Introduction to Analytic Number Theory

Ian Petrow

E-mail address: ian.petrow@math.ethz.ch

Based on and translated from course notes by Philippe Elbaz-Vincent and Philippe Michel

Contents

CHAPTER 1

Counting Prime Numbers

1. Introduction

It has been known since the time of Euclid that there are infinitely many prime numbers. Arguing by contradiction, suppose that there were only finitely many primes $p_1, \dots, p_n$. Then the number $p_1 \cdots p_n + 1$ must have a prime divisor not equal to any of $p_1, \dots, p_n$. In this course we will be interested in quantifying the infinitude of prime numbers. To do so, we define the prime counting function

$$\pi(x) = \#\{p \in \mathscr{P} : p \leqslant x\}.$$

Euclid's theorem therefore says that $\pi(x) \to \infty$ as $x \to \infty$, but the question is

at what rate?

One can do experiments with prime numbers relatively easily. To do so, the first step is to produce the list of prime numbers up to a certain limit. A simple and systematic method is given by the *sieve of Eratosthenes*:

(1) Write the list of all integers up to X.
(2) Cross out 1 (which isn't prime).
(3) Keep 2, and cross out all proper multiples of 2.
(4) Keep 3, and cross out all proper multiples of 3 that aren't already eliminated.
(5) etc.
(6) The first number not crossed out by the preceding steps is automatically prime. Keep it and cross out all its proper multiples.
(7) etc.

For example, for $X = 30$ we get

2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30

2 3 5 7 9 11 13 15 17 19 21 23 25 27 29

2 3 5 7 11 13 17 19 23 25 29

2 3 5 7 11 13 17 19 23 29

…

2 3 5 7 11 13 17 19 23 29

Remark 1.1. In the above example, note that we get the list of prime numbers $\leqslant X$ at the third step (where we crossed out multiples of 5). This phenomenon is explained simply in the following criteria:

An integer $n \geqslant 2$ is composite if and only if it has a divisor d satisfying $1 < d \leqslant \sqrt{n}$.

PROOF. If n has a divisor $1 < d \leqslant \sqrt{n}$, then $n > 1$ and $d \neq n, 1$, since $\sqrt{n} < n$. Thus n is composite. In the opposite direction, if n is composite and $d \neq n, 1$ is a divisor of n, then n/d is also a divisor of n with $1 < n/d < n$. (The divisor n/d of n is called the complimentary divisor of d.) Set

$$d' = \min(d, n/d).$$

Then we have $1 < d' \leqslant \sqrt{n}$. □

Thus, after the third step (where we cross out all the multiples of 5) we have in fact found all of the prime numbers < 49, since $7^2 = 49$.

Notice that the spacing between consecutive prime numbers seems to grow, which suggests that the set of primes $\mathscr{P}$ becomes less and less dense. Following much more intensive numerical experiments, Gauss and Legendre around the year 1795 gave a conjectural asymptotic formula for the function $\pi(x)$: *the prime number conjecture*. About a century later in 1896, this conjecture was proven by Hadamard and de la Vallée-Poussin, and thus became *the prime number theorem*:

THEOREM (Prime Number Theorem (PNT)). *As $x \to \infty$ we have*

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

Remark 1.2. Let us specify some notation once and for all. Let f, g be two functions on $\mathbb{R}$ with g non-zero for x sufficiently large. We write $f \sim g$ if and only if

$$\frac{f(x)}{g(x)} \to 1$$

as $x \to \infty$. If g is non-negative the notation $f = O(g)$ means that there exists an absolute constant C such that

$$|f(x)| \leqslant Cg(x)$$

for all x in the domain of f and g. We may also write $f \ll g$, which means the exact same thing as $f = O(g)$. From time to time we might write $f = O_\varepsilon(g)$ or $f \ll_\varepsilon g$, which means that the constant C is also allowed to depend on ε, i.e. $C = C(\varepsilon)$. Lastly, we write $f = o(g)$ if for all $\varepsilon > 0$ there exists a constant $N > 0$ such that

$$|f(x)| \leqslant \varepsilon g(x)$$

for all $x \geqslant N$. Vaguely, this means that f is *strictly* bounded by g.

Remark 1.3. In the 1830s, Dirichlet formulated the prime number conjecture in a slightly different form: introducing the following function, called the *logarithmic integral*

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

he conjectured that

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

Seeing as $\mathrm{Li}(x) \sim \frac{x}{\log x}$ (do an integration by parts) this formulation of the conjecture is equivalent to the original. However, Dirichlet's version is better: as we will see later, the proof of the PNT in fact gives the following asymptotic formula

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

for some absolute constant $c>0$. The celebrated *Riemann Hypothesis* predicts that

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

2. Euler's Method

With some care, Euclid's proof of the infinitude of primes can be modified to produce an explicit lower bound on $\pi(x)$. Such a bound is extremely bad. Euler came up with another proof of the infinitude of primes, which gives much better result and opened the line of attack which eventually led to the proof of the prime number theorem. In some sense, Euler's method is the starting point for all of analytic number theory. Euler's method is combinatorial and analytic in nature, and is based on the zeta function that he introduced himself: for $s>1$ one considers the convergent series

$$\zeta(s) = \sum_{n\geqslant 1} \frac{1}{n^s}.$$

By comparison with an integral we have

$$\zeta(s) = \sum_{n\geqslant 1} \frac{1}{n^s} \geqslant \int_1^\infty \frac{dx}{x^s} = \frac{1}{s-1}, \tag{1.1}$$

which tends to $+\infty$ as $s\to 1^+$ from the right. The fundamental observation of Euler is that the fundamental theorem of arithmetic allows one to express $\zeta(s)$ in terms of the prime numbers: for $s>1$ consider for each prime number p the series

$$\zeta_p(s) = 1 + \frac{1}{p^s} + \frac{1}{(p^2)^s} + \cdots + \frac{1}{(p^\alpha)^s} + \cdots,$$

that is to say the series ζ_p is the series ζ restricted to powers of p. The series $\zeta_p(s)$ is a geometric series, and therefore we have

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

By the unique factorization of integers,

$$\zeta_2(s)\zeta_3(s) = \sum_{\alpha_2\geqslant 0}\sum_{\alpha_3\geqslant 0} \frac{1}{(2^{\alpha_2}3^{\alpha_3})^s}$$

is the series $\zeta(s)$ restricted to the integers whose prime factorizations contain only powers of 2 and 3. Likewise, $\zeta_2(s)\zeta_3(s)\zeta_5(s)$ is the series $\zeta(s)$ restricted to integers whose prime factorizations only contain powers of 2,3, and 5, and so on.

Supposing that

$$\mathscr{P} = \{2,3,5,\ldots,p_{\max}\}$$

is finite, we get the identity

$$\zeta(s) = \sum_{n\geqslant 1} \frac{1}{n^s} = \prod_{p\in\mathscr{P}} \zeta_p(s) = \prod_{p\in\mathscr{P}} (1-\frac{1}{p^s})^{-1}. \tag{1.2}$$

As the product $\prod_{p\in\mathscr{P}}(1-\frac{1}{p^s})^{-1}$ is finite, and for each prime number p the series $\zeta_p(s)$ is well defined at $s=1$ (it takes the value $\zeta_p(1) = (1-\frac{1}{p})^{-1}$), we see that $\zeta(s)$ should have a finite limit as $s\to 1^+$. This contradicts (1.1), therefore $\mathscr{P}$ is infinite.

In fact we shall soon see that even though $\mathscr{P}$ is infinite, the identity (1.2) holds for all $s>1$ (that is to say the infinite product converges and is equal to $\zeta(s)$). Then, Euler's method allows

for precise quantitative results on counting prime numbers. Indeed, taking the logarithm of (1.2) we see that for $s>1$

$$\log(\zeta(s)) = -\sum_p \log(1-\frac{1}{p^s}) = \sum_p (\frac{1}{p^s} + O(\frac{1}{p^{2s}})) = \sum_p \frac{1}{p^s} + O(1).$$

On the other hand, we saw in (1.1) that

$$\log(\zeta(s)) \geqslant \log(\frac{1}{s-1}) = -\log(s-1).$$

Taking the limit as s tends to 1, we have (by the monotone convergence theorem) the following.

THEOREM. *The series*

$$\sum_{p\in\mathscr{P}} \frac{1}{p}$$

is divergent.

3. Chebyshev's Method

We begin with the follow result of Chebyshev (circa 1850), which only uses elementary methods and which gives the correct order of magnitude for the function $\pi(x)$.

Theorem 1.4. *There exist constants $0<c<C$ such that for $x\geqslant 2$ one has*

$$c\frac{x}{\log x} \leqslant \pi(x) \leqslant C\frac{x}{\log x}.$$

PROOF. Let $n\geqslant 1$ and consider its factorial

$$n! = \prod_{1\leqslant k\leqslant n} k.$$

Chebyshev's method is based on the fact that this number $n!$ is divisible by and only divisible by all of the prime numbers $\leqslant n$. Let us recall the following definition.

Definition 1.5 (p-adic valuation)**.** *For $n\in\mathbb{Z}-\{0\}$ and p a prime number, the p-adic valuation of n, written $v_p(n)$, is the largest integer $\alpha\geqslant 0$ such that p^α divides n. That is to say, such that $p^\alpha \mid n$ and $p^{\alpha+1}\nmid n$. In particular, one has*

$$n = \prod_{p\mid n} p^{v_p(n)} = \prod_{p\in\mathscr{P}} p^{v_p(n)}.$$

If $n=0$ we set $v_p(0)=\infty$.

Notation: We will denote by $\mathbb{N}=\{0,1,2,\ldots\}$ the set of non-negative integers.

Note that $v_p(mn)=v_p(n)+v_p(m)$ for all $n,m\in\mathbb{N}$. Therefore we have

$$n! = \prod_{p\leqslant n} p^{v_p(n!)},$$

and moreover that $v_p(n!)\geqslant 1$ for all $p\leqslant n$. Therefore, taking the logarithm of this expression we have

$$\log(n!) = \sum_{p\leqslant n} v_p(n!)\log p.$$

We proceed by evaluating the two sides of this equations by different means. Consider first the left hand side

$$\log(n!) = \sum_{k\leqslant n} \log k.$$

Such a sum can be evaluated by comparison against the integral

$$\int_1^n \log t\, dt = n\log n - n + 1$$

(for more details, see Chapter 2) and we find

$$\log(n!) = \sum_{k\leqslant n} \log k = n\log n - n + O(\log n). \tag{1.3}$$

Thus, we get

$$\sum_{p\leqslant n} v_p(n!)\log p = n\log n - n + O(\log n).$$

Now we need to evaluate the valuation $v_p(n!)$. Let k be an integer. Then the valuation $v_p(k)$ is

$$v_p(k) = \max\{\alpha \geqslant 0 : p^\alpha \mid k\} = \sum_{\substack{\alpha\geqslant 1\\ p^\alpha\mid k}} 1$$

and so

$$v_p(n!) = \sum_{1\leqslant k\leqslant n} v_p(k) = \sum_{1\leqslant k\leqslant n}\sum_{\substack{\alpha\geqslant 1\\ p^\alpha\mid k}} 1 = \sum_{\alpha\geqslant 1}\sum_{\substack{1\leqslant k\leqslant n\\ p^\alpha\mid k}} 1 = \sum_{\alpha\geqslant 1} \lfloor \frac{n}{p^\alpha}\rfloor,$$

where

$$x \mapsto \lfloor x\rfloor = \sum_{1\leqslant k\leqslant x} 1 = x - \{x\}$$

is the *integer part* function of x (and $\{x\}$ designates the *fractional part* of x). Above, we used the identity

$$\sum_{\substack{1\leqslant k\leqslant n\\ p^\alpha\mid k}} 1 = \sum_{1\leqslant k'\leqslant n/p^\alpha} 1 = \lfloor \frac{n}{p^\alpha}\rfloor.$$

We have therefore that

$$\sum_{p\leqslant n}\log p \sum_{\substack{\alpha\geqslant 1\\ p^\alpha\leqslant n}} \lfloor \frac{n}{p^\alpha}\rfloor = n\log n - n + O(\log n). \tag{1.4}$$

Now we evaluate the sum

$$\sum_{\substack{\alpha\geqslant 1\\ p^\alpha\leqslant n}} \lfloor \frac{n}{p^\alpha}\rfloor.$$

Seeing as $\lfloor x\rfloor \leqslant x$, this sum is bounded by

$$\sum_{\substack{\alpha\geqslant 1\\ p^\alpha\leqslant n}} \lfloor \frac{n}{p^\alpha}\rfloor \leqslant \sum_{\alpha\geqslant 1}\frac{n}{p^\alpha} = \frac{n}{p(1-1/p)} = \frac{n}{p}(1+O(\frac{1}{p})) = \frac{n}{p} + O(\frac{n}{p^2}).$$

Therefore we have

$$n\log n - n + O(\log n) \leqslant n\sum_{p\leqslant n}\log p(\frac{1}{p} + O(\frac{1}{p^2})) = n\sum_{p\leqslant n}\frac{\log p}{p} + O(n)$$

since

$$\sum_{p\leqslant n}\frac{\log p}{p^2} = O(1).$$

Dividing by n we find that

$$\sum_{p\leqslant n}\frac{\log p}{p} \geqslant \log n + O(1),$$

which shows again that $\mathscr{P}$ is infinite.

3.1. The binomial coefficient. Unfortunately it is not possible to extract the prime counting function from the above approach. The problem is, essentially, that the "weights" $v_p(n!)$ vary too much as p runs between 2 and n. To fix this shortcoming, Chebyshev considered instead of the integer $n!$ the binomial coefficient $\binom{2n}{n} = (2n)!/(n!)^2$.

First of all, by (1.3) we find

$$\log\binom{2n}{n} = (2n\log 2n - 2n + O(\log n)) - 2(n\log n - n + O(\log n)) = (\log 4)n + O(\log 2n).$$

Remark 1.6. It is possible to get this asymptotic formula in a more elementary way. We have

$$\sum_{k=0}^{2n}\binom{2n}{k} = (1+1)^{2n}$$

and we know that among the $2n+1$ terms in the above sum, $\binom{2n}{n}$ is the largest. Thus

$$\frac{(1+1)^{2n}}{2n+1} \leqslant \binom{2n}{n} \leqslant (1+1)^{2n}.$$

So, taking logarithms, we find

$$\log\binom{2n}{n} = 2n(\log 2) + O(\log n).$$

Note that $\binom{2n}{n}$ is divisible by all of the prime numbers in the interval $(n, 2n]$. Setting

$$\theta(x) = \sum_{p\leqslant x}\log p,$$

we therefore have

$$\log\binom{2n}{n} = \sum_{p\leqslant 2n} v_p(\binom{2n}{n})\log p \geqslant \sum_{n<p\leqslant 2n}\log p = \theta(2n) - \theta(n). \tag{1.5}$$

We get that for all $n \geqslant 2$

$$\theta(2n) - \theta(n) \leqslant (\log 2)2n.$$

Given a real number $x \geqslant 2$ there always exists an even number $2n$ such that $0 \leqslant x - 2n \leqslant 2$. For such a choice of $2n$ we have

$$\theta(x) - \theta(2n) \leqslant \log x,$$

and

$$0 \leqslant \theta(x/2) - \theta(n),$$

so that we see for all $x \geqslant 2$

$$\theta(x) - \theta(x/2) \leqslant (\log 2)x + O(\log x).$$

Using this inequality with $x, x/2, x/4, \ldots$ etc. (in fact $O(\log x)$ times) we deduce that

$$\theta(x) = \sum_{k\geqslant 0}\theta(x/2^k) - \theta(x/2^{k+1}) \leqslant \sum_{0\leqslant k\ll\log x}(\log 2)\frac{x}{2^k} \leqslant (2\log 2)x + O((\log x)^2).$$

From this last formula we can easily deduce an upper bound for $\pi(x)$ of the correct order of magnitude. We have

$$\pi(x) = \sum_{p \leqslant x} 1 = \sum_{p \leqslant x^{1/2}} 1 + \sum_{x^{1/2} < p \leqslant x} 1.$$

The first term is bounded by $x^{1/2}$, and the second term satisfies

$$\sum_{x^{1/2} < p \leqslant x} 1 \leqslant \frac{1}{\log(x^{1/2})} \sum_{x^{1/2} < p \leqslant x} \log p = \frac{2}{\log x}(\theta(x) - \theta(x^{1/2})) \leqslant 2(\log 4)\frac{x}{\log x} + O(x^{1/2})$$

and therefore

$$\pi(x) \leqslant 2(\log 4)\frac{x}{\log x} + O(x^{1/2}).$$

Remark 1.7. By a slightly more developed argument (integration by parts, see Chapter 2), one can show that

$$\pi(x) \leqslant (\log 4)\frac{x}{\log x}(1 + O(\frac{1}{\log x})),$$

as $x \to \infty$.

3.2. The lower bound. To get a lower bound, we need to be more precise about the value of $v_p(\binom{2n}{n})$. We have

$$v_p(\binom{2n}{n}) = v_p(2n!) - 2v_p(n!) = \sum_{\alpha \geqslant 1} \lfloor \frac{2n}{p^\alpha} \rfloor - 2\lfloor \frac{n}{p^\alpha} \rfloor = \sum_{\alpha \geqslant 1} \varpi(\frac{n}{p^\alpha})$$

where

$$\varpi(x) = \lfloor 2x \rfloor - 2\lfloor x \rfloor = 2\{x\} - \{2x\}.$$

The function $\varpi(x)$ is periodic with period 1 and is given by

$$\varpi(x) = \begin{cases} 0 & \text{if } x \in [0, 1/2) \\ 1 & x \in [1/2, 1). \end{cases}$$

In particular, notice that $\varpi(x)$ varies much less than the function $x \mapsto \lfloor x \rfloor$. Note that for $p \leqslant 2n$ the number of non-zero terms in the sum

$$\sum_{\alpha \geqslant 1} \varpi(\frac{n}{p^\alpha})$$

is bounded by $\alpha \leqslant \log 2n / \log p$, and seeing as $\varpi(n/p^\alpha) \leqslant 1$, we get the upper bound

$$\log\binom{2n}{n} = (\log 4)n + O(\log n) = \sum_{p \leqslant 2n} \log p \sum_{\alpha \geqslant 1} \varpi(\frac{n}{p^\alpha}) \leqslant \sum_{p \leqslant 2n} \log p \frac{\log 2n}{\log p} = (\log 2n)\pi(2n). \tag{1.6}$$

Thus we get the lower bound

$$(\log 2)\frac{2n}{\log 2n} + O(1) \leqslant \pi(2n)$$

and more generally, for every integer $x \geqslant 2$

$$(\log 2)\frac{x}{\log x}(1 + O(\frac{\log x}{x})) \leqslant \pi(x).$$

From this we deduce a lower bound for the function $\theta(x)$.

$$\theta(x) = \sum_{p \leqslant x^{1/2}} \log p + \sum_{x^{1/2} < p \leqslant x} \log p \geqslant (\log x^{1/2})(\pi(x) - \pi(x^{1/2})) + O(x^{1/2}\log x) = \frac{1}{2}(\log x)\pi(x) + O(x^{1/2}\log x),$$

and therefore

$$\theta(x) \geqslant \frac{\log 2}{2} x(1 + O(\frac{\log x}{x^{1/2}})). \tag{1.7}$$

Thus, we have shown the existence of constants $0 < c < C$ such that for all $x \geqslant 2$

$$cx \leqslant \theta(x) \leqslant Cx, \quad \text{and} \quad c\frac{x}{\log x} \leqslant \pi(x) \leqslant C\frac{x}{\log x}. \tag{1.8}$$

□

Remark 1.8. With a bit more care, we can show that the constants c and C can be taken to be arbitrarily close to $\log 2$ and $2\log 2$, respectively.

From Chebyshev's theorem, we can deduce the following useful asymptotic formula.

THEOREM (Mertens). *We have*

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

PROOF. Recall the formula (1.4): for all $n \geqslant 2$

$$\sum_{p \leqslant n} \log p \sum_{\substack{\alpha \geqslant 1 \\ p^\alpha \leqslant n}} \lfloor \frac{n}{p^\alpha} \rfloor = n \log n - n + O(\log n).$$

The contribution of the $\alpha \geqslant 2$ to the left hand side of the above is bounded by

$$\sum_{p \leqslant n} \log p \sum_{\substack{\alpha \geqslant 2 \\ p^\alpha \leqslant n}} \lfloor \frac{n}{p^\alpha} \rfloor \leqslant \sum_{p \leqslant n} \log p \sum_{2 \leqslant \alpha} \frac{n}{p^\alpha} \ll n \sum_{p \leqslant n} \frac{\log p}{p^2} = O(n).$$

Considering the remainder, and writing $\lfloor n/p \rfloor = n/p + O(1)$ we get

$$\sum_{p \leqslant n} \log p \frac{n}{p} + O(\sum_{p \leqslant n} \log p) = n \log n - n + O(\log n) + O(n).$$

Chebyshev's theorem gives us that $\theta(n) \ll n$, and so we find that

$$\sum_{p \leqslant n} \log p \frac{n}{p} = n \log n + O(n).$$

□

CHAPTER 2

Sums of arithmetic functions

In this chapter, we will present several basic methods to evaluate sums over the integers. The terms of these sums will be called *arithmetic function.*

Definition 2.1. *An arithmetic function is a complex-valued function on the positive integers, $f : \mathbb{N}_{\geqslant 1} \to \mathbb{C}$. We write $\mathscr{A}$ for the $\mathbb{C}$-vector space of arithmetic functions.*

Example(s) 2.2. (1) The constant function

$$1 : n \mapsto 1,$$

(2) The delta function at 1

$$\delta : n \mapsto \begin{cases} 1 & n = 1 \\ 0 & n \neq 1, \end{cases}$$

(3) The characteristic function of the set of prime numbers

$$1_{\mathscr{P}} : n \mapsto \begin{cases} 1 & n = p \in \mathscr{P} \\ 0 & n \notin \mathscr{P}, \end{cases}$$

(4) The same function weight by the logarithm

$$\log.1_{\mathscr{P}} : n \mapsto \log(n) 1_{\mathscr{P}}(n),$$

(5) The *von Mangolt* function

$$\Lambda : n \mapsto \begin{cases} \log p & n = p^\alpha, \alpha \geqslant 1 \\ 0 & n \neq p^\alpha. \end{cases}$$

Despite its artificial appearance, the last function above arises naturally in the study of prime numbers, where it plays a fundamental role.

Definition 2.3. *Let f be an arithmetic function. The* summation function *of f is the function defined on $\mathbb{R}_{\geqslant 0}$ by*

$$x \mapsto M_f(x) = \sum_{1 \leqslant n \leqslant x} f(n).$$

The summation function of f is a piecewise constant function, and in this chapter, we will present methods to study the following question:

Problem. Given an arithmetic function f, determine the behavior of $M_f(x)$ as $x \to \infty$.

Example(s) 2.4. (1) $M_1(x) = \sum_{1 \leqslant n \leqslant x} 1 = \lfloor x \rfloor = x + O(1)$.
(2) $\pi(x) = M_{1_{\mathscr{P}}}(x) = \sum_{p \leqslant x} 1$.
(3) $\theta(x) = M_{\log.1_{\mathscr{P}}}(x) = \sum_{p \leqslant x} \log p$.
(4) $\psi(x) = M_\Lambda(x) = \sum_{1 < p^\alpha \leqslant x} \log p$.

Guided by the above examples, we will compare $\theta(x)$ and $\psi(x)$. Observe first of all that

$$\theta(x) \leqslant \psi(x).$$

More precisely, we have

$$\psi(x) = \theta(x) + \sum_{\substack{p^\alpha \leqslant x \\ \alpha \geqslant 2}} \log p.$$

We have

$$\begin{aligned}\sum_{\substack{p^\alpha \leqslant x \\ \alpha \geqslant 2}} \log p &= \sum_{p \leqslant \sqrt{x}} \log p \sum_{\substack{p^\alpha \leqslant x \\ \alpha \geqslant 2}} 1 \leqslant \sum_{p \leqslant x^{1/2}} \log p \lfloor \frac{\log x}{\log p} \rfloor \\ &\leqslant \log x \sum_{p \leqslant x^{1/2}} 1 \\ &\ll x^{1/2} \log x.\end{aligned}$$

Thus

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

We saw that by Chebyshev's method $\theta(x) \gg x$ (see (1.7)) and we deduce that

$$\psi(x) \sim \theta(x). \tag{2.1}$$

1. Approximation by integrals, integration by parts

If f is the restriction to $\mathbb{N}_{\geqslant 1}$ of a continuous function on $\mathbb{R}$, then $M_f(x)$ is often well approximated by

$$\int_1^x f(t)\, dt.$$

For example, if f is *monotone* we have

Theorem 2.5 (Monotone comparison)**.** *If f is monotone we have*

$$M_f(x) = \int_1^x f(t)\, dt + O(|f(1)| + |f(x)|). \tag{2.2}$$

Proof. Suppose that f is monotone increasing. The result is deduced by summing the following inequality over $n \geqslant 2$:

$$\int_{n-1}^n f(t)\, dt \leqslant f(n) \leqslant \int_n^{n+1} f(t)\, dt.$$

□

For example, we have

$$M_{\log}(x) = \sum_{n \leqslant x} \log(n) = \int_1^x \log(t)\, dt + O(\log(x)) = x \log x - x + O(\log x). \tag{2.3}$$

The following result allows one to evaluate the summation function of a product of an arithmetic function by a "smooth" function:

Theorem 2.6 (Integration by parts)**.** *Let g be an arithmetic function. Let $a < b \in \mathbb{R}_{>0}$ and $f : [a,b] \to \mathbb{C}$ a $C^1([a,b])$ function. We have*

$$M_{fg}(b) - M_{fg}(a) = \sum_{a<n\leqslant b} f(n)g(n) = [f(t)M_g(t)]_{t=a}^{t=b} - \int_a^b M_g(t)f'(t)\,dt$$
$$= f(b)M_g(b) - f(a)M_g(a) - \int_a^b M_g(t)f'(t)\,dt$$

Proof. Suppose first of all that $b = a+1 = n+1 \in \mathbb{N}_{\geqslant 1}$. Then

$$[f(x)M_g(x)]_{x=a}^{x=b} - \int_a^b M_g(x)f'(x)\,dx = M_g(n+1)f(n+1) - M_g(n)f(n) - M_g(n)(f(n+1)-f(n))$$
$$= f(n+1)g(n+1)$$
$$= M_{fg}(n+1) - M_{fg}(n).$$

By summing the above formula up, we extend to arbitrary $a \leqslant b \in \mathbb{N}_{\geqslant 1}$, and then to a, b general real numbers. □

Remark 2.7. Note the analogy between the above formula and the classical integration by parts formula: if f is C^1 and g is continuous, we have

$$\int_a^b f(t)g(t)\,dt = [f(t)G(t)]_{t=a}^{t=b} - \int_a^b f'(t)G(t)\,dt,$$

where

$$G(x) = \int_1^x g(t)\,dt$$

is an anti-derivative (or, a primitive) of g.

