an entire function to be the infimum of such α. We have the following elementary fact [4, p. 74] for entire functions of finite order which are nowhere zero.

Proposition 4.2. *If $f(z)$ is an entire function of finite order, and $f(z) \neq 0$ for all $z \in \mathbb{C}$, then f is necessarily of the form $e^{g(z)}$, where $g(z)$ is a polynomial. The order of f is therefore the degree of g.*

Proof. We begin by defining $g(z) = \log f(z)$ so as to be single valued, and noting that $g(z)$ is also holomorphic away from some branch cut since f is non-zero. On any large enough circle $|z| = R$, we have the bound

$$\Re g(z) = \log|f(z)| < 2R^{\alpha},$$

where α is the order of the entire function f, using the property that $f(z) = O(e^{|z|^{\alpha}})$. Away from the branch cut, $g(z)$ can be defined by a power series, written

$$g(z) = \sum_{n=0}^{\infty} (a_n + ib_n) z^n.$$

Now for $z = Re^{i\theta}$, we have

$$\Re g(Re^{i\theta}) = \sum_{n=0}^{\infty} a_n R^n \cos(n\theta) - \sum_{n=1}^{\infty} b_n R^n \sin(n\theta).$$

This is a Fourier series for $\Re g(Re^{i\theta})$ as a function of θ, so bounding its Fourier coefficients:

$$
\begin{aligned}
\pi |a_n| R^n = \pi \left| \frac{1}{2\pi} \int_0^{2\pi} \Re g(Re^{i\theta}) \cos(n\theta) \mathrm{d}\theta \right| \\
\leq \int_0^{2\pi} \left| \Re g(Re^{i\theta}) \right| \mathrm{d}\theta \\
= \int_0^{2\pi} \left| \Re g(Re^{i\theta}) \right| + \Re g(Re^{i\theta}) \mathrm{d}\theta - 2\pi a_0 \\
\leq C\pi R^\alpha,
\end{aligned}
$$

where C is some absolute constant independent of R. Therefore,

$$|a_n| < CR^{\alpha - n},$$

so we can let $R \to \infty$ to conclude $a_n = 0$ for $n > \alpha$. We can bound the b_n similarly, so we conclude that $g(z)$ is a polynomial of degree less than or equal to α. Hence the order of f is simply the degree of g. $\qquad\square$

We now consider the case where an entire function $f(z)$ has zeros at $z_1, z_2, \ldots$ and we wish to relate the distribution of its zeros to its order, say ρ. It will be important when writing f as an infinite product that its zeros are sufficiently far apart. The easiest way to answer the question of the distribution of its zeros is by a formula due to Jensen.

Proposition 4.3. *(Jensen's Formula) If f is an entire function with $f(0) \neq 0$ and zeros at $z_1, z_2, \ldots$, none of which are on $|z| = R$, then*

$$\frac{1}{2\pi} \int_0^{2\pi} \log\left| f(Re^{i\theta}) \right| \mathrm{d}\theta = \int_0^R r^{-1} n(r) \mathrm{d}r + \log|f(0)|,$$

where $n(r)$ is the number of zeros of f with radius less than r.

Proof. We flesh out the proof in [7, p. 13]. Firstly, using the assumption that $f(0) \neq 0$,

$$
\begin{aligned}
\frac{1}{2\pi} \int_0^{2\pi} \log\left| f(Re^{i\theta}) \right| \mathrm{d}\theta - \log|f(0)| &= \frac{1}{2\pi} \int_0^{2\pi} \Re \log f(Re^{i\theta}) \mathrm{d}\theta - \log|f(0)| \\
&= \frac{1}{2\pi} \int_0^{2\pi} \left(\Re \int_0^R \frac{\mathrm{d}}{\mathrm{d}r} \log f(re^{i\theta}) \mathrm{d}r \right) \mathrm{d}\theta \\
&= \frac{1}{2\pi} \Re \int_0^{2\pi} \int_0^R \frac{f'(re^{i\theta}) e^{i\theta}}{f(re^{i\theta})} \mathrm{d}r \mathrm{d}\theta \\
&= \Re \int_0^R \frac{1}{2\pi i r} \int_0^{2\pi} \frac{f'(re^{i\theta}) i r e^{i\theta}}{f(re^{i\theta})} \mathrm{d}\theta \mathrm{d}r \\
&= \Re \int_0^R r^{-1} \left(\frac{1}{2\pi i} \oint_{|z|=r} \frac{f'(z)}{f(z)} \mathrm{d}z \right) \mathrm{d}r.
\end{aligned}
$$

Now, by the argument principle and the fact that f is entire so has no poles,

$$\frac{1}{2\pi i} \oint_{|z|=r} \frac{f'(z)}{f(z)} dz = n(r),$$

where $n(r)$ represents the number of zeros of f inside the ball of radius r centred at zero. We therefore conclude that

$$\frac{1}{2\pi} \int_0^{2\pi} \log\left|f(Re^{i\theta})\right| d\theta - \log|f(0)| = \int_0^R r^{-1} n(r) dr.$$

$\square$

We now put this formula to immediate use, following the remainder of the argument in [4, Chapter 11]. Suppose as before that f is an entire function of order ρ with $f(0) \neq 0$. By the bound in the proof of Proposition 4.2, we have $\log\left|f(Re^{i\theta})\right| < R^\alpha$ on some large enough R for any $\alpha > \rho$. Invoking Jensen's formula gives

$$\int_0^R r^{-1} n(r) dr \ll \frac{1}{2\pi} \int_0^{2\pi} \log\left|f(Re^{i\theta})\right| d\theta < 2R^\alpha.$$

Since $n(r)$ is an increasing function of r, we have

$$\int_R^{2R} r^{-1} n(r) dr \geq n(R) \int_R^{2R} r^{-1} dr = n(R) \log 2.$$

From this, it follows that

$$n(R) \leq \frac{1}{\log 2} \int_R^{2R} r^{-1} n(r) dr \leq \frac{1}{\log 2} \int_0^{2R} r^{-1} n(r) dr \ll R^\alpha,$$

so we conclude that

$$n(R) = O(R^\alpha) \tag{4.2}$$

for any $\alpha > \rho$. This tells us that an entire function $f(z)$ must have its zeros sufficiently far apart. Consequently, if each zero z_n has radius r_n, we know that for any $\beta > \alpha$, $\sum_{n=1}^\infty r_n^{-\beta}$ converges. To see this, we split the sum into different annuli:

$$\sum_{n=1}^\infty r_n^{-\beta} = \sum_{k=0}^\infty \left(\sum_{2^k \leq r_n < 2^{k+1}} r_n^{-\beta} \right) \leq \sum_{k=0}^\infty \left(\sum_{2^k \leq r_n < 2^{k+1}} 2^{-\beta k} \right).$$

The number of terms in each of the sums is $O(2^{\alpha k})$ by (4.2). Thus,

$$\sum_{n=1}^\infty r_n^{-\beta} \ll \sum_{k=0}^\infty 2^{k(\alpha-\beta)} < \infty,$$

since we assumed $\beta > \alpha$. This therefore holds for any $\beta > \rho$, since α can be made arbitrarily close to ρ. We will now be chiefly concerned with entire functions $f(z)$ of order $\rho = 1$, and zeros at points $z_1, z_2, \ldots$. By the previous results, we know that $\sum_{n=1}^\infty r_n^{-1-\varepsilon}$ converges for any $\varepsilon > 0$, so in particular $\sum_{n=1}^\infty r_n^{-2}$ converges. We prove the following:

Theorem 4.4. *(The Hadamard Product Formula) For an entire function $f(z)$ of order 1, with zeros at $z_1, z_2, \ldots$ and $f(0) \neq 0$, there exist constants A, B such that*

$$f(z) = e^{A+Bz} \prod_{n=1}^{\infty} (1 - z/z_n) e^{z/z_n}.$$

Before proceeding, we must first check that the infinite product term has nice properties. Firstly, the infinite product

$$P(z) = \prod_{n=1}^{\infty} (1 - z/z_n) e^{z/z_n} \tag{4.3}$$

is absolutely convergent, and is uniformly convergent in compact sets containing no z_n. To show this, note that if $\log P(z)$ is convergent, then so is $P(z)$. We have

$$\log P(z) = \sum_{n=1}^{\infty} \log(1 - z/z_n) + z/z_n,$$

so by the Taylor expansion of $\log(1 - z)$, we have

$$\log P(z) = \sum_{n=1}^{\infty} z/z_n + \left(-(z/z_n) - \frac{(z/z_n)^2}{2} - \cdots \right)$$
$$= \sum_{n=1}^{\infty} \sum_{k=2}^{\infty} \frac{-(z/z_n)^k}{k}$$
$$= z^{-2} \sum_{n=1}^{\infty} z_n^{-2} \sum_{k=0}^{\infty} \frac{-(z/z_n)^k}{k+2}. \tag{4.4}$$

The absolute value of the second infinite sum is always bounded by $\log|1 - z/z_n|$, so we have

$$|\log P(z)| \leq |z|^{-2} \sum_{n=1}^{\infty} r_n^{-2} \log|1 - z/z_n| < \infty,$$

whenever z is not equal to any of the z_n. Therefore, $\log P(z)$ is absolutely convergent which implies $P(z)$ is convergent as required. We have therefore found an entire function $P(z)$ with exactly the same zeros (with multiplicity) as $f(z)$. Now, write

$$f(z) = F(z)P(z). \tag{4.5}$$

We have that $F(z)$ is also an entire function, and crucially one without zeros. In order to invoke Proposition 4.2, we must prove that $F(z)$ is finite order. We proceed by finding a lower bound for $|P(z)|$, and hence an upper bound for $|F(z)|$ on a sequence of circles $|z| = R$.

We must keep the values of R away from the zeros z_n. Since $\sum r_n^{-2}$ converges, the total length of all the intervals $(r_n - r_n^{-2}, r_n + r_n^{-2})$ is finite. Hence it cannot cover any infinite part of the real line. In particular, for each $r > 0$, there is an $R > r$ such that R does not lie in one of the intervals. In other words, there are arbitrarily large values of R such that

$$|R - r_n| > r_n^{-2}.$$

for every $n \in \mathbb{N}$. We proceed by fixing a large value R, and splitting the infinite product $P(z)$ into subproducts, depending on the radius of each zero. Define

$$P(z) = P_1(z)P_2(z)P_3(z),$$

where each P_i is the product over the following sets of z_n:

$$P_1(z): \quad |z_n| < \frac{1}{2}R,$$
$$P_2(z): \quad \frac{1}{2}R \leq |z_n| \leq 2R,$$
$$P_3(z): \quad |z_n| > 2R.$$

For the terms of $P_1(z)$, on $|z| = R$ we have

$$\left| (1 - z/z_n)e^{z/z_n} \right| \geq (|z/z_n| - 1)e^{\Re(z/z_n)} \geq (|z/z_n| - 1)e^{-|z/z_n|} > e^{-R/r_n},$$

using the fact that $|z/z_n| > 2$, so that the multiplying factor of the exponential term is greater than 1. We also have that

$$\sum_{r_n < R/2} r_n^{-1} = \sum_{r_n < R/2} r_n^{-1-\varepsilon}\, r_n^{\varepsilon} < \left(\frac{1}{2}R \right)^{\varepsilon} \sum_{r_n < R/2} r_n^{-1-\varepsilon} = C \left(\frac{1}{2}R \right)^{\varepsilon},$$

for some constant C, since the (possibly infinite) sum will always converge for any $\varepsilon > 0$. Therefore,

$$|P_1(z)| = \prod_{r_n < R/2} \left| (1 - z/z_n)e^{z/z_n} \right| > \exp\left(-R \left(\sum_{r_n < R/2} r_n^{-1} \right) \right) > \exp\left(-R^{1+2\varepsilon} \right),$$

for ε small enough. Now, for the terms in $P_2(z)$, we have

$$\left| (1 - z/z_n)e^{z/z_n} \right| = \frac{|z_n - z|}{|z_n|} e^{\Re(z/z_n)} \geq \frac{e^{-2}|z_n - z|}{2R} > CR^{-3},$$

for some constant $C > 0$, where the last inequality uses the fact that we chose our z to be greater than r_n^{-2} in distance away from every z_n. Therefore, the least distance between any z_n and z is proportional to $|z|^{-2} = R^{-2}$. Using (4.2), we have that the

29

number of zeros between $R/2$ and $2R$, and hence the number of terms in the product $P_2(z)$, is $O(R^{1+\varepsilon})$. Therefore,

$$|P_2(z)| > \exp\left(-R^{1+\varepsilon}(\log CR^3)\right) > \exp\left(-R^{1+2\varepsilon}\right),$$

for R large enough. Finally, for the terms of $P_3(z)$, we have

$$\left|(1 - z/z_n)e^{z/z_n}\right| \geq |1 - |z/z_n||e^{\Re(z/z_n)} > \frac{1}{2}e^{\Re(z/z_n)} > e^{-c(R/r_n)^2},$$

for some $c > 0$ which absorbs the factor of $\frac{1}{2}$, and the extra factor of R/r_n in the exponent, which is bounded above by $\frac{1}{2}$. We also have

$$\sum_{r_n > 2R} r_n^{-2} = \sum_{r_n > 2R} r_n^{-1-\varepsilon} r_n^{-1+\varepsilon} < (2R)^{-1+\varepsilon} \sum_{r_n > 2R} r_n^{-1-\varepsilon} = C(2R)^{-1+\varepsilon},$$

for some constant $C > 0$, which finally gives

$$\begin{aligned}
|P_3(z)| &> \prod_{r_n > 2R} e^{-c(R/r_n)^2} \\
&= \exp\left(-cR^2 \sum_{r_n > 2R} r_n^{-2}\right) \\
&> \exp\left(-C'R^2(2R)^{-1+\varepsilon}\right) \\
&> \exp\left(-R^{1+2\varepsilon}\right),
\end{aligned}$$

for R sufficiently large. Therefore, we can conclude that

$$|P(z)| = |P_1(z)||P_2(z)||P_3(z)| > \exp\left(-3R^{1+2\varepsilon}\right) > \exp\left(-R^{1+3\varepsilon}\right),$$

again for R large enough. Since $f(z)$ was order 1 by assumption, we have $f(z) < \exp\left(R^{1+\varepsilon}\right)$ for each $\varepsilon > 0$, R sufficiently large. Then we have that

$$|F(z)| = |f(z)||P(z)|^{-1} < \exp\left(R^{1+4\varepsilon}\right).$$

Since ε can be made arbitrarily small, we conclude that $F(z)$ is in fact entire of order 1 with no zeros. Invoking Proposition 4.2, we have that $F(z) = e^{A+Bz}$ for some constants A, B, leading us to Hadamard's formula,

$$f(z) = e^{A+Bz} \prod_{n=1}^{\infty} (1 - z/z_n)e^{z/z_n}. \tag{4.6}$$

4.2 Infinite Product Representation of $\xi(s)$ and $\xi(s, \chi)$

The formula (4.6) turns out to be very useful indeed. It allows us to write the ξ functions, and hence L-functions, in terms of their zeros. This is key to studying their distribution, and hence vital in proof of the prime number theorem. Owing to the pole of $\zeta(s)$ at $s = 1$, we must treat $\xi(s)$ and $\xi(s, \chi)$ separately as in [4, Chapter 12]. It is convenient in this section to define $\xi(s)$ as

$$\xi(s) = \frac{1}{2}s(s-1)\pi^{-s/2}\Gamma(s/2)\zeta(s), \tag{4.7}$$

so that $\xi(s)$ is entire. Note that the relation $\xi(s) = \xi(1 - s)$ holds as before. Therefore,

$$\xi(0) = \xi(1) = \frac{1}{2}\pi^{-1/2}\Gamma(1/2)\{\text{Res}_{s=1}\zeta(s)\} = \frac{1}{2}.$$