Example(s) 2.8. We have

$$\theta(x) = \sum_{p\leqslant x} \log(p) = \sum_{2\leqslant n\leqslant x} \log(n)1_{\mathscr{P}}(n)$$
$$= \pi(x)\log(x) - \int_{1.5}^x \pi(t)\,\frac{dt}{t}$$
$$= \pi(x)\log(x) + O(1) + O\left(\int_2^x \frac{dt}{\log(t)}\right)$$
$$= \pi(x)\log(x) + O\left(\frac{x}{\log x}\right),$$

where the second to last equality is due to Chebyshev's theorem. In particular, as $\pi(x) \gg x/\log x$ we see that

$$\theta(x) = \pi(x)\log(x)(1+o(1)).$$

Thus by (2.1) we have the following equivalent statements of the prime number theorem

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

In fact, it's the last asymptotic formula that Hadamard and de la Vallée-Poussin showed in their proof of the prime number theorem.

Corollary 2.9 (Euler-MacLaurin Formula)**.** *Let f be a C^1 function on $\mathbb{R}_{>0}$, and let*

$$\psi_1(x) = x - \lfloor x \rfloor - 1/2 = \{x\} - 1/2.$$

We have for all $x > 1$ that

$$M_f(x) = \int_1^x f(t)\,dt + \int_1^x \psi_1(t) f'(t)\,dt - \psi_1(1) f(1) - \psi_1(x) f(x),$$

in particular

$$\left| M_f(x) - \int_1^x f(t)\,dt \right| \leqslant \int_1^x |f'(t)|\,dt + |f(1)| + |f(x)|.$$

Proof. Exercise. □

Remark 2.10. We have $\psi_1(x) = B_1(\{x\})$, where $B_1(x) = x - 1/2$. This polynomial is called the *first Bernoulli polynomial.*

Remark 2.11. The important thing about this formula is that the function f is not necessarily monotone.

2. Dirichlet Convolution

The Dirichlet convolution is a composition law on the set of arithmetic functions that realizes the multiplicative structure of the integers.

Let $f, g \in \mathscr{A}$, and define $f * g \in \mathscr{A}$ by setting

$$f * g(n) = \sum_{ab=n} f(a) g(b) = \sum_{d|n} f(d) g(n/d).$$

Proposition 2.12. *The triple $(\mathscr{A}, +, *)$ has the structure of a commuative, associative $\mathbb{C}$-algebra with multiplicative identity element δ (recall example 2.2). The set of invertible element of the algebra $\mathscr{A}$ is*

$$\mathscr{A}^\times = \{f \in \mathscr{A} : f(1) \neq 0\}.$$

For $f \in \mathscr{A}^\times$, we write $f^{(-1)}$ for its convolution inverse.

Proof. We only verify certain parts of this proposition, the remainder is left to the reader.

- Commutativity

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

- Identity element

$$f * \delta(n) = \sum_{d|n} f(n/d)\delta(d) = f(n/1) = f(n)$$

- Units: If f is invertible with inverse $f^{(-1)}$, then

$$f * f^{(-1)}(1) = \delta(1) = 1 = f(1) f^{(-1)}(1).$$

  Therefore $f(1) \neq 0$. Let f be such that $f(1) \neq 0$. We seek a $g \in \mathscr{A}$ that satisfies $f * g = \delta$, in particular $f * g(1) = f(1) g(1) = 1$. Therefore $g(1) = 1/f(1)$. For any $n > 1$

$$f * g(n) = \delta(n) = 0 = \sum_{d|n} f(d) g(n/d) = g(n) f(1) + \sum_{\substack{d|n \\ d>1}} f(d) g(n/d)$$

Thus

$$g(n) = -\frac{1}{f(1)} \sum_{\substack{d|n \\ d>1}} f(d)g(n/d).$$

As $n/d < n$ since $d > 1$, the value $g(n)$ is determined by the values of g at integers $\leqslant n/2$. Therefore the function g is defined by recurrence.

□

NOTATION. *If f is an arithmetic function, and $k \in \mathbb{N}$, we will denote the k-fold iterated convolution by*

$$f^{(*k)} = \underbrace{f * f * \cdots * f}_{k \text{ times}},$$

and if f is invertible, we extend the above notation to all $k \in \mathbb{Z}$ in the evident manner:

$$f^{(*-k)} = (f^{(-1)})^{(*k)}.$$

Example(s) 2.13.

(1) $1 * 1(n) = \sum_{d|n} 1 = d(n)$ is the number of divisors of n.

(2) $1 * 1 * 1(n) = \sum_{abc=n} 1 = d_3(n)$ is the number of representations of n as a product of three integers. More generally, we write

$$d_k(n) = 1^{(*k)}(n) = 1 * \cdots * 1(n) = \sum_{d_1 \cdots d_k = n} 1$$

for the number of representations of n as a product of k integers.

(3) We have

$$\log = \Lambda * 1, \qquad \text{i.e.} \qquad \log(n) = \sum_{d|n} \Lambda(d).$$

Indeed, if $n = \prod_p p^{\alpha_p}$ then

$$\begin{aligned} \log(n) = \log\left(\prod_p p^{\alpha_p}\right) &= \sum_p \alpha_p \log(p) \\ &= \sum_p \sum_{1 \leqslant \alpha \leqslant \alpha_p} \log(p) \\ &= \sum_{p^\alpha | n} \log(p) = \sum_{d|n} \Lambda(d). \end{aligned}$$

Möbius inversion formula: The *Möbius* function is by definition the inverse of the constant function 1:

$$\mu = 1^{(-1)}, \mu(1) = 1, \mu(n) = -\sum_{\substack{d|n \\ d<n}} \mu(d) \text{ for } n \geqslant 2.$$

In the following section we will show that

(1) If n is divisible by a square not equal to 1 (i.e. there exists a prime p such that $p^2 \mid n$), then $\mu(n) = 0$.

(2) If n is square-free, and has r prime factors (i.e. $n = p_1 \cdots p_r$), then $\mu(n) = (-1)^r$.

Theorem 2.14 (Möbius inversion formula)**.** *Let $f, g \in \mathscr{A}$. The following identities are equivalent*

(1) *For all n, $f(n) = \sum_{d|n} g(d)$.*

(2) *For all n, $g(n) = \sum_{d|n} \mu(d) f(n/d)$.*

PROOF. The first identity is equivalent to $f = g*1$ and the second is equivalent to $g = f*\mu$. Then it is clear that

$$f = g*1 \iff f*\mu = g*1*\mu \iff f*\mu = g*(1*\mu) \iff f*\mu = g.$$

□

Thus, we have for example

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

3. Application to counting prime numbers

Theorem 2.15 (Mertens)**.** *We have*

$$\sum_{n\leqslant x} \frac{\Lambda(n)}{n} = \log(x) + O(1) \tag{2.4}$$

$$\sum_{p\leqslant x} \frac{\log p}{p} = \log(x) + O(1), \tag{2.5}$$

$$\sum_{p\leqslant x} \frac{1}{p} = \log\log(x) + O(1). \tag{2.6}$$

PROOF. We have already seen a proof of (2.5) of the statement of the theorem, but Dirichlet convolution will give us a very short way to show it. We have already seen that

$$\sum_{\substack{p^\alpha\leqslant x\\ \alpha\geqslant 2}} \frac{\log p}{p^\alpha} = O(1),$$

and as

$$\sum_{n\leqslant x} \frac{\Lambda(n)}{n} = \sum_{p\leqslant x} \frac{\log p}{p} + \sum_{\substack{p^\alpha\leqslant x\\ \alpha\geqslant 2}} \frac{\log p}{p^\alpha},$$

we have that (2.4) holds if and only if (2.5) holds. The formula (2.6) follows from (2.5) by integration by parts. So it suffices to show (2.4). To show (2.4), we evaluate the sum

$$M_{\log}(x) = \sum_{n\leqslant x} \log(n)$$

in two different ways. On the one hand we have already seen that

$$\sum_{n\leqslant x} \log(n) = x\log x + O(x).$$

On the other hand, $\log = \Lambda * 1$ and

$$\begin{aligned}\sum_{n\leqslant x} \log(n) &= \sum_{n\leqslant x}\sum_{d|n} \Lambda(d)\\ &= \sum_{d\leqslant x} \Lambda(d) \sum_{m\leqslant x/d} 1\\ &= \sum_{d\leqslant x} \Lambda(d)\lfloor \frac{x}{d}\rfloor\\ &= \sum_{d\leqslant x} \Lambda(d)\frac{x}{d} + O\left(\sum_{d\leqslant x}\Lambda(d)\right).\end{aligned}$$

By Chebychev's theorem,

$$\sum_{d\leqslant x}\Lambda(d)=\sum_{p\leqslant x}\log p+\sum_{p\leqslant x^{1/2}}\sum_{p^\alpha\leqslant x}\log p=O(x)+O(x^{1/2}\log x)=O(x).$$

Therefore,

$$x\log x+O(x)=\sum_{n\leqslant x}\log(n)=x\sum_{d\leqslant x}\frac{\Lambda(d)}{d}+O(x),$$

which proves the theorem. □

4. Multiplicative functions

Definition 2.16. *A non-zero arithmetic function f is called multiplicative if and only if for all $m,n\geqslant 1$ with $(m,n)=1$ we have $f(mn)=f(m)f(n)$. A non-zero arithmetic function is called completely multiplicative if for all $m,n\geqslant 1$ we have $f(mn)=f(m)f(n)$.*

In particular, a multiplicative function satisfies $f(1)=1$, and it is determined completely by its values on prime powers. A completely multiplicative function is determined by its values on the primes. Let us write $p^\alpha\| n$ if $p^\alpha\mid n$ but $p^{\alpha+1}\nmid n$. That is to say, $p^\alpha\| n$ if and only if $v_p(n)=\alpha$. If $n=\prod_{p^{\alpha_p}\| n}p^{\alpha_p}$, then we have

$$f(n)=f(\prod_{p^{\alpha_p}\| n}p^{\alpha_p})=\prod_{p^{\alpha_p}\| n}f(p^{\alpha_p}).$$

If f is completely multiplicative, then we have

$$f(n)=\prod_{p^{\alpha_p}\| n}f(p)^{\alpha_p}.$$

Proposition 2.17. *If f and g are multiplicative, then $f*g$ and $f^{(-1)}$ are as well.*

Proof. If $(m,n)=1$, then the set of divisors of mn, i.e. $\{d\geqslant 1:d\mid mn\}$, is in bijection with the set $\{(d_1,d_2):d_1\mid m,d_2\mid n\}$ of pairs of divisors of m and n. The bijection is given the following two maps, which are inverse to each other

$$d\mapsto((d,m),(d,n))$$
$$(d_1,d_2)\mapsto d_1d_2.$$

Given two multiplicative functions f,g and m,n relatively prime, we have

$$\begin{aligned}f*g(mn)&=\sum_{d\mid mn}f(d)g(\frac{mn}{d})=\sum_{d_1\mid m}\sum_{d_2\mid n}f(d_1d_2)g(\frac{m}{d_1}\frac{n}{d_2})\\&=\sum_{d_1\mid m}\sum_{d_2\mid n}f(d_1)f(d_2)g(\frac{m}{d_1})g(\frac{n}{d_2})=\left(\sum_{d_1\mid m}f(d_1)g(\frac{m}{d_1})\right)\left(\sum_{d_2\mid n}f(d_2)g(\frac{n}{d_2})\right).\end{aligned}$$

In similar fashion, let $g=f^{(-1)}$, and suppose f is multiplicative. We show g is multiplicative by induction. We have $f(1)=1$ and $g(1)=1$. Let $m,n>1$ be relatively prime, and suppose as the induction hypothesis that for all m',n' relatively prime such that $m'n'<mn$ we have $g(m'n')=g(m')g(n')$. Then we have

$$\begin{aligned}g(mn)=f(1)g(mn)=-\sum_{\substack{d\mid mn\\ d>1}}f(d)g(\frac{mn}{d})&=-\sum_{d_1\mid m}\sum_{\substack{d_2\mid n\\ d_1d_2>1}}f(d_1)f(d_2)g(\frac{m}{d_1})g(\frac{n}{d_2})\\&=-f*g(m)f*g(n)+f(1)g(m)f(1)g(n)=g(m)g(n),\end{aligned}$$

since $f*g(m)=f*g(n)=0$. □

Remark 2.18. On the other hand, if f is completely multiplicative, it is not necessarily the case that $f^{(-1)}$ is also completely multiplicative. Consider, for example, 1 and μ.

4.1. Examples. Seeing as the functions 1 and Id are multiplicative, $\mu = 1^{(-1)}$ is also multiplicative, and so are also $d = 1 * 1$, $d_k = d_{k-1} * 1$, and $\varphi = \mathrm{Id} * \mu$. Therefore to determine these functions, it suffices to calculate them on prime powers: for $\alpha \geqslant 0$, the set of divisors of p^α is of the form $\{p^\beta : 0 \leqslant \beta \leqslant \alpha\}$. We therefore find that

$$d(p^\alpha) = |\{\beta : 0 \leqslant \beta \leqslant \alpha\}| = \alpha + 1,$$

and more generally

$$d_k(p^\alpha) = |\{(\beta_1, \dots, \beta_k) \in \mathbb{N}^k : \beta_1 + \cdots + \beta_k = \alpha\}|.$$

The relation $\mu * 1 = \delta$ gives for $\alpha \geqslant 1$

$$\mu(1) = 1, \quad \text{and} \quad \sum_{0 \leqslant \beta \leqslant \alpha} \mu(p^\beta) = 0.$$

Therefore we have that

$$\mu(1) = 1, \quad \mu(p) = -1, \quad \mu(p^\alpha) = 0, \text{ for all } \alpha \geqslant 2.$$

Using the equality $\varphi = \mathrm{Id} * \mu$, we have

$$\varphi(1) = 1, \quad \varphi(p^\alpha) = p^\alpha - p^{\alpha-1} = p^\alpha(1 - \frac{1}{p}).$$

Therefore for $n = \prod_{p^\alpha \| n} p^\alpha$, we have

$$d(n) = \prod_{p^\alpha \| n} (\alpha + 1), \quad \mu(n) = \prod_{p^\alpha \| n} \mu(p^\alpha), \quad \varphi(n) = \prod_{p^\alpha \| n} (p^\alpha - p^{\alpha-1}) = n \prod_{p \mid n} (1 - \frac{1}{p}).$$

CHAPTER 3

Dirichlet Series

1. Review of Power Series

Given a sequence $a = (a_n)_{n\geqslant 0}$, its associated power series is the series

$$F(a,q) = \sum_{n\geqslant 0} a_n q^n.$$

Let ρ be its radius of convergence, which we suppose is strictly positive (in particular, $|a_n| = o((\frac{1.01}{\rho})^n)$, say, as $n \to \infty$). The series $F(a,q)$ therefore defines a holomorphic function in the open disk $\{q \in \mathbb{C}, |q| < \rho\}$ and the data of this function $F(a,q)$ determines the sequence a by the Cauchy integral formula. Let b be another sequence with associated power series

$$F(b,q) = \sum_{n\geqslant 0} b_n q^n$$

of radius of convergence $\rho' > 0$, so that the product $F(a,q)F(b,q)$ defines a power series with radius of convergence $\rho'' \geqslant \min(\rho, \rho')$ given by

$$F(c,q) = F(a,q)F(b,q) = \sum_{n\geqslant 0} c_n q^n, \quad \text{where} \quad c_n = \sum_{\ell+m=n} a_\ell b_m.$$

This type of relation opens the door to the study of arithmetic problems by analytic means. For example, consider $a = (r_\square(n))_{n\geqslant 0}$:

$$a_n = r_\square(n) = |\{m \in \mathbb{Z} : m^2 = n\}| = \begin{cases} 2 & \text{if } n \text{ is a square } \geqslant 1, \\ 1 & \text{if } n = 0, \\ 0 & \text{else.} \end{cases}$$

The associated power series is

$$F(q) = 1 + 2\sum_{m\geqslant 1} q^{m^2} = \sum_{m\in\mathbb{Z}} q^{m^2},$$

and its radius of convergence is 1. The powers of this series are of the form (when $k \geqslant 1$)

$$F(q)^k = \sum_{n\geqslant 0} r_k(n) q^n,$$

where

$$r_k(n) = |\{(m_1, \dots, m_k) \in \mathbb{Z} : m_1^2 + \cdots + m_k^2 = n\}|$$

is the number of ways of writing n as a sum of k integer squares. The function $F(q)$ has nice analytic properties (it is an example of a modular form) that permit one to study the number of representations $r_k(n)$. On the other hand, as we will see later, this function (up to a change of variables) is strongly connected to the Riemann zeta function $\zeta(s)$.

2. Dirichlet Series

Dirichlet series are to arithmetic functions as power series are to sequences of numbers. Let $f \in \mathscr{A}$ be an arithmetic function. The *Dirichlet series* associated to f is the series in the complex variable s given by

$$s \mapsto L(s,f) = \sum_{n\geqslant 1} \frac{f(n)}{n^s}.$$

Definition 3.1. *An arithmetic function $f : \mathbb{N}_{\geqslant 1} \to \mathbb{C}$ is of polynomial growth if it satisfies one of the following equivalent conditions.*

- *There exists a constant $A \in \mathbb{R}$ (depending on f) such that $|f(n)| = O(n^A)$.*
- *There exists $\sigma \in \mathbb{R}$ such that the series $L(\sigma, f)$ is absolutely convergent.*

In this case we write

$$\sigma_f = \inf\{\sigma \in \mathbb{R} : L(\sigma,f) \text{ converges absolutely }\} \in \mathbb{R} \cup \{-\infty\};$$

The number σ_f is called the abscissa of convergence of $L(s,f)$.

We leave the equivalence of the above two conditions as an exercise.

The abscissa of convergence is the analogue for Dirichlet series of the radius of convergence for power series.

Proposition 3.2. *Let f be an arithmetic function with polynomial growth, and let σ_f be its abscissa of convergence. For all $\sigma > \sigma_f$, the series $L(s,f)$ converges absolutely and uniformly in the half-plane $\{s \in \mathbb{C} : \mathrm{Re}(s) \geqslant \sigma\}$. In this domain, the derivative of $L(s,f)$ is the Dirichlet series of the arithmetic function*

$$-\log.f : n \mapsto -\log(n) f(n),$$

that is to say,

$$L'(s,f) = L(s, -\log.f) = \sum_{n\geqslant 1} \frac{-\log(n) f(n)}{n^s},$$

which has abscissa of convergence σ_f as well.

Proof. Let $\sigma > \sigma_f$. Then uniformly for $\mathrm{Re}(s) \geqslant \sigma$, we have

$$\frac{|f(n)|}{|n^s|} \leqslant \frac{|f(n)|}{n^\sigma}$$

and thus the sum of holomorphic functions $\sum_{n\geqslant 1} f(n) n^{-s}$ is uniformly and absolutely convergent. It is bounded above by the absolutely convergent series $\sum_{n\geqslant 1} |f(n)| n^{-\sigma}$, and so converges uniformly in $\mathrm{Re}(s) \geqslant \sigma$, and defines a holomorphic function on $\mathrm{Re}(s) \geqslant \sigma$. It follows from the Cauchy integral formula that the derivative $L'(s,f)$ is given by the sum of the derivatives, that is

$$L'(s,f) = \sum_{n\geqslant 1} f(n) \frac{d}{ds}(n^{-s}) = \sum_{n\geqslant 1} \frac{-\log(n) f(n)}{n^s} = L(s, -\log.f).$$

This Dirichlet series is therefore convergent for s such that $\mathrm{Re}(s) > \sigma_f$. Applying the same reasoning as above to the arithmetic function $n \mapsto -\log(n) f(n)$, we see that $L(s, -\log.f)$ is absolutely convergent for $\mathrm{Re}(s) > \sigma_f$. In other words, $\sigma_{-\log.f} \leqslant \sigma_f$. On the other hand, as $|f(n)| \log(n) \geqslant |f(n)|$ for $n \geqslant 3$, $\sigma_{-\log.f} \geqslant \sigma_f$. Therefore $\sigma_f = \sigma_{-\log.f}$. □

The following lemma is very useful: it shows that an arithmetic function of polynomial growth is determined by its Dirichlet series.

Lemma 3.3. *Let f,g be two arithmetic functions of polynomial growth. We suppose that for all s contained in a non-empty open subset of the half-plane $\{s : \mathrm{Re}(s) > \max(\sigma_f,\sigma_g)\}$ we have $L(s,f) = L(s,g)$. Then $f = g$.*

PROOF. We may assume without loss of generality that $g \equiv 0$ by replacing f by $f-g$ and g by 0. Therefore, we may assume that $L(s,f) = 0$ for s in an open subset of the half-plane $\{s : \mathrm{Re}(s) > \sigma_f\}$. As this function is holomorphic, it is identically zero in the half-plane. By replacing f by

$$f.\mathrm{Id}^{-(\sigma_f-1)} : n \mapsto f(n)n^{-(\sigma_f-1)}$$

we may assume without loss of generality that $\sigma_f \leqslant 1$ and in particular that $f(n) = O(1)$. Suppose that $f \not\equiv 0$ and let n_0 be the smallest integer such that $f(n_0) \neq 0$. We have for $\mathrm{Re}(s) > 2$ that

$$0 = \sum_{n \geqslant n_0} \frac{f(n)}{n^s} = \frac{f(n_0)}{n_0^s}\left(1 + \frac{n_0^s}{f(n_0)} \sum_{n \geqslant n_0+1} \frac{f(n)}{n^s}\right),$$

so we have

$$0 = 1 + \frac{n_0^s}{f(n_0)} \sum_{n \geqslant n_0+1} \frac{f(n)}{n^s}$$

because $f(n_0)/n_0^s$ does not vanish. Since f is bounded, we see (by comparison with an integral) that for $\mathrm{Re}(s) > 2$

$$\sum_{n \geqslant n_0+1} \frac{f(n)}{n^s} = O\left(\frac{1}{\mathrm{Re}(s-1)(n_0+1)^{\mathrm{Re}(s)-1}}\right)$$

and so

$$0 = 1 + \frac{n_0^s}{f(n_0)} \sum_{n \geqslant n_0+1} \frac{f(n)}{n^s} = 1 + O\left(\frac{1}{\mathrm{Re}(s-1)}\right),$$

which is a contradiction when $\mathrm{Re}(s) \to \infty$. □

The main reason to introduce Dirichlet series is the following.

Theorem 3.4. *Let $f,g \in \mathscr{A}$, with $\sigma_f, \sigma_g < \infty$. Then, $\sigma_{f*g} \leqslant \max(\sigma_f,\sigma_g)$, and for $\mathrm{Re}(s) > \max(\sigma_f,\sigma_g)$ we have*

$$L(s, f*g) = L(s,f)L(s,g).$$

PROOF. Let $\mathrm{Re}(s) > \max(\sigma_f,\sigma_g)$, so that

$$\begin{aligned}
\sum_{n \geqslant 1} \frac{|f*g(n)|}{|n^s|} &= \sum_{n \geqslant 1} \frac{\left|\sum_{ab=n} f(a)g(b)\right|}{n^{\mathrm{Re}(s)}} \\
&\leqslant \sum_{n \geqslant 1} \sum_{ab=n} \frac{|f(a)||g(b)|}{(ab)^{\mathrm{Re}(s)}} \\
&= \sum_{a,b \geqslant 1} \frac{|f(a)||g(b)|}{(ab)^{\mathrm{Re}(s)}} = \sum_{a \geqslant 1} \frac{|f(a)|}{a^{\mathrm{Re}(s)}} \sum_{b \geqslant 1} \frac{|f(b)|}{b^{\mathrm{Re}(s)}} < \infty.
\end{aligned}$$

All of the above identities and swaps of order of summation above are justified by the fact that we are summing positive terms. We have thus shown that $\sigma_{f*g} \leqslant \max(\sigma_f,\sigma_g)$. Moreover, for $\mathrm{Re}(s) > \max(\sigma_f,\sigma_g)$, we have by absolute convergence that we can regroup the terms arbitrarily, and so we have

$$L(s,f)L(s,g) = L(s, f*g).$$

□

Corollary 3.5. *Suppose that f is invertible for the Dirichlet convolution, and that σ_f, $\sigma_{f^{(-1)}} < \infty$, so that for $\mathrm{Re}(s) > \max(\sigma_f, \sigma_{f^{(-1)}})$ we have*

$$L(s,f)L(s,f^{(-1)}) = 1.$$

In particular, $L(s,f)$ does not vanish in the half-plane $\{s : \mathrm{Re}(s) > \max(\sigma_f, \sigma_{f^{(-1)}})\}$.

2.1. Examples.

- For $f \equiv 1$ we have
$$L(s,1) = \sum_{n\geqslant 1} \frac{1}{n^s} = \zeta(s),$$
and $\sigma_1 = 1$. The Möbius function is the inverse of the 1 function. We saw that $|\mu(n)| \leqslant 1$ and so $\sigma_\mu \leqslant 1$. Moreover, we have
$$\sum_{n\geqslant 1} \frac{|\mu(n)|}{n} \geqslant \sum_{p\geqslant 2} \frac{1}{p} = \infty.$$
So we have that $\sigma_\mu = 1$. We therefore see that for $\mathrm{Re}(s) > 1$, the function $\zeta(s)$ does not vanish, and that
$$L(s,\mu) = \sum_{n\geqslant 1} \frac{\mu(n)}{n^s} = \frac{1}{\zeta(s)}.$$
- Let d be the divisor function: we have that $d = 1 * 1$, and so $\sigma_d \leqslant 1$. Since $d(n) \geqslant 1$, we have $\sigma_d = 1$. For $\mathrm{Re}(s) > 1$ we have
$$L(s,d) = \zeta(s)^2.$$
- The Euler function φ is $\mu * \mathrm{Id}$ and as $\sigma_{\mathrm{Id}} = 2$, and $\sigma_\mu = 1$, we have $\sigma_\varphi \leqslant 2$. For $\mathrm{Re}(s) > 2$ we have
$$L(s,\varphi) = L(s,\mu)L(s,\mathrm{Id}) = \frac{\zeta(s-1)}{\zeta(s)}.$$
- The von Mangolt function $\Lambda = \log * \mu$. We have
$$L(s,\Lambda) = L(s,\log)L(s,\mu) = -\frac{\zeta'(s)}{\zeta(s)}.$$
We have $\sigma_\Lambda \leqslant 1$ and seeing as the series $\sum_{n\geqslant 1} \frac{\Lambda(n)}{n}$ diverges, $\sigma_\Lambda = 1$. As $\zeta(s)$ doesn't vanish for $\mathrm{Re}(s) > 1$, we have $\frac{\zeta'(s)}{\zeta(s)}$ is holomorphic on that domain. Since $\{s : \mathrm{Re}(s) > 1\}$ is simply connected, the primitive of $\frac{\zeta'(s)}{\zeta(s)}$ exists and is holomorphic, and this function is the logarithm $\log\zeta(s)$. We have
$$(\log\zeta(s))' = \frac{\zeta'(s)}{\zeta(s)},$$
i.e. the logarithmic derivative of $\zeta(s)$.

3. Dirichlet series and multiplicative functions

Theorem 3.6. *Let $f \in \mathscr{A}$ be a multiplicative function of polynomial growth, then for all $\sigma > \sigma_f$ we have*

(1) *For all p prime, the series*

$$L_p(s,f) := \sum_{\alpha \geqslant 0} \frac{f(p^\alpha)}{p^{\alpha s}}$$

converges absolutely and uniformly in the half plane $\mathrm{Re}(s) \geqslant \sigma$. We call $L_p(s,f)$ the local factor of f at p.

(2) *Moreover, we have*

$$L(s,f) = \prod_p L_p(s,f) = \lim_{P \to \infty} \prod_{p \leqslant P} L_p(s,f),$$

and the convergence is uniform in this half-plane.

(3) *More precisely, if we write*

$$L^{>P}(s,f) = \lim_{P' \to \infty} \prod_{P < p \leqslant P'} L_p(s,f),$$

then as $P \to \infty$ we have

$$L^{>P}(s,f) \to 1$$

uniformly in every half-plane $\mathrm{Re}(s) \geqslant \sigma$, $\sigma > \sigma_f$.

(4) *Conversely, if f is an arithmetic function such that $\sigma_f < \infty$ and $f(1) = 1$ and if $L(s,f)$ satisfies*

$$L(s,f) = \prod_p L_p(s,f) = \lim_{P \to \infty} \prod_{p \leqslant P} L_p(s,f)$$

for s sufficiently large, then f is multiplicative.

PROOF. (1) Let $\mathrm{Re}(s) \geqslant \sigma$, then

$$\sum_{\alpha \geqslant 0} \frac{|f(p^\alpha)|}{|p^{\alpha s}|} \leqslant \sum_{n \geqslant 1} \frac{|f(n)|}{n^\sigma} < \infty,$$

from which follows the absolute and uniform convergence of $L_p(s,f)$.

(2) For $P \geqslant 2$ write

$$p_1 < p_2 < \cdots < p_k \leqslant P$$

for the set of prime numbers $\leqslant P$, so

$$\prod_{p \leqslant P} L_p(s,f) = \sum_{\alpha_1, \ldots, \alpha_k \geqslant 0} \frac{f(p_1^{\alpha_1}) \cdots f(p_k^{\alpha_k})}{(p_1^{\alpha_1} \cdots p_k^{\alpha_k})^s} = \sum_{\alpha_1, \ldots, \alpha_k \geqslant 0} \frac{f(p_1^{\alpha_1} \cdots p_k^{\alpha_k})}{(p_1^{\alpha_1} \cdots p_k^{\alpha_k})^s} = \sum_{\substack{n \geqslant 1 \\ p \mid n \Rightarrow p \leqslant P}} \frac{f(n)}{n^s}.$$

Thus, an integer that doesn't appear in the previous sum has at least one prime divisor $> P$ and so

$$\left| L(s,f) - \prod_{p \leqslant P} L_p(s,f) \right| \leqslant \sum_{n > P} \frac{|f(n)|}{n^\sigma} \to 0,$$

as $P \to \infty$ with uniform convergence.

(3) We will now show that as $P \to \infty$ that

$$L^{>P}(s,f) = \prod_{p > P} L_p(s,f)$$

tends to 1 uniformly in every half-plane $\mathrm{Re}(s) \geqslant \sigma > \sigma_f$. We have $\prod_{p>P} L_p(s,f)$ is the Dirichlet series of the multiplicative function f_P, which is 0 if n has a prime factor $\leqslant P$ and $f(n)$ if not (if $n = 1$, then $f(1) = 1$ since 1 doesn't have any prime factors,

and in particular has no prime factors $\leqslant P$). It is clear that $\sigma_{f_P} \leqslant \sigma_f$. We have $L^{>P}(s,f) = L(s,f_P)$ and

$$L(s,f_P) = \prod_{p>P} L_p(s,f) = \sum_{\substack{n\geqslant 1\\ p|n\Rightarrow p>P}} \frac{f(n)}{n^s} = 1 + \sum_{\substack{n>1\\ p|n\Rightarrow p>P}} \frac{f(n)}{n^s}$$

and so

$$\left|\prod_{p>P} L_p(s,f) - 1\right| \leqslant \sum_{\substack{n>1\\ p|n\Rightarrow p>P}} \frac{|f(n)|}{n^\sigma} \leqslant \sum_{n>P} \frac{|f(n)|}{n^\sigma} \to 0.$$

(4) Conversely, let $\widetilde{f}$ be the unique multiplicative function defined by

$$\widetilde{f}(n) = \prod_{p^\alpha \| n} f(p^\alpha).$$

Then $\sigma_{\widetilde{f}} \leqslant \sigma_f + 1$. Indeed, for $\sigma > \sigma_f$, we have $|f(n)|/n^\sigma = o(1)$ and thus for n sufficiently large (say, $n > N$) we have $|f(n)|/n^\sigma < 1$. From this we deduce

$$\frac{|\widetilde{f}(n)|}{n^\sigma} = \prod_{p^\alpha\|n} \frac{|f(p^\alpha)|}{p^{\alpha\sigma}} = \prod_{\substack{p^\alpha\|n\\ p^\alpha\leqslant N}} \frac{|f(p^\alpha)|}{p^{\alpha\sigma}} \prod_{\substack{p^\alpha\|n\\ p^\alpha> N}} \frac{|f(p^\alpha)|}{p^{\alpha\sigma}} < \prod_{\substack{p^\alpha\|n\\ p^\alpha\leqslant N}} \frac{|f(p^\alpha)|}{p^{\alpha\sigma}} = O_f(1),$$

since there are only finitely many (the number depends on $N = N(f)$) values of $f(p^\alpha)/p^{\alpha\sigma}$ as factors in this last product. Thus, $L(s,\widetilde{f})$ converges absolutely for $\mathrm{Re}(s) > \sigma + 1$ and so

$$L(s,\widetilde{f}) = \prod_p L_p(s,\widetilde{f}) = \prod_p L_p(s,f).$$

We therefore have for s with $\mathrm{Re}(s)$ sufficiently large that

$$L(s,\widetilde{f}) = \sum_{n\geqslant 1} \frac{\widetilde{f}(n)}{n^s} = L(s,f) = \sum_{n\geqslant 1} \frac{f(n)}{n^s},$$

which by Lemma 3.3 implies that for all $n \geqslant 1$ that $f(n) = \widetilde{f}(n)$.

□

Corollary 3.7. *If f is completely multiplicative, then for $\mathrm{Re}(s) > \sigma_f$ we have*

$$L(s,f) = \prod_p (1 - \frac{f(p)}{p^s})^{-1}.$$

These factorization results into Euler products allow a further localization of the zeros of $L(s,f)$.

Proposition 3.8. *Let $f \in \mathscr{A}^\times$ be a multiplicative function. Let $\sigma > \sigma_f$. The number of local factors $L_p(s,f)$ that have a zero in the half-plane $\mathrm{Re}(s) \geqslant \sigma$ is finite and the zeros of $L(s,f)$ in this half-plane are exactly the zeros of the local factors $L_p(s,f)$. More precisely, there exists P such that*

$$\prod_{p>P} L_p(s,f) = \frac{L(s,f)}{\prod_{p\leqslant P} L_p(s,f)}$$

does not vanish in $\mathrm{Re}(s) \geqslant \sigma$. Thus, the zeros of $L(s,f)$ in this half-plane of absolute convergence are exactly the zeros of the local factors $L_p(s,f)$.

PROOF. Let $\sigma > \sigma_f$. We saw that if p is sufficiently large, for all s with $\mathrm{Re}(s) \geqslant \sigma$ that

$$|L_p(s,f) - 1| \leqslant \frac{1}{2},$$

thus the number of local factors that vanish in the half-plane $\mathrm{Re}(s) \geqslant \sigma$ is finite. On the other hand, we saw previously that if P is sufficiently large and $\mathrm{Re}(s) \geqslant \sigma$ that

$$|L^{>P}(s,f) - 1| \leqslant \sum_{n>P} \frac{|f(n)|}{n^\sigma} < 1/2,$$

so $L^{>P}(s,f)$ does not vanish in the half-plane $\mathrm{Re}(s) \geqslant \sigma$. □

3.1. Examples. For $\mathrm{Re}(s) > 1$, we have

$$\zeta(s) = \sum_{n\geqslant 1} \frac{1}{n^s} = \prod_p (1 - \frac{1}{p^s})^{-1}$$

and

$$\frac{1}{\zeta(s)} = \sum_{n\geqslant 1} \frac{\mu(n)}{n^s} = \prod_p (1 - \frac{1}{p^s}).$$

The function $\zeta(s)$ does not vanish for $\mathrm{Re}(s) > 1$. Moreover, we have for $\mathrm{Re}(s) > 1$ that

$$L(s,d) = \sum_{n\geqslant 1} \frac{d(n)}{n^s} = \zeta^2(s) = \prod_p (1 - \frac{1}{p^s})^{-2},$$

and

$$L(s, d^{(-1)}) = \zeta(s)^{-2} = \prod_p (1 - \frac{2}{p^s} + \frac{1}{p^{2s}}).$$

Now, if $\mathrm{Re}(s) > 2$ we have

$$L(s,\varphi) = \frac{\zeta(s-1)}{\zeta(s)} = \prod_p (\frac{1-p^{-s}}{1-p^{1-s}}) = \prod_p (1 + (1 - \frac{1}{p}) \sum_{\alpha\geqslant 1} \frac{1}{p^{\alpha(s-1)}}).$$

Here is a more interesting example. Let $\mu_2 \in \mathscr{A}^\times$ be the multiplicative function defined on prime powers by

$$\mu_2(p^\alpha) = \begin{cases} 1 & \text{if } \alpha = 0 \\ -1 & \text{if } \alpha = 1 \text{ and } p \neq 2 \\ -4 & \text{if } \alpha = 1 \text{ and } p = 2 \\ 0 & \text{if } \alpha \geqslant 2. \end{cases}$$

That is to say, μ_2 is similar to the Möbius function, but where the value at 2 has been modified. Note that $|\mu_2(n)| \leqslant 4$ for all $n \geqslant 1$, so that $\sigma_{\mu_2} \leqslant 1$. We have

$$L_2(s,\mu_2) = 1 - \frac{4}{2^s},$$

which has a zero at $s = 2$, and whenever $\mathrm{Re}(s) > 1$ that

$$L(s,\mu_2) = \sum_{n\geqslant 1} \frac{\mu_2(n)}{n^s} = (1 - \frac{4}{2^s}) \prod_{p\neq 2} (1 - \frac{1}{p^s}).$$

Therefore the only zero of $L(s,\mu_2)$ in $\mathrm{Re}(s)>1$ is at $s=2$. Let us now check that this does not contradict Corollary 3.5. We compute $\mu_2^{(-1)}$. By Proposition 2.17, $\mu_2^{(-1)}$ is multiplicative, so it suffices to compute it on prime powers. Computing using the recursive definition we find

$$\mu_2^{(-1)}(p^\alpha)=\begin{cases}1 & \text{if } p\neq 2\\ 4^\alpha & \text{if } p=2.\end{cases}$$

Let $\sigma>2$. We have for s with $\mathrm{Re}(s)\geqslant\sigma$ that

$$|L(s,\mu_2^{(-1)})|\leqslant\sum_{n\geqslant 1}\frac{|\mu_2^{(-1)}(n)|}{n^\sigma}=(1-\frac{4}{2^\sigma})^{-1}\prod_{p\neq 2}(1-\frac{1}{p^\sigma})^{-1}=\frac{(1-\frac{1}{2^\sigma})}{(1-\frac{4}{2^\sigma})}\zeta(\sigma)<\infty,$$

so $\sigma_{\mu_2^{(-1)}}\leqslant 2$. On the other hand,

$$\sum_{n\geqslant 1}\frac{|\mu_2^{(-1)}(n)|}{n^2}\geqslant\sum_{\alpha\geqslant 0}\frac{\mu_2^{(-1)}(2^\alpha)}{2^{2\alpha}}=\sum_{\alpha\geqslant 0}1=\infty.$$

So $\sigma_{\mu_2^{(-1)}}=2$. Therefore the example of μ_2 does not contradict Corollary 3.5.

CHAPTER 4

Primes in Arithmetic Progressions

Definition 4.1. *An arithmetic progression is a doubly-infinite subset of $\mathbb{Z}$ satisfying the following property: There exists a positive integer $q > 0$ such that the distance between two consecutive integers of this subset is always q. The integer q is called the modulus of the arithmetic progression.*

It is easy to see that arithmetic progressions of modulus q are of the form

$$L_{q,a} = a + q\mathbb{Z} \subseteq \mathbb{Z},$$

where a is an integer. We remark that if $a \equiv a' \pmod{q}$, that we have $L_{q,a} = L_{q,a'}$. Thus arithmetic progressions of modulus q are indexed by the congruence classes modulo q (i.e. by the ring $\mathbb{Z}/q\mathbb{Z}$). There are therefore q of them.

Thus n belongs to $L_{q,a}$ if and only if $n \equiv a \pmod{q}$. The class $\overline{a}$ of a modulo q is called the class of the arithmetic progression. We will write indifferently $L_{q,a}$ or $L_{q,\overline{a}}$ and by abuse of language, we will talk about the integer a of the class of the arithmetic progresssion.

As $L_{q,a}$ is infinite, it is natural to ask oneself if its intersection with the prime numbers $\mathscr{P}$ is as well. That is almost always the case.

Theorem 4.2 (Dirichlet's theorem on primes in arithmetic progressions)**.** *Let $a, q > 0$ be two relatively prime integers. Then, the set*

$$\mathscr{P}_{q,a} = \mathscr{P} \cap L_{q,a}$$

is infinite. Said differently, there exist infinitely many prime numbers $p \equiv a \pmod{q}$.

Remark 4.3. The condition that a and q be relatively prime is necessary. Indeed, if $(a, q) \neq 1$, then there exists at most one prime p of the form $a + qn$, and the only possibility is $p = (a, q)$. In other words the congruence classes containing an infinite number of primes are exactly those of $(\mathbb{Z}/q\mathbb{Z})^\times$.

In the vein of the prime number theorem, we can pose more precise questions on the density of the set $\mathscr{P}_{q,a}$. We therefore set

$$\pi(x; q, a) = |\mathscr{P}_{q,a} \cap [1, x]| = |\{p \leqslant x : p \equiv a \pmod{q}\}| = \sum_{\substack{p \equiv a \pmod{q} \\ p \leqslant x}} 1$$

the counting function of the primes $p \equiv a \pmod{q}$. At the beginning of the 20th century, Landau showed this generalization of the prime number theorem:

Theorem 4.4 (Landau)**.** *Let $a, q > 0$ be relatively prime integers. Then*

$$\pi(x; q, a) = \frac{1}{\varphi(q)}\pi(x)(1 + o(1)) = \frac{1}{\varphi(q)}\frac{x}{\log x}(1 + o(1)).$$

If $p > q$, then p necessarily occurs in some arithmetic progression modulo q for which the class is relatively prime to q. Dirichlet's theorem says that each of these congruence classes is attained infinitely often. Landau's theorem says that the asymptotic proportion of prime numbers falling into each of these classes does not depend on that class. That is to say, asymptotically there are no "privileged" congruence classes. In this chapter, we will not prove Landau's theorem, but instead we will more simply give the analogue of Merten's theorem.

Theorem 4.5. *We have*

$$\sum_{\substack{n\leqslant x\\ n\equiv a \pmod q}} \frac{\Lambda(n)}{n} = \frac{1}{\varphi(q)}\log(x) + O(1) \tag{4.1}$$

$$\sum_{\substack{p\leqslant x\\ p\equiv a \pmod q}} \frac{\log p}{p} = \frac{1}{\varphi(q)}\log(x) + O(1), \tag{4.2}$$

$$\sum_{\substack{p\leqslant x\\ p\equiv a \pmod q}} \frac{1}{p} = \frac{1}{\varphi(q)}\log\log(x) + O(1). \tag{4.3}$$

The method (due to Dirichlet) will consist of finding a nice expression (from an analytic point of view) for the congruence condition

$$n \equiv a \pmod q,$$

then treating it in the sum over prime numbers. The crucial point is the group structure of the set $(\mathbb{Z}/q\mathbb{Z})^\times$.

1. Characters of a finite abelian group

Let G be a finite abelian group, we write e for the identity element. Let

$$\mathscr{C}(G) = \mathbb{C}^G = \{f : G \to \mathbb{C}\},$$

the vector space of functions from G to $\mathbb{C}$. It is a complex vector space (and even an algebra under multiplication of two functions) of dimension $|G|$. A $\mathbb{C}$-basis of $\mathscr{C}(G)$ is given by the set

$$\mathscr{B}_0 = \{\delta_g : g \in G\}, \text{ where } \delta_g(g') = \begin{cases} 1 & g = g' \\ 0 & g \neq g'. \end{cases}$$

The point is that for any $f \in \mathscr{C}(G)$, we have the decomposition

$$f = \sum_g f(g)\delta_g,$$

so $\mathscr{B}_0$ spans $\mathscr{C}(G)$. It is also easy to see that $\mathscr{B}_0$ is linearly independent.

The space $\mathscr{C}(G)$ is equipped with a hermitian scalar product

$$\langle f, f'\rangle = \frac{1}{|G|}\sum_{g\in G} f(g)\overline{f'}(g) \tag{4.4}$$

for which the basis $\mathscr{B}_0$ is orthogonal:

$$\langle \delta_g, \delta_{g'}\rangle = \frac{1}{|G|}\delta_{g=g'} = \frac{1}{|G|}\begin{cases} 1 & g = g' \\ 0 & g \neq g'. \end{cases}$$

Note that the above facts hold for G any finite set. That is to say, we have not yet used the group structure of G at all. On the other hand, we will soon see that for G an abelian group, $\mathscr{C}(G)$ possesses a canonical orthonormal basis coming from the structure of the group G.

As G is a group, it acts on itself by right translations. This action induces an action of G on $\mathscr{C}(G)$: to each element g of G we associate the endomorphism $T_g \in \mathrm{End}(\mathscr{C}(G))$ defined by

$$T_g : f \mapsto T_g f, \text{ where } T_g f : g' \mapsto T_g f(g') = f(g'g).$$

We can verify easily that

$$T_g \circ T_{g'} = T_{gg'}, \text{ and } T_e = \mathrm{Id}_{\mathscr{C}(G)},$$

from which we deduce that T_g is invertible, the inverse being

$$(T_g)^{-1} = T_{g^{-1}}$$

and thus the operator

$$T : g \mapsto T_g$$

in fact defines a group homomorphism

$$T : G \to \mathrm{Aut}(\mathscr{C}(G)).$$

Moreover, as G is abelian, the elements T_g, $g \in G$ commute:

$$T_g \circ T_{g'} = T_{gg'} = T_{g'g} = T_{g'} \circ T_g.$$

What is more, we have

$$\langle T_g f, T_g f' \rangle = \frac{1}{|G|} \sum_{g' \in G} f(g'g)\overline{f'}(g'g) = \frac{1}{|G|} \sum_{g \in G} f(g)\overline{f'}(g) = \langle f, f' \rangle.$$

Thus $\{T_g : g \in G\}$ is a set of commuting isometries for the scalar product (4.4). Recall the spectral theorem.

THEOREM (Spectral Theorem). *Let $(V, \langle , \rangle)$ be a finite-dimensional hermitian vector space, and $\mathscr{T} \subset \mathrm{End}(V)$ a set of pairwise commuting endomorphisms. Then there exists an orthonormal basis of V given by simultaneous eigenvectors of all of the elements of $\mathscr{T}$ if and only if each $T \in \mathscr{T}$ is normal (that is to say, $TT^* = T^*T$).*

In the situation at hand, the spectral theorem implies that $\mathscr{C}(G)$ possesses an orthonormal basis of eigenvectors of all of the T_g. In fact, up to permutation, this basis is unique.

Theorem 4.6. *There exists a unique orthonormal basis $\widehat{G}$ of $\mathscr{C}(G)$ consisting of eigenvectors for all of the T_g, such that for all $\chi \in \widehat{G}$ we have $\chi(e) = 1$. We have an equality*

$$\widehat{G} = \mathrm{Hom}_{\mathbb{Z}}(G, \mathbb{C}^\times),$$

where $\mathrm{Hom}_{\mathbb{Z}}(G, \mathbb{C}^\times)$ designates the set of group homomorphisms from G to $\mathbb{C}^\times$. This set is a group: the group of characters of G, or also the dual of G.

PROOF. The set $\{T_g : g \in G\}$ is a family of unitary operators, so $T_g^* = T_{g^{-1}}$, and therefore normal and pairwise commuting. By the spectral theorem, this family is diagonalizable with respect to an orthonormal basis. Let $\mathscr{B}$ be such a basis. The $\psi \in \mathscr{B}$ are thus non-zero eigenvectors of all of the T_g and we have for all g that

$$T_g \psi(x) = \psi(xg) = \chi_\psi(g)\psi(x), \tag{4.5}$$

where $\chi_\psi(g)$ denotes the eigenvalue of T_g associated to the eigenvalue ψ. We therefore associate to every ψ a function $\chi_\psi \in \mathscr{C}(G)$ defined by

$$\chi_\psi : g \mapsto \chi_\psi(g).$$

I claim that $\widehat{G} = \{\chi_\psi\}_{\psi \in \mathscr{B}}$ is the basis that we are looking for. We have

$$T_e\psi = \mathrm{Id}(\psi) = \chi_\psi(e)\psi = \psi,$$

$$T_g \circ T_{g'}\psi = T_{gg'}\psi = \chi_\psi(g)\chi_\psi(g')\psi = \chi_\psi(gg')\psi$$

$$T_g^{-1}\psi = T_{g^{-1}}\psi = \chi_\psi(g^{-1})\psi = (\chi_\psi(g))^{-1}\psi,$$

thus (since $\psi \neq 0$) we have for all $g, g' \in G$ that

$$\chi_\psi(e) = 1, \quad \chi_\psi(gg') = \chi_\psi(g)\chi_\psi(g'), \quad \chi_\psi(g^{-1}) = \chi_\psi(g)^{-1}.$$

In other words, χ_ψ is a group homomorphism from G to $\mathbb{C}^\times$. We still need to show that $\widehat{G}$ is an orthonormal basis of eigenvectors for $\mathscr{C}(G)$. Up to this point, we only know that $\mathscr{B}$ is such a basis.