Furthermore, the trivial zeros of $\zeta(s)$ are cancelled by the poles of $\Gamma(s/2)$, so the only zeros of $\xi(s)$ are the non-trivial zeros of $\zeta(s)$. Now, when $\frac{1}{2} < \sigma < 2$, it is clear that $\xi(s)$ is bounded, since the only pole of any of its components is of $\zeta(s)$ at $s = 1$, at which point $\xi(s)$ is finite anyway. When $\sigma \geq 2$, by definition we have $\zeta(s) = O(1)$, while Stirling's formula (Theorem A5) gives $\Gamma(s/2) = O(\exp|s|\log|s|)$. Therefore,

$$\xi(s) = O(\exp(|s|\log|s|)).$$

This implies that $\xi(s)$ is of order at most 1, with zeros at the non-trivial zeros of the zeta-function. Applying Theorem 4.4, we have

$$\xi(s) = e^{A+Bz}\prod_{\rho}(1 - s/\rho)e^{s/\rho},$$

where the ρ are the non-trivial zeros of $\zeta(s)$, and A, B are constants. Now, we can take appropriate branches of logarithm which yields

$$\log\xi(s) = A + Bz + \sum_{\rho}\left(\log(1 - s/\rho) + s/\rho\right).$$

We already proved in generality that the logarithm of the infinite product representation is uniformly convergent for integral functions of order 1 in compact sets containing no zeros ρ, so we can differentiate termwise to give

$$\frac{\xi'(s)}{\xi(s)} = B + \sum_{\rho}\left(\frac{1}{s - \rho} + \frac{1}{\rho}\right).$$

On the other hand, logarithmic differentiation from the definition of $\xi(s)$ in (4.7) gives

$$\frac{\xi'(s)}{\xi(s)} = \frac{\mathrm{d}}{\mathrm{d}s}\left(-\frac{s}{2}\log\pi + \log(s - 1) + \log\Gamma(s/2 + 1) + \log\zeta(s)\right)$$

$$= -\frac{1}{2}\log\pi + \frac{1}{s - 1} + \frac{1}{2}\frac{\Gamma'(s/2 + 1)}{\Gamma(s/2 + 1)} + \frac{\zeta'(s)}{\zeta(s)},$$

31

where we absorb the factor of $s/2$ into the Γ-function. This implies that

$$\frac{\zeta'(s)}{\zeta(s)} = -\frac{1}{s-1} + B + \frac{1}{2}\log\pi - \frac{1}{2}\frac{\Gamma'(s/2+1)}{\Gamma(s/2+1)} + \sum_\rho \left(\frac{1}{s-\rho} + \frac{1}{\rho}\right). \qquad (4.8)$$

In the case of $\xi(s,\chi)$ for a primitive character χ modulo q, the function is already entire, so there is no need to cancel any poles. Thus, there is an analogous formula for L-functions derived in exactly the same way, namely

$$\frac{L'(s,\chi)}{L(s,\chi)} = -\frac{1}{2}\log\frac{q}{\pi} + B(\chi) - \frac{1}{2}\frac{\Gamma'(s/2+a/2)}{\Gamma(s/2+a/2)} + \sum_\rho \left(\frac{1}{s-\rho} + \frac{1}{\rho}\right), \qquad (4.9)$$

where ρ are the non-trivial zeros of $L(s,\chi)$. It is through these formulae where we begin to see the relevance of their zeros to the primes: we may write L-functions as infinite products over the primes, which consequently through logarithmic differentiation yields an expression directly involving their non-trivial zeros.

4.3   A Zero-Free Region for $\zeta(s)$

Recall the formula (2.13). Estimating $\psi(x)$ well, and hence $\pi(x)$, essentially boils down to two things. The first is making the contribution of each individual non-trivial zero, i.e. $|x^\rho|$, as small as possible. To ensure this, we aim to find minimum distance ρ must lie from the line $\sigma = 1$ in the critical strip, so that each term is lower order than x. The second way of ensuring the contribution of the zeros is minimised is to quantify how many terms there are in the sum. In this section, we will be concerned with the former. Such a minimum distance of each zero from the 1-line is known as a *zero-free region*. The case of trivial and non-trivial characters must be treated separately, owing to the pole at $s = 1$ for the trivial character.

In any case, our work so far has dealt mainly with primitive characters, so for the trivial character we refer to $\zeta(s)$. Recall (2.6), where

$$\log\zeta(s) = \sum_p \sum_{m=1}^\infty m^{-1}p^{-ms} \quad (\sigma > 1).$$

Since the right hand side is uniformly convergent on $\sigma > 1$, differentiating termwise gives

$$\frac{\zeta'(s)}{\zeta(s)} = -\sum_p \log p \left(\sum_{m=1}^\infty p^{-ms}\right) = -\sum_{n=1}^\infty \Lambda(n)n^{-s}, \quad (\sigma > 1)$$

where $\Lambda(n)$ is the von Mangoldt function defined in chapter 2. Therefore,

$$-\Re\frac{\zeta'(s)}{\zeta(s)} = \sum_{n=1}^\infty \Lambda(n)n^{-\sigma}\cos(t\log n), \quad (\sigma > 1). \qquad (4.10)$$

Now, consider the inequality, valid for all real θ,

$$0 \le 2(1 + \cos\theta)^2 = 3 + 4\cos\theta + \cos 2\theta.$$

Applied to (4.10), it implies that for all t,

$$3\left(-\frac{\zeta'(\sigma)}{\zeta(\sigma)}\right) + 4\left(-\Re\frac{\zeta'(\sigma + it)}{\zeta(\sigma + it)}\right) + \left(-\Re\frac{\zeta'(\sigma + 2it)}{\zeta(\sigma + 2it)}\right) \ge 0. \tag{4.11}$$

We now wish to bound each term by above in terms of the non-trivial zeros. Firstly, since $\zeta(s)$ has a simple pole at $s = 1$ of residue 1, it follows that for $1 < \sigma \le 2$:

$$-\frac{\zeta'(\sigma)}{\zeta(\sigma)} < \frac{1}{\sigma - 1} + A_1,$$

for some positive constant A_1. For complex s with $1 < \sigma \le 2$, $t \ge 2$, we refer back to the equality (4.8). Since Stirling's formula implies $\log\Gamma(s) \sim s\log s$, we have

$$\left|\frac{1}{2}\frac{\Gamma'(s/2 + 1)}{\Gamma(s/2 + 1)}\right| < A_2\log t,$$

for a positive constant A_2. Since all other terms in (4.8) except the sum are bounded,

$$-\Re\frac{\zeta'(s)}{\zeta(s)} < A_2\log t - \sum_\rho \Re\left(\frac{1}{s - \rho} + \frac{1}{\rho}\right). \tag{4.12}$$

Furthermore, since

$$\Re\frac{1}{s - \rho} = \frac{\sigma - \beta}{|s - \rho|^2}, \quad \text{and} \quad \Re\frac{1}{\rho} = \frac{\beta}{|\rho|^2},$$

the sum over zeros is strictly positive. We can therefore throw away any number of terms in (4.12), and the inequality will still hold. Choose t such that t coincides with the imaginary part of a non-trivial zero, say $\rho_1 = \beta + i\gamma$. Then we certainly have

$$-\Re\frac{\zeta'(\sigma + 2it)}{\zeta(\sigma + 2it)} < A_2\log t,$$

and

$$-\Re\frac{\zeta'(\sigma + it)}{\zeta(\sigma + it)} < A_2\log t - \frac{1}{\sigma - \beta}, \tag{4.13}$$

where the second inequality throws away all but the $(s - \rho)^{-1}$ term in the sum corresponding to the zero ρ_1. Putting all of these inequalities into (4.11), we obtain

$$\frac{4}{\sigma - \beta} < \frac{3}{\sigma - 1} + A\log t,$$

where A is a positive constant. Let $\sigma = 1 + \delta/\log t$. Then, sparing tedious algebraic details, we have

$$\beta < 1 - \frac{\delta(1 - A\delta)}{(3 + A\delta)\log t},$$

33

which, upon choosing δ small enough in relation to A, gives $c > 0$ such that

$$\beta < 1 - \frac{c}{\log t}.$$

Since $\beta + i\gamma$ was an arbitrary zero, we may therefore find a c for any $t \geq 2$ such that any zero satisfies this property. It follows from the functional equation that if ρ is a non-trivial zero of $\zeta(s)$, then so is $\overline{\rho}$. Therefore, the above region holds upon replacing t by $|t|$. Hence, we have found an explicit zero free region for $\zeta(s)$, and proved the following.

Theorem 4.5. *For all $s = \sigma + it$ satisfying*

$$\sigma \geq 1 - \frac{c}{\log|t|}, \quad |t| \geq 2,$$

we have $\zeta(s) \neq 0$.

In order to extend this theorem to all t, we must prove that there are no zeros arbitrarily close to $\sigma = 1$ where $|t| < 2$. This is implied by the statement that $\zeta(1+it) \neq 0$ for all t. To show this, note that an identical argument as before involving the (double) infinite sum representation of $\log \zeta(s)$ in (2.6) gives

$$3 \log \zeta(\sigma) + 4\Re \log \zeta(\sigma + it) + \Re \log \zeta(\sigma + 2it) \geq 0,$$

so that exponentiation implies

$$\left| \zeta^3(\sigma)\zeta^4(\sigma + it)\zeta(\sigma + 2it) \right| \geq 1, \tag{4.14}$$

all of this when $\sigma > 1$. Suppose $\zeta(1 + it) = 0$. As $\sigma \to 1$, $\zeta(\sigma) \sim (\sigma - 1)^{-1}$, while $\zeta(\sigma + it) \sim (\sigma - 1)$ (by assumption), so that the left hand side of (4.14) goes to zero. This clearly contradicts (4.14), so we conclude that $\zeta(1 + it)$ is non-zero. Thus, we may extend the zero free region to all s satisfying

$$\sigma \geq 1 - \frac{c}{\log(|t| + 2)}.$$

4.4   A Zero-Free Region for $L(s, \chi)$

We now consider L-functions of a non-trivial character, since by (2.5) a zero-free region for those of trivial character follows immediately from that of $\zeta(s)$. An added complication in this case is the value of q, as such regions will depend on the size of this modulus. Following [4, Chapter 14], suppose χ is a non-trivial character of modulus $q \geq 3$, and that $t \geq 0$. It is enough to consider non-negative t, since any zero of $L(s, \chi)$ with $t < 0$ is a zero of $L(s, \overline{\chi})$ with $t > 0$. Logarithmic differentiation of the product formula gives

$$-\frac{L'(s, \chi)}{L(s, \chi)} = \sum_{n=1}^{\infty} \Lambda(n) n^{-\sigma} \chi(n) e^{-it \log n}, \tag{4.15}$$

which is obtained in the same way as for $\zeta(s)$, but with an added factor of $\chi(n)$. Owing to the fact that $\chi(n)$ is a $\phi(q)$-th root of unity (as long as n and q are coprime), we may write $\Re\chi(n)e^{-it\log n}$ as $\cos\theta$, for some θ. Replacing χ by χ^2 doubles the contribution to the argument from χ (as it is a root of unity). Therefore, $\Re\chi^2(n)e^{-2it\log n} = \cos 2\theta$. Moreover, replacing χ by χ_0 and t by zero produces $\cos 0 = 1$, so we may write an analogous inequality to (4.11) involving L-functions: namely

$$3\left[-\frac{L'(\sigma,\chi_0)}{L(\sigma,\chi_0)}\right] + 4\left[-\Re\frac{L'(\sigma+it,\chi)}{L(\sigma+it,\chi)}\right] + \left[-\Re\frac{L'(\sigma+2it,\chi^2)}{L(\sigma+2it,\chi^2)}\right] \geq 0. \qquad (4.16)$$

Notice that if χ is a real character, $\chi^2 = \chi_0$, and this can cause serious problems. For now, we assume χ is complex so that this issue is avoided. Recall (4.9), which implies

$$-\Re\frac{L'(s,\chi)}{L(s,\chi)} = \frac{1}{2}\log\frac{q}{\pi} + \frac{1}{2}\Re\frac{\Gamma'(s/2+a/2)}{\Gamma(s/2+a/2)} - \Re B(\chi) - \Re\sum_\rho\left(\frac{1}{s-\rho} + \frac{1}{\rho}\right).$$

By logarithmic differentiation of the Hadamard product for $\xi(s,\chi)$, followed by application of the functional equation, we have

$$B(\chi) = \frac{\xi'(0,\chi)}{\xi(0,\chi)} = -\frac{\xi'(1,\overline{\chi})}{\xi(1,\overline{\chi})} = -B(\overline{\chi}) - \sum_{\overline{\rho}}(\frac{1}{1-\overline{\rho}} + \frac{1}{\overline{\rho}}),$$

where $\overline{\rho}$ are the non-trivial zeros of $L(s,\overline{\chi})$. Moreover, it is clear that $B(\overline{\chi}) = \overline{B(\chi)}$, so

$$2\Re B(\chi) = -\sum_{\overline{\rho}}(\Re\frac{1}{1-\overline{\rho}} + \Re\frac{1}{\overline{\rho}}).$$