Note also that as G is finite, we have by Lagrange's theorem that for all $\chi \in \mathrm{Hom}_{\mathbb{Z}}(G, \mathbb{C}^\times)$,

$$\chi(g^{|G|}) = \chi(e) = 1 = \chi(g)^{|G|},$$

and so χ takes values in $\mu_{|G|} \subset \mathbb{C}^\times$, the set of $|G|$th roots of unity. In particular, $\chi_\psi(g)$ is a complex number of modulus 1 for all g and

$$\langle \chi_\psi, \chi_\psi \rangle = \frac{1}{|G|} \sum_{g \in G} \chi_\psi(g)\overline{\chi_\psi(g)} = \frac{1}{|G|} \sum_{g \in G} 1 = 1. \tag{4.6}$$

As ψ is non-zero, there exists g be such that $\psi(g) \neq 0$. By (4.5), we have for all g that $\psi(g) = \chi_\psi(g)\psi(e)$ and so $\psi(e) \neq 0$, thus

$$\chi_\psi(g) = \frac{1}{\psi(e)}\psi(g), \text{ and so } \chi_\psi = \frac{1}{\psi(e)}\psi.$$

That is to say, χ_ψ belongs to the subspace $\mathbb{C}.\psi$ generated by ψ. Seeing as all of the χ_ψ are non-zero (recall $\chi_\psi(e) = 1$), the family $\{\chi_\psi\} = \widehat{G}$ is an orthonormal basis of $\mathscr{C}(G)$ which is contained in $\mathrm{Hom}(G, \mathbb{C}^\times)$.

Now we show conversely that every element of $\mathrm{Hom}_{\mathbb{Z}}(G, \mathbb{C}^\times)$ belongs to $\widehat{G}$. Let $\psi \in \mathrm{Hom}_{\mathbb{Z}}(G, \mathbb{C}^\times)$. Then $\psi \in \mathscr{C}(G)$ and as $\psi \not\equiv 0$ (we have $\psi(e) = 1$), there exists $\chi \in \widehat{G}$ such that $\langle \psi, \chi \rangle \neq 0$. Thus, we have for all g

$$0 \neq \langle \psi, \chi \rangle = \langle T_g\psi, T_g\chi \rangle = \frac{1}{|G|} \sum_{g' \in G} \psi(g'g)\overline{\chi}(g'g) = \psi(g)\overline{\chi}(g)\langle \psi, \chi \rangle,$$

from which it follows that for all g we have $\psi(g)\overline{\chi(g)} = 1$. That is, $\psi(g) = \chi(g)$. □

1.1. Properties of characters. Let us extract from the above proof the following important properties.

1.1.1. *Group structure.* The set $\widehat{G} = \mathrm{Hom}_{\mathbb{Z}}(G, \mathbb{C}^\times)$ has a natural group structure (called the dual group of G) via multiplication of functions. For all $\chi, \chi' \in \widehat{G}$, we define

$$\chi.\chi' : g \mapsto \chi(g)\chi'(g), \quad \chi^{-1} : g \mapsto \chi(g)^{-1}.$$

The functions $\chi.\chi'$ and χ^{-1} are clearly themselves characters. The identity element of $\widehat{G}$ is the constant function 1, which we write

$$\chi_0 : g \mapsto 1$$

and which we call the *trivial character*.

Said differently, in a pedantic fashion, $\widehat{G}$ is a subgroup of the group of units of the algebra $\mathscr{C}(G)$ with the structure of multiplication of functions $f, f' \in \mathscr{C}(G)$ given by

$$f.f' : g \mapsto f(g)f'(g).$$

1.1.2. *Unitary structure.* We showed that

$$|\widehat{G}| = \dim \mathscr{C}(G) = |G|.$$

In particular, (Lagrange's theorem) for all $\chi \in \widehat{G}$, we have $\chi^{|G|} = \chi_0$, that is to say that for all $g \in G$

$$\chi(g)^{|G|} = 1,$$

said otherwise, characters take their values in the $|G|$th roots of unity

$$\mu_{|G|} = \{\zeta \in \mathbb{C} : \zeta^{|G|} = 1\}.$$

We say that they are *unitary*. In particular, for all g

$$\chi(g)^{-1} = \overline{\chi(g)}, \text{ i.e. } \chi^{-1} = \overline{\chi},$$

the complex conjugate of χ.

1.1.3. *Orthogonality relations.*

Proposition 4.7. *We have*

$$\delta_{\chi=\chi'} = \frac{1}{|G|} \sum_{g' \in G} \chi(g')\overline{\chi'}(g') \tag{4.7}$$

$$\delta_{g=g'} = \frac{1}{|G|} \sum_{\chi \in \widehat{G}} \overline{\chi}(g)\chi(g') \tag{4.8}$$

Proof. The first relation is the orthogonality

$$\langle \chi, \chi' \rangle = \delta_{\chi=\chi'}.$$

To show the second relation, let us consider the decomposition of δ_g in the basis $\widehat{G}$:

$$\delta_g = \sum_\chi \langle \delta_g, \chi \rangle \chi = \frac{1}{|G|} \sum_\chi \left(\sum_{g' \in G} \delta_g(g')\overline{\chi}(g') \right).\chi = \frac{1}{|G|} \sum_\chi \overline{\chi}(g)\chi. \tag{4.9}$$

Evaluating this identity at g', we obtain the relation (4.8). □

Remark 4.8. The expression (4.9) is the *Fourier decomposition* of the function δ_g. This expression will be very useful to us in what follows.

1.2. Characters of a cyclic group. Let $G = \langle g \rangle$ be a cyclic group generated by g. Let q be its order. A character χ of G is completely determined by its value on g, i.e. $\chi(g)$. Indeed, for all $g' \in G$, write $g' = g^m$ with $m \in \mathbb{Z}$ and $\chi(g') = \chi(g)^m$. Note that this does not depend on the choice of $m \in \mathbb{Z}$: if $g' = g^{m'}$ then $m' \equiv m \pmod{q}$ and

$$\chi(g^{m'}) = \chi(g)^m \chi(g)^{m'-m} = \chi(g^m),$$

since $\chi(g)$ is a qth root of unity. Thus all qth roots of unity ζ define uniquely a character of G by taking

$$\chi(g^m) = \zeta^m.$$

In other words, we have the following proposition.

Proposition 4.9. *Let G be a cyclic group of order q. The choice of a generator of G determines an isomorphism between the group of qth roots of unities and the group $\widehat{G}$. This isomorphism is given by*

$$\zeta \in \mu_q \mapsto \chi, \text{ where } \chi : g^m \mapsto \chi(g^m) = \zeta^m.$$

In the case of the group $\mathbb{Z}/q\mathbb{Z}$, the characters (called additive characters) are given explicitly by the functions

$$\psi_n : x \pmod{q} \mapsto e^{2\pi i n \frac{x}{q}}.$$

Note that ψ_n only depends on the class of n modulo q and

$$n \pmod{q} \mapsto \psi_n$$

is the searched for isomorphism.

Remark 4.10. Note that μ_q is a cyclic group of order q generated by the complex exponential

$$\zeta_q = e(\frac{1}{q}), \text{ where } e(x) = e^{2\pi i x}.$$

We therefore have

$$\widehat{G} \simeq \mu_q \simeq \mathbb{Z}/q\mathbb{Z} \simeq G.$$

Note that this isomorphism is not canonical since it depends on the choice of generator g. More generally, for any finite abelian group G we have $G \simeq \widehat{G}$.

2. Dirichlet Characters

Now we apply the above theory to the multiplicative groups $(\mathbb{Z}/q\mathbb{Z})^\times$.

Definition 4.11. *Let $q \geqslant 1$ be an integer. The characters of the abelian group $(\mathbb{Z}/q\mathbb{Z})^\times$ are called Dirichlet characters of modulus q. The trivial character (i.e. the constant function equal to 1) is denoted χ_0.*

2.1. The arithmetic function associated to a Dirichlet character. Let $\chi \in \widehat{(\mathbb{Z}/q\mathbb{Z})^\times}$ be a character. We extend this character by 0 to be a function on $\mathbb{Z}/q\mathbb{Z}$, and extend this function by periodicity to all of $\mathbb{Z}$. That is to say, we define an arithmetic function χ by

$$\chi(n) = \begin{cases} \chi(n \pmod{q}) & (n,q) = 1 \\ 0 & (n,q) > 1. \end{cases}$$

By abuse of language, the arithmetic function χ thus obtained is also called a Dirichlet character. In fact, when mathematicians talk about Dirichlet characters they most often mean

the arithmetic function and not the character of $(\mathbb{Z}/q\mathbb{Z})^\times$. In particular, $\chi \in \mathscr{A}$ is periodic of period q, vanishes on integers not relatively prime to q, and is completely multiplicative:

$$\chi(mn) = \chi(m)\chi(n) \quad \text{for all } m, n \in \mathbb{N}_{\geqslant 1}.$$

The orthogonality relations (4.7) and (4.8) are therefore written

$$\delta_{\chi=\chi'} = \frac{1}{\varphi(q)} \sum_{a \,(\mathrm{mod}\, q)} \chi(a)\overline{\chi'(a)}, \tag{4.10}$$

and for $(ab, q) = 1$

$$\delta_{a\equiv b \,(\mathrm{mod}\, q)} = \frac{1}{\varphi(q)} \sum_{\chi \,(\mathrm{mod}\, q)} \overline{\chi}(a)\chi(b), \tag{4.11}$$

where the sum runs over all Dirichlet characters modulo q.

Next we consider the Dirichlet series associated to a Dirichlet character

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

As $|\chi(n)| \leqslant 1$, this series converges absolutely for $\mathrm{Re}(s) > 1$, and in this domain we have

$$L(s, \chi) = \prod_p (1 - \frac{\chi(p)}{p^s})^{-1}.$$

Proposition 4.12. *Let $\chi \,(\mathrm{mod}\, q)$ be a Dirichlet character.*

- *If $\chi = \chi_0$, we have for $\mathrm{Re}(s) > 1$*

$$L(s, \chi_0) = \prod_{p|q} (1 - \frac{1}{p^s})\zeta(s),$$

  which admits an analytic continuation to the half plane $\mathrm{Re}(s) > 0$ with a simple pole at $s = 1$.
- *If $\chi \neq \chi_0$, the series $L(s, \chi)$ converges uniformly on compacta in the half-plane $\mathrm{Re}(s) > 0$ and thus defines a holomorphic function in this domain. More precisely, we have for $\mathrm{Re}(s) > 0$*

$$\sum_{1 \leqslant n \leqslant X} \frac{\chi(n)}{n^s} = L(s, \chi) + O(\frac{q|s|}{\sigma} X^{-\sigma}). \tag{4.12}$$

Proof. If $\chi = \chi_0$, then

$$L(s, \chi_0) = \prod_p (1 - \frac{\chi_0(p)}{p^s})^{-1} = \prod_{(p,q)=1} (1 - \frac{1}{p^s})^{-1} = \prod_{p|q} (1 - \frac{1}{p^s})\zeta(s).$$

Therefore the result follows from the analytic continuation of $\zeta(s)$ to the domain $\mathrm{Re}(s) > 0$ (see exercises).

If χ is non-trivial, we have for every interval I of the form $[\cdot, \cdot)$ of length exactly q

$$\sum_{n \in I} \chi(n) = 0.$$

Indeed, I contains exactly one time each congruence class of $\mathbb{Z}/q\mathbb{Z}$ since it is length q and the equality above follows from the orthogonality relation (4.10). Therefore for all t we have

$$|M_\chi(t)| \leqslant q.$$

By integration by parts we therefore have

$$\sum_{1<n\leqslant X}\frac{\chi(n)}{n^s}=X^{-s}M_\chi(X)-1+s\int_1^X M_\chi(t)t^{-s-1}\,dt.$$

Since $M_\chi(t)$ is bounded, we have that the first term tends to 0 when $\sigma=\mathrm{Re}(s)>0$. Moreover, the integral converges absolutely as $X\to\infty$ for s in the same domain. Thus, the series converges. By the same argument, we have

$$L(s,\chi)-\sum_{1\leqslant n\leqslant X}\frac{\chi(n)}{n^s}=\sum_{n>X}\frac{\chi(n)}{n^s}=\left[M_\chi(t)t^{-s}\right]_X^\infty+s\int_X^\infty M_\chi(t)t^{-s-1}\,dt\ll qX^{-\sigma}+\frac{|s|}{\sigma}qX^{-\sigma},$$

which tends to 0 as $X\to\infty$. This proves (4.12). This upper bound proves that $\sum_{1\leqslant n\leqslant X}\frac{\chi(n)}{n^s}$ converges uniformly towards $L(s,\chi)$ on compact subsets of the half-plane $\mathrm{Re}(s)>0$. We deduce from this the holomorphy of $L(s,\chi)$ in that domain. □

3. Beginning of the proof of Mertens theorem in arithmetic progressions

We already saw in our final proof of the classic Mertens theorem 2.15 that it suffices to show that

$$\sum_{\substack{n\leqslant x\\ n\equiv a\,(\mathrm{mod}\,q)}}\frac{\Lambda(n)}{n}=\frac{1}{\varphi(q)}\log x+O(1).$$

By (4.11) we have

$$\sum_{\substack{n\leqslant x\\ n\equiv a\,(\mathrm{mod}\,q)}}\frac{\Lambda(n)}{n}=\frac{1}{\varphi(q)}\sum_{\chi\,(\mathrm{mod}\,q)}\overline{\chi}(a)S_\chi(x),\tag{4.13}$$

where

$$S_\chi(x)=\sum_{n\leqslant x}\frac{\chi(n)}{n}\Lambda(n).$$

If $\chi=\chi_0$,

$$S_{\chi_0}(x)=\sum_{\substack{n\leqslant x\\ (n,q)=1}}\frac{\Lambda(n)}{n}=\sum_{n\leqslant x}\frac{\Lambda(n)}{n}-\sum_{p\mid q}\log p\sum_{\substack{\alpha\geqslant 1\\ p^\alpha\leqslant x}}\frac{1}{p^\alpha}=\log x+O(\log q),$$

and thus the contribution of this term to the sum (4.13) is of size

$$\log x+O_q(1).$$

To finish the proof of Mertens theorem in arithmetic progressions, it suffices to show that for all $\chi\neq\chi_0$

$$S_\chi(x)=\sum_{n\leqslant x}\frac{\chi(n)}{n}\Lambda(n)=O_q(1).$$

In order to calculate $S_\chi(x)$, we consider the sum

$$T_\chi(x)=\sum_{n\leqslant x}(\log n)\frac{\chi(n)}{n}.$$

By integration by parts and using the fact that $M_\chi(x)$ is bounded if $\chi\neq\chi_0$, we see that

$$T_\chi(x)=O(q).$$

On the other hand, using the equality $\log = 1 * \Lambda$, we have by (4.12)

$$\begin{aligned} T_\chi(x) &= \sum_{a\leqslant x} \frac{\chi(a)}{a}\Lambda(a) \sum_{b\leqslant x/a} \frac{\chi(b)}{b} \\ &= \sum_{a\leqslant x} \frac{\chi(a)}{a}\Lambda(a)\Big(L(1,\chi) + O(q\frac{a}{x})\Big) \\ &= L(1,\chi)S_\chi(x) + O_q(1) \end{aligned}$$

by Chebyshev's theorem. Thus, if we show that

$$L(1,\chi) \neq 0,$$

we will have shown

$$S_\chi(x) = O_q(1), \tag{4.14}$$

which will conclude the proof of Mertens theorem. The non-vanishing of $L(s,\chi)$ at $s=1$ is in fact the key point in the proof of Dirichlet's theorem on primes in arithmetic progressions.

4. Non-vanishing of Dirichlet L-functions at the point $s=1$

Theorem 4.13 (Dirichlet)**.** *Let $\chi \pmod q$ be a non-trivial Dirichlet character. Then*

$$L(1,\chi) \neq 0.$$

We will give two different proofs of this theorem. One only uses real analysis (given in the exercises), and the other goes by complex analysis. Whatever the proof is, a key point will always be an argument relying on positivity.

4.1. Proof by complex analysis. The idea is the following. We consider the product of all of the Dirichlet L-functions:

$$\prod_{\chi \pmod q} L(s,\chi) = L(s,\chi_0) \prod_{\chi\neq\chi_0} L(s,\chi) =: L_q(s).$$

The function $L_q(s)$ thus defined is actually a Dirichlet series. The associated arithmetic function $n \mapsto a_q(n)$ is multiplicative and given by the Dirichlet convolution of all of the Dirichlet characters modulo q, i.e. $a_q = *_{\chi \pmod q}\chi$. We have

$$L_q(s) = \sum_{n\geqslant 1} \frac{a_q(n)}{n^s},$$

and we will soon see that the coefficients $a_q(n)$ satisfy $a_q(n) \geqslant 0$ for all $n \geqslant 1$. We will also show for all n relatively prime to q that

$$a_q(n^{\varphi(q)}) \geqslant 1.$$

From these facts we deduce that the abscissa of convergence σ_q of a_q satisfies $\sigma_q \geqslant 1/\varphi(q)$. We also have à priori that $\sigma_q \leqslant 1$ since a_q is a convolution of arithmetic functions whose abscissae of convergence are equal to 1. Indeed, setting $\sigma = \frac{1}{\varphi(q)}$ we have

$$\sum_{n\geqslant 1} \frac{a_q(n)}{n^\sigma} \geqslant \sum_{\substack{m\geqslant 1\\(m,q)=1}} \frac{a_q(m^{\varphi(q)})}{m^{\varphi(q)\sigma}} \geqslant \sum_{\substack{m\geqslant 1\\(m,q)=1}} \frac{1}{m} = \infty.$$

Lemma 4.14 (Landau)**.** *Let*

$$L(s) = \sum_{n \geqslant 1} \frac{f(n)}{n^s}$$

be a Dirichlet series of abscissa of convergence $\sigma_f < \infty$ and such that the coefficients $f(n)$ are non-negative. Then $L(s)$ does not admit an analytic continuation in a neighborhood of the point $s = \sigma_f$.

By Landau's lemma, $L_q(s)$ does *not* admit analytic continuation to some neighborhood of σ_q, which itself satisfies $1/\varphi(q) \leqslant \sigma_q \leqslant 1$. But by proposition 4.12, $L_q(s)$ does admits a meromorphic continuation to the half-plane $\mathrm{Re}(s) > 0$ in which the only possible pole in this domain is at $s = 1$. If there existed a character χ for which $L(s, \chi)$ vanished at $s = 1$, then the function $L_q(s)$ will be holomorphic at $s = 1$ and therefore have an analytic continuation to the whole half-plane $\{s : \mathrm{Re}(s) > 0\}$, since the pole of $L(s, \chi_0)$ at 1 is simple. Contradiction!

So, the proof of $L(1, \chi) \neq 0$ reduces to:

(1) showing that $a_q(n) \geqslant 0$ for all $n \geqslant 1$,
(2) showing that $a_q(n^{\varphi(q)}) \geqslant 1$ for all $(n, q) = 1$,
(3) proving Landau's lemma.

We start with the third of these.

Proof of Landau's lemma. Without loss of generality, we can suppose that $\sigma_f = 0$ by replacing $L(s)$ with $L(s - \sigma_f)$. Let us suppose that $L(s)$ admits an analytic continuation in an open disk D centered at $\sigma_f = 0$. Thus $L(s)$ and all of its derivatives define holomorphic functions in the domain $D \cup \{s : \mathrm{Re}(s) > 0\}$.

Our strategy will be to show that the series $L(\sigma)$ converges absolutely for $\sigma = 0$ and in fact for $\sigma \in D$ with $\sigma < 0$. This will be a contradiction with the fact that $\sigma_f = 0$. First we show that

$$\sum_{n \geqslant 1} f(n) = L(0).$$

This is a consequence of the monotone convergence theorem, and we recall the argument now. As $f(n) \geqslant 0$, the function

$$\sigma \in (0, 1] \mapsto L(\sigma) = \sum_{n \geqslant 1} \frac{f(n)}{n^\sigma}$$

is decreasing and is bounded above by its limit as $\sigma \to 0^+$ which is $L(0)$. This isn't quite what we want, however. The series $\sum_n f(n)$ is defined as $\lim_{N \to \infty} \sum_{n \leqslant N} f(n)$. But we have for all $N \geqslant 1$ that

$$\sum_{n \leqslant N} f(n) = \lim_{\sigma \to 0^+} \sum_{n \leqslant N} \frac{f(n)}{n^\sigma} \leqslant \lim_{\sigma \to 0^+} L(\sigma) = L(0).$$

Thus we have shown that the sum $\sum_{n \geqslant 1} f(n)$ converges by the monotone convergence theorem (it has non-negative terms). We calculate the limit by noting that for all $\sigma > 0$ we have

$$L(\sigma) = \sum_{n \geqslant 1} \frac{f(n)}{n^\sigma} \leqslant \sum_{n \geqslant 1} f(n),$$

and by letting σ tend to 0.

Now we show that $L(\sigma)$ converges for $\sigma < 0$ and $\sigma \in D$. Let us consider the kth derivative of $L(s)$: By Proposition 3.2, for $\mathrm{Re}(s) > 0$ it is given by

$$L^{(k)}(s) = (-1)^k \sum_{n \geqslant 1} \frac{f(n)(\log n)^k}{n^s}.$$

As $f(n)(\log n)^k$ is non-negative, the same argument as before give that

$$(-1)^k \sum_{n\geqslant 1} f(n)(\log n)^k = L^{(k)}(0).$$

If $\sigma \in D$, then $L(\sigma)$ is calculated by its Taylor series at 0:

$$L(\sigma) = \sum_{k\geqslant 0} \frac{L^{(k)}(0)}{k!}\sigma^k = \sum_{k\geqslant 0} \frac{(-\sigma)^k}{k!} \sum_{n\geqslant 1} f(n)(\log n)^k.$$

Suppose $\sigma < 0$, then as $(-\sigma)^k \geqslant 0$ the terms of this double sum are all non-zero and we can permute them as we like:

$$L(\sigma) = \sum_{n\geqslant 1} f(n) \sum_{k\geqslant 0} \frac{(-\sigma\log n)^k}{k!} = \sum_{n\geqslant 1} f(n)\exp(-\sigma\log n) = \sum_{n\geqslant 1} \frac{f(n)}{n^\sigma}.$$

Thus the series converges for $\sigma < 0$, which contradicts the fact that $\sigma_f = 0$. □

4.2. Positivity of the $a_q(n)$. Since the function $n \mapsto a_q(n)$ is multiplicative, it suffices to show that $a_q(p^k) \geqslant 0$ for p prime and $k \geqslant 0$. It therefore suffices to show that the coefficients of

$$L_{q,p}(s) = \prod_\chi L_p(s,\chi) = \prod_\chi (1-\chi(p)p^{-s})^{-1} = \sum_{k\geqslant 0} \frac{a_q(p^k)}{p^{ks}}$$

are non-negative. If $p \mid q$ then $L_{q,p}(s) = 1$. We can therefore suppose that p is relatively prime to q. Setting $z = p^{-s}$ (note $|z| < 1$ if $\mathrm{Re}(s) > 0$), we consider the convergent power series

$$E(z) = \prod_\chi (1-\chi(p)z)^{-1} = \sum_{k\geqslant 0} a_q(p^k) z^k.$$

Taking the logarithm, we have

$$\log E(z) = \sum_\chi \log((1-\chi(p)z)^{-1}) = \sum_\chi \sum_{k\geqslant 1} \chi(p)^k \frac{z^k}{k} = \sum_\chi \sum_{k\geqslant 1} \chi(p^k)\frac{z^k}{k}$$
$$= \sum_{k\geqslant 1} \frac{z^k}{k} \sum_\chi \chi(p^k) = \varphi(q) \sum_{\substack{k\geqslant 1 \\ p^k \equiv 1 \,(\mathrm{mod}\, q)}} \frac{z^k}{k}.$$

We thus see that $\log(E(z))$ is a power series with non-negative coefficients. Since

$$E(z) = \exp(\log(E(z))) = 1 + \log(E(z)) + \frac{\log(E(z))^2}{2} + \cdots + \frac{(\log(E(z)))^\ell}{\ell!} + \cdots$$

we see that the coefficients of $E(z)$ are themselves also non-negative.

Now we show that $a_q(n^{\varphi(q)}) \geqslant 1$. Note that for $k = \varphi(q)$ we have

$$p^k = p^{\varphi(q)} \equiv 1 \,(\mathrm{mod}\, q)$$

by Lagrange's theorem (here the group is $(\mathbb{Z}/q\mathbb{Z})^\times$). We see that the $\varphi(q)$th coefficient of $\log E(z)$ is equal to 1, and from this we deduce that the $\varphi(q)$th coefficient of $E(z)$ is $\geqslant 1$. This already permits us to deduce that the abscissa of convergence of $L_q(s)$ is $\geqslant 1/\varphi(q)$.

We can do a little better and completely calculate $L_q(s)$. Let e_p be the order of $p \,(\mathrm{mod}\, q)$ in the group $(\mathbb{Z}/q\mathbb{Z})^\times$. We have

$$\log(E(z)) = \frac{\varphi(q)}{e_p} \sum_{k\geqslant 1} \frac{z^{e_p k}}{k} = \frac{\varphi(q)}{e_p} \log((1-z^{e_p})^{-1}) = \log((1-z^{e_p})^{-\frac{\varphi(q)}{e_p}})$$

and so

$$E(z) = \frac{1}{(1-z^{e_p})^{\frac{\varphi(q)}{e_p}}} = (1 + z^{e_p} + z^{2e_p} + \cdots)^{\frac{\varphi(q)}{e_p}},$$

and so

$$L_{q,p}(s) = \frac{1}{(1-p^{-e_p s})^{\frac{\varphi(q)}{e_p}}} = \sum_{k \geqslant 0} \frac{a_q(p^k)}{p^{ks}}.$$

We deduce from this that the $a_q(p^k)$ are *non-negative* integers such that $a_q(p^k) \geqslant 1$ for all k which are multiples of $\varphi(q)$ and so for all n prime with q we have

$$a_q(n^{\varphi(q)}) \geqslant 1.$$

CHAPTER 5

Riemann's Memoir

More than 150 years ago, in 1859, Riemann published his celebrated memoir *Über die Anzahl der Primzahlen unter einer gegebenen Grösse*. In this memoir, Riemann presented the foundations for the study of the analytic properties of the zeta function

$$\zeta(s) = \sum_{n \geqslant 1} \frac{1}{n^s} = \prod_p \left(1 - \frac{1}{p^s}\right)^{-1}, \quad \mathrm{Re}(s) > 1 \tag{5.1}$$

as a function of a complex variable s, and the relation between these and the problem of counting prime numbers.