Since the zeros of L-functions are symmetrically distributed around the critical line, we may change $1 - \overline{\rho}$ to ρ with no effect, as this simply permutes the terms. Therefore,

$$\Re B(\chi) = -\frac{1}{2}\sum_\rho(\Re\frac{1}{\rho} + \Re\frac{1}{\overline{\rho}}) = -\sum_\rho\Re\frac{1}{\rho}.$$

Substituting this into (4.17), combined with the fact that the Γ'/Γ term is $O\left(\log(2+t)\right)$ as in the previous section, we have

$$-\Re\frac{L'(s,\chi)}{L(s,\chi)} < c_1\mathcal{L} - \sum_\rho\Re\frac{1}{s-\rho}, \quad \text{where} \quad \mathcal{L} := \log q + \log(t+2), \qquad (4.17)$$

which holds for any $\sigma > 1$ and any primitive character χ (we have not yet used the assumption that χ is complex). Since

$$\Re\frac{1}{s-\rho} = \frac{\sigma-\beta}{|s-\rho|^2} \geq 0, \quad (\sigma > 1),$$

we may as before throw away any number of the terms in the sum without changing the inequality. It is not immediately obvious that the inequality (4.17) holds for $L(s,\chi^2)$,

since χ^2 is non-trivial, but not necessarily primitive. However, suppose χ^2 is induced by a character χ_1. Then by (2.4), we have

$$\log L(s, \chi^2) = \log L(s, \chi_1) + \sum_{p|q} \log(1 - \chi_1(p)p^{-s}),$$

which implies

$$\frac{L'(s, \chi^2)}{L(s, \chi^2)} = \frac{L'(s, \chi_1)}{L(s, \chi_1)} + \sum_{p|q} \frac{-\chi_1(p)p^{-s} \log p}{1 - \chi_1(p)p^{-s}}.$$

Therefore,

$$\left| \frac{L'(s, \chi^2)}{L(s, \chi^2)} - \frac{L'(s, \chi_1)}{L(s, \chi_1)} \right| \leq \sum_{p|q} \frac{p^{-\sigma} \log p}{1 - p^{-\sigma}} \leq \sum_{p|q} \log p \leq \log q.$$

The first inequality is just the triangle inequality, the second inequality may be shown by the fact that $|x(1 - x)^{-1}| \leq 1$ for all $x \in [0, 1/2]$ (just check the derivative), while the third is since the sum of logarithms is the logarithm of the product. Thus, the inequality (4.17) also holds for $L'(s, \chi^2)/L(s, \chi^2)$. We proceed as in the previous section, omitting the entire series in (4.17) for $L'(\sigma + 2it, \chi^2)/L(\sigma + 2it, \chi^2)$. We choose t to coincide with the imaginary part of a non-trivial zero of $L(s, \chi)$, and keep only the term in the series corresponding to this zero, which gives

$$-\Re \frac{L'(\sigma + it, \chi)}{L(\sigma + it, \chi)} < c_1 \mathcal{L} - \frac{1}{\sigma - \beta}.$$

Finally, the term for the trivial character has the same bound as in the previous section for $\zeta(\sigma)$, say $(\sigma - 1)^{-1} + c_2$. It follows from (4.16) that

$$\frac{4}{\sigma - \beta} < \frac{3}{\sigma - 1} + c_3 \mathcal{L}.$$

Setting $\sigma = 1 + c_4 \mathcal{L}^{-1}$, we have as in the previous section (choosing an appropriate c_4)

$$\beta < 1 - \frac{c_5}{\mathcal{L}} \tag{4.18}$$

for an absolute constant c_5. This inequality is also true for complex non-primitive characters. Indeed, if the character χ is induced by a primitive χ_1, then any zeros of $L(s, \chi)$ are either zeros of $L(s, \chi_1)$ or any of the terms $(1 - \chi_1(p)p^{-s})$ in the product part of (2.4). Any zero of those terms must lie on $\sigma = 0$, so the relation (4.18) still holds for the zeros of $L(s, \chi)$.

Suppose now that χ is a real primitive character. We now have the issue that $\chi^2 = \chi_0$. Note that this is the same as having χ^2 induced by the trivial character of modulus 1, so we may refer to our previous argument, which implies that

$$\left| \frac{L'(s, \chi_0)}{L(s, \chi_0)} - \frac{\zeta'(s)}{\zeta(s)} \right| \leq \log q$$

for $\sigma > 1$. With regard to ζ'/ζ, we refer to (4.8), which with our usual estimates, and not assuming that t is large, gives

$$-\Re\frac{\zeta'(s)}{\zeta(s)} < \Re\frac{1}{s-1} + c_6 \log t.$$

Thus, combining the previous two inequalities using the triangle inequality gives

$$-\Re\frac{L'(\sigma+2it,\chi_0)}{L(\sigma+2it,\chi_0)} < \Re\frac{1}{\sigma-1+2it} + c_7\mathcal{L},$$

with $\mathcal{L}$ as before. We therefore replace the inequality used for complex χ by this one, and by (4.16) we have

$$\frac{4}{\sigma-\beta} < \frac{3}{\sigma-1} + \Re\left(\frac{1}{\sigma-1+2it}\right) + c_8\mathcal{L},$$

where in this case $t = \gamma$ so as to coincide with the imaginary part of a non-trivial zero of $L(s,\chi)$. Take $\sigma = 1 + \delta/\mathcal{L}$, and require that $\delta/\mathcal{L} \le \gamma$, so that

$$\frac{4}{\sigma-\beta} < \frac{3\mathcal{L}}{\delta} + \frac{\mathcal{L}}{5\delta} + c_8\mathcal{L}.$$

Therefore, with some rearranging we have

$$\beta < 1 - \frac{4-5c_8\delta}{16+5c_8\delta}\frac{\delta}{\mathcal{L}}.$$

This gives the same type of bound for a complex character χ, subject to the condition that $\gamma \ge \delta/\mathcal{L}$, which is certainly satisfied when $\gamma \ge \delta/\log q$. As before, χ need not be primitive. It now remains to consider when $|t| < \delta/\log q$.

Proposition 4.6. *There is at most one zero of $L(s,\chi)$ for a real non-trivial character χ satisfying $\beta > 1 - \delta/\log q$.*

It is a consequence of this proposition that any such zero must be real. Indeed, for a real character χ, we have $\overline{\chi} = \chi$, so zeros of $L(s,\chi)$ are consequently symmetric about the real axis. Thus, existence of a complex zero in this region guarantees the existence of at least two - a contradiction. Therefore, assume that $L(s,\chi)$ has zeros at $\beta \pm i\gamma$ for $\gamma \ne 0$. Now, (4.17) may be adapted for the case of $s = \sigma > 1$ to give

$$-\frac{L'(\sigma,\chi)}{L(\sigma,\chi)} < c_{10}\log q - \sum_\rho \frac{1}{\sigma-\rho},$$

where the sum is real since each zero occurs in a conjugate pair (or is itself real). Note that this inequality also requires χ to be primitive. We may assume this, since we may adapt our argument as previously to induced characters. Keeping only the terms corresponding to the zeros $\beta \pm i\gamma$, we have

$$-\frac{L'(\sigma,\chi)}{L(\sigma,\chi)} < c_{10}\log q - \frac{2(\sigma-\beta)}{(\sigma-\beta)^2+\gamma^2}.$$

37

For the left hand side, we may bound by

$$-\frac{L'(\sigma,\chi)}{L(\sigma,\chi)} = \sum_{n=1}^{\infty} \chi(n)\Lambda(n)n^{-\sigma} \geq -\sum_{n=1}^{\infty} \Lambda(n)n^{-\sigma} = \frac{\zeta'(\sigma)}{\zeta(\sigma)} > \frac{-1}{\sigma-1} - c_{11},$$

so that

$$-\frac{1}{\sigma-1} < c_{12}\log q - \frac{2(\sigma-\beta)}{(\sigma-\beta)^2+\gamma^2}.$$

Taking $\sigma = 1 + 2\delta/\log q$, we have by assumption on γ that

$$|\gamma| < \frac{\delta}{\log q} = \frac{1}{2}(\sigma-1) < \frac{1}{2}(\sigma-\beta).$$

Substituting into the previous inequality, we have

$$-\frac{\log q}{2\delta} = -\frac{1}{\sigma-1} < c_{12}\log q - \frac{8}{5(\sigma-\beta)}.$$

Rearranging this gives

$$\beta < 1 - \frac{2(8-5(1+2\delta c_{12}))}{5(1+2\delta c_{12})}\frac{\delta}{\log q}.$$

Choosing δ well in relation to c_{12} therefore implies $\beta < 1 - \delta/\log q$ (it is important to note that this is the δ of the previous result). It suffices therefore to consider the case that we have 2 real zeros (or a double real zero). The argument for this case is essentially the same, with $\gamma = 0$. Therefore, we have proved that the only zero satisfying

$$\beta > 1 - \delta/\log q, \quad |\gamma| < \frac{\delta}{\log q},$$

is a single real zero. In essence, we have now proved an (almost) zero free region for all L-functions, which may be stated in the following theorem.

Theorem 4.7. *There exists a positive constant c such that if χ is a complex character modulo q, there is no zero of $L(s,\chi)$ in the region*

$$\sigma \geq \left\{ \begin{array}{ll} 1 - c/\log q|t| & \text{if } |t| \geq 1, \\ 1 - c/\log q & \text{if } |t| < 1. \end{array} \right\}. \tag{4.19}$$

Moreover, if χ is a real non-trivial character, there is at most one zero in this region, in which case it is a single real zero.

There is one final fact that we must show with regard to real zeros: for a given q, there is at most one real character whose L-function has a real zero satisfying (4.19). We adapt a more general result of Landau detailed in [4, p. 93], which is enough for our purposes.

Proposition 4.8. *Suppose χ_1, χ_2 are distinct primitive real characters of modulus q. If their corresponding L-functions have real zeros at β_1, β_2, then they satisfy*

$$\min(\beta_1, \beta_2) < 1 - \frac{\delta'}{\log q}.$$

where δ' is a positive constant.

Proof. By the properties of Dirichlet characters, $\chi_1\chi_2$ is also a real character of modulus q. Moreover, it is non-trivial as $\chi_1\chi_2(n) = 1$ for all n coprime to q would contradict $\chi_1 \neq \chi_2$. It is not necessarily primitive, but by a similar argument to that used when considering χ^2, the inequality (4.17) holds for $\chi_1\chi_2$. In particular, for $\sigma > 1$,

$$-\frac{L'(\sigma, \chi_1\chi_2)}{L(\sigma, \chi_1\chi_2)} < c_{13} \log q.$$

Furthermore, we have the earlier bounds

$$-\frac{L'(\sigma, \chi_i)}{L(\sigma, \chi_i)} < c_{14} \log q - \frac{1}{\sigma - \beta_i}, \qquad -\frac{\zeta'(\sigma)}{\zeta(\sigma)} < \frac{1}{\sigma - 1} + c_{15}.$$

Since

$$-\frac{\zeta'(\sigma)}{\zeta(\sigma)} - \frac{L'(\sigma, \chi_1)}{L(\sigma, \chi_1)} - \frac{L'(\sigma, \chi_2)}{L(\sigma, \chi_2)} - \frac{L'(\sigma, \chi_1\chi_2)}{L(\sigma, \chi_1\chi_2)} = \sum_{n=1}^{\infty} \Lambda(n)[1 + \chi_1(n)][1 + \chi_2(n)]n^{-\sigma} \geq 0,$$

we therefore have the inequality

$$\frac{1}{\sigma - \beta_1} + \frac{1}{\sigma - \beta_2} < \frac{1}{\sigma - 1} + c_{16} \log q,$$

where c_{16} is an absolute constant. Setting $\sigma = 1 + \delta / \log q$, we conclude that the distance of at least one of the β_i from 1 is bounded in terms of $1/\log q$. For if both are arbitrarily close to 1, our choice of σ implies $\delta^{-1} < c_{16}$, which is a contradiction if δ is small enough. Thus, there exists $\delta' > 0$ such that at least one of the β_i is less than $1 - \delta'/\log q$. $\qquad\square$

The possible existence of real zeros close to $s = 1$ is a real issue, and bounding the distance they must lie from $s = 1$ is difficult. This shall be deferred until later when we study Siegel's theorem.

4.5 The Numbers $N(T)$ and $N(T, \chi)$

We now turn our attention to formulas of the type stated by Riemann and proved by von Mangoldt. These concern the number of zeros of L-functions below a given height in the critical strip, estimating how frequently the non-trivial zeros occur. We first study the function $N(T, \chi)$, defined as the number of zeros of $L(s, \chi)$ in the rectangle $0 < \sigma < 1$, $|t| < T$. It suffices to study the case of χ primitive and for now we shall assume $q > 1$.

We proceed by the principle of argument on the function $\xi(s,\chi)$, over the rectangle R with vertices at

$$5/2 - iT, \quad 5/2 + iT, \quad -3/2 + iT, \quad -3/2 - iT,$$

traversed anti-clockwise. Since $\xi(s,\chi)$ is holomorphic inside this region, its change in argument is precisely the number of zeros of the function (with multiplicity). Moreover, since the only zeros of this function lie at non-trivial zeros of $L(s,\chi)$, we have that

$$2\pi N(T,\chi) = \Delta_R \arg \xi(s,\chi), \tag{4.20}$$

where the right hand represents the total change in argument of $\xi(s,\chi)$ over the rectangle R. Split the rectangle into two halves along the line $\sigma = 1/2$, and call the right half L. By the functional equation, we have

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

so that the change in argument along each half of the rectangle is the same (c here is independent of s). Thus, it suffices to consider L. We have

$$\Delta_L \arg \left(\frac{q}{\pi}\right)^{(s+a)/2} = \arg \left(\frac{q}{\pi}\right)^{iT/2} - \arg \left(\frac{q}{\pi}\right)^{-iT/2} = T \log \left(\frac{q}{\pi}\right).$$

In order to estimate the change in argument of the $\Gamma(s/2 + a/2)$ term, we must use Stirling's formula (Theorem A5) which states that

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

We have that

$$\Delta_L \arg \Gamma \left(\frac{1}{2}(s/2 + a/2)\right) = \Im \log \Gamma \left(\frac{\alpha}{2}\right) - \Im \log \Gamma \left(\frac{\overline{\alpha}}{2}\right),$$

upon setting $\alpha = \frac{1}{2} + a + iT$, by the properties of complex logarithm. Notice that changing s for $\overline{s}$ in Stirling's formula simply gives its complex conjugate (with the same error term), so that

$$\log \Gamma(z) - \log \Gamma(\overline{z}) = 2i\Im \left((z - \frac{1}{2}) \log z - z + \frac{1}{2} \log 2\pi\right) + O(|z|^{-1}).$$

Setting $z = \alpha/2$ into this expression, we have

$$\Delta_L \arg \Gamma(s/2 + a/2) = \Im \left[(\alpha - 1) \log \frac{\alpha}{2} - \alpha + \log 2\pi\right] + O(1)$$

$$= T \log \frac{T}{2} - T + O(1). \tag{4.21}$$

It therefore remains to estimate $\Delta_L \arg L(s,\chi)$ and we aim to prove that it is $O(\log qT)$. To do this, we will make use of the following lemma.

Lemma 4.9. *If $\rho = \beta + i\gamma$ are the non-trivial zeros of $L(s,\chi)$ where χ is primitive, then for all real t, we have*

$$\sum_{\rho} \frac{1}{1 + (t - \gamma)^2} = O(\mathcal{L})$$

where $\mathcal{L} := \log\left(q(|t| + 2)\right)$.

Proof. Recall (4.17) of the previous section, where we had

$$-\Re\frac{L'(s,\chi)}{L(s,\chi)} < c_1\mathcal{L} - \sum_{\rho} \Re\frac{1}{s - \rho}.$$

Set $s = 2 + it$. On this line $|L'/L|$ is bounded by an absolutely convergent sum independent of t, so this term may be dispensed of in the equation above to give

$$\sum_{\rho} \Re\frac{1}{(2 + it) - \rho} \ll \mathcal{L}.$$

Moreover, we have

$$\Re\frac{1}{(2 + it) - \rho} = \frac{(2 - \beta)}{(2 - \beta)^2 + (t - \gamma)^2} \gg \frac{1}{1 + (t - \gamma)^2}.$$

In conclusion,

$$\sum_{\rho} \frac{1}{1 + (t - \gamma)^2} = O(\mathcal{L}).$$

$\square$

Now, two conclusions are immediate from this lemma. First, the number of terms satisfying $|t - \gamma| < 1$ is of order at most $\mathcal{L}$, since if the terms of the sum are all $O(1)$, the number of terms must be at most $O(\mathcal{L})$. Secondly, the sum over zeros with $|t - \gamma| \geq 1$ is also $O(\mathcal{L})$ for the same reason. Now, it follows that

$$\frac{L'(s,\chi)}{L(s,\chi)} = \sum_{\rho:\ |t-\gamma|<1} \frac{1}{s - \rho} + O(\mathcal{L}). \tag{4.22}$$

To see this, we use the equation (4.9). Setting $s = 2 + it$ and subtracting from the original equation, we have

$$\frac{L'(s,\chi)}{L(s,\chi)} - \frac{L'(2 + it,\chi)}{L(2 + it,\chi)} = F(s) + \sum_{\rho} \left(\frac{1}{s - \rho} - \frac{1}{2 + it - \rho}\right),$$

where $F(s)$ is a term involving Γ'/Γ, which by Stirling's formula is essentially $O(\mathcal{L})$. As in the proof of the previous lemma, $L'(2 + it,\chi)/L(2 + it,\chi)$ is $O(1)$, so we have

$$\frac{L'(s,\chi)}{L(s,\chi)} = \sum_{\rho} \left(\frac{1}{s - \rho} - \frac{1}{2 + it - \rho}\right) + O(\mathcal{L}).$$

41

It remains to estimate the sum. First, consider the case when $|t - \gamma| \geq 1$. We have

$$|s - \rho|^2 \geq (t - \gamma)^2, \quad |2 + it - \rho|^2 \geq (t - \gamma)^2,$$

and we may take square roots while maintaining the inequality. Thus,

$$\left| \frac{1}{s - \rho} - \frac{1}{2 + it - \rho} \right| = \frac{2 - \sigma}{|s - \rho||2 + it - \rho|} \leq \frac{3}{(t - \gamma)^2} \ll \frac{1}{1 + (t - \gamma)^2},$$

so by the previous lemma, the sum over these terms is $O(\mathcal{L})$. When $|t - \gamma| < 1$, we have $|2 + it - \rho| \geq 1$, and the number of terms in the sum is $O(\mathcal{L})$ as discussed before. Therefore,

$$\sum_{\rho : |t - \gamma| < 1} \left(\frac{1}{s - \rho} - \frac{1}{2 + it - \rho} \right) = \sum_{\rho : |t - \gamma| < 1} \frac{1}{s - \rho} + O(\mathcal{L}).$$

We therefore put everything together to give (4.22). We may now estimate $\Delta_L \arg L(s, \chi)$. For the vertical line from $5/2 - it$ to $5/2 + it$, by (4.15) we have

$$\Delta \arg L(s, \chi) = \Im \int_{5/2 - it}^{5/2 + it} \frac{L'(s, \chi)}{L(s, \chi)} \mathrm{d}s = \Im \left[\sum_{n=1}^{\infty} \frac{\chi(n)\Lambda(n)}{n^s \log n} \right]_{5/2 - it}^{5/2 + it} = O(1).$$

For the remaining line segments, we use (4.22), giving

$$\Delta \arg L(s, \chi) = \Im \int_{1/2 \pm it}^{5/2 \pm it} \frac{L'(s, \chi)}{L(s, \chi)} \mathrm{d}s = \sum_{\rho : |t - \gamma| < 1} \Im \int_{1/2 \pm it}^{5/2 \pm it} (s - \rho)^{-1} \mathrm{d}s + O(\mathcal{L}).$$

Each integral inside the sum is given by $\Delta \arg(s - \rho)^{-1}$ along each of the segments, which is $O(1)$ as it is bounded by π. There are also $O(\mathcal{L})$ terms in the sum by the above, so the change in argument of L'/L over these segments is $O(\mathcal{L})$. Putting everything together, we have

$$\Delta_L \arg L(s, \chi) = O(\mathcal{L}). \tag{4.23}$$

We now restrict to a fixed $t = T \geq 2$. Then, putting together (4.21), (4.21), and (4.23), we have

$$\Delta_L \arg \xi(s, \chi) = T \log \frac{qT}{2\pi} - T + O(\log qT).$$

In conclusion, using (4.20), and the fact that the contribution over R is twice that over L, we have

$$\frac{1}{2} N(T, \chi) = \frac{T}{2\pi} \log \frac{qT}{2\pi} - \frac{T}{2\pi} + O(\log qT). \tag{4.24}$$

The argument is easily modified for $q = 1$, and the function $N(T)$, defined as the number of zeros of $\zeta(s)$ in the region $0 < \sigma < 1$, $0 \leq t \leq T$. Note that it is sufficient to consider the upper half-plane in this case, as $\zeta(s)$ is symmetric about the real-axis. The only modification required is a multiplicative factor of $s(s - 1)/2$ in the function $\xi(s, \chi)$ so that it is entire, and its change in argument gives only zeros (and not poles). The change

of argument of this factor over the right-half of the rectangle is easily calculated to be essentially

$$\pi + O(T^{-1}),$$

so can be ignored, as it is lower order than the error term in (4.24), with which we recover the famous Riemann von-Mangoldt formula,

$$N(T) = \frac{T}{2\pi} \log \frac{T}{2\pi} - \frac{T}{2\pi} + O(\log T),$$

for $T \geq 2$. Note that the factor of $1/2$ in the more general formula above is for ease of comparison with this well-established formula for $\zeta(s)$. We now have sufficient information with regard to the non-trivial zeros of L-functions (and in particular $\zeta(s)$) in order to proceed, and prove at least one landmark result regarding the distribution of primes. Our aim is now to expose the connection of these non-trivial zeros and the primes, studying the method behind formulas such as (2.13).

Chapter 5

The Distribution of Prime Numbers

5.1 The Prime Number Theorem

We now move on to showing the connections of $\psi(x)$ and $\psi(x, \chi)$ to non-trivial zeros of certain L-functions. Then, with the knowledge of these zeros gained in the previous section, we will produce asymptotic formulas for these functions, and hence asymptotic formulas for the prime counting functions. The general idea is to use the following lemma [5, Lemma 4].

Lemma 5.1. *Let $y > 0$, $c > 1$ and $T \geq 1$. Defining*

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

we have

$$I(y, T) = \left\{ \begin{array}{ll} 0, & 0 < y < 1 \\ 1, & y > 1 \end{array} \right\} + O\left(\frac{y^c}{T|\log y|} \right).$$

Proof. We begin with the case $0 < y < 1$. Here, we evaluate the integral using the path $c - iT \to N - iT \to N + iT \to c + iT$, letting $N \to \infty$. This is equivalent, since the integrand is holomorphic on the right half-plane. Consider the line segment from $N - iT \to N + iT$ first. Using the estimation lemma, we have

$$\left| \int_{N-iT}^{N+iT} \frac{y^s}{s} \mathrm{d}s \right| \leq 2T \sup_{\theta \in [0,1]} \left| \frac{y^{N+iT(2\theta-1)}}{N + iT(2\theta - 1)} \right|$$

$$\leq 2T \frac{y^N}{N - T} \to 0 \text{ as } N \to \infty.$$

Therefore, it is enough to consider the other two segments as $N \to \infty$. We have

$$\left| \int_{c\pm iT}^{N\pm iT} \frac{y^s}{s} \mathrm{d}s \right| \ll \int_c^N \frac{y^\theta}{T} \mathrm{d}\theta \ll \frac{y^c}{T|\log y|}.$$

which implies the first result. Now suppose $y > 1$. Consider the closed contour Γ, defined to be the rectangle with vertices at $-N-iT, -N+iT, c+iT, c-iT$. By Cauchy's residue theorem, we have that

$$\frac{1}{2\pi i} \oint_\Gamma \frac{y^s}{s} \mathrm{d}s = \mathrm{Res}_{s=0}\left(\frac{y^s}{s}\right) = 1.$$

Denote the path $c - iT \to -N - iT \to -N + iT \to c + iT$ by γ, then rearranging gives

$$\frac{1}{2\pi i} \int_{c-iT}^{c+iT} \frac{y^s}{s} \mathrm{d}s = 1 + \frac{1}{2\pi i} \int_\gamma \frac{y^s}{s} \mathrm{d}s.$$

The integral over γ when $y > 1$ is completely analogous to the case when $0 < y < 1$, so

$$\int_\gamma \frac{y^s}{s} \mathrm{d}s = O\left(\frac{y^c}{T|\log y|}\right),$$

and this gives the required result. $\qquad \square$

The idea now is to use a slightly modified function as the integrand and apply this lemma. From now on we assume that χ is a primitive character modulo q. Using (4.15), we have for $x \notin \mathbb{N}$,

$$\begin{aligned}
\frac{1}{2\pi i} \int_{c-iT}^{c+iT} \left\{-\frac{L'(s,\chi)}{L(s,\chi)}\right\} \frac{x^s}{s} \mathrm{d}s &= \frac{1}{2\pi i} \int_{c-iT}^{c+iT} \sum_{n=1}^\infty \chi(n)\Lambda(n) \frac{(x/n)^s}{s} \mathrm{d}s \\
&= \sum_{n=1}^\infty \chi(n)\Lambda(n) I(x/n, T) \\
&= \psi(x,\chi) + O\left(\sum_{n=1}^\infty \chi(n)\Lambda(n) \frac{(x/n)^c}{T|\log x/n|}\right). \quad (5.1)
\end{aligned}$$

Termwise integration is justified by the uniform convergence of the infinite sum representation of L'/L, and the last step applies the previous lemma: only terms with $n \leq x$ have $x/n > 1$, so the integral evaluates to 1 on exactly these terms, with the same error for all terms in the sum. We now find a more useful estimate for the error term. Firstly, for $n < x/2$ or $n > 2x$, we have that $|\log x/n| > \log 2$. Moreover, $\chi(n)$ always has modulus at most 1, so for these terms

$$\sum_{\substack{n<x/2 \\ n>2x}} \chi(n)\Lambda(n) \frac{(x/n)^c}{T|\log x/n|} \ll \frac{x^c}{T} \sum_{n=1}^\infty \Lambda(n) n^{-c} = -\frac{x^c}{T}\frac{\zeta'(c)}{\zeta(c)}. \quad (5.2)$$

We now set $c = 1 + (\log x)^{-1}$, and use the estimate (4.12). This implies that (5.2) is of order

$$\frac{x^c}{T}\frac{1}{c-1} \ll \frac{x\log x}{T},$$

45

where we used that $x^c = x^{1+1/\log x} = ex \ll x$. Now, when $x/2 < n < 2x$, we must find a different estimate for the $|\log x/n|$ term. Consider the function

$$f(y) = \log y + \frac{1}{y} - 1,$$

defined for positive y. It is easy to check that its derivative is strictly negative when $0 < y < 1$, and strictly positive for $y > 1$. Thus, it attains its minimum at $f(1) = 0$, so is non-negative for all positive y. It follows that

$$\log y \geq 1 - \frac{1}{y}$$

for all $y > 0$. We therefore apply this inequality. When $x/2 < n < x$, we have

$$|\log x/n| = \log x/n \geq 1 - \frac{n}{x} = \frac{x - n}{x} \geq \frac{|x - n|}{2x}.$$

When $x < n < 2x$, we have

$$|\log x/n| = \log n/x \geq 1 - \frac{x}{n} = \frac{n - x}{n} \geq \frac{|x - n|}{2x}.$$