In his memoir, Riemann mentioned an "explicit" formula, shown later by von Mangolt, relating the summation function of the function Λ

$$M_\Lambda(x) = \sum_{n \leqslant x} \Lambda(n)$$

to a sum over the zeros of $\zeta(s)$. Sufficient information on the location of the zeros of $\zeta(s)$ then implies an asymptotic formula for $M_\Lambda(x)$ as $x \to \infty$. Riemann showed that there are an infinity of "non-trivial" zeros of ζ, and gave an asymptotic formula for their number. He also formulated in his memoir the celebrated *Riemann hypothesis*, which predicts that all of the non-trivial zeros of $\zeta(s)$ are situated on the line $\mathrm{Re}(s) = 1/2$. The proof of this hypothesis is one of the most important problems in mathematics. The Riemann hypothesis is one of the famous Millenium Prize Problems, and the Clay mathematics institute has offered a 1000000 USD reward for its solution.

Riemann showed the following in his memoir.

Analytic continuation. *The function $\zeta(s)$ admits a meromorphic continuation to $\mathbb{C}$. It is holomorphic everywhere except for $s = 1$ where it has a simple pole of residue equal to 1.*

Functional equation. *Let*

$$\xi(s) = \zeta_\infty(s)\zeta(s) \quad \text{where} \quad \zeta_\infty(s) = \pi^{-s/2}\Gamma(s/2).$$

Then $\xi(s)$ admits a meromorphic continuation to $\mathbb{C}$ with two simple poles at $s = 0, 1$ and satisfies for $s \neq 0, 1$

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

In this statement, $\Gamma(s)$ is the Euler gamma function defined for $\mathrm{Re}(s) > 0$ by

$$\Gamma(s) = \int_0^\infty e^{-t} t^s \frac{dt}{t}.$$

The analytic properties of $\Gamma(s)$ are well-understood, and you will work some of them out on exercise sheet 8.

We saw that, taking into account the Euler product (5.1), that $\zeta(s)$ does not vanish for $\mathrm{Re}(s) > 1$. As $\Gamma(s)$ doesn't vanish for any $s \in \mathbb{C}$, we deduce from the functional equation that

$\xi(s)$ has no zeros in $\mathrm{Re}(s) < 0$ or $\mathrm{Re}(s) > 1$. On the other hand, since $\Gamma(s/2)$ has simple poles at negative even numbers, $\zeta(s)$ vanishes to order 1 at $s = -2, -4, \ldots$. These zeros are called *trivial zeros* of ζ, and the non-trivial zeros of ζ are the same as the zeros of $\xi(s)$ and are contained in the critical strip

$$\{s \in \mathbb{C} : \mathrm{Re}(s) \in [0,1]\}.$$

Riemann gave an asymptotic formula for their number:

COUNT OF ZEROS. *For $T \geqslant 0$, let*

$$N(T) = |\{\rho = \beta + it : \zeta(\rho) = 0, \beta \in [0,1], |t| \leqslant T\}|$$

be the number of non-trivial zeros of $\zeta(s)$ of height $\leqslant T$. We have

$$N(T) = \frac{T}{\pi} \log\left(\frac{T}{2\pi e}\right) + O(\log(2 + |T|)).$$

This formula for the count of zeros can be deduced from the following formula which relates the zeros of ζ to prime numbers.

EXPLICIT FORMULA. *Let $f \in \mathscr{C}_c^\infty(\mathbb{R}_{>0})$ and let*

$$\widetilde{f}(s) = \int_0^\infty f(x) x^s \frac{dx}{x}$$

be its Mellin transform. Let $\check{f}(x) = x^{-1} f(x^{-1})$. We have the identity

$$\sum_{n \geqslant 1} (f(n) + \check{f}(n)) \Lambda(n) = \widetilde{f}(1) + \frac{1}{2\pi i} \int_{\mathrm{Re}(s)=1/2} \left(\frac{\zeta_\infty'}{\zeta_\infty}(s) + \frac{\zeta_\infty'}{\zeta_\infty}(1-s) \right) \widetilde{f}(s)\, ds - \sum_{\substack{\zeta(\rho)=0 \\ \mathrm{Re}(\rho) \in [0,1]}} \widetilde{f}(\rho).$$

In the next several chapters, we prove the above facts from Riemann's memoir, and generalize them to Dirichlet L-functions $L(s, \chi)$.

CHAPTER 6

The functional equation

1. Some integral transforms

The theory of the Fourier transform describes the structure of vector spaces of functions on the real line $\mathbb{R}$, taking into account the fact that $(\mathbb{R},+)$ is a commutative group under addition. It permits one to write functions (or at least sufficiently "nice" ones) on $\mathbb{R}$ as "linear combinations" of functions adapted to the group structure of $\mathbb{R}$, and these functions are the characters which we saw examples of before in the case of finite abelian groups.

1.1. The characters of $\mathbb{R}$.

Definition 6.1. *Let $(G,+)$ be an abelian topological group, that is an abelian group equipped with the structure of a topological space for which addition $+:(x_1,x_2)\mapsto x_1+x_2$ and inversion $[-1]:x\mapsto -x$ are continuous maps. The set of character of G is the set $\mathrm{Hom}_c(G,\mathbb{C}^\times)$ of continuous group homomorphisms from G to $(\mathbb{C}^\times,\times)$. This set is an abelian group under multiplication of functions.*

A character is called unitary if it takes its values on the unit circle $S^1=\{z\in\mathbb{C}^\times : |z|=1\}$. We write $\widehat{G}$ for the subgroup of unitary characters.

Remark 6.2. If G is compact, then every character is unitary. Indeed, let $\psi\in\mathrm{Hom}_c(G,\mathbb{C}^\times)$ be a character. Then $\psi(G)$ is a compact subgroup of $\mathbb{C}^\times$ and is thus contained in S^1. (If there existed $z\in\psi(G)$ such that $|z|\neq 1$ then for all $n\in\mathbb{Z}$, $z^n\in\psi(G)$, we either have $z^n\to 0$ or ∞ according to $|z|<1$ or $|z|>1$.)

We write $e(x)=\exp(2\pi i x)$.

Theorem 6.3. *The map*

$$y\in\mathbb{C}\mapsto\psi_y : x\mapsto e(xy)$$

is an isomorphism of groups

$$(\mathbb{C},+)\simeq\mathrm{Hom}_c(\mathbb{R},\mathbb{C}^\times).$$

The restriction of this map to $\mathbb{R}$ is an isomorphism of groups

$$(\mathbb{R},+)\simeq\widehat{\mathbb{R}}.$$

Proof. It is not hard to see that this map is an injective group homomorphism: if y is such that for all x, $x\mapsto e(xy)=1$, then y must be 0, since it derivative (in x) is $2\pi i y e(xy)=0$. Now we show that the map is surjective. Let $\psi\in\mathrm{Hom}_c(\mathbb{R},\mathbb{C}^\times)$ and

$$\Psi(x)=\int_0^x\psi(t)\,dt$$

its anti-derivative (ψ is continuous, hence integrable). We have

$$\Psi(x+y)=\int_0^{x+y}\psi(t)\,dt=\int_0^x\psi(t)\,dt+\int_0^y\psi(x+t)\,dt=\Psi(x)+\psi(x)\Psi(y).$$

We fix y such that $\Psi(y) \neq 0$ (such a y exists, since otherwise $\Psi'(y) = \psi(y) = 0$ for all y, which is impossible). We deduce from the previous identity that $x \mapsto \psi(x)$ is differentiable and that

$$\psi(x+y) = \psi(x)\psi(y) = \psi(x) + \psi'(x)\Psi(y)$$

and thus

$$\frac{\psi'(x)}{\psi(x)} = \frac{\psi(y)-1}{\Psi(y)}.$$

Setting $u = \frac{\psi(y)-1}{\Psi(y)}$ we see that

$$\psi(x) = C\exp(ux)$$

and moreover we have $C = 1$ by evaluating both sides of the preceding equality at $x = 0$. □

Corollary 6.4. *The map*

$$n \in \mathbb{Z} \mapsto \psi_n : x \mapsto e(nx)$$

is an isomorphism of groups

$$(\mathbb{Z}, +) \simeq \mathrm{Hom}_c(\mathbb{R}/\mathbb{Z}, \mathbb{C}^\times) = \widehat{\mathbb{R}/\mathbb{Z}}.$$

Proof. The equality $\mathrm{Hom}_c(\mathbb{R}/\mathbb{Z}, \mathbb{C}^\times) = \widehat{\mathbb{R}/\mathbb{Z}}$ follows from the fact that $\mathbb{R}/\mathbb{Z}$ is compact. The fact that the map is an injective group homomorphism is the same as in the proof of Theorem 6.3. We show the surjectivity. Let $\psi \in \widehat{\mathbb{R}/\mathbb{Z}}$, then ψ defines a unitary character on $\mathbb{R}$ which is 1 on $\mathbb{Z}$ and thus by Theorem 6.3 is of the form $\psi(x) = e(nx)$ with $n \in \mathbb{R}$. As $\psi(1) = 1 = \exp(2\pi i n)$ we have that $n \in \mathbb{Z}$. □

1.2. The Schwartz class. We say that a function $f : \mathbb{R} \to \mathbb{C}$ belongs to the Schwartz class $\mathscr{S}(\mathbb{R})$ of functions if $f \in \mathscr{C}^\infty(\mathbb{R})$ and if f and all its derivatives are of rapid decrease. Stated differently, f is Schwartz class if for all $A \geqslant 0$ and every integer $j \geqslant 0$ we have

$$f^{(j)}(x) \ll_{j,A} (1+|x|)^{-A}.$$

1.3. Fourier transform. Let $f \in \mathscr{S}(\mathbb{R})$. The Fourier transform of f is the function

$$\widehat{f} : y \mapsto \int_{\mathbb{R}} f(x)e(-xy)\,dx.$$

As f is of rapid decrease, we see that the preceding integral converges absolutely and uniformly on $\mathbb{R}$ and that it defines an infinitely differentiable function with derivatives

$$\widehat{f}^{(j)}(y) = \int_{\mathbb{R}} (-2\pi i x)^j f(x)e(-xy)\,dx = \widehat{M^j f}(y), \quad \text{with} \quad Mf : x \mapsto (-2\pi i x)f(x). \tag{6.1}$$

Also note that by integration by parts, we have for all $y \neq 0$ that

$$\widehat{f}(y) = \frac{1}{2\pi i y}\int_{\mathbb{R}} f'(x)e(-xy)\,dx = \frac{1}{2\pi i y}\widehat{f'}(y), \tag{6.2}$$

and iterating this, we obtain for all $j \geqslant 1$ that

$$\widehat{f}(y) = \left(\frac{1}{2\pi i y}\right)^j \widehat{f^{(j)}}(y).$$

This calculation is justified by the rapid decrease of the derivatives of f. Thus, for $|y| \geqslant 1$, we have for all $j \geqslant 0$

$$\widehat{f}(y) \ll_j |y|^{-j}\int_{\mathbb{R}} |f^{(j)}(x)|\,dx \ll_{j,f} |y|^{-j},$$

so $\widehat{f}(y)$ is of rapid decrease. Seeing as for all $j \geqslant 0$, $M^j f \in \mathscr{S}(\mathbb{R})$, using (6.1), we see that its derivatives are also of rapid decrease. In other words, we have proved the following.

Proposition 6.5. *The Fourier transform is a linear map from $\mathscr{S}(\mathbb{R})$ to $\mathscr{S}(\mathbb{R})$.*

1.3.1. *Behavior of the Fourier transform under translations.* Let $h \in \mathbb{R}$. We write $[+h]$ for the translation map on the space of functions on $\mathbb{R}$ defined by

$$[+h]f : x \mapsto f(x+h).$$

The map $[+h]$ is an invertible linear map $\mathscr{S}(\mathbb{R}) \to \mathscr{S}(\mathbb{R})$. By changing variables, we have

$$\widehat{[+h]f}(y) = e(-hy)\widehat{f}(y).$$

In particular, by considering the Fourier transform of the quotient

$$\frac{[+h]f(x) - f(x)}{h} = \frac{f(x+h) - f(x)}{h},$$

and passing to the limit, we have another proof of (6.2)

$$: \widehat{f}'(y) = (2\pi i y)\widehat{f}(y), \tag{6.3}$$

and for all $j \geqslant 0$ that

$$\widehat{f^{(j)}}(y) = (2\pi i y)^j \widehat{f}(y). \tag{6.4}$$

1.3.2. *Behavior of the Fourier transform under dilations.* Let $\lambda \in \mathbb{R}^\times$ and denote by $[\times\lambda]$ the "dilation by λ" on the space of functions on $\mathbb{R}$

$$[\times\lambda]f : x \mapsto f(\lambda x).$$

Thus, by changing variables, we have for all $f \in \mathscr{S}(\mathbb{R})$

$$\widehat{[\times\lambda]f}(y) = \frac{1}{|\lambda|}\widehat{f}(y/\lambda) = \frac{1}{|\lambda|}[\times\lambda^{-1}]\widehat{f}(y).$$

1.3.3. *Fourier inversion formula.* This is perhaps the most important result in the theory of the Fourier transform. For $f \in \mathscr{S}(\mathbb{R})$,

$$\widehat{\widehat{f}} = [\times -1]f \quad \text{ that is, } \quad \widehat{\widehat{f}}(x) = f(-x) \tag{6.5}$$

1.3.4. *Fourier transform of the Gaussian.* Let $f(x) = e^{-\pi x^2}$. Then

$$\widehat{f}(y) = f(y). \tag{6.6}$$

1.4. Poisson Summation.

Proposition 6.6. *Let $f \in \mathscr{S}(\mathbb{R})$. For all $u \in \mathbb{R}$ we have*

$$\sum_{n\in\mathbb{Z}} f(n+u) = \sum_{n\in\mathbb{Z}} \widehat{f}(n)e(nu).$$

Proof. Let

$$f_{\mathbb{Z}}(u) = \sum_{n\in\mathbb{Z}} f(n+u).$$

Since $f \in \mathscr{S}(\mathbb{R})$, the series that defines $f_{\mathbb{Z}}$ converges uniformly, as well as the series constructed from the derivatives of f. Thus, we deduce that $f_{\mathbb{Z}} \in \mathscr{C}^\infty(\mathbb{R})$. Moreover, $f_{\mathbb{Z}}$ is periodic of period 1. Thus by decomposition into Fourier series we have

$$f_{\mathbb{Z}}(u) = \sum_{n\in\mathbb{Z}} c_{f_{\mathbb{Z}}}(n)e(nu),$$

with

$$\begin{aligned} c_{f_{\mathbb{Z}}}(n) &= \int_0^1 f_{\mathbb{Z}}(u)e(-nu)\,du = \int_0^1 \left(\sum_{m\in\mathbb{Z}} f(m+u)\right)e(-nu)\,du \\ &= \int_0^1 \left(\sum_{m\in\mathbb{Z}} f(m+u)e(-n(u+m))\right)du = \sum_{m\in\mathbb{Z}}\int_0^1 f(m+u)e(-n(u+m))\,du \\ &= \sum_{m\in\mathbb{Z}}\int_m^{m+1} f(u)e(-nu)\,du = \widehat{f}(n). \end{aligned}$$

□

We can use this formula to analyze the summation function of a function in arithmetic progressions.

Corollary 6.7. *Let $q, a \in \mathbb{Z}$, $q \neq 0$. We have*

$$\sum_{n\equiv a \pmod q} f(n) = \frac{1}{q}\sum_{n\in\mathbb{Z}} \widehat{f}(\frac{n}{q})e(\frac{an}{q}).$$

Proof. We have

$$\begin{aligned} \sum_{n\equiv a \pmod q} f(n) &= \sum_{n\in\mathbb{Z}} f(qn+a) = \sum_{n\in\mathbb{Z}} f(q(n+\frac{a}{q})) = \sum_{n\in\mathbb{Z}} [\times q]f(n+\frac{a}{q}) \\ &= \sum_{n\in\mathbb{Z}} \widehat{[\times q]f}(n)e(\frac{an}{q}) = \frac{1}{q}\sum_{n\in\mathbb{Z}} \widehat{f}(\frac{n}{q})e(\frac{an}{q}). \end{aligned}$$

□

2. The Mellin transform

Let $f \in \mathscr{C}^\infty(\mathbb{R}_{\geqslant 0})$ be a function that decays rapidly as $x \to \infty$ along with all its derivatives. The Mellin transform of f is defined for $\mathrm{Re}(s) > 0$ by

$$\widetilde{f}(s) = \int_{\mathbb{R}_{\geqslant 0}} f(x)x^s\,d^\times x,$$

where we have set

$$d^\times x = \frac{dx}{x}.$$

It is the measure on $(\mathbb{R}_{>0}, \times)$ which is invariant by translations $x \mapsto yx$. The integral $\widetilde{f}(s)$ converges at 0 because $\mathrm{Re}(s) > 0$ and at ∞ because f is of rapid decay. The convergence is uniform on compacts subsets of $\{s : \mathrm{Re}(s) > 0\}$ and so defines a holomorphic function of s in that domain. Note that if we further assume that $f(x) = O(x^N)$ as $x \to 0$ then $\widetilde{f}(s)$ defines a holomorphic function in the domain $\mathrm{Re}(s) > -N$.

By integration by parts, if $\mathrm{Re}(s) > 1$ we have

$$\widetilde{f}'(s) = -(s-1)\widetilde{f}(s-1). \tag{6.7}$$

Define

$$g(s) = -\frac{1}{s}\widetilde{f}'(s+1)$$

in the domain $\mathrm{Re}(s) > -1$, since f' decays rapidly as $x \to \infty$. In the domain $\mathrm{Re}(s) > 0$, the function $g(s)$ matches $\widetilde{f}(s)$, and so by analytic continuation $g(s)$ is the unique analytic function extending $\widetilde{f}(s)$. Thus we define $\widetilde{f}(s)$ to be equal to $g(s)$ in the domain $\mathrm{Re}(s) \in (-1, 0]$.

The residue at $s=0$ of $\widetilde{f}(s)$ is $-\widetilde{f}'(1) = -\int_{\mathbb{R}_{>0}} f'(y)\,dy = f(0)$. If moreover f vanishes at 0 then $\widetilde{f}(s)$ is holomorphic for $\mathrm{Re}(s) > -1$ and more generally if all of the derivatives of f up to order $n \geqslant 0$ vanish at 0 then $\widetilde{f}(s)$ is holomorphic for $\mathrm{Re}(s) > -(n+1)$.

Furthermore, iterating (6.7) we have

$$\widetilde{f}(s) = \frac{(-1)^n}{s(s+1)\cdots(s+n-1)}\widetilde{f^{(n)}}(s+n), \tag{6.8}$$

which implies that for all $n \geqslant 0$, $|\mathrm{Re}(s)| \leqslant \sigma$ and $|\mathrm{Im}(s)| \geqslant 1$ we have

$$|\widetilde{f}(s)| \ll_{\sigma,f,n} \frac{1}{|s|^n}. \tag{6.9}$$

2.1. The Gamma Function. The Gamma function Γ is by definition the Mellin transform of the function $x \mapsto e^{-x}$.

$$\Gamma(s) = \widetilde{e^{-x}}(s) = \int_0^\infty e^{-x} x^s\, d^\times x.$$

This function verifies for all $\mathrm{Re}(s) > 0$

$$s\Gamma(s) = \Gamma(s+1), \quad \text{and} \quad \Gamma(n+1) = n!.$$

Thus $\Gamma(s)$ admits a meromorphic continuation to $\mathbb{C}$ with simple poles at $s = -n$, $n \in \mathbb{N}$ of residues $(-1)^n/n!$.

2.2. Dilations. Let $\lambda > 0$. We have

$$\widetilde{[\times\lambda]f}(s) = \lambda^{-s}\widetilde{f}(s).$$

2.3. The Mellin Transform and the Fourier transform. Let $f \in \mathscr{C}_c^\infty(\mathbb{R}_{>0})$, we define $g \in \mathscr{C}_c^\infty(\mathbb{R})$ by changing variables

$$f(y) = g(\log y), \quad g(x) = f(\exp(x))$$

In particular $\widetilde{f}(s)$ defines a holomorphic function on $\mathbb{C}$. We have, setting $s = \sigma + it$

$$\begin{aligned}
\widetilde{f}(s) &= \int_0^\infty g(\log y)\exp(s\log y)\,\frac{dy}{y} \\
&= \int_{-\infty}^\infty g(x)\exp(sx)\,dx \\
&= \int_{-\infty}^\infty g(x)\exp(\sigma x)e(\frac{xt}{2\pi})\,dx \\
&= \widehat{g\exp(\sigma x)}(-\frac{t}{2\pi}).
\end{aligned}$$

In particular, by Fourier inversion we have that

$$\int_{-\infty}^\infty \widetilde{f}(\sigma+it)\,dt = \int_{-\infty}^\infty \widehat{g\exp(\sigma x)}(-\frac{t}{2\pi})\,dt = 2\pi g(0) = 2\pi f(1),$$

which we may write as

$$\frac{1}{2\pi i}\int_{(\sigma)} \widetilde{f}(s)\,ds = f(1).$$

Replacing f by $[\times y]f$ we obtain the *Mellin inversion formula*: For all $y > 0$ and every $\sigma \in \mathbb{R}$

$$\frac{1}{2\pi i}\int_{(\sigma)} \widetilde{f}(s)y^{-s}\,ds = f(y).$$

Note that if $f \in \mathscr{S}(\mathbb{R})$ instead, the Mellin inversion formula remains valid for $\sigma > 0$.

3. The functional equation of the Riemann zeta function

Theorem 6.8. *Let*

$$\xi(s) = \zeta_\infty(s)\zeta(s), \qquad \textit{where} \qquad \zeta_\infty(s) = \pi^{-s/2}\Gamma(s/2), \quad \mathrm{Re}(s) > 1.$$

This function has a meromorphic continuation to $\mathbb{C}$*, is holomorphic on* $\mathbb{C} - \{0,1\}$*, and has simple poles at* $s = 0, 1$*. It satisfies the functional equation*

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

In particular, the function $\zeta(s) = \sum_{n\geqslant 1} n^{-s}$*, defined initially for* $\mathrm{Re}(s) > 1$ *extends to a meromorphic function on* $\mathbb{C}$*, holomorphic in* $\mathbb{C} - \{1\}$*, with a simple pole at* $s = 1$ *with residue* 1*. It satisfies the functional equation*

$$\zeta(s) = 2^s \pi^{s-1} \sin\Big(\frac{\pi s}{2}\Big)\Gamma(1-s)\zeta(1-s).$$

Proof. The second part of the Theorem concerning $\zeta(s)$ is derived from the first and the fact that $\zeta_\infty(s)$ does not vanish on $\mathbb{C}$, and has a simple pole at $s = 0$. The idea of the proof of the first part consists of making $\zeta(s)$ appear via the invariance property of the Mellin transform under dilations. More precisely, let $f \in \mathscr{S}(\mathbb{R})$ and $\mathrm{Re}(s) > 1$. We have for all $n \geqslant 1$

$$\widetilde{[\times n]f}(s) = \widetilde{f}(s)n^{-s}$$

and so

$$\widetilde{f}(s)\zeta(s) = \sum_{n\geqslant 1} \widetilde{[\times n]f}(s) = \sum_{n\geqslant 1}\int_0^\infty f(nx)x^s\, d^\times x = \int_0^\infty \Big(\sum_{n\geqslant 1} f(nx)\Big) x^s\, d^\times x.$$

In the last equality, the interchange of the sum and the integral is justified since for all $x > 0$ and all $A \geqslant 2$

$$\sum_{n\geqslant 1} |f(nx)| \ll_A \sum_{n\geqslant 1} (1+nx)^{-A} \ll_A \min(x^{-1}, x^{-A})$$

as we may see by considering the cases $x < 1$ and $x \geqslant 1$ separately. Thus the double sum is absolutely and uniformly convergent in the vertical strip $\{s \in \mathbb{C} : \mathrm{Re}(s) \in [a,b]\}$ with $1 < a \leqslant b$.