Therefore, we have the estimate $|\log x/n| \geq |x - n|/2x$ for all $x/2 < n < 2x$. The error term in this case is therefore of order

$$\sum_{x/2 < n < 2x} \frac{x\Lambda(n)}{T|x - n|} \ll \sum_{x/2 < n < 2x} \frac{x \log x}{T|x - n|} \ll \frac{x \log x}{T} \int_1^x \frac{\mathrm{d}t}{t} = \frac{x \log^2 x}{T},$$

where we use estimation of the sum by an integral of the same type. Thus, putting all these estimates together gives

$$\psi(x, \chi) = \frac{1}{2\pi i} \int_{c-iT}^{c+iT} \left\{ -\frac{L'(s, \chi)}{L(s, \chi)} \right\} \frac{x^s}{s} \mathrm{d}s + O\left(\frac{x \log^2 x}{T} \right), \tag{5.3}$$

where $x \notin \mathbb{N}$, $c = 1 + 1/\log x$, and $T \geq 2$. The strategy to derive a formula such as the one in (2.13) is to use the residue theorem to estimate this integral by the residues of the integrand, within a rectangle with vertices at

$$c - iT, \quad c + iT, \quad -U + iT, \quad -U - iT,$$

where U is a large non-integer value. We then 'add back' estimates for the integrals we over-counted, namely the integrals along the path $c - iT \to -U - iT \to -U + iT \to c + iT$. It is therefore clear how formulas such as (2.13) are derived: the over-counted integrals are essentially lower order than $x \log^2 x/T$, so can be ignored, which just leaves the residues of the integrand inside the rectangle of height T. We will not go into the details of this, but rather use a more refined method which leads more directly to the prime number theorem. Instead of extending the path of integration infinitely left, we only

46

extend it the length of the zero-free region of $L(s, \chi)$ past the line $\sigma = 1$. This has the effect of only picking up the residues of the integrand at either $s = 1$, or at an 'exceptional' real zero of an L-function of real character, as mentioned at the end of section 4.5. To avoid the problem of a pole of this type, we shall only consider $\psi(x)$ for now. We have the analogous formula

$$\psi(x) = \frac{1}{2\pi i} \int_{\alpha - iT}^{\alpha + iT} \left\{ -\frac{\zeta'(s)}{\zeta(s)} \right\} \frac{x^s}{s} \mathrm{d}s + O\left(\frac{x \log^2 x}{T} \right), \tag{5.4}$$

for $x \notin \mathbb{N}$, $T \geq 2$ and $\alpha = 1 + (\log x)^{-1}$. Set $\mu = 1 - c/\log T$ as in Theorem 4.5, and integrate anti-clockwise over the rectangle with vertices at $\mu \pm iT$, $\alpha \pm iT$. By virtue of Theorem 4.5, $\zeta(s)$ is never zero inside this rectangle, so the only residue we pick up is at its pole at $s = 1$, which is

$$x \operatorname{Res}_{s=1} \left\{ -\frac{\zeta'(s)}{\zeta(s)} \right\} = x.$$

Integrating on this closed contour, we 'over-count' the integrals over the path $\alpha - iT, \mu - iT, \mu + iT, \alpha + iT$. Therefore, by the residue theorem,

$$\psi(x) = x + O\left(\int_{\alpha \pm iT}^{\mu \pm iT} \left\{ -\frac{\zeta'(s)}{\zeta(s)} \right\} \frac{x^s}{s} \mathrm{d}s \right) + O\left(\int_{\mu - iT}^{\mu + iT} \left\{ -\frac{\zeta'(s)}{\zeta(s)} \right\} \frac{x^s}{s} \mathrm{d}s \right) + O\left(\frac{x \log^2 x}{T} \right).$$

It is now a matter of estimating these integrals in terms of T and x. The key to doing this effectively is to estimate $-\zeta'/\zeta$ on this region. Recall (4.22) from the previous chapter. A similar formula holds for $\zeta(s)$ in the rectangle $-1 < \sigma < 2$, as long as $s = \sigma + iT$ does not coincide with a zero of $\zeta(s)$, namely

$$\frac{\zeta'(s)}{\zeta(s)} = \sum_{\rho: \, |T - \gamma| < 1} \frac{1}{s - \rho} + O(\log T).$$

By the results of the previous chapter restricted to the case of $\zeta(s)$, there are $O(\log T)$ terms in the sum. Ensuring that T is chosen such that $|\gamma - T| \gg (\log T)^{-1}$ for each zero $\rho = \beta + i\gamma$, we have that each term is also $O(\log T)$. This gives the useful estimate

$$\frac{\zeta'(s)}{\zeta(s)} = O(\log^2 T), \tag{5.5}$$

for $-1 < \sigma < 2$. We first apply this to the integrals along the horizontal paths, giving

$$\int_{\alpha \pm iT}^{\mu \pm iT} \left\{ -\frac{\zeta'(s)}{\zeta(s)} \right\} \frac{x^s}{s} \mathrm{d}s \ll \frac{\log^2 T}{T} \left| \int_{\alpha}^{\mu} x^\sigma \mathrm{d}\sigma \right| \ll \frac{x^\alpha \log^2 T}{T} \ll \frac{x \log^2 T}{T}.$$

Here, the second inequality uses the fact that the length of the path of integration is certainly less than 1, while the integrand achieves its maximum value at $x^\alpha = O(x)$. For the vertical path, we have

$$\int_{\mu - iT}^{\mu + iT} \left\{ -\frac{\zeta'(s)}{\zeta(s)} \right\} \frac{x^s}{s} \mathrm{d}s \ll \log^2 T \int_{-T}^{T} \frac{x^\mu}{|\mu + it|} \mathrm{d}t \ll x^\mu \log^3 T.$$

Therefore, putting these estimates together gives

$$\psi(x) = x + O(\frac{x \log^2 x}{T}) + O(\frac{x \log^2 T}{T}) + O(x^\mu \log^3 T).$$

Ensuring that $T \leq x$ as we may allows us to combine the error terms to give

$$\psi(x) = x + O\left(x \log^3 x \{T^{-1} + x^{\mu-1}\}\right).$$

We now make a specific choice of T: namely $T = \exp(\sqrt{\log x}) < x$. Note that this choice of T gives

$$x^{\mu-1} = x^{-c/\log T} = \exp\left(-\frac{c \log x}{\log T}\right) = \exp\left(-c\sqrt{\log x}\right),$$

so that we have

$$x \log^3 x \{T^{-1} + x^{\mu-1}\} \ll x \exp\{-\min(1, c/2)\sqrt{\log x} + 3 \log \log x\} \ll x \exp\{-C\sqrt{\log x}\},$$

for some constant C. We have therefore proved the existence of an absolute constant C such that

$$\psi(x) = x + O\left(x \exp\{-C\sqrt{\log x}\}\right). \tag{5.6}$$

Now, recall Lemma 2.9 from Chapter 2 - the whole reason we aimed to estimate $\psi(x)$. This allows us to relate this estimate of $\psi(x)$ to $\pi(x)$, through

$$\pi(x) = \frac{\psi(x)}{\log x} + \int_2^x \frac{\psi(t)}{t \log^2 t} dt + O(x^{1/2}).$$

Define

$$\mathrm{Li}(x) = \int_2^x \frac{dt}{\log t},$$

and note that integration by parts gives

$$\int_2^x \frac{dt}{\log t} = \frac{x}{\log x} + \int_2^x \frac{dt}{\log^2 t} - \log 2,$$

so that substituting (5.6) into (5.7) gives

$$\pi(x) = \frac{x}{\log x} + \int_2^x \frac{dt}{\log^2 t} + O(x \exp\{-C\sqrt{\log x}\})$$
$$= \mathrm{Li}(x) + O(x \exp\{-C\sqrt{\log x}\}).$$

In conclusion, we have proved the following.

Theorem 5.2. *(The Prime Number Theorem) The number of primes less than a given value x, $\pi(x)$, is given by the asymptotic formula*

$$\pi(x) = \mathrm{Li}(x) + O(x \exp\{-C\sqrt{\log x}\}),$$

where C is an absolute constant. In particular,

$$\pi(x) \sim \mathrm{Li}(x).$$

48

We now wish to prove a similar formula as (5.6) for $\psi(x, \chi)$. The method used for $\zeta(s)$ will actually work for almost every function $L(s, \chi)$ of a given modulus - with the exception of at most one function of real character. The possible issue here, is that this L-function of real character may have a real zero arbitrarily close to $s = 1$, so when integrating on a similar rectangle to the case of $\zeta(s)$, we pick up an extra residue corresponding to this problematic zero. It is therefore of vital importance to bound the distance of such a zero away from $s = 1$, so that its contribution is essentially negligible compared to the residue x of the trivial L-function at $s = 1$. The problem of bounding this distance is one we can solve using Siegel's theorem.

5.2 Siegel's Theorem

In this section, our aim is to find a lower bound for $L(1, \chi)$ in terms of the modulus q, for real characters χ. In doing this, we immediately get a lower bound for the distance of any real zero away from $s = 1$ by the mean value theorem. Siegel's theorem is stated as follows.

Theorem 5.3. *For any $\varepsilon > 0$, there exists $C_1(\varepsilon)$ such that for real primitive characters χ of modulus q,*

$$L(1, \chi) > C_1(\varepsilon)q^{-\varepsilon}.$$

We follow the proof in [9], and consider the function

$$F(s) = \zeta(s)L(s, \chi_1)L(s, \chi_2)L(s, \chi_1\chi_2), \tag{5.7}$$

where χ_1 and χ_2 are real primitive characters of modulus q_1 and q_2 respectively. We shall assume that q_1 and q_2 are different throughout the proof. Furthermore, $\chi_1\chi_2$ is the real character modulo q_1q_2 formed by the product of χ_1 and χ_2. It is not necessarily a primitive character, however is certainly real and non-trivial. If it were trivial, this would imply that χ_1 and χ_2 induce the same character modulo q_1q_2, which contradicts the fact that their periodicities are by definition different. We know by the argument in Dirichlet's theorem that the three L-functions are non-zero at $s = 1$, so $F(s)$ is analytic everywhere except $s = 1$, where it has a residue of

$$\rho = L(1, \chi_1)L(1, \chi_2)L(1, \chi_1\chi_2).$$

Key to the proof of the theorem is the following lemma.

Lemma 5.4. *There exists a positive constant c such that*

$$F(s) > \frac{1}{2} - \frac{c\rho}{1 - s}(q_1q_2)^{8(1-s)} \tag{5.8}$$

for real $7/8 < s < 1$.

Proof. Multiplying the Euler products of the constituent factors of $F(s)$ gives

$$\sum_{n=1}^{\infty} a_n n^{-s} = F(s) = \prod_p g(s, p),$$

$$g(s, p) = \{(1 - p^{-s})(1 - \chi_1(p)p^{-s})(1 - \chi_2(p)p^{-s})(1 - \chi_1\chi_2(p)p^{-s})\}^{-1}.$$

First note that $a_1 = 1$. Now, differentiating the geometric series formula gives

$$(1 - x)^{-1} = \sum_{k=0}^{\infty} x^k,$$

$$(1 - x)^{-2} = \sum_{k=0}^{\infty} (k + 1)x^k,$$

$$(1 - x)^{-4} = \frac{1}{6} \sum_{k=0}^{\infty} (k + 1)(k + 2)(k + 3)x^k, \quad (|x| < 1).$$

Thus, for $\chi_1 = \chi_2 = 1$, we have $g(s, p) = (1 - p^{-s})^{-4} = \frac{1}{6} \sum_{k=0}^{\infty} (k+1)(k+2)(k+3)p^{-ks}$. If $\chi_1(p) = -1$ and $\chi_2(p) = \pm 1$ or vice versa, we have $g(s, p) = (1 - p^{-2s})^{-2} = \sum_{k=0}^{\infty} (k + 1)p^{-2ks}$. If one of the χ_i is zero, and the other is $1, -1$, or 0, we may write $g(s, p)$ in a similar way involving derivatives of the geometric series, all with positive terms. So, in all cases, $g(s, p) = \sum_{k=0}^{\infty} h(k, p)p^{-ms}$ where $h(k, p) \geq 0$. It follows that $a_n \geq 0$ for all n. Since $F(s)$ is analytic on $\sigma > 1$, we may use an identical argument to that in the proof of $L(1, \chi) \neq 0$ to write $F(s) = \sum_{m=0}^{\infty} b_m(2 - s)^m$ for $|s - 2| < 1$, with all $b_m \geq 0$. Moreover, $b_0 = F(2) \geq a_0 = 1$. Also, we have

$$\frac{\rho}{s - 1} = \frac{\rho}{1 - (2 - s)} = \rho \sum_{m=0}^{\infty} (2 - s)^m, \quad |2 - s| < 1.$$

It therefore follows that

$$F(s) - \frac{\rho}{s - 1} = \sum_{m=0}^{\infty} (b_m - \rho)(2 - s)^m, \quad |2 - s| < 1.$$

Since the left hand side effectively removes the pole of $F(s)$ at $s = 1$ of residue ρ, the left hand side is therefore analytic on the larger region $\sigma > 0$, so the above formula is also valid for $|2 - s| < 2$. Now consider the circle $|2 - s| = 3/2$. Firstly, $\zeta(s)$ is bounded for these values, and from the continuation of $L(s, \chi)$ to the right half-plane, we have

$$|L(s, \chi)| \leq q|s| \int_1^{\infty} x^{-\sigma-1} < 2q|s|, \quad (\sigma > 0).$$

So, we have the bounds

$$|L(s, \chi_1)| < c_1 q_1, \quad |L(s, \chi_2)| < c_1 q_2, \quad |L(s, \chi_1\chi_2)| < c_1 q_1 q_2.$$

50

For an absolute constant c_1. It follows that

$$|F(s)| < c_2 q_1^2 q_2^2, \qquad |2 - s| = \frac{3}{2}.$$

The ratio $\rho/(s-1)$ also satisfies this upper bound (assuming a good choice of c_2), since ρ is the product of three L-functions satisfying the upper bounds above, while $1/(s-1)$ is bounded on the circle $|s - 2| = 3/2$. Now, we estimate Cauchy's form for the Taylor coefficients $(b_m - \rho)$, giving

$$|b_m - \rho| = \left| \frac{1}{2\pi i} \oint_\gamma \frac{F(s) - \rho/(s-1)}{(s-2)^{m+1}} ds \right| \leq \left(\frac{3}{2} \right)^{-m-1} \sup_\gamma \left| F(s) - \frac{\rho}{s-1} \right| < 2c_2 q_1^2 q_2^2 \left(\frac{2}{3} \right)^m,$$

where γ denotes the circle $|2 - s| = 3/2$. Thus, for $7/8 < s < 1$, we have

$$\sum_{m=M}^\infty |b_m - \rho|(2-s)^m \leq \sum_{m=M}^\infty 2c_2 q_1^2 q_2^2 \left[\frac{2}{3}(2-s) \right]^m$$

$$\leq 2c_2 q_1^2 q_2^2 \sum_{m=M}^\infty \left(\frac{3}{4} \right)^m < c_3 q_1^2 q_2^2 \left(\frac{3}{4} \right)^M < c_3 q_1^2 q_2^2 e^{-M/4},$$

using the fact that $e^{-1/4} \approx 0.77 > 3/4$. By the fact that $b_0 \geq 1$ and $b_m \geq 1$, we get

$$F(s) - \frac{\rho}{s-1} > 1 - \rho \sum_{m=0}^{M-1} (2-s)^m - c_3 q_1^2 q_2^2 e^{-M/4} = 1 - \rho \frac{(2-s)^M - 1}{1 - s} - c_3 q_1^2 q_2^2 e^{-M/4}.$$

We choose M to satisfy $\frac{1}{2} e^{-1/4} < c_3 q_1^2 q_2^2 e^{-M/4} < \frac{1}{2}$. The upper bound implies that

$$F(s) > \frac{1}{2} - \frac{\rho}{1-s}(2-s)^M,$$

while taking the logarithm of the lower bound gives

$$M < 8 \log q_1 q_2 + c_4.$$

Therefore,

$$(2-s)^M = \exp[M \log(1 + 1 - s)] < \exp[M(1-s)] < c_5 (q_1 q_2)^{8(1-s)},$$

and it follows that there exists a constant c such that

$$F(s) > \frac{1}{2} - \frac{c\rho}{1-s}(q_1 q_2)^{8(1-s)}.$$

$\square$

Siegel's theorem will now follow quickly from this lemma. For a given $\varepsilon > 0$, we split into two cases: the first is that there exists a real primitive character χ such that $L(s, \chi)$ has a real zero β_1 in the open interval $(\frac{\varepsilon}{16}, 1)$; the second is that no such character exists. If the former holds, we choose χ_1 to be such a character, and it follows that $F(\beta_1) = 0$ since $F(s)$ is analytic in the given open interval. If the latter holds, we choose χ_1 to be any real primitive character, and β_1 any value in $(\frac{\varepsilon}{16}, 1)$. We claim that $F(\beta_1) < 0$ in this case. Indeed, from the continuation of $\zeta(s)$ to the right half-plane in (3.3), we have that $\zeta(\beta_1) < 0$. The three L-functions are positive for values greater than 1 by the product formula, and are non-zero at $s = 1$ by the argument in Dirichlet's theorem. Moreover, they do not vanish in the interval $(\frac{\varepsilon}{16}, 1)$ by assumption, so must be positive in this interval. Thus, $F(\beta_1) < 0$ as claimed. In either case, $F(\beta_1) \leq 0$, so the lemma implies that (for $\varepsilon < 1/2$)

$$c\rho > \frac{1}{2}(1 - \beta_1)(q_1 q_2)^{-8(1-\beta_1)}.$$

The quantity ρ can be bounded above using a useful upper bound for $L(s, \chi)$ for real values $1 - \frac{1}{\log q} \leq s \leq 1$. There is an analogous bound for $L'(s, \chi)$ which we use later, and they are given by

$$|L(s, \chi)| \leq c_6 \log q, \tag{5.9}$$

$$\left|L'(s, \chi)\right| \leq c_7 \log^2 q, \tag{5.10}$$

where χ is a real primitive character modulo $q \geq 3$. We prove the latter first. Since $L'(s, \chi)$ is analytic on $\sigma > 0$, we start from the formula

$$L'(s, \chi) = -\sum_{n=1}^{\infty} \chi(n)(\log n)n^{-s}, \quad (\sigma > 0).$$

Note that when $1 - \frac{1}{\log q} \leq s \leq 1$, and $n \leq q$, we have

$$n^{-s} = e^{-s \log n} \leq e^{-(1-1/\log q)\log n} \leq e^{1-\log n} = en^{-1},$$

so that

$$\left|\sum_{n=1}^{q} \chi(n)(\log n)n^{-s}\right| \leq e \sum_{n=1}^{q}(\log n)n^{-1} \leq c_7 \log q \int_1^q \frac{\mathrm{d}t}{t} \leq c_7 \log^2 q.$$

When $n > q$, $(\log n)n^{-s}$ is a decreasing function in n, so

$$\left|\sum_{n=q+1}^{\infty} \chi(n)(\log n)n^{-s}\right| \leq (\log q)q^{-s} \max_N \left|\sum_{n=q+1}^{N} \chi(n)\right| \leq (\log q)eq^{-1}q \leq c_7 \log q,$$

where the penultimate step here uses that any sum over $\chi(n)$ is at most q, since a sum over every residue modulo q is zero. Thus, the conclusion follows that

$$\left|L'(s, \chi)\right| \leq c_7 \log^2 q.$$

For the bound on $L(s, \chi)$, the argument is identical, except the original summation formula is without the factor of $\log n$, which in turn removes a factor of $\log q$ in the upper bound. Now, (5.9) implies

$$\rho < (c_6 \log q_1) L(1, \chi_2)(c_6 \log q_1 q_2). \tag{5.11}$$

We now take q_1 and χ_1 to be fixed, so that terms involving q_1 may be absorbed into any constants - adding a dependency upon the choice of χ_1. We let χ_2 be any real primitive character of modulus $q_2 > q_1$. Thus, inserting (5.11) into (5.9), we have

$$L(1, \chi_2) > C q_2^{-8(1-\beta_1)}(\log q_2)^{-1},$$

where C is a constant depending only on the choice of χ_1, and hence only on ε - essentially, if ε is small enough, we are forced into picking χ_1 to be a real primitive character whose L-function has a zero in $(\frac{\varepsilon}{16}, 1)$. Since $8(1 - \beta_1) < \frac{\varepsilon}{2}$, we have

$$L(1, \chi_2) > C(\varepsilon) q_2^{-\varepsilon/2}(\log q_2)^{-1} > C(\varepsilon) q_2^{-\varepsilon},$$

provided q_2 is large enough, as we may suppose. This gives Siegel's theorem. An important corollary for our uses is the following.

Corollary 5.5. *Let χ be a real primitive character modulo q. Then for any $\varepsilon > 0$, there exists a positive $C(\varepsilon)$ such that any real zero β_1 of $L(s, \chi)$ satisfies*

$$\beta_1 < 1 - C(\varepsilon) q^{-\varepsilon}.$$

Proof. Let $\varepsilon > 0$ and χ a real primitive character modulo q be given. We assume for a contradiction that there exists a real zero $\beta > 1 - C(\varepsilon) q^{-\varepsilon}$, where $C(\varepsilon)$ is an absolute constant. The bound for $L'(s, \chi)$ in (5.10) gives

$$\left| L'(s, \chi) \right| < c_7 \log^2 q.$$

The mean value theorem implies there exists an $\alpha \in (\beta, 1)$ such that

$$(1 - \beta) L'(\alpha, \chi) = L(1, \chi) - L(\beta, \chi),$$

so

$$L(1, \chi) = L(1, \chi) - L(\beta, \chi) = (1 - \beta) L'(\alpha, \chi) \leq C'(\varepsilon)(\log^2 q) q^{-\varepsilon}.$$

Repeating this argument with $\varepsilon/2$, say, makes the contribution of $\log^2 q$ negligible, and contradicts Siegel's theorem. $\qquad \square$

Thus, we have now a bound in terms of q of any real zero away from $s = 1$, and hence a complete zero-free region for L-functions. We may now use a modified argument as in the prime number theorem, and prove a similar result for arithmetic progressions.

5.3 The Prime Number Theorem for Arithmetic Progressions

We use the equation (5.3) as a starting point once again, but this time for general L-functions of modulus $q \geq 3$. Let $T \geq 2$ be a fixed value, and let $\mu = 1 - c/\log qT$, such that L-functions of all characters, with the exception of at most one real character, have no zeros in the region $\sigma \geq \mu$. We denote the exceptional real character by χ_1 (if it exists), and such an exceptional real zero as β_1. For all characters χ that are not χ_1 or χ_0, the integrand in (5.3) has no residues inside the rectangle with vertices at $\mu \pm iT$, $\alpha \pm iT$, where α is once again $1 + (\log x)^{-1}$. Therefore, for χ not equal to χ_1 or χ_0:

$$\psi(x,\chi) = O\left(\int_{\alpha \pm iT}^{\mu \pm iT} \left\{-\frac{L'(s,\chi)}{L(s,\chi)}\right\} \frac{x^s}{s} \mathrm{d}s\right) + O\left(\int_{\mu - iT}^{\mu + iT} \left\{-\frac{L'(s,\chi)}{L(s,\chi)}\right\} \frac{x^s}{s} \mathrm{d}s\right) + O\left(\frac{x\log^2 x}{T}\right).$$

For $\chi = \chi_1$, we pick up a residue of $-x^{\beta_1}/\beta_1$, corresponding to the real zero of $L(s,\chi)$ which lies inside the rectangle given. It then remains to estimate the 'over-counted' integrals as in the proof of the prime number theorem. The arguments are completely analogous, with the exception of the slightly different estimate for L'/L, given for $-1 < \sigma < 2$ as

$$-\frac{L'(s,\chi)}{L(s,\chi)} = O(\log^2 qT),$$

which is deduced from (4.22) in the same way as the estimate for ζ'/ζ. Note that this only applies for primitive χ. Thus, we have the analogous result to previously (valid only for primitive χ) that

$$\psi(x,\chi) = O(\frac{x\log^2 x}{T}) + O(\frac{x\log^2 qT}{T}) + O(x^\mu \log^3 qT), \qquad (5.12)$$

with the addition of a residue of $-x^{\beta_1}/\beta_1$ if $\chi = \chi_1$. Now, this will also hold for non-primitive χ. Indeed, if χ is induced by the character χ', we have

$$\left|\psi(x,\chi) - \psi(x,\chi')\right| \ll \sum_{\substack{n \leq x \\ (n,q) > 1}} \Lambda(n) = \sum_{p|q} \sum_{k:\, p^k \leq x} \log p \ll (\log x) \sum_{p|q} \log p \ll (\log x)(\log q),$$

which is negligible in comparison to the error in (5.12). Note that this also holds for the error in estimating $\psi(x,\chi_0)$ by $\psi(x)$, since χ_0 is induced by the trivial character of modulus 1. Therefore, (5.12) also holds for $\chi = \chi_0$, with the addition of the x term in the formula for $\psi(x)$ (from the derivation of the prime number theorem). We now ensure that q is negligible compared to the effect of T, by ensuring $q \leq \exp(C\sqrt{\log x})$ for some positive constant C, then setting $T = \exp(C\sqrt{\log x})$. It then follows in exactly the same manner as before that

$$\psi(x,\chi) = \left\{\begin{array}{ll} x & \text{if } \chi = \chi_0, \\ -\frac{x^{\beta_1}}{\beta_1} & \text{if } \chi = \chi_1, \\ 0 & \text{otherwise.} \end{array}\right\} + O(x\exp\{-C'\sqrt{\log x}\}),$$

where C' is an absolute constant depending only on C. Now, recall the relation in (2.15), that for a coprime to q,

$$\psi(x; q, a) = \frac{1}{\phi(q)} \sum_\chi \overline{\chi}(a) \psi(x, \chi).$$

Substituting the asymptotic formula for $\psi(x, \chi)$, we therefore have

$$\psi(x; q, a) = \frac{x}{\phi(q)} - \frac{\overline{\chi}_1(a) x^{\beta_1}}{\phi(q)\beta_1} + O(x \exp\{-C'\sqrt{\log x}\}).$$

We now invoke the corollary to Siegel's theorem on the zero β_1, in order to absorb it into the error term. For a given $\varepsilon > 0$, we have a constant $C(\varepsilon)$ such that

$$\beta_1 < 1 - C(\varepsilon)q^{-\varepsilon}.$$

This gives us that

$$x^{\beta_1} < x x^{-C(\varepsilon)q^{-\varepsilon}} = x \exp\{-C(\varepsilon)q^{-\varepsilon}(\log x)\}. \tag{5.13}$$

Now, choose a large value of N such that

$$q \le (\log x)^N, \tag{5.14}$$

and take $\varepsilon = (2N)^{-1}$. Then, $-q^{-\varepsilon} \le -(\log x)^{-1/2}$, which substituted into (5.13) gives

$$x^{\beta_1} \ll x \exp\{-C(N)\sqrt{\log x}\},$$

where C is a constant depending only on N. We therefore may absorb the term involving x^β into an error term which depends only on N, giving the key result

$$\psi(x; q, a) = \frac{x}{\phi(q)} + O(x \exp\{-C'(N)\sqrt{\log x}\}). \tag{5.15}$$

As in the case of $\psi(x)$, we have the relation

$$\pi(x; q, a) = \frac{\psi(x; q, a)}{\log x} + \int_2^x \frac{\psi(t; q, a)}{t \log^2 t} dt + O(x^{1/2}),$$

which upon substituting (5.15) gives us the following theorem.

Theorem 5.6. *(The Prime Number Theorem for Arithmetic Progressions) If a is coprime to a given modulus q, then subject to the condition (5.14), we have*

$$\pi(x; q, a) = \frac{\mathrm{Li}(x)}{\phi(q)} + O(x \exp\{-C(N)\sqrt{\log x}\}),$$

where C is an absolute constant depending only on N. In particular,

$$\pi(x; q, a) \sim \frac{\mathrm{Li}(x)}{\phi(q)}.$$

We have therefore completed the main goal of this report, characterising the 'global' behaviour of prime numbers in arithmetic progressions in a concrete way. An important thing to notice is that no distinction is made between different residue classes coprime to q: asymptotically, they all grow at the same rate, suggesting that the set of all primes is effectively evenly distributed among arithmetic progressions. However, sometimes asymptotic formulas do not tell the whole story and are merely an indication of the global behaviour of these progressions. Indeed, despite the 'even-splitting' nature among arithmetic progressions, we will see that further analysis shows a certain bias of primes towards particular progressions over others.