In fact, the above integral is absolutely convergent at $+\infty$ for all $s \in \mathbb{C}$ but could diverge at 0 if $\mathrm{Re}(s) < 1$. Thus we divide the integral into two

$$\int_0^\infty \cdots = \int_0^1 \cdots + \int_1^\infty \cdots$$

For the first, we make the change of variables $x \leftrightarrow 1/x$ and thus obtain

$$\widetilde{f}(s)\zeta(s) = \int_1^\infty \Big(\sum_{n\geqslant 1} f(n/x)\Big) x^{-s}\, d^\times x + \int_1^\infty \Big(\sum_{n\geqslant 1} f(nx)\Big) x^s\, d^\times x.$$

Suppose that f is even (thus $\widehat{f}$ is also even). We have then

$$\sum_{n\geqslant 1} f(n/x) = \frac{1}{2}\sum_{n\in\mathbb{Z}} f(n/x) - \frac{1}{2}f(0)$$

applying the Poisson summation formula to the first term and applying the formula for the behavior of the Fourier transform under dilations, we get

$$\sum_{n\geqslant 1} f(n/x) = \frac{1}{2}\big(\widehat{f}(0)x - f(0)\big) + x\sum_{n\geqslant 1}\widehat{f}(nx).$$

We have

$$\int_1^\infty \frac{1}{2}\left(\widehat{f}(0)x - f(0)\right)x^{-s}\,d^\times x = -\frac{1}{2}\left(\frac{\widehat{f}(0)}{1-s} + \frac{f(0)}{s}\right)$$

and so

$$\widetilde{f}(s)\zeta(s) = -\frac{1}{2}\left(\frac{\widehat{f}(0)}{1-s} + \frac{f(0)}{s}\right) + \int_1^\infty \left(\sum_{n\geqslant 1} f(nx)\right)x^s\,d^\times x + \int_1^\infty \left(\sum_{n\geqslant 1}\widehat{f}(nx)\right)x^{1-s}\,d^\times x. \tag{6.10}$$

Since f and $\widehat{f}$ are of rapid decay as $x\to\infty$, we have for all $A\geqslant 0$ and all $x\geqslant 1$ that

$$\sum_{n\geqslant 1} f(nx), \quad \sum_{n\geqslant 1}\widehat{f}(nx) \ll_A x^{-A}$$

and that the two preceding integrals converge absolutely and uniformly in every vertical strip $\{s\in\mathbb{C} : \mathrm{Re}(s)\in[a,b]\}$ with $a<b\in\mathbb{R}$. These integrals therefore define holomorphic functions on $\mathbb{C}$. Thus the function

$$s\mapsto \widetilde{f}(s)\zeta(s)$$

admits a meromorphic continuation to $\mathbb{C}$, holomorphic on $\mathbb{C}-\{0,1\}$ and with at most two simple poles at $s=0,1$.

On the other hand, replacing f by $\widehat{f}$ and s by $1-s$ in the above identity, and we use

$$\widehat{\widehat{f}}(x) = f(-x) = f(x),$$

we obtain the *functional equation*:

Theorem 6.9. *Let $f\in\mathscr{S}(\mathbb{R})$ be an even function. Then the function defined for $\mathrm{Re}(s)>1$ by*

$$s\mapsto \widetilde{f}(s)\zeta(s)$$

admits a meromorphic continuation to $\mathbb{C}$ with at most two simple poles at $s=0,1$ of residues at $0,1$ given by

$$-\frac{f(0)}{2}, \quad \text{and} \quad \frac{\widehat{f}(0)}{2}$$

and satisfies the functional equation

$$\widetilde{f}(s)\zeta(s) = \widetilde{\widehat{f}}(1-s)\zeta(1-s).$$

In particular, if f is such that $\widehat{f}=f$ then we obtain

$$\widetilde{f}(s)\zeta(s) = \widetilde{f}(1-s)\zeta(1-s).$$

An example of such a function is the *Gaussian*

$$f(x) = \exp(-\pi x^2) \tag{6.11}$$

whose Mellin transform is

$$\widetilde{f}(s) = \frac{1}{2}\pi^{-s/2}\widetilde{e^{-y}}(s/2) = \frac{1}{2}\pi^{-s/2}\Gamma(s/2).$$

□

Remark: there are many functions such that $f=\widehat{f}$. Another example is $(\cosh(\pi x))^{-1}$. See [GR7, 3.523.3] for its Mellin transform.

4. Primitive Characters, Gauss Sums, and Dirichlet L-functions

4.1. Primitive Characters. Let χ be a Dirichlet character modulo q. There is a natural number associated to χ called its conductor.

Definition 6.10. *The minimal $q^\star \mid q$ such that $\chi = \chi_0\chi^\star$ with $\chi^\star$ a Dirichlet character modulo $q^\star$ and χ_0 the trivial character modulo q is called the conductor of χ modulo q. If $q^\star = q$ then χ is called primitive.*

Given χ modulo q, the character $\chi^\star$ modulo $q^\star$ in the above factorization is unique, and is called the primitive character inducing χ. Note that we have

$$\chi(a) = \chi^\star(a) \quad \text{if } (a,q) = 1.$$

However, the two characters could take different values if $(a,q^\star) = 1$ but $(a,q) \neq 1$.

4.2. Gauss sums. Let χ modulo q be a Dirichlet character.

Definition 6.11. *The quantity*

$$\tau(\chi) = \sum_{b \,(\mathrm{mod}\, q)} \chi(b) e\left(\frac{b}{q}\right)$$

is called the Gauss sum of χ.

We have

$$e\left(\frac{a}{q}\right) = \frac{1}{\varphi(q)} \sum_{\chi \,(\mathrm{mod}\, q)} \overline{\chi}(a)\tau(\chi) \quad \text{if } (a,q) = 1, \tag{6.12}$$

by the orthogonality relations for Dirichlet characters. The equation (6.12) is the Fourier expansion of the additive character as a function on the group $(\mathbb{Z}/q\mathbb{Z})^\times$. Similarly,

$$\chi(a)\tau(\overline{\chi}) = \sum_{b \,(\mathrm{mod}\, q)} \overline{\chi}(b) e\left(\frac{ab}{q}\right) \quad \text{if } (a,q) = 1, \tag{6.13}$$

since we have by a change of variables $b = na^{-1} \,(\mathrm{mod}\, q)$

$$\begin{aligned}
\sum_{b \,(\mathrm{mod}\, q)} \overline{\chi}(b) e\left(\frac{ab}{q}\right) &= \sum_{b \in (\mathbb{Z}/q\mathbb{Z})^\times} \overline{\chi}(b) e\left(\frac{ab}{q}\right) \\
&= \sum_{n \in (\mathbb{Z}/q\mathbb{Z})^\times} \overline{\chi}(na^{-1}) e\left(\frac{n}{q}\right) \\
&= \chi(a) \sum_{n \in (\mathbb{Z}/q\mathbb{Z})^\times} \overline{\chi}(n) e\left(\frac{n}{q}\right) \\
&= \chi(a) \sum_{n \,(\mathrm{mod}\, q)} \overline{\chi}(n) e\left(\frac{n}{q}\right) = \chi(a)\tau(\overline{\chi}).
\end{aligned}$$

If $\tau(\overline{\chi}) \neq 0$, then we get the expansion of $\chi(a)$ in terms of additive characters $e(\cdot)$.

Lemma 6.12. *If χ is primitive, then* (6.13) *holds for all a modulo q.*

Proof. Suppose that $(a,q) > 1$, and let

$$\frac{a}{q} = \frac{a_1}{q_1}$$

with $(a_1, q_1) = 1$ and $q_1 \mid q$, $q_1 < q$. If a is a multiple of q then (6.13) holds by the orthogonality of characters anyway, so we may also suppose that $q_1 > 1$.

Our goal is to show

$$\sum_{b \,(\mathrm{mod}\, q)} \overline{\chi}(b) e\left(\frac{a_1 b}{q_1}\right) = 0.$$

Let $q = q_1 q_2$ and let $b = uq_1 + v$ where

$$0 < u \leqslant q_2 \quad \text{and} \quad 0 < v \leqslant q_1.$$

Then

$$\sum_{v=1}^{q_1} e\left(\frac{v a_1}{q_1}\right) S(v),$$

where

$$S(v) = \sum_{u=1}^{q_2} \overline{\chi}(uq_1 + v).$$

Note that $S(v)$ is periodic modulo q_1. So let c be such that

$$(c, q) = 1, \quad \text{and} \quad c \equiv 1 \,(\mathrm{mod}\, q_1).$$

Then

$$\overline{\chi}(c) S(v) = \sum_{u=1}^{q_2} \overline{\chi}(cuq_1 + cv) = \sum_{u=1}^{q_2} \overline{\chi}(uq_1 + cv) = S(v).$$

If we can construct such a c for which $\chi(c) \neq 1$, then it will show that $S(v) = 0$ for all v, which will finish the proof. The construction of c is where we use primitive. Indeed, there exist c_1, c_2 such that

$$(c_1 c_2, q) = 1, \quad c_1 \equiv c_2 \,(\mathrm{mod}\, q_1), \quad \text{and} \quad \chi(c_1) \neq \chi(c_2),$$

since if $\chi(c_1) = \chi(c_2)$ for all such c_1, c_2, then χ would not be primitive. Finally, we let $c = c_1 c_2^{-1} \,(\mathrm{mod}\, q)$, which finishes the proof. □

Lemma 6.13. *If χ is primitive, then we have*

$$|\tau(\chi)| = \sqrt{q}.$$

Proof. We have

$$|\chi(a)|^2 |\tau(\chi)|^2 = \sum_{m_1 \,(\mathrm{mod}\, q)} \sum_{m_2 \,(\mathrm{mod}\, q)} \chi(m_1) \overline{\chi}(m_2) e\left(\frac{a(m_1 - m_2)}{q}\right).$$

Then we take the sum over all $a \,(\mathrm{mod}\, q)$ of both sides and use the fact that

$$\delta_{m_1 \equiv m_2 \,(\mathrm{mod}\, q)} = \frac{1}{q} \sum_{a \,(\mathrm{mod}\, q)} e\left(\frac{a(m_1 - m_2)}{q}\right)$$

to conclude that

$$|\tau(\chi)|^2 \sum_{a \,(\mathrm{mod}\, q)} |\chi(a)|^2 = q \sum_{m \,(\mathrm{mod}\, q)} |\chi(m)|^2,$$

from which the Lemma follows, since the sum is non-zero. □

4.3. The functional equation for Dirichlet L-functions. Let

$$\kappa = \frac{1}{2}(1-\chi(-1)) = \begin{cases} 0 & \text{if } \chi \text{ is even} \\ 1 & \text{if } \chi \text{ is odd.} \end{cases}$$

Theorem 6.14. *Let χ be a primitive character modulo $q > 1$. Let*

$$\Lambda(s,\chi) = L_\infty(s,\chi)L(s,\chi), \qquad \text{where} \qquad L_\infty(s,\chi) = \Big(\frac{q}{\pi}\Big)^{s/2}\Gamma\Big(\frac{s+\kappa}{2}\Big).$$

Then $\Lambda(s,\chi)$ admits a holomorphic continuation to all of $\mathbb{C}$ and satisfies

$$\Lambda(s,\chi) = \epsilon(\chi)\Lambda(1-s,\overline{\chi})$$

where

$$\epsilon(\chi) = i^{-\kappa}\frac{\tau(\chi)}{\sqrt{q}}$$

is called the root number of χ, and satisfies $|\epsilon(\chi)| = 1$.

Proof. Let $f \in \mathscr{S}(\mathbb{R})$ be an even function if χ is even, and let f be an odd function if χ is odd. Let $\mathrm{Re}(s) > 1$. Then since $\widetilde{[\times n]f}(s) = n^{-s}\widetilde{f}(s)$, we have

$$\widetilde{f}(s)L(s,\chi) = \sum_{n\geqslant 1}\chi(n)\widetilde{[\times n]f}(s) = \sum_{n\geqslant 1}\int_0^\infty \chi(n)f(nx)x^s\,d^\times x = \int_0^\infty\Big(\sum_{n\geqslant 1}\chi(n)f(nx)\Big)x^s\,d^\times x.$$

As before in the proof of Theorem 6.8, we have that

$$\sum_{n\geqslant 1}\chi(n)f(nx) \ll_A \frac{1}{x}\min(1,x^{-A}),$$

so the interchange of summation and integration is justified for $\mathrm{Re}(s) > 1$. Then, following the previous proof, we have

$$\widetilde{f}(s)L(s,\chi) = \int_1^\infty\Big(\sum_{n\geqslant 1}\chi(n)f(nx)\Big)x^s\,d^\times x + \int_1^\infty\Big(\sum_{n\geqslant 1}\chi(n)f(n/x)\Big)x^{-s}\,d^\times x. \tag{6.14}$$

We have in either case of χ being even or odd that

$$\sum_{n\geqslant 1}\chi(n)f(n/x) = \frac{1}{2}\sum_{n\in\mathbb{Z}}\chi(n)f(n/x),$$

since $\chi(0) = 0$. Splitting over residue classes, we have

$$\begin{aligned}
\sum_{n\geqslant 1}\chi(n)f(n/x) &= \frac{1}{2}\sum_{a\,(\mathrm{mod}\,q)}\chi(a)\sum_{n\equiv a\,(\mathrm{mod}\,q)}f(n/x) \\
&= \frac{1}{2}\sum_{a\,(\mathrm{mod}\,q)}\chi(a)\sum_{n\equiv a\,(\mathrm{mod}\,q)}[\times x^{-1}]f(n) \\
&= \frac{1}{2}\sum_{a\,(\mathrm{mod}\,q)}\chi(a)\frac{1}{q}\sum_{n\in\mathbb{Z}}\widehat{[\times x^{-1}]f}\Big(\frac{n}{q}\Big)e\Big(\frac{an}{q}\Big) \\
&= \frac{1}{2}\frac{x}{q}\sum_{a\,(\mathrm{mod}\,q)}\chi(a)\sum_{n\in\mathbb{Z}}\widehat{f}\Big(\frac{nx}{q}\Big)e\Big(\frac{an}{q}\Big),
\end{aligned}$$

by Corollary 6.7 and the rule for the the behavior of Fourier transforms under dilations. Now using the fact that χ is primitive, we have by Lemma 6.12 that

$$\begin{aligned}\sum_{n\geqslant 1}\chi(n)f(n/x) &= \frac{1}{2}\frac{x}{q}\sum_{n\in\mathbb{Z}}\widehat{f}\left(\frac{nx}{q}\right)\sum_{a\,(\mathrm{mod}\,q)}\chi(a)e\left(\frac{an}{q}\right)\\ &= \frac{1}{2}\frac{x}{q}\tau(\chi)\sum_{n\in\mathbb{Z}}\overline{\chi}(n)\widehat{f}\left(\frac{nx}{q}\right).\end{aligned}$$

Since the Fourier transform of an even function is even and the Fourier transform of an odd function is odd, we have

$$\sum_{n\geqslant 1}\chi(n)f(n/x) = \frac{x}{q}\tau(\chi)\sum_{n\geqslant 1}\overline{\chi}(n)\widehat{f}\left(\frac{nx}{q}\right).$$

Thus we have shown that when $\mathrm{Re}(s)>1$ and χ primitive that

$$\widetilde{f}(s)L(s,\chi) = \frac{\tau(\chi)}{q}\int_1^\infty\left(\sum_{n\geqslant 1}\overline{\chi}(n)\widehat{f}\left(\frac{nx}{q}\right)\right)x^{1-s}\,d^\times x + \int_1^\infty\left(\sum_{n\geqslant 1}\chi(n)f(nx)\right)x^s\,d^\times x. \tag{6.15}$$

The right hand side of (6.15) is actually holomorphic in all of $\mathbb{C}$, and so defines the unique analytic continuation of $\widetilde{f}(s)L(s,\chi)$ to $\mathbb{C}$. We have

$$\begin{aligned}&\frac{\tau(\chi)}{q^s}\widetilde{\widehat{f}}(1-s)L(1-s,\overline{\chi}) =\\ &\qquad\frac{\tau(\chi)\tau(\overline{\chi})}{q^s q^{1-s}}\int_1^\infty\left(\sum_{n\geqslant 1}\chi(n)\widetilde{\widehat{f}}(nx)\right)x^s\,d^\times x + \frac{\tau(\chi)}{q^s}\int_1^\infty\left(\sum_{n\geqslant 1}\overline{\chi}(n)\widehat{f}(nx)\right)x^{1-s}\,d^\times x.\end{aligned}$$

Now, by an easy change of variables, we have $\overline{\tau(\chi)} = \chi(-1)\tau(\overline{\chi})$, so by Fourier inversion and Lemma 6.13, we have that

$$\frac{\tau(\chi)}{q^s}\widetilde{\widehat{f}}(1-s)L(1-s,\overline{\chi}) = \widetilde{f}(s)L(s,\chi). \tag{6.16}$$

Let

$$f(x) = \begin{cases} e^{-\pi x^2} & \text{if } \chi \text{ is even}\\ xe^{-\pi x^2} & \text{if } \chi \text{ is odd.}\end{cases} \tag{6.17}$$

We can compute

$$\widehat{f}(y) = \begin{cases} f(y) & \text{if } \chi \text{ is even}\\ -if(y) & \text{if } \chi \text{ is odd,}\end{cases}$$

and

$$\widetilde{f}(s) = \begin{cases} \frac{1}{2}\pi^{-s/2}\Gamma(\frac{s}{2}) & \text{if } \chi \text{ is even}\\ \frac{1}{2\sqrt{\pi}}\pi^{-s/2}\Gamma(\frac{s+1}{2}) & \text{if } \chi \text{ is odd.}\end{cases}$$

Then we derive from (6.16) the formula in the Theorem. □

CHAPTER 7

The Hadamard Factorization

In this chapter, we will establish a formula (called the Hadamard factorization) which expresses a holomorphic function on $\mathbb{C}$ as an infinite product indexed by its zeros.

1. Functions of bounded order

Definition 7.1. *Let $\alpha \geqslant 0$. A holomorphic function $f : s \mapsto f(s)$ on $\mathbb{C}$ is said to be of bounded order if there exists a constant $\alpha \geqslant 0$ such that for all $s \in \mathbb{C}$ and all $\varepsilon > 0$, we have*

$$|f(s)| \ll_{\varepsilon,f} \exp(|s|^{\alpha+\varepsilon}).$$

We will say that such a function f is of order $\leqslant \alpha$. We say that f is of order α if it is of order $\leqslant \alpha$ but not of order $\leqslant \alpha'$ for any $\alpha' < \alpha$.

Example(s) 7.2. A polynomial is a function of order 0. The exponential function $s \mapsto \exp(s)$ is of order 1. More generally, a function of the form $s \mapsto \exp(P(s))$ with P a polynomial is of order $\deg(P)$.

This last example is a function of order $\alpha = \deg(P)$ which does not vanish on $\mathbb{C}$. We have the following converse statement.

Lemma 7.3. *A function of order $\leqslant \alpha$ which does not vanish on $\mathbb{C}$ is of the form*

$$f(s) = \exp(P(s))$$

with P a polynomial of degree $\leqslant \alpha$.

Proof. Since $f(s)$ doesn't vanish, the logarithm $g(s) = \log(f(s))$ is well-defined and holomorphic on $\mathbb{C}$ (indeed a holomorphic function on a simply-connected domain has a well-defined primitive). Consider its Taylor series development around $s = 0$

$$g(s) = \sum_{n \geqslant 0} c_n s^n.$$

We will show that $c_n = 0$ if $n > \alpha$, which will give us the result. We have for $R > 0$ by the Cauchy integral formula

$$c_n = \frac{1}{R^n} \int_0^1 g(R.e(x)) e(-nx)\, dx.$$

By hypothesis, we have for some positive $C_{\varepsilon,f}$ that

$$\mathrm{Re}(g(s)) = \log|f(s)| \leqslant C_{\varepsilon,f} R^{\alpha+\varepsilon}, \quad \text{where} \quad |s| \leqslant R.$$

Note this is a weaker assertion than $\ll_{\varepsilon,f}$, since the left hand side could be negative. Suppose that we had the stronger upper bound

$$|g(s)| \ll_{\varepsilon,f} R^{\alpha+\varepsilon} \quad \text{where} \quad |s| \leqslant R,$$

then we would have that

$$|c_n| \ll_{\varepsilon,f} R^{\alpha+\varepsilon-n},$$

which, taking $R \to \infty$ implies the conclusion of the Lemma.

Unfortunately, we only have an upper bound for the real part of $g(s)$, and so we are forced to perform several contortions to get around this fact. We decompose c_n into its real and imaginary part, $c_n = a_n + ib_n$. Setting $s = R.e(x)$, we have

$$\mathrm{Re}(g(s)) = \sum_{n\geqslant 0} a_n R^n \cos(2\pi nx) - \sum_{n\geqslant 1} b_n R^n \sin(2\pi nx)$$

and thus

$$a_n R^n = \begin{cases} 2\int_0^1 \mathrm{Re}(g(s))\cos(2\pi nx)\,dx & \text{if } n \geqslant 1 \\ \int_0^1 \mathrm{Re}(g(s))\,dx \text{ if } n = 0 \end{cases}$$

and so

$$|a_n|R^n \leqslant 2\int_0^1 |\mathrm{Re}(g(R.e(x)))|\,dx = 2\int_0^1 \left(|\mathrm{Re}(g(R.e(x)))| + \mathrm{Re}(g(R.e(x)))\right)dx - 2a_0,$$

since

$$2\int_0^1 \mathrm{Re}(g(R.e(x)))\,dx = 2a_0.$$

We note that

$$|\mathrm{Re}(g(R.e(x)))| + \mathrm{Re}(g(R.e(x))) = 2\max(0, \mathrm{Re}(g(R.e(x)))) \leqslant C_{\varepsilon,f} R^{\alpha+\varepsilon},$$

and so, for $n > \alpha$

$$|a_n| \leqslant (R^{\alpha+\varepsilon-n} + 2|a_0|R^{-n}) \to 0, \quad R \to \infty.$$

The same reasoning (replacing $\cos(2\pi nx)$ by $\sin(2\pi nx)$) shows that for $n > \alpha$ we have $b_n = 0$ and thus $c_n = 0$. □

Remark 7.4. Note that the conclusion of Lemma 7.3 remains valid if the following condition (à priori weaker) on f is satisfied:

There exists a sequence of positive real numbers $(R_n)_{n\geqslant 0}$ *satisfying* $R_n \to \infty$, $n \to \infty$ *and such that for all* $n \geqslant 0$, *every* $\varepsilon > 0$ *and all* $s \in \mathbb{C}$ *of modulus* $|s| = R_n$ *we have*

$$f(s) \ll_{f,\varepsilon} \exp(|s|^{\alpha+\varepsilon}).$$

In fact, by the maximum principle, this last condition implies that f is of order $\leqslant \alpha$.

2. First estimation of zeros

For $R > 0$ we write

$$Z(f,R) = \{\rho \in \mathbb{C} : f(\rho) = 0, |\rho| \leqslant R\} \subseteq Z(f) = \{\rho \in \mathbb{C} : f(\rho) = 0\},$$

for the set of zeros of f contained in the disk of radius R, and the set of all zeros of f in $\mathbb{C}$, respectively. In the following, the following convention will be useful: given $k : Z(f) \to \mathbb{C}$ a function defined on $Z(f)$, the expressions

$$\sum_{\rho\in Z(f,R)} k(\rho), \quad \text{and} \quad \prod_{\rho\in Z(f,R)} k(\rho)$$

will be used as shorthand for the sum and product

$$\sum_{\rho\in Z(f,R)} m(\rho)k(\rho), \quad \text{and} \quad \prod_{\rho\in Z(f,R)} k(\rho)^{m(\rho)},$$

where $m(\rho)$ is the order of vanishing of f at ρ (otherwise known as the multiplicity of ρ). In particular, the notation

$$N(f,R) := \sum_{\rho \in Z(f,R)} 1$$

will designate the number of zeros of f of modulus $\leqslant R$, counted with multiplicity.

Theorem 7.5. *Let f be a function of order $\leqslant \alpha$, for all R and all $\varepsilon > 0$, we have the upper bound*

$$N(f,R) \ll_{\varepsilon,f} R^{\alpha+\varepsilon} + \begin{cases} 1 & \text{if } f(0) = 0 \\ 0 & \text{if } f(0) \neq 0 \end{cases}.$$

In particular, the series

$$\sum_{\rho \in Z(f)} \frac{1}{1+|\rho|^{\alpha+\varepsilon}}$$

converges.

Note that the second assertion of Theorem 7.5 follows from the first assertion. Indeed, we have

$$\sum_{\rho \in Z(f,R)} \frac{1}{1+|\rho|^{\alpha+\varepsilon}} = \int_0^R N(f,r) \frac{d}{dr}\left(\frac{1}{1+r^{\alpha+\varepsilon}}\right) dr = (\alpha+\varepsilon) \int_0^R \frac{N(f,r) r^{\alpha+\varepsilon}}{r(1+r^{\alpha+\varepsilon})^2}\, dr$$

by integration by parts. Then, by the first part of the Theorem we have

$$N(f,r) \ll_{f,\varepsilon} r^{\alpha+\frac{1}{2}\varepsilon},$$

so that

$$\sum_{\rho \in Z(f,R)} \frac{1}{1+|\rho|^{\alpha+\varepsilon}} \ll_{f,\varepsilon} \int_0^R \frac{r^{2\alpha+\frac{3}{2}\varepsilon}}{r(1+r^{\alpha+\varepsilon})^2}\, dr \ll_{f,\varepsilon} 1$$

as $R \to \infty$. So the sum converges absolutely.

The proof of the first assertion of Theorem 7.5 uses the following formula.

Proposition 7.6 (Jensen)**.** *Let $R > 0$ and f be a holomorphic function in a neighborhood of the disk $\{s \in \mathbb{C} : |s| \leqslant R\}$. Suppose that f does not vanish at 0, nor on the circle of radius R, that is $\{s \in \mathbb{C} : |s| = R\}$. We have the equation*

$$\int_0^1 \log|f(R.e(t))/f(0)|\, dt = \log \prod_\rho \frac{R}{|\rho|} = \sum_\rho \log \frac{R}{|\rho|}$$

where ρ runs over the set of zeros of f of modulus $\leqslant R$ (counted with multiplicity).

Proof. We factor $f(s)$ as

$$f(s) = F(s) \prod_\rho (s-\rho)$$

with $F(s)$ a holomorphic function that does not vanish for $|s| \leqslant R$. So we have

$$\log|f(s)/f(0)| = \log|F(s)/F(0)| + \sum_\rho \log|(s-\rho)/\rho|,$$

and by integrating each term it suffices to prove the proposition to show

(1)

$$\int_0^1 \log\left|\frac{F(R.e(t))}{F(0)}\right| dt = 0,$$

(2)
$$\int_0^1 \log\left|\frac{R.e(t)-\rho}{\rho}\right| = \log\frac{R}{|\rho|}.$$

For the first assertion we have $\log|F(s)/F(0)| = \mathrm{Re}(\log F(s)/F(0))$ and $\log F(s)/F(0)$ is a holomorphic function in a neighborhood of the disk of radius R that vanishes at the point $s=0$ (since, again, $F(s)/F(0)$ is holomorphic and has no zeros in a simply-connected domain). Then,

$$\int_0^1 \log(F(R.e(t))/F(0))\,dt = \frac{1}{2\pi i}\int_{C_R} \log\frac{F(z)}{F(0)}\,\frac{dz}{z} = 0,$$

where C_R is the circle of radius R, by the change of variables $R.e(t) = z$ and the Cauchy integral formula. For the second assertion, we note that

$$\left|\frac{R.e(t)-\rho}{\rho}\right| = \frac{R}{|\rho|}|e(t)-\frac{\rho}{R}| = \frac{R}{|\rho|}|1-\frac{\overline{\rho}}{R}e(t)|$$

and so

$$\log\left|\frac{R.e(t)-\rho}{\rho}\right| = \log(R/|\rho|) + \log\left|1-\frac{\overline{\rho}}{R}e(t)\right|.$$

Then, the fact that

$$\int_0^1 \log\left|1-\frac{\overline{\rho}}{R}e(t)\right| dt = 0$$

comes from the fact that the function $s \mapsto \log(1-\frac{\overline{\rho}}{R}s)$ is a holomorphic function in a neighborhood of the unit disk vanishing at 0, and the second assertion follows by the Cauchy integral formula in the same fashion as above. □

Proof of Theorem 7.5. Suppose for now that $f(0) \neq 0$ and that f has no zeros in the circle of radius $2R$. Let $\rho \in Z(f,R)$, then $\log(2R/|\rho|) \geqslant \log 2$ and thus Jensen's formula gives us that

$$(\log 2)N(f,R) \leqslant \sum_{|\rho|\leqslant R} \log(2R/|\rho|) \leqslant \sum_{|\rho|<2R} \log(2R/|\rho|) = \int_0^1 \log|f(2R.e(t))/f(0)|\,dt.$$

As f is of order $\leqslant \alpha$, this last integral is bounded above by $\ll_\varepsilon (2R)^{\alpha+\varepsilon}$ for all $\varepsilon > 0$.

Now consider the case that f has a zero on the circle of radius $2R$. We can always find a $R' > R$ such that $R' - R \in [0,1]$ and f has no zero on the circle of radius $2R'$. Note that we may assume that $R \gg_f 1$, since there exists some sufficiently small neighborhood of 0 where f has no zeros. Then

$$N(f,R) \leqslant N(f,R') \ll_{\varepsilon,f} (R')^{\alpha+\varepsilon} \ll_f R^{\alpha+\varepsilon},$$

since $R \gg_f 1$. Now suppose f vanishes to order m at 0. We have

$$N(f,R) = N(f(s)/s^m, R) + m \ll_{\varepsilon,f} R^{\alpha+\varepsilon} + 1$$

by the previous result. □

Theorem 7.7 (Hadamard Factorization)**.** *Let f be an entire function of order at most 1 such that $f(0) \neq 0$. Then we have*

$$f(s) = A\exp(bs)\prod_{\rho\in Z(f)}(1-\frac{s}{\rho})e^{s/\rho}$$

with $A = f(0)$, $b = f'(0)/f(0)$, and the product is uniformly convergent on compacts of $\mathbb{C}$.

PROOF. Let $K \subseteq \mathbb{C}$ be a compact set. Then $K \cap Z(f)$ is finite, and for $s \in K$ and $\rho \notin K$ a zero of f, we have

$$(1 - \frac{s}{\rho})e^{s/\rho} = 1 + O_K(\frac{1}{|\rho|^2}).$$

Since the series $\sum_\rho \frac{1}{|\rho|^2}$ converges (we can ignore finitely many zeros when showing convergence), we deduce the uniform convergence of the infinite product on any compact $K \subset \mathbb{C}$.

Consider the infinite product

$$h(s) = \prod_{\rho \in Z(f)} (1 - \frac{s}{\rho})e^{s/\rho}.$$

By construction the function $f(s)/h(s)$ is holomorphic and does not vanish on $\mathbb{C}$. We will now show that the quotient is a function of order at most 1, from which by Lemma 7.3 we can conclude that $f(s)/h(s) = \exp(a + bs)$, which will (almost) finish the theorem. More precisely, we will show that f/h satisfies the condition of remark 7.4.

According to Theorem 7.5, there exists an n_0 such that for all $n \geqslant n_0$ there exists $R_n \in [n, n+1)$ satisfying

$$||\rho| - R_n| \geqslant 1/(4n^2) \tag{7.1}$$

for all zeros $\rho \in Z(f)$.

To see (7.1), it suffices to decompose the annulus of width 1

$$\{s \in \mathbb{C} : |s| \in [n, n+1)\}$$

in $2(n+1)^2$ disjoint sub-annuli of thickness $1/(2(n+1)^2)$. If each of these annuli contained a zero of f, the disk of radius $n+1$ would contain $\geqslant 2(n+1)^2$ zeros, which would contradict Theorem 7.5 for n sufficiently large.

Let $n \geqslant n_0 + 100$ and s such that $|s| = R_n$. We would like to give a lower bound for $|h(s)|$, and to do so it suffices to give an upper bound for

$$\log(|h(s)|^{-1}) = -\sum_\rho \log|(1 - \frac{s}{\rho})e^{s/\rho}|.$$

We decompose the sum in three parts

$$\log(|h(s)|^{-1}) = -\sum_{|\rho| \leqslant |s|-1} \cdots - \sum_{|\rho|-|s| \in (-1,1]} \cdots - \sum_{|\rho| \geqslant |s|+1} \cdots.$$

Let ρ be one of the ρ appearing in the first term. We have

$$-\log|(1 - \frac{s}{\rho})e^{s/\rho}| = -\log|1 - \frac{s}{\rho}| - \mathrm{Re}(\frac{s}{\rho}) \leqslant \log R_n + \frac{|s|}{|\rho|}$$

and so the first term is bounded by

$$\sum_{|\rho| \leqslant R_n - 1} \log R_n + \frac{R_n}{|\rho|} \ll_\varepsilon N(f, R_n) \log R_n + \sum_{|\rho| \leqslant R_n - 1} \frac{R_n^{1+\varepsilon}}{|\rho|^{1+\varepsilon}} \ll_\varepsilon R_n^{1+\varepsilon}.$$

In this upper bound we have used the fact that for all $\varepsilon > 0$

$$|\frac{R_n}{\rho}| \geqslant 1 \Rightarrow |\frac{R_n}{\rho}|^{1+\varepsilon} \geqslant |\frac{R_n}{\rho}|.$$

The second term is bounded above in absolute value by

$$\sum_{|\rho|-|s| \in (-1,1]} \log(R_n^3) \ll_{\varepsilon, f} R_n^{1+\varepsilon},$$

since for such ρ we have by (7.1)

$$R_n^{-3} \ll |(1-\frac{s}{\rho})e^{s/\rho}| \ll 1.$$

The third term is bounded above in absolute value by

$$\sum_{|\rho|\geqslant 1+R_n} \log(1+O(\frac{|s|^2}{|\rho|^2})) \ll \sum_{|\rho|\geqslant 1+R_n} \frac{R_n^2}{|\rho|^2} \leqslant R_n^{1+\varepsilon} \sum_{|\rho|\geqslant 1+R_n} \frac{1}{|\rho|^{1+\varepsilon}} \ll_\varepsilon R_n^{1+\varepsilon},$$

since as $|s|/|\rho|<1$ we see by the Taylor series expansion that $(1-\frac{s}{\rho})e^{s/\rho} = 1+O(|s|^2/|\rho|^2)$. Thus we get the upper bound

$$|h(s)|^{-1} \ll_{\varepsilon,f} \exp(R_n^{1+\varepsilon})$$

for all $\varepsilon>0$.

Therefore there exists a sequence $(R_n)_{n\geqslant 0}$ with $R_n\to\infty$ such that for all n, and every $\varepsilon>0$ and s of modulus R_n we have

$$|\frac{f(s)}{h(s)}| \ll_\varepsilon \exp(R_n^{1+\varepsilon}).$$

By remark 7.4 we get that (since $f(s)/h(s)$ does not have any zeros in $\mathbb{C}$)

$$f(s)/h(s) = \exp(a+bs)$$

for $a,b\in\mathbb{C}$.

Note that for all ρ

$$(1-\frac{s}{\rho})e^{s/\rho}|_{s=0} = 1, \quad \text{and} \quad \frac{d}{ds}(1-\frac{s}{\rho})e^{s/\rho}|_{s=0} = 0$$

and so

$$h(0)=1, \quad h'(0)=0, \quad e^a = f(0), \quad b = f'(0)/e^a = f'(0)/f(0).$$

□

Corollary 7.8. *For all $s\in\mathbb{C}-Z(f)$, we have*

$$\frac{d}{ds}(\log f(s)) := \frac{f'(s)}{f(s)} = b + \sum_{\rho\in Z(f)} \frac{1}{\rho} - \frac{1}{\rho-s}.$$

Moreover, the convergence of the series on the right is uniform on compacts $K\subseteq\mathbb{C}$ such that $K\cap Z(f)=\varnothing$.

Proof. Since the infinite product in Theorem 7.7 converges uniformly on compacts, we have by the Cauchy integral formula

$$f'(s) = bf(s) + A\exp(bs)\sum_\rho ((1-\frac{s}{\rho})e^{s/\rho})' \prod_{\rho'\neq\rho}(1-\frac{s}{\rho})e^{s/\rho'} = bf(s) + A\exp(bs)h(s)\sum_\rho \frac{((1-\frac{s}{\rho})e^{s/\rho})'}{(1-\frac{s}{\rho})e^{s/\rho}}$$

and so for all $s\notin Z(f)$

$$\frac{f'(s)}{f(s)} = b + \sum_\rho \frac{((1-\frac{s}{\rho})e^{s/\rho})'}{(1-\frac{s}{\rho})e^{s/\rho}} = b + \sum_{\rho\in Z(f)} \frac{1}{\rho} - \frac{1}{\rho-s}.$$

Pick $K\cap Z(f)=\varnothing$ a compact, and let $s\in K$. We have

$$\frac{1}{\rho} - \frac{1}{\rho-s} = O_K(\frac{1}{|\rho|^2}),$$

so the series for $f'(s)/f(s)$ converges uniformly on compacts which avoid $Z(f)$. □

CHAPTER 8

The explicit formula

Let us write

$$\xi_0(s) = s(1-s)\xi(s),$$

which is an entire function on $\mathbb{C}$.

We we apply the theory of entire functions that we have just developed to the functions

$$\xi_0(s), \quad \text{and} \quad \Lambda(s,\chi),$$

for χ primitive modulo $q > 1$.

Proposition 8.1. *The function $\xi_0(s)$ is of growth order $\leqslant 1$.*

Proof. By the functional equation we have

$$\xi_0(s) = \xi_0(1-s),$$

and so we are free to suppose that $\mathrm{Re}(s) \geqslant 1/2$ in the below calculations. By (6.10) taking $f(x) = e^{-\pi x^2}$ (see (6.11)), we have

$$\xi_0(s) = -1 + 2s(1-s)\int_1^\infty \Big(\sum_{n\geqslant 1} e^{-\pi n^2 x^2}\Big) x^s \, d^\times x + 2s(1-s)\int_1^\infty \Big(\sum_{n\geqslant 1} e^{-\pi n^2 x^2}\Big) x^{1-s} \, d^\times x. \tag{8.1}$$

For $x \geqslant 1$, we have

$$\sum_{n\geqslant 1} e^{-\pi n^2 x^2} \ll e^{-\pi x^2}$$

and for $\mathrm{Re}(s) \geqslant 1/2$ we have

$$\int_1^\infty \Big(\sum_{n\geqslant 1} e^{-\pi n^2 x^2}\Big) x^s \, d^\times x \ll \int_1^\infty e^{-\pi x^2} x^\sigma \, d^\times x \ll \pi^{-\sigma/2}\Gamma(\frac{\sigma}{2}) \ll_\varepsilon \exp(|s|^{1+\varepsilon}),$$

by exercise 2 from exercise sheet 8 (this is called Stirling's approximation for the gamma function). On the other hand, we have for $\mathrm{Re}(s) \geqslant 1/2$ that

$$\int_1^\infty \Big(\sum_{n\geqslant 1} e^{-\pi n^2 x^2}\Big) x^{1-s} \, d^\times x \ll 1,$$

which finishes the proof. □

Proposition 8.2. *For any χ primitive modulo $q > 1$, the function $\Lambda(s,\chi)$ is of growth order $\leqslant 1$.*

Proof. Exactly the same as the proof of Proposition 8.1, but using (6.15) and (6.17) in place of (6.10) and (6.11). □

1. Application to counting zeros of $\zeta(s)$

Recall that the set $Z(\xi_0)$ of zeros of $\xi_0(s)$ is exactly the set of non-trivial zeros of $\zeta(s)$ (the zeros which are not negative even integers). Note that by the functional equation, the fact that $\Gamma(s)$ does not vanish on $\mathbb{C}$ and the fact that $\zeta(s)$ does not vanish for $\mathrm{Re}(s) > 1$, we have

$$Z(\xi_0) \subset \{s, \mathrm{Re}(s) \in [0,1]\}.$$

We will use the following generic notation for such a zero

$$\rho = \beta + i\gamma, \quad \beta \in [0,1], \quad \gamma \in \mathbb{R}.$$

Taking the logarithmic derivative, we see that

$$\frac{\xi_0'}{\xi_0}(s) = \frac{1}{s} + \frac{1}{s-1} + \frac{\zeta_\infty'}{\zeta_\infty}(s) + \frac{\zeta'(s)}{\zeta(s)} = b + \sum_{\rho \in Z(\xi_0)} \frac{1}{\rho} - \frac{1}{\rho - s}. \tag{8.2}$$

Note that

$$\xi_0(s) = \overline{\xi_0(\overline{s})},$$

since $\xi_0(s)$ takes real values on the real line and is holomorphic. It follows that the multi-set $Z(\xi_0)$ is stable under complex conjugation ($\rho \to \overline{\rho}$ preserves the set of zeros and their multiplicity). Thus

$$\sum_{\rho \in Z(\xi_0)} \frac{1}{\rho} - \frac{1}{\rho - s} = \frac{1}{2}\left(\sum_{\rho \in Z(\xi_0)} \frac{1}{\rho} + \frac{1}{\overline{\rho}} - \frac{1}{\rho - s} - \frac{1}{\overline{\rho} - s} \right).$$

In this last sum, we isolate the terms $1/\rho + 1/\overline{\rho}$: we have

$$\frac{1}{\rho} + \frac{1}{\overline{\rho}} = \frac{2\beta}{|\rho|^2}, \quad \beta \in [0,1]$$

and thus

$$\sum_{\rho \in Z(\xi_0)} \frac{1}{\rho} + \frac{1}{\overline{\rho}} \ll \sum_{\rho} \frac{1}{|\rho|^2} < \infty.$$

It follows that the series

$$\frac{1}{2}\left(\sum_{\rho \in Z(\xi_0)} \frac{1}{s-\rho} + \frac{1}{s-\overline{\rho}} \right)$$

is uniformly convergent on compacts of $\mathbb{C} - Z(\xi_0)$. Thus we have shown

Proposition 8.3. *We have for all $s \notin Z(\xi_0)$,*

$$\frac{\xi_0'}{\xi_0}(s) = \sum_{\rho \in Z(\xi_0)}{}^{*} \frac{1}{s-\rho} + O(1),$$

and moreover

$$-\frac{\zeta'}{\zeta}(s) = \frac{1}{s} + \frac{1}{s-1} - \sum_{\rho \in Z(\xi_0)}{}^{*} \frac{1}{s-\rho} + \frac{\zeta_\infty'}{\zeta_\infty}(s) + O(1).$$

In the above summation, the $*$ indicates the convention that we group the terms of the sum in pairs $(\rho, \overline{\rho})$:

$$\sum_{\rho \in Z(\xi_0)}{}^{*} \frac{1}{s-\rho} = \frac{1}{2} \sum_{\rho \in Z(\xi_0)} \frac{1}{s-\rho} + \frac{1}{s-\overline{\rho}},$$

since otherwise the sums would not converge absolutely.

We will now make use of this identity to more finely count the zeros of ξ_0. Let

$$N(T) = \#\{\rho = \beta + i\gamma \in Z(\xi_0) : |\gamma| \leqslant T\}.$$

Corollary 8.4. *For every $T \geqslant 1$,*

$$N(T+1) - N(T) \ll \log(2+T)$$

and

$$N(T) \ll T\log(2+T).$$

Proof. Let $s = 2 + iT$. By Proposition 8.3, the Dirichlet series for ζ'/ζ, and exercise 4 from exercise sheet 11,

$$\sum_{\rho\in Z(\xi_0)}^{*} \frac{1}{s-\rho} \ll \log(2+T),$$

Taking the real part of this identity we have

$$\operatorname{Re}\left(\frac{1}{s-\rho}\right) = \frac{2-\beta}{|s-\rho|^2} \geqslant \frac{1}{|2-\beta|^2 + |\gamma - T|^2} \geqslant \frac{1}{4+|\gamma-T|^2} \geqslant \frac{1}{5}\delta_{|\gamma-T|\leqslant 1}$$

and so

$$N(T+1) - N(T-1) \leqslant 10 \sum_{\rho\in Z(\xi_0)}^{*} \operatorname{Re}\left(\frac{1}{s-\rho}\right) \ll \log(2+T).$$

The second part follows from the first by summing up. □

Let us also note the following two byproducts of this proof:

$$\sum_{\rho\in Z(\xi_0)}^{*} \frac{1}{4+|\gamma-T|^2} \ll \log(2+T),$$

and

$$\frac{\xi_0'}{\xi_0}(\sigma + iT) = \sum_{\substack{\rho\in Z(\xi_0)\\ |\gamma\pm T|\leqslant 1}}^{*} \frac{1}{s-\rho} + O(\log(2+|T|)), \tag{8.3}$$

for all $-1 \leqslant \sigma \leqslant 2$.

2. Application to counting zeros of $L(s,\chi)$

Suppose that χ is primitive modulo $q > 1$. Recall that $\kappa = \frac{1}{2}(1-\chi(-1))$ and

$$L_\infty(s,\chi) = \left(\frac{\pi}{q}\right)^{-s/2} \Gamma\left(\frac{s+\kappa}{2}\right)$$

$$\Lambda(s,\chi) = L_\infty(s,\chi)L(s,\chi).$$

We have in similar fashion to the zeta function that

$$Z(\Lambda) \subset \{s \in \mathbb{C} : \operatorname{Re}(s) \in [0,1]\}.$$

We have, applying the Hadamard factorization, that

$$\frac{\Lambda'}{\Lambda}(s,\chi) = \frac{L_\infty'}{L_\infty}(s,\chi) + \frac{L'}{L}(s,\chi) = b(\chi) + \sum_{\rho\in Z(\Lambda)} \left(\frac{1}{\rho} - \frac{1}{\rho - s}\right).$$

Here note that the constant $b(\chi)$ now depends on χ and

$$b(\chi) = \frac{\Lambda'}{\Lambda}(0,\chi).$$

It seems to be a difficult problem to estimate $b(\chi)$ as a function of q. However, we can get around this problem by exploiting the functional equation for Λ.

Since χ may take complex values, we no longer have the symmetry that the zeros of Λ are preserved by $\rho \mapsto \overline{\rho}$. However, recalling the functional equation for Λ we have

$$\Lambda(s,\chi) = \varepsilon(\chi)\Lambda(1-s,\overline{\chi}) = \varepsilon(\chi)\overline{\Lambda(1-\overline{s},\chi)},$$

by the Dirichlet series for $L(s,\chi)$ and analytic continuation. Therefore the set $Z(\Lambda)$ is preserved by $\rho \mapsto 1-\overline{\rho}$.

We can use this to eliminate $b(\chi)$ from the above formulas as follows. Note that $\overline{b(\chi)} = b(\overline{\chi})$ since $L(\overline{s},\overline{\chi}) = \overline{L(s,\chi)}$, and also

$$b(\chi) = \frac{\Lambda'}{\Lambda}(0,\chi) = -\frac{\Lambda'}{\Lambda}(1,\overline{\chi}) = -b(\overline{\chi}) - \sum_{\overline{\rho}\in Z(\Lambda(\overline{\chi}))}\left(\frac{1}{1-\overline{\rho}} + \frac{1}{\overline{\rho}}\right)$$

by the functional equation. Putting these facts together, we find

$$2\,\mathrm{Re}(b(\chi)) = -\sum_{\overline{\rho}\in Z(\Lambda(\overline{\chi}))}\left(\frac{1}{1-\overline{\rho}} + \frac{1}{\overline{\rho}}\right).$$

But since the zeros are preserved under $\rho \mapsto 1-\overline{\rho}$, we may write

$$\mathrm{Re}(b(\chi)) = -\frac{1}{2}\sum_{\rho\in Z(\Lambda(\chi))}\left(\frac{1}{\rho} + \frac{1}{\overline{\rho}}\right) = -\sum_{\rho\in Z(\Lambda(\chi))}\mathrm{Re}(\frac{1}{\rho}).$$

This re-arrangement is justified because all of the re-arranged terms are non-negative, and re-arrangement of non-negative terms does not alter the sum.

Remark 8.5. Note that if χ is a real-valued character, then $b(\chi)$ is negative. If χ had a zero very close to 1, then $b(\chi)$ could be extremely large.

We have therefore derived the following:

Proposition 8.6. *We have for all $s \notin Z(\Lambda)$,*

$$\mathrm{Re}(\frac{\Lambda'}{\Lambda}(s,\chi)) = \sum_{\rho\in Z(\Lambda)}^{*}\mathrm{Re}(\frac{1}{s-\rho}),$$

and

$$-\mathrm{Re}(\frac{L'}{L}(s,\chi)) = -\sum_{\rho\in Z(\Lambda)}^{*}\mathrm{Re}(\frac{1}{s-\rho}) + \mathrm{Re}(\frac{L'_\infty}{L_\infty}(s,\chi)),$$

where the $*$ on the sum means we follow the same convention as in Proposition 8.3.

We can use this proposition to count the zeros of $\Lambda(s,\chi)$. Let

$$N(T,\chi) = \#\{\rho = \beta + i\gamma \in Z(\Lambda) : |\gamma| \leqslant T\}.$$

Corollary 8.7. *For every $T \geqslant 1$,*

$$N(T+1,\chi) - N(T,\chi) \ll \log(q(2+T))$$

and

$$N(T,\chi) \ll T\log(q(2+T)).$$

PROOF. Let $s = 2 + iT$. Exactly as before, by Proposition 8.6, and exercise 4 from exercise sheet 11,

$$\sum_{\rho \in Z(\Lambda)}^{*} \mathrm{Re}(\frac{1}{s-\rho}) \ll \log(q(2+T)),$$

and

$$\mathrm{Re}\left(\frac{1}{s-\rho}\right) \geqslant \frac{1}{4+|\gamma - T|^2} \geqslant \frac{1}{5}\delta_{|\gamma - T| \leqslant 1}$$

and so

$$N(T+1,\chi) - N(T-1,\chi) \leqslant 10 \sum_{\rho \in Z(\Lambda)}^{*} \mathrm{Re}\left(\frac{1}{s-\rho}\right) \ll \log(q(2+T)).$$

□

We also have

$$\sum_{\rho \in Z(\Lambda)}^{*} \frac{1}{4+|\gamma - T|^2} \ll \log(q(2+T)),$$

and

$$\frac{\Lambda'}{\Lambda}(\sigma + iT) = \sum_{\substack{\rho \in Z(\Lambda) \\ |\gamma \pm T| \leqslant 1}}^{*} \frac{1}{s-\rho} + O(\log(q(2+|T|))), \tag{8.4}$$

for all $-1 \leqslant \sigma \leqslant 2$.

3. Weil's explicit formula

In this section we prove the following theorem.

Theorem 8.8. *Let $f \in \mathscr{C}_c^{\infty}(\mathbb{R}_{>0})$ and*

$$\widetilde{f}(s) = \int_0^{\infty} f(x) x^s \frac{dx}{x}$$

its Mellin transform. Let $\check{f}(x) = x^{-1} f(x^{-1})$. We have the identity

$$\sum_{n \geqslant 1} (f(n) + \check{f}(n))\Lambda(n) = \widetilde{f}(1) + \widetilde{f}(0) + \frac{1}{2\pi i}\int_{(1/2)} \left(\frac{\zeta_\infty'}{\zeta_\infty}(s) + \frac{\zeta_\infty'}{\zeta_\infty}(1-s)\right)\widetilde{f}(s)\, ds - \sum_{\rho \in Z(\xi_0)} \widetilde{f}(\rho).$$

Recall that the notation $\int_{(c)}$ denotes the integral $\int_{c-i\infty}^{c+i\infty}$ along the vertical line, oriented counter-clockwise.