5.4 Further Topics in Analytic Number Theory

Recall the observation in the very first chapter that the number of digits in the error approximating $\pi(x)$ by $\mathrm{Li}(x)$ was approximately half that of x. However, the error term we proved in the prime number theorem is not quite of the order $x^{1/2}$. So, is it possible to improve this? Note that the error term effectively corresponds to how good the zero-free region of $\zeta(s)$ is. With more advanced methods, introduced by Vinogradov and Korobov [7, Chapter 6], one can prove a slightly sharper zero-free region,

$$\sigma \geq 1 - C(\log t)^{-2/3}(\log \log t)^{-1/3}$$

for an absolute constant C. Using this zero-free region, one can prove as in [7, Chapter 12] that

$$\pi(x) = \mathrm{Li}(x) + O\left(x \exp\{-C(\log x)^{3/5}(\log \log x)^{-1/5}\}\right).$$

Now, it is of vital importance to mention the Riemann Hypothesis (RH) - one of the most significant open problems in mathematics. In his 1859 memoir, Riemann conjectured that every non-trivial zero of $\zeta(s)$ lies on the 'critical line', $\sigma = 1/2$. Assuming this gives essentially the strongest possible version of the prime number theorem, which we show using the method from [4, p. 113]. Starting with the explicit formula (2.13), for $T \geq 2$,

$$\psi(x) = x - \sum_{|\rho| < T} \frac{x^\rho}{\rho} + O\left(\frac{x \log^2 x}{T}\right).$$

Then, since necessarily $|x^\rho| = x^{1/2}$, and $\sum_{|\rho| < T} \frac{1}{\rho} = O(\log^2 T)$, we set $T = x^{1/2}$ to give

$$\psi(x) = x + O(x^{1/2} \log^2 x),$$
$$\pi(x) = \mathrm{Li}(x) + O(x^{1/2} \log^2 x).$$

Note that this coincides with our observation that the error was approximately of order $x^{1/2}$! Therefore, a proof of the Riemann Hypothesis would give us the ability to give highly accurate estimates of how many primes there are less than a given value. As of yet, sadly, no such proof has been found despite the best efforts of some of the very best mathematicians. There is a much stronger analogous hypothesis in the case of Dirichlet

L-functions, namely that all L-functions also have all their non-trivial zeros on the critical line - known as the Generalised Riemann Hypothesis (GRH). This then gives the same error estimate for $\pi(x; q, a)$ under GRH as for $\pi(x)$ under RH (only subject to the condition that $q \leq x^{1/2}$!) [4, p. 124].

We now return to the 'bias' of primes to certain residue classes which was alluded to at the end of the previous section, and use data and results from [10]. Our aim is to pit the different arithmetic progressions of a given modulus against each other in a 'prime number race'. As a first example, consider the primes mod 4. With the exception of 2, primes are either congruent to 1 or 3 modulo 4, and by the prime number theorem for arithmetic progressions, we have

$$\pi(x; 4, 1) \sim \pi(x; 4, 3) \sim \frac{1}{2}\text{Li}(x).$$

Now, let us consider the empirical evidence.

x	$\pi(x; 4, 3)$	$\pi(x; 4, 1)$
100	13	11
200	24	21
300	32	29
400	40	37
500	50	44
1000	87	80

Table 5.1: A comparison of $\pi(x; 4, 3)$ and $\pi(x; 4, 1)$

Despite growing at the same rate, $\pi(x; 4, 3)$ 'stays ahead in the race' against $\pi(x; 4, 1)$. Or, more formally $\pi(x; 4, 3) - \pi(x; 4, 1)$ stays positive for every value of x shown. Is this mere coincidence? Let us consider $\pi(x; 4, 3) - \pi(x; 4, 1)$ plotted over many more values of x in the figure below[1]. Remarkably, $\pi(x; 4, 3)$ stays ahead for all but one of the first $100,000$ values.
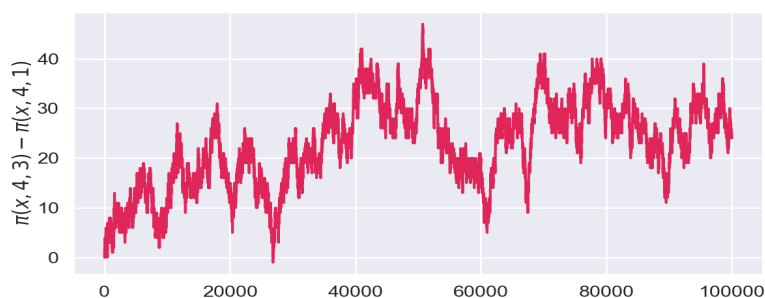

Figure 5.1: The difference $\pi(x; 4, 3) - \pi(x; 4, 1)$ for $0 \leq x \leq 10^5$

[1]Plot made using Python 3.7.4 and the library Seaborn 0.10.0.

This bias in favour of the form $4n + 3$ among primes is known as Chebyshev's bias, who first remarked upon it. However, $\pi(x; 4, 3)$ does not always stay ahead in this race. In fact, Littlewood (1914) showed that $\pi(x; 4, 3) - \pi(x; 4, 1)$ changes sign infinitely often. Despite this, the first time that $\pi(x; 4, 1)$ takes the lead is at $x = 26,861$, before immediately losing it at $x = 26,863$! The case of modulus 4 is by no means an isolated case either. Indeed, consider the mod 10 race (the last digit of the primes). Numbers coprime to 10 are 1, 3, 7 and 9, so consider the growth of the number of primes in these arithmetic progressions.

x	$\pi(x; 10, 1)$	$\pi(x; 10, 3)$	$\pi(x; 10, 7)$	$\pi(x; 10, 9)$
100	5	7	6	5
200	10	12	12	10
500	22	24	24	23
1000	40	42	46	38
2000	73	78	77	73
5000	163	172	169	163
10,000	306	310	308	303
20,000	563	569	569	559
50,000	1274	1290	1288	1279
100,000	2387	2402	2411	2390

Table 5.2: The mod 10 prime race for values up to 10^5

The pattern here is perhaps less easy to see, but after some examination we spot that the two middle columns seem to remain slightly ahead - there is essentially a small bias towards primes with a last digit of 3 and 7 compared to 1 and 9. So, why do these biases occur? Another prime race helps to reveal the pattern, namely the mod 8 race.

x	$\pi(x; 8, 1)$	$\pi(x; 8, 3)$	$\pi(x; 8, 5)$	$\pi(x; 8, 7)$
1,000	37	44	43	43
5,000	161	168	168	171
10,000	295	311	314	308
50,000	1,257	1,295	1,292	1,288
100,000	2,384	2,409	2,399	2,399
1,000,000	19,552	19,653	19,623	19,669

Table 5.3: The mod 8 race for values up to 10^6

There is no clear winner in this race, but there is a clear loser - namely primes of the form $8n + 1$. So far, the losers in the races have been the following.

- 1 mod 3,

- 1 and 9 mod 10,

- 1 mod 8.

Now, these residue classes are precisely the quadratic residues that lie in the multiplicative group of each modulus. This gives us a heuristic idea for why these biases occur: numbers in the progressions corresponding to quadratic residues are somehow 'less likely' to be prime, as *all* the squares of primes lie in these residue classes. So, what results have been proved (or disproved) relating to this phenomenon? These biases were studied comprehensively in a paper of Rubinstein and Sarnak (1994) [11]. A big conjecture in this area was that of Knapowski and Turán (1962) [10, p. 3]:

Conjecture 5.7. *As $X \to \infty$,*

$$\frac{1}{X} \#\{x \leq X \ s.t. \ \pi(x; 4, 3) - \pi(x; 4, 1) > 0\} \to 1.$$

In essence, this states that Chebyshev's bias holds for almost all values of x. Although empirical evidence suggests this conjecture may be true, it was disproved by Sarnak and Knapowski independently (1994 and 1993 respectively). However, Rubinstein and Sarnak had the idea of counting $1/x$ for each $x \leq X$, rather than simply 1. Then, since $\sum_{x \leq X} x^{-1} \approx \log X$, we normalise by a factor of $\log X$ rather than X. Their inspired idea had the following result [10, p. 18]:

$$\frac{1}{\log X} \sum_{\substack{x \leq X \\ \pi(x;4,3) - \pi(x;4,1) > 0}} \frac{1}{x} \to 0.9959\ldots \text{ as } X \to \infty.$$

To count the bias in this way is termed the 'logarithmic measure'. Thus, instead of Chebyshev's bias holding almost always, it actually holds approximately 99.59% of the time! The way this result was derived relies heavily on results we proved regarding Dirichlet L-functions. The caveat to this result, is that it relies on two (as of yet) unproven conjectures. The first is GRH, while the other is known as the Grand Simplicity Hypothesis, which states that all the zeros on the critical line are linearly independent. They also argued in greater generality than just the case of $q = 4$. Indeed, they proved that if a, b are coprime to q, with a a square mod q and b not square, then a bias necessarily exists towards b, but holds less than 100% of the time. More precisely, there exists $1/2 < c < 1$ such that

$$\frac{1}{\log X} \sum_{\substack{x \leq X \\ \pi(x;q,b) - \pi(x;q,a) > 0}} \frac{1}{x} \to c \text{ as } X \to \infty.$$

Moreover, they proved that if a and b are either both squares or both non-squares mod q, then no such bias exists, and this limit is exactly $1/2$. In a nutshell, there is a bias if and only if exactly one of a and b is a quadratic residue. These biases may also be studied in greater generality: more recent work by Lemke Oliver and Soundararajan (2016) [12] studied biases in *sequences* of primes in arithmetic progressions. They observe that the size of the biases between different tuples of reduced residue classes can be even

more vast. As a basic example, consider the last digits of all pairs of consecutive prime numbers (the case of modulus 10). We define

$$\pi(x; 10, (a, b)) = \#\{p_n \leq x \text{ s.t. } p_n \equiv a \ (10) \text{ and } p_{n+1} \equiv b \ (10)\}.$$

Then, if x_0 is such that $\pi(x_0) = 10^8$, we have the following data for $\pi(x_0; 10, (a, b))$.

	$a = 1$	$a = 3$	$a = 7$	$a = 9$
$b = 1$	4,623,042	6,010,382	6,373,981	7,991,431
$b = 3$	7,429,438	4,442,562	6,755,195	6,372,941
$b = 7$	7,504,612	7,043,695	4,439,355	6,012,739
$b = 9$	5,442,345	7,502,896	7,431,870	4,622,916

Table 5.4: The value of $\pi(x_0; 10, (a, b))$ for different values of a and b, where $\pi(x_0) = 10^8$

Notice the fact that the diagonals are vastly different from the other numbers in the table: essentially, if you pick a prime at random less than x_0, the chance that the next prime will have the same digit is actually much smaller than 25 percent! This phenomenon was explained only heuristically in [12], and relies on the Hardy-Littlewood k-tuples conjecture, which is essentially a result of the type of the prime number theorem but for arbitrary sequences of primes. To give an idea of the strength of this conjecture, one special case is a strengthening of the (as of yet unproven) twin prime conjecture, which states there are infinitely many primes that are 2 apart. The study of these biases in sequences of primes less than a given value therefore 'blur the lines' between problems of global and local type - a broad categorisation of problems relating to the primes which we discussed earlier. We shall not go into the details of the results of [12], as they are beyond the level of this report. In summary, we have now explored the data relating to primes in arithmetic progressions, and given even more detail regarding their behaviour in specific residue classes. This is in stark contrast to the result of the prime number theorem for arithmetic progressions, which makes no distinction between different residue classes of a given modulus.

Chapter 6

Conclusions

To summarise, we have studied the frequency of prime numbers in detail, particularly those in arithmetic progressions. We began by using elementary properties of Dirichlet characters and L-functions around $s = 1$ to prove Dirichlet's theorem, the fact that the arithmetic progression $qn + a$, with a and q coprime contains infinitely many primes.

Following this, we found the analytic continuation of L-functions to the whole complex plane, using methods discovered by Riemann. Using this analytic continuation, we wrote the logarithmic derivative of L-functions as infinite sums in terms of their nontrivial zeros, which exploited the Hadamard product formula. In addition to elementary bounds on L-functions, this yielded information on the position of zeros in the critical strip, as well as their frequency. Armed with this information, it was then possible to use a clever contour integration trick on the logarithmic derivative of L-functions, which related the non-trivial zeros of L-functions to the function $\psi(x, \chi)$ through Euler's product formula.

We first followed this procedure in the special case of $\zeta(s)$, which gave us an asymptotic formula for $\psi(x)$ and hence for $\pi(x)$ - the number of primes less than a given value - completing our proof of the prime number theorem. We proceeded to the more general case of L-functions, which gave an asymptotic formula for $\psi(x, \chi)$ with the caveat of a term involving a problematic real zero of a particular L-function. Thus, we dealt with the problem of possible real zeros near $s = 1$ through Siegel's theorem, which allowed us to complete the formula for $\psi(x, \chi)$. This then gave us the prime number theorem for arithmetic progressions: the number of primes less than a given value x in the arithmetic progression $qn + a$ is asymptotically $\mathrm{Li}(x)/\phi(q)$ (as long as q and a are coprime).

Following the proof of this significant result, it was important to note that it didn't tell us the whole story. Firstly, it didn't give us the best possible error term, and we showed that the assumption of GRH gives us a much stronger result. We also noted that this formula was only asymptotic: its more local behaviour revealed certain biases towards certain arithmetic progressions, namely those where a was non-square modulo

q. Furthermore, we noticed how these biases become even more wild when considering sequences of consecutive primes.

The main takeaway from this report is that although the primes seem well-behaved in the long run, if you zoom in enough to smaller scales, the behaviour becomes more and more bizarre. Let us return to Quote 1.1: was Euler right in what he said? Perhaps we have found some order on a large scale, but we have seen that delving further and further into the primes only adds more and more mystery; whether this mystery will ever be solved is uncertain. Consequently, there is no doubt that the study of primes will be at the forefront of mathematical research for many years to come.

Appendix

1 Tests for Convergence

The following convergence tests show uniform convergence of functions defined either by an infinite series or improper integral.

Theorem A1. *(Weierstrass M-test) Let $f(z) = \sum_n f_n(z)$ be a series of complex valued functions on a domain $D \subset \mathbb{C}$. If $|f_n(z)| \leq M_n$ for all $z \in D$ and a series of constants M_n with $\sum_n M_n < \infty$, then $f(z)$ is uniformly convergent on D.*

Theorem A2. *(M-test for integrals) Suppose $f(z,t)$ is integrable on $[a, \infty)$, and for each $z \in D \subset \mathbb{C}$. Suppose $|f(z,t)| \leq M(t)$ for all $t \in [a, \infty)$, $z \in D$. Then if*

$$\int_a^\infty M(t)\mathrm{d}t < \infty,$$

$F(z) = \int_a^\infty f(z,t)\mathrm{d}t$ is absolutely and uniformly convergent on D.

The former result is well-known from any introductory course in complex analysis. The latter is perhaps less familiar, and is adapted from the result in [13]. Both shall be used extensively. A key consequence of uniform convergence is the following.

Proposition A3. *If $f(z)$ is defined by a uniformly convergent series or improper integral, where the summands or integrand is holomorphic on a domain $D \subset \mathbb{C}$, then f is holomorphic on D.*

Proof. We consider the case where $f(z) = \sum_n f_n(z)$, with each f_n holomorphic on some domain D. Choose any closed contour $\gamma \subset D$. Then by Cauchy's theorem, we have $\oint_\gamma f_n(z)\mathrm{d}z = 0$ for all n. Then by uniform convergence,

$$\oint_\gamma f(z)\mathrm{d}z = \sum_n \oint_\gamma f_n(z)\mathrm{d}z = 0,$$

so by Morera's theorem, $f(z)$ is holomorphic on D. The case where f is defined by an improper integral is analogous, where we change the order of integration instead of interchanging summation and integration (all justified by uniform convergence). $\qquad\square$

In a nutshell: uniform limits of holomorphic functions are holomorphic.

2 Euler's Gamma Function

For a complex variable s with $\sigma > 0$, Euler's Gamma function is defined as

$$\Gamma(s) = \int_0^\infty x^{s-1}e^{-x}\mathrm{d}x. \tag{A1}$$

Results in the section on convergence tests imply that the integral converges absolutely and uniformly to a holomorphic function on this region, since

$$|\Gamma(s)| \leq \int_0^1 x^{\sigma-1}\mathrm{d}x + C\int_1^\infty e^{-x/2}\mathrm{d}x < \infty,$$

where C is some absolute constant. Moreover, integration by parts gives the key relation

$$s\Gamma(s) = \Gamma(s+1). \tag{A2}$$

Since $\Gamma(1) = 1$ (just by computing the integral), it follows by induction that $\Gamma(n) = (n-1)!$ for positive integers n. In this way, the Gamma function is a generalisation of the factorial function. We may also use this relation repeatedly to give the analytic continuation of $\Gamma(s)$. Suppose $\Re(s) \leq 0$, and s is not an integer. Then there exists n such that $\Re(s+n) > 0$, and by induction of (A2) we have

$$\Gamma(s) = \frac{\Gamma(s+n)}{s(s+1)\dots(s+n-1)}.$$

Since s is not a non-positive integer, the right hand side is analytic, hence so is the left side. By (A2), we have that Gamma has a simple pole of residue 1 at $s = 0$, and it follows that it also has poles at all negative integers n, each with residue $(-1)^n/n!$. We now state a useful formula involving the Gamma function, and provide a sketch of the proof for brevity.

Theorem A4. *(The Reflection Formula) For $s \in \mathbb{C} \setminus \mathbb{Z}$,*

$$\Gamma(s)\Gamma(1-s) = \frac{\pi}{\sin \pi s}.$$

Proof. There is an equivalent definition of $\Gamma(s)$ introduced by Weierstrass [14, Theorem 1.1.2]. We have

$$\Gamma(s) = s^{-1}e^{-\gamma s}\prod_{n=1}^{\infty}(1+s/n)^{-1}e^{s/n}, \tag{A3}$$

where γ is the Euler-Mascheroni constant, defined as

$$\gamma = \lim_{n\to\infty}\left(\sum_{1}^{n}\frac{1}{k} - \log n\right).$$

Since $\sin(s)/s$ is an entire function of order 1 (since $\sin(s) = O(e^{|s|})$) with zeros at non-zero integer multiples of π, we may use Theorem 4.4 to write it as

$$\frac{\sin(s)}{s} = e^{A+Bs}\prod_{n\neq 0}\left(1 - \frac{s}{n\pi}\right)e^{s/n\pi} = e^{A+Bs}\prod_{n=1}^{\infty}\left(1 - \frac{s^2}{n^2\pi^2}\right),$$

combining each positive and negative term. Since the left hand side tends to 1 as $s \to 0$, it follows that $A = 0$. Moreover, since $\sin(s)/s$ is an even function, we also must have $B = 0$ so that the right hand side is even. We conclude that

$$\sin(\pi s) = \pi s\prod_{n=1}^{\infty}\left(1 - \frac{s^2}{n^2}\right).$$

65

Therefore, using (A3) we have

$$\frac{1}{\Gamma(s)\Gamma(-s)} = -s^2 \prod_{n=1}^{\infty} \left(1 - \frac{s^2}{n^2}\right) = -s \, \frac{\sin(\pi s)}{\pi}.$$

Thus,

$$\Gamma(s)\Gamma(1-s) = -s\Gamma(s)\Gamma(-s) = \frac{\pi}{\sin(\pi s)}.$$

$\square$

It follows from this formula that $\Gamma(s)$ is never zero. The right hand side is never zero, so any zero of $\Gamma(s)$ must coincide with a pole of $\Gamma(1-s)$. However, any such pole would be at a non-positive integer, which implies any zero would also have to be an integer. Γ is non-zero at the integers, so is therefore non-zero everywhere. Finally, we have Stirling's approximation to the Gamma function, as given in [4, Chapter 10].

Theorem A5. *(Stirling's Formula) The formula*

$$\log \Gamma(s) = \left(s - \frac{1}{2}\right) \log s - s + \frac{1}{2} \log 2\pi + O(|s|^{-1})$$

is valid as $|s| \to \infty$, where $-\pi + \delta < \arg s < \pi - \delta$ for some fixed $\delta > 0$.

3 Gauss Sums

We follow [4, Chapter 9]. Consider a character χ to modulus q. The *Gauss sum* of χ is then defined as

$$\tau(\chi) := \sum_{m=1}^{q} \chi(m)e^{2\pi i m/q}. \tag{A4}$$

The following Lemma is an important property of Gauss sums.

Lemma A6. *For all $n \in \mathbb{N}$ and for all **primitive** characters χ,*

$$\chi(n)\tau(\overline{\chi}) = \sum_{h=1}^{q} \overline{\chi}(h)e^{2\pi i n h/q} \tag{A5}$$

Proof. For n coprime to q, this property is clear:

$$\chi(n)\tau(\overline{\chi}) = \sum_{m=1}^{q} \overline{\chi}(m)\chi(n)e^{2\pi i m/q} = \sum_{m=1}^{q} \overline{\chi}(mn^{-1})e^{2\pi i m/q} = \sum_{h=1}^{q} \overline{\chi}(h)e^{2\pi i n h/q},$$

where the last step sets $h = mn^{-1}$. If n and q are not coprime, the proof of this fact is rather more delicate, as the left hand side will be zero. It also requires χ to be primitive. However, proving that the right hand side is zero is not particularly enlightening, so the reader is referred to [4] for details. $\square$

This Lemma has a useful Corollary. Multiplying both sides of (A5) by their conjugates gives

$$|\chi(n)|^2|\tau(\overline{\chi})|^2 = \left(\sum_{h_1=1}^{q} \overline{\chi}(h_1)e^{2\pi inh_1/q}\right)\overline{\left(\sum_{h_2=1}^{q} \overline{\chi}(h_2)e^{2\pi inh_2/q}\right)}$$

$$= \sum_{h_1=1}^{q}\sum_{h_2=1}^{q} \overline{\chi}(h_1)\chi(h_2)e^{2\pi in(h_1-h_2)/q}.$$

Now sum over all residues modulo q. $|\chi(n)|^2$ takes the value 1 at precisely $\phi(q)$ values, and is otherwise zero. Meanwhile, the sum of $e^{2\pi in(h_1-h_2)/q}$ over all n is zero by symmetry, unless $h_1 \equiv h_2$ (q), in which case it always takes the value 1. Thus,

$$\phi(q)\ |\tau(\overline{\chi})|^2 = q\sum_{h}|\chi(h)|^2 = q\ \phi(q),$$

which implies

$$|\tau(\chi)| = q^{1/2} \tag{A6}$$

for all primitive characters χ, upon swapping χ and $\overline{\chi}$.

4 The Poisson Summation Formula

Definition A7. A smooth function is *Schwartz* if for every $c \in \mathbb{R}$, $n \in \mathbb{N} \cup \{0\}$:

$$\left|f^{(n)}(x)\right| = o(|x|^c),$$

where $f^{(n)}$ denotes the nth derivative of f. In other words, every derivative of the function essentially decays exponentially.

We now have the following [15, Theorem 8.37].

Theorem A8. *(Poisson Summation Formula) Let $f : \mathbb{R} \to \mathbb{C}$ be a Schwartz function, and let $\hat{f}$ be its Fourier transform:*

$$\hat{f}(y) = \int_{-\infty}^{\infty} e^{2\pi ixy}f(x)\mathrm{d}y.$$

Then

$$\sum_{m\in\mathbb{Z}} f(m) = \sum_{n\in\mathbb{Z}} \hat{f}(n),$$

with the sums converging absolutely.

Proof. Define $F(x) = \sum_{m \in \mathbb{Z}} f(x+m)$, and note that since f is Schwartz, we have uniform convergence of the sum to a continuous function. To see this, we can bound each term by some arbitrarily large reciprocal power of $|m|$ multiplied by some large constant, since $f(x+m) = o(|x+m|^c)$ for all real numbers c. F is 1-periodic, so consider the Fourier coefficients of F:

$$\int_0^1 e^{-2\pi inx} F(x)\mathrm{d}x = \int_0^1 e^{-2\pi inx} \left(\sum_{m \in \mathbb{Z}} f(x+m) \right) \mathrm{d}x$$

$$= \sum_{m \in \mathbb{Z}} \int_0^1 e^{-2\pi inx} f(x+m)\mathrm{d}x \qquad \text{(By uniform convergence)}$$

$$= \sum_{m \in \mathbb{Z}} \int_m^{m+1} e^{-2\pi in(y-m)} f(y)\mathrm{d}y \qquad \text{(Substitute } x = y - m)$$

$$= \sum_{m \in \mathbb{Z}} \int_m^{m+1} e^{-2\pi iny} f(y)\mathrm{d}y \qquad (e^{-2\pi in(y-m)} = e^{-2\pi iny})$$

$$= \int_{-\infty}^{\infty} e^{-2\pi iny} f(y)\mathrm{d}y = \hat{f}(n).$$

So we have

$$F(x) = \sum_{n \in \mathbb{Z}} \hat{f}(n)e^{inx}.$$

Now note that $F(0) = \sum_{m \in \mathbb{Z}} f(m)$ by definition. Therefore,

$$\sum_{m \in \mathbb{Z}} f(m) = F(0) = \sum_{n \in \mathbb{Z}} \hat{f}(n).$$

$\square$

This theorem may be applied to $f(n) = e^{-(n+\alpha)^2 \pi/x}$ as in [4, p. 63-64]. Each derivative is dominated by an exponentially decaying function, so is clearly Schwartz. Thus,

$$\hat{f}(n) = \int_{-\infty}^{\infty} e^{2\pi int} f(t)\mathrm{d}t$$

$$= \int_{-\infty}^{\infty} e^{2\pi int - (t+\alpha)^2 \frac{\pi}{x}}\mathrm{d}t$$

$$= e^{-n^2 \pi x - 2\pi in\alpha} \int_{-\infty}^{\infty} e^{-\frac{\pi}{x}(t - inx + \alpha)^2}\mathrm{d}t$$

$$= \left(\frac{x}{\pi}\right)^{1/2} e^{-n^2 \pi x - 2\pi in\alpha} \int_{-\infty}^{\infty} e^{-v^2}\mathrm{d}v \quad \left(\text{sub. } v = (x/\pi)^{1/2}(t - inx + \alpha)\right)$$

$$= x^{1/2} e^{-n^2 \pi x - 2\pi in\alpha}$$

By the Poisson summation formula, we therefore have the relation

$$\sum_{n=-\infty}^{\infty} e^{-(n+\alpha)^2 \pi/x} = x^{1/2} \sum_{n=-\infty}^{\infty} e^{-n^2 \pi x + 2\pi in\alpha}. \tag{A7}$$

Bibliography

[1] G. F. Simmons, *Calculus Gems*, p. 198. American Mathematical Society, 2019.

[2] H. Riesel, *Prime Numbers and Cryptography*, pp. 226–238. Birkhäuser Boston, 1994.

[3] O. Ore, *Number Theory and Its History*, p. 65. Dover Publications, 2012.

[4] H. Davenport and H. L. Montgomery, *Multiplicative number theory*. Springer, 2000.

[5] D. R. Heath-Brown, *Prime number theory and the Riemann zeta-function*. London Mathematical Society Lecture Note Series, Cambridge University Press, 2005.

[6] K. F. Ireland and M. I. Rosen, *A classical introduction to modern number theory*, ch. 16. Springer, second ed., 1990.

[7] A. Ivić, *The Riemann zeta-function*. Dover, 2003.

[8] J.-P. Serre and L. L. Scott, *Linear representations of finite groups*, pp. 3–25. Graduate texts in mathematics 42, Springer, second ed., 1996.

[9] T. Estermann, "On Dirichlet's L Functions," *Journal of the London Mathematical Society*, vol. s1-23, no. 4, pp. 275–279, 1948.

[10] A. Granville and G. Martin, "Prime number races," *The American Mathematical Monthly*, vol. 113, no. 1, pp. 1–33, 2006.

[11] M. Rubinstein and P. Sarnak, "Chebyshev's bias," *Experimental Mathematics*, vol. 3, pp. 173–197, 1994.

[12] R. J. Lemke Oliver and K. Soundararajan, "Unexpected biases in the distribution of consecutive primes," *Proceedings of the National Academy of Sciences*, vol. 113, no. 31, pp. E4446–E4454, 2016.

[13] V. A. Zorich and R. Cooke, *Mathematical Analysis II*, ch. 17, p. 420. Universitext, Springer, first ed., 2004.

[14] G. E. Andrews, R. Askey, and R. Roy, *The Gamma and Beta Functions*, p. 1–60. Encyclopedia of Mathematics and its Applications, Cambridge University Press, 1999.

[15] M. C. Pereyra and L. A. Ward, *Harmonic Analysis: From Fourier to Wavelets*, ch. 8. American Mathematical Society, 2012.

COVID19 IMPACT SHEET
Project 3/4
Department of Mathematical Sciences

Student Name: Oliver Welch

Year group (3/4): 3

Project Topic: Introduction to Analytic Number Theory

Project supervisor(s): Pankaj Vishe

Did Covid19 prevent you from completing part of your project report (Yes/No):
No

If 'Yes', please indicate what it prevented you from doing (max 100 words):
N/A

Please summarise the action taken in response (max 100 words): N/A