PROOF. We calculate the following integral in two different ways:

$$\frac{1}{2\pi i}\int_{(3/2)} \widetilde{f}(s) \frac{\xi_0'}{\xi_0}(s)\, ds.$$

For $\mathrm{Re}(s) > 1$ we have

$$\frac{\xi_0'}{\xi_0}(s) = \frac{1}{s} + \frac{1}{s-1} + \frac{\zeta_\infty'}{\zeta_\infty}(s) - \sum_{n \geqslant 1} \frac{\Lambda(n)}{n^s}$$

so that inverting summation and integration we get thanks to the Mellin inversion formula

$$\begin{aligned} \frac{1}{2\pi i}\int_{(3/2)} \widetilde{f}(s) \frac{\xi_0'}{\xi_0}(s)\, ds &= \frac{1}{2\pi i}\int_{(3/2)} \left(\frac{1}{s} + \frac{1}{s-1} + \frac{\zeta_\infty'}{\zeta_\infty}(s)\right)\widetilde{f}(s)\, ds - \sum_{n \geqslant 1} \Lambda(n) \frac{1}{2\pi i}\int_{(3/2)} \widetilde{f}(s) n^{-s}\, ds \\ &= \frac{1}{2\pi i}\int_{(3/2)} \left(\frac{1}{s} + \frac{1}{s-1} + \frac{\zeta_\infty'}{\zeta_\infty}(s)\right)\widetilde{f}(s)\, ds - \sum_{n \geqslant 1} \Lambda(n) f(n). \end{aligned} \tag{8.5}$$

Now we calculate the integral on the left hand side of (8.5) by deforming the contour. Let $T>0$ and R_T the rectangle whose corners are at $3/2\pm iT$ and $-1/2\pm iT$. We can choose T as large as we would like so that for all $\rho\in Z(\xi_0)$, we have

$$|T\pm\gamma|\gg\frac{1}{\log(2+T)}.$$

Choosing this special T and applying it in (8.3), we find for this special T and all $-1\leqslant\sigma\leqslant 2$

$$\frac{\xi_0'}{\xi_0}(\sigma+iT)\ll\log(2+|T|)^2.$$

We have (integrating counterclockwise)

$$\frac{1}{2\pi i}\int_{R_T}\widetilde{f}(s)\frac{\xi_0'}{\xi_0}(s)\,ds=\sum_{\substack{\zeta(\rho)=0\\ \beta\in[0,1],|\gamma|\leqslant T}}\operatorname{res}_{s=\rho}\widetilde{f}(s)\frac{\xi_0'}{\xi_0}(s)=\sum_{\substack{\zeta(\rho)=0\\ \beta\in[0,1],|\gamma|\leqslant T}}\widetilde{f}(\rho).$$

As $T\to\infty$, the integrals along the horizontal segments tend to 0. We therefore get

$$\sum_{\substack{\zeta(\rho)=0\\ \beta\in[0,1]}}\widetilde{f}(\rho)=\frac{1}{2\pi i}\int_{(3/2)}\widetilde{f}(s)\frac{\xi_0'}{\xi_0}(s)\,ds-\frac{1}{2\pi i}\int_{(-1/2)}\widetilde{f}(s)\frac{\xi_0'}{\xi_0}(s)\,ds.$$

Note that the series converges, since $\widetilde{f}(s)\ll(1+|\operatorname{Im}(s)|)^{-A}$ for all $A\gg 0$ by (6.9).

Now we do the change of variables $s\leftrightarrow 1-s$ and we use the functional equation

$$\frac{\xi_0'}{\xi_0}(s)=-\frac{\xi_0'}{\xi_0}(1-s),$$

which gives

$$\begin{aligned}\sum_{\substack{\zeta(\rho)=0\\ \beta\in[0,1]}}\widetilde{f}(\rho)&=\frac{1}{2\pi i}\int_{(3/2)}\big(\widetilde{f}(s)+\widetilde{f}(1-s)\big)\frac{\xi_0'}{\xi_0}(s)\,ds\\&=\frac{1}{2\pi i}\int_{(3/2)}\left(\frac{1}{s}+\frac{1}{s-1}+\frac{\zeta_\infty'}{\zeta_\infty}(s)\right)\big(\widetilde{f}(s)+\widetilde{f}(1-s)\big)\,ds-\sum_{n\geqslant 1}\Lambda(n)\big(f(n)+\check{f}(n)\big),\end{aligned}$$

since

$$\check{f}(n)=\frac{1}{2\pi i}\int_{(3/2)}\widetilde{f}(1-s)n^{-s}\,ds.$$

It remains to calculate

$$\frac{1}{2\pi i}\int_{(3/2)}\left(\frac{1}{s}+\frac{1}{s-1}\right)\big(\widetilde{f}(s)+\widetilde{f}(1-s)\big)\,ds=\widetilde{f}(1)+\widetilde{f}(0)+\frac{1}{2\pi i}\int_{(1/2)}\left(\frac{1}{s}+\frac{1}{s-1}\right)\big(\widetilde{f}(s)+\widetilde{f}(1-s)\big)\,ds$$

by moving the contour (we pass a pole at $s=1$ of residue $\widetilde{f}(1)+\widetilde{f}(0)$). Making the change of variables $s\leftrightarrow 1-s$ we get

$$\frac{1}{2\pi i}\int_{(1/2)}\left(\frac{1}{s}+\frac{1}{s-1}\right)\big(\widetilde{f}(s)+\widetilde{f}(1-s)\big)\,ds=\frac{1}{2\pi i}\int_{(1/2)}\left(\frac{1}{1-s}-\frac{1}{s}\right)\big(\widetilde{f}(s)+\widetilde{f}(1-s)\big)\,ds=0.$$

□

The same method of proof can also be used to show the following generalization.

Theorem 8.9. *With the same hypotheses as Theorem 8.8, we have the identity*

$$\sum_{n\geqslant 1}\left(\chi(n)f(n)+\overline{\chi}(n)\check{f}(n)\right)\Lambda(n)=\frac{1}{2\pi i}\int_{(1/2)}\left(\frac{L'_\infty(s,\chi)}{L_\infty(s,\chi)}+\frac{L'_\infty(1-s,\overline{\chi})}{L_\infty(1-s,\overline{\chi})}\right)\widetilde{f}(s)\,ds-\sum_{\substack{L(\rho,\chi)=0\\ \mathrm{Re}(\rho)\in[0,1]}}\widetilde{f}(\rho).$$

Proof. See exercise 2 of sheet 13. □

We will now show how a zero-free region for the zeta function implies the prime number theorem with an explicit error term. In the next chapter, we will show the following theorem.

Theorem 8.10 (Hadamard and de la Vallée-Poussin)**.** *There exists an absolute constant $c>0$ such that $\zeta(s)$ does not vanish in the region $(s=\sigma+it)$*

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

Remark: In fact, $c=1/5.69693$ is admissible (Kadiri 2005).

Corollary 8.11. *There exists $C>0$, such that for $f\in\mathscr{C}_c^\infty(\mathbb{R}_{>0})$ and $X\geqslant 2$*

$$\sum_{n\geqslant 1}\Lambda(n)f(\frac{n}{X})=X\int_{\mathbb{R}}f(x)\,dx+O_f\left(X\exp(-C\sqrt{\log X})\right).$$

Proof. Let $f_X(x)=f(x/X)$. We have

$$\widetilde{f_X}(s)=\widetilde{f}(s)X^s$$

and the explicit formula gives for X sufficiently large (so that $f(Xn)=0$ for all $n\geqslant 1$)

$$\sum_{n\geqslant 1}\Lambda(n)f(\frac{n}{X})=X\int_{\mathbb{R}}f(x)\,dx+\int_{\mathbb{R}}f(x)\frac{dx}{x}+\frac{1}{2\pi i}\int_{(1/2)}\left(\frac{\zeta'_\infty}{\zeta_\infty}(s)+\frac{\zeta'_\infty}{\zeta_\infty}(1-s)\right)\widetilde{f}(s)X^s\,ds-\sum_{\substack{\zeta(\rho)=0\\ \mathrm{Re}(\rho)\in[0,1]}}\widetilde{f}(\rho)X^\rho.$$

The second term above is $O_f(1)$, the third is $O_f(X^{1/2})$ (since $|X^s|=X^{1/2}$ for $\mathrm{Re}(s)=1/2$) and for the fourth we have (writing $\rho=\beta+i\gamma$)

$$\sum_{\substack{\zeta(\rho)=0\\ \mathrm{Re}(\rho)\in[0,1]}}|\widetilde{f}(\beta+i\gamma)|X^\beta\leqslant X\sum_{\substack{\zeta(\rho)=0\\ \mathrm{Re}(\rho)\in[0,1]}}|\widetilde{f}(\beta+i\gamma)|X^{-\frac{c}{\log(2+|\gamma|)}}=X\sum_{\substack{\zeta(\rho)=0\\ \mathrm{Re}(\rho)\in[0,1]}}|\widetilde{f}(\beta+i\gamma)|\exp(-\frac{c\log X}{\log(2+|\gamma|)}).$$

Next we split the sum over zeros as follows

$$\sum_{\substack{\zeta(\rho)=0\\ \mathrm{Re}(\rho)\in[0,1]}}(\cdots)=\sum_{\substack{\zeta(\rho)=0\\ \mathrm{Re}(\rho)\in[0,1]\\ \log(2+|\gamma|)\leqslant\sqrt{\log X}}}(\cdots)+\sum_{\substack{\zeta(\rho)=0\\ \mathrm{Re}(\rho)\in[0,1]\\ \log(2+|\gamma|)>\sqrt{\log X}}}(\cdots)\tag{8.6}$$

The first term is bounded by

$$\ll\exp(-c\sqrt{\log X})\sum_{\substack{\zeta(\rho)=0\\ \mathrm{Re}(\rho)\in[0,1]}}|\widetilde{f}(\beta+i\gamma)|\ll\exp(-c\sqrt{\log X}).$$

For the second, we use the fact that

$$|\widetilde{f}(\beta+i\gamma)|\ll(1+|\gamma|)^{-2}$$

as $\gamma \to \infty$ to see that the second term in (8.6) is bounded by

$$\ll \sum_{\substack{\zeta(\rho)=0 \\ \mathrm{Re}(\rho)\in[0,1] \\ \log(2+|\gamma|)>\sqrt{\log X}}} |\widetilde{f}(\beta+i\gamma)| \ll \sum_{\substack{\zeta(\rho)=0 \\ \mathrm{Re}(\rho)\in[0,1] \\ \log(2+|\gamma|)>\sqrt{\log X}}} \frac{\exp(-\frac{1}{2}\sqrt{\log X})}{(1+|\gamma|)^{3/2}} \ll \exp(-\frac{1}{2}\sqrt{\log X}).$$

□

Remark 8.12. The above corollary counts prime numbers p in a window of size $\sim X$ when they are weighted by $(\log p) f(p/X)$ for f a fixed smooth function with compact support. This choice permits us to use the rapid decay of $|\widetilde{f}(\beta+i\gamma)|$, which comes from integration by parts applied two times to the Mellin transform. The implicit constant in the term $O_f(X\exp(-C\sqrt{\log X}))$ depends thus directly on the size of f and its first two derivatives. Choosing a sequence of functions f depending on X, we can by approximation arguments replace the term $f(n/X)$ by the characteristic function of the interval $[1, X]$ and obtain

$$\sum_{n\leqslant X} \Lambda(n) = X + O(X\exp(-C'\sqrt{\log X})), \tag{8.7}$$

which is the original formulation of the prime number theorem. The passage from Corollary 8.11 to (8.7) is an exercise in classical analysis; all of the number-theoretic ideas are already contained in the former. We leave the derivation of (8.7) to the exercise sheets.

CHAPTER 9

The theorem of Hadamard and de la Vallée-Poussin

As we saw, the zeros of $\xi_0(s)$ are all situated in the critical strip

$$\{s \in \mathbb{C} : \mathrm{Re}(s) \in [0,1]\}.$$

This follows from the fact that $\zeta(s)$ does not vanish for $\mathrm{Re}(s) > 1$, the fact that $\Gamma(s)$ does not vanish on $\mathbb{C}$, and the functional equation for $\xi(s)$. In this chapter, we will improve this zone of non-vanishing and show that ξ_0 does not vanish for s slightly inside the critical strip.

Theorem 9.1. *There exists a constant $c > 0$ such that*

$$Z(\xi_0) \subset \{s \in \mathbb{C} : \mathrm{Re}(s) \leqslant 1 - \frac{c}{\log(2 + |\mathrm{Im}(s)|)}\}.$$

Thus $\zeta(s)$ does not vanish for

$$\mathrm{Re}(s) > 1 - \frac{c}{\log(2 + |\mathrm{Im}(s)|)}. \tag{9.1}$$

1. Warm-up: Qualitative zero free region

In the previous chapter we showed that Theorem 9.1 of Hadamard and de la Vallée-Poussin in its quantitative form above implies the quantitative form of the proof of the prime number theorem 8.11, and in exercises you saw that it even implies (8.7). In fact, we even have that the "qualitative" zero-free region

$$\zeta(\beta + i\gamma) = 0 \quad \Rightarrow 1 - \beta > 0$$

implies the "qualitative" prime number theorem in the form

$$\sum_{n \leqslant X} \Lambda(n) = X + o(X),$$

by the same proof as before.

As a warm-up to the quantitative zero-free region of Theorem 9.1 in the next section, we first prove the qualitative version in this section.

Theorem 9.2 (Qualitative zero-free region)**.** *For all $t \in \mathbb{R}$ we have that $\zeta(1 + it) \neq 0$.*

Proof. First, recall that ζ has a simple pole at $s = 1$, which implies that the negative of the logarithmic derivative $-\frac{\zeta'}{\zeta}$ has a simple pole at $s = 1$ of residue 1. For $\sigma > 1$ we have

$$\sum_{n \geqslant 1} \frac{\Lambda(n)}{n^\sigma} = -\frac{\zeta'}{\zeta}(\sigma) = \frac{1}{\sigma - 1} + O(1) \tag{9.2}$$

which tends to infinity as $\sigma \to 1^+$.

The function ζ is holomorphic. For $t \neq 0$, suppose ζ had a zero at $s = 1 + it$. Let $\ell \geqslant 0$ denote its order of vanishing. Then $-\frac{\zeta'}{\zeta}$ has a simple pole of residue $-\ell$. Thus for $\sigma > 1$ we have

$$\sum_{n \geqslant 1} n^{-it} \frac{\Lambda(n)}{n^\sigma} = -\frac{\zeta'}{\zeta}(\sigma + it) = \frac{-\ell}{\sigma - 1} + O_t(1) \leqslant C_t \tag{9.3}$$

for some positive function C_t of t, since $-\ell/(\sigma - 1) \leqslant 0$. Note that $|n^{-it}| = 1$ and $\Lambda(n) \geqslant 0$, so that we also have

$$\left| \sum_{n \geqslant 1} n^{-it} \frac{\Lambda(n)}{n^\sigma} \right| \leqslant \frac{1}{\sigma - 1} + O(1)$$

by (9.2). This implies that $|\ell| \leqslant 1$, i.e. that $\ell \in \{0, 1\}$.

Suppose that $\ell = 1$. Then

$$\sum_{n \geqslant 1} n^{-it} \frac{\Lambda(n)}{n^\sigma} = \frac{-1}{\sigma - 1} + O_t(1). \tag{9.4}$$

Informally, combining (9.2) and (9.4) we find that if n is such that $\Lambda(n) \neq 0$ then $n^{-it} \approx -1$, so that we would have $n^{-2it} \approx 1$, and so (9.2) suggests

$$\sum_{n \geqslant 1} n^{-2it} \frac{\Lambda(n)}{n^\sigma} = \frac{1}{\sigma - 1} + O(1).$$

But then also we have by (9.3)

$$\sum_{n \geqslant 1} n^{-2it} \frac{\Lambda(n)}{n^\sigma} \leqslant C_t, \tag{9.5}$$

some positive constant depending on t. This is a contradiction.

How do we make this argument precise? By the continuity of $z \mapsto z^2$ at $z = -1$ we have that for all $\varepsilon > 0$ there exists $\delta > 0$ such that $\mathrm{Re}(n^{-it}) + 1 \leqslant \delta$ implies that $1 - \mathrm{Re}(n^{-2it}) \leqslant \varepsilon$. Then

$$\sum_{\substack{n \geqslant 1 \\ 1 - \mathrm{Re}(n^{-2it}) > \varepsilon}} \frac{\Lambda(n)}{n^\sigma} \leqslant \sum_{\substack{n \geqslant 1 \\ \mathrm{Re}(n^{-it}) + 1 > \delta}} \frac{\Lambda(n)}{n^\sigma} \leqslant \sum_{n \geqslant 1} \frac{\mathrm{Re}(n^{-it}) + 1}{\delta} \frac{\Lambda(n)}{n^\sigma} = O_t(\frac{1}{\delta}), \tag{9.6}$$

where the last equality follows from (9.2) and (9.4) since

$$\frac{1}{\sigma - 1} + \frac{-1}{\sigma - 1} = 0.$$

This then implies that

$$\sum_{n \geqslant 1} \Big(1 - \mathrm{Re}(n^{-2it})\Big) \frac{\Lambda(n)}{n^\sigma} = \sum_{\substack{n \geqslant 1 \\ 1 - \mathrm{Re}(n^{-2it}) > \varepsilon}} (\cdots) + \sum_{\substack{n \geqslant 1 \\ 1 - \mathrm{Re}(n^{-2it}) \leqslant \varepsilon}} (\cdots) \leqslant C_t \left(\frac{1}{\delta} + \frac{\varepsilon}{\sigma - 1} \right),$$

where the last inequality follows from (9.2) and (9.6). (Side note: we could just as well have written $\ll_t$ instead of $\leqslant C_t$ here.) Then we take $\varepsilon = \frac{1}{10 C_t}$, and σ close enough to 1 so that $\sigma - 1 \leqslant \frac{\delta}{10 C_t}$. Then we get

$$\frac{2/10}{\sigma - 1} \geqslant \sum_n \Big(1 - \mathrm{Re}(n^{-2it})\Big) \frac{\Lambda(n)}{n^\sigma} = \sum_n \frac{\Lambda(n)}{n^\sigma} - \mathrm{Re}\left(\sum_n n^{-2it} \frac{\Lambda(n)}{n^\sigma} \right) \geqslant \frac{1}{\sigma - 1} + O_t(1),$$

by (9.3). Taking $\sigma \to 1^+$ gives us a contradiction. This establishes the qualitative zero-free region, hence the qualitative prime number theorem. □

2. Quantitative zero free region

We need to clean up the above argument and make explicit how everything depends on t. Recall from (8.3) that

$$-\frac{\zeta'}{\zeta}(s) = \frac{1}{s-1} - \sum_{\substack{\rho\in Z(\xi_0)\\ |s-\rho|\leqslant 1}} \frac{1}{s-\rho} + O(\log(2+|t|)). \tag{9.7}$$

Note that for $|s-1| \gg (\log(2+|t|)^{-1})$, that is for $|t| \gg (\log(2+|t|)^{-1})$, we can absorb the first term $1/(s-1)$ into the error term.

Suppose henceforth that $\sigma > 1$. Since each ρ has real part at most 1, it follows that each fraction $1/(s-\rho)$ has positive real part when $s = \sigma + it$. Thus for $|t| \gg (\log(2+|t|)^{-1})$, we have

$$-\operatorname{Re}\left(\frac{\zeta'}{\zeta}(\sigma+it)\right) \leqslant C_t(\log(2+|t|)^{-1}) \tag{9.8}$$

for some positive constant $C_t > 0$. Now suppose that $\zeta(\beta+it) = 0$. We aim to show that $1-\beta \gg (\log(2+|t|)^{-1}$. Since ζ has a pole at $s=1$, we know that it has no zeros in that region. So we may suppose that $|t| \gg 1$. We improve on (9.3) by taking into account the contribution from $\rho = \beta + it$:

$$\sum_n \operatorname{Re}(n^{-it})\frac{\Lambda(n)}{n^\sigma} = -\operatorname{Re}\left(\frac{\zeta'}{\zeta}(\sigma+it)\right) \leqslant -\frac{1}{\sigma-\beta} + O(\log(2+|t|)). \tag{9.9}$$

As before, we use (9.2) to write this as

$$\sum_n \left(1+\operatorname{Re}(n^{-it})\right)\frac{\Lambda(n)}{n^\sigma} \leqslant \frac{1}{\sigma-1} - \frac{1}{\sigma-\beta} + O(\log(2+|t|)). \tag{9.10}$$

We also note by (9.2) and (9.9) applied to $2t$ that

$$\sum_n \left(1-\operatorname{Re}(n^{-2it})\right)\frac{\Lambda(n)}{n^\sigma} \geqslant \frac{1}{\sigma-1} + O(\log(2+|t|)). \tag{9.11}$$

Now we aim to implement more quantitatively the idea that $n^{-it} \approx -1$ implies $n^{-2it} \approx 1$. Let us write $n^{-it} = e^{i\theta}$ so that $\operatorname{Re}(n^{-it}) + 1 = \cos(\theta) + 1$ and $1 - \operatorname{Re}(n^{-2it}) = 1 - \cos(2\theta)$. These quantities are related by the formula $\cos(2\theta) = 2\cos^2(\theta) - 1$, from which we deduce that

$$1-\cos(2\theta) = 2(1-\cos^2(\theta)) = 2(1-\cos(\theta))(1+\cos(\theta)) \leqslant 4(1+\cos(\theta)),$$

or

$$1 - \operatorname{Re}(n^{-2it}) \leqslant 4(1+\operatorname{Re}(n^{-it})). \tag{9.12}$$

Insert this into (9.10) and (9.11) to get

$$\frac{1}{\sigma-1} \leqslant 4\left(\frac{1}{\sigma-1} - \frac{1}{\sigma-\beta}\right) + O(\log(2+|t|)), \tag{9.13}$$

which can be re-arranged as

$$\frac{4}{\sigma-\beta} - \frac{3}{\sigma-1} \leqslant C(\log(2+|t|)), \tag{9.14}$$

for some positive constant $C > 0$. We choose σ so that $\sigma - 1 = (1-\beta)/\varepsilon$ for some small enough $\varepsilon > 0$. Note that if β is close enough to 1, then $\sigma < 10$. Then $\sigma - \beta = (1+1/\varepsilon)(1-\beta)$, so that

(9.14) reads

$$\frac{1}{1-\beta}\left(\frac{4}{1+1/\varepsilon}-3\varepsilon\right)\leqslant C(\log(2+|t|)).$$

Taking $\varepsilon=1/4$, say, leads to the required estimate $1-\beta\gg(\log(2+|t|))^{-1}$.

3. Zero-free region for Dirichlet L-functions

Let χ be a primitive Dirichlet character modulo $q>1$, and $t\in\mathbb{R}$.

Theorem 9.3. *We have that $L(\beta+i\gamma,\chi)=0$ implies that $1-\beta\gg(\log(q(2+|\gamma|)))^{-1}$, except possibly when $\chi^2=\chi_0$, i.e. χ is "quadratic", in which case there is at most one simple zero $\beta\in\mathbb{R}$, $L(\beta,\chi)=0$ with $1-\beta\ll(\log(q(2+|\gamma|)))^{-1}$.*

Proof. As before, we have

$$\sum_{n\geqslant 1}\mathrm{Re}(n^{-it}\chi(n))\frac{\Lambda(n)}{n^{\sigma}}=-\mathrm{Re}\left(\frac{L'}{L}(\sigma+it,\chi)\right)\leqslant\frac{-1}{\sigma-\beta}+O(\log(q(2+|t|))).$$

Setting

$$\eta=\begin{cases}1 & \text{if } \chi^2 \text{ is trivial}\\ 0 & \text{otherwise,}\end{cases}$$

we also have

$$\sum_{n\geqslant 1}\mathrm{Re}(n^{-2it}\chi^2(n))\frac{\Lambda(n)}{n^{\sigma}}=-\mathrm{Re}(\frac{L'}{L}\left(\sigma+it,\chi^2\right)\leqslant\frac{\eta}{\sigma+it-1}+O(\log(q(2+|t|))). \tag{9.15}$$

If χ^2 is non-trivial, then $\eta=0$, and we have that (9.15) is $\ll\log(q(2+|t|))$. If $|t|\gg\log(q(2+|t|))^{-1}$, then we also have that $(\sigma+it-1)^{-1}\ll\log(q(2+|t|))$, so (9.15) is likewise $\ll\log(q(2+|t|))$.

So, we argue as before and it gives the required conclusion in these cases. It remains to consider the case that χ is non-trivial, and χ^2 is trivial and $|t|\ll\log(q(2+|t|))^{-1}$. So, suppose we can find two zeros

$$\beta_1+it_1,\text{ and }\beta_2+it_2\text{ of }L(s,\chi)$$

with $1-\beta_j,\ t_j\ll\log(q(2+|t|))^{-1}$. Then for $\sigma>1$, we have (Proposition 8.6)

$$-\sum_{n\geqslant 1}\chi^2(n)\frac{\Lambda(n)}{n^{\sigma}}\leqslant\sum_{n\geqslant 1}\chi(n)\frac{\Lambda(n)}{n^{\sigma}}=-\mathrm{Re}\left(\frac{L'}{L}(\sigma,\chi)\right)\leqslant-\mathrm{Re}\left(\frac{1}{\sigma-\beta_1-it_1}\right)-\mathrm{Re}\left(\frac{1}{\sigma-\beta_2-it_2}\right)+O(\log(q(2+|t|))), \tag{9.16}$$

by dropping negative terms, and

$$-\sum_{n\geqslant 1}\chi^2(n)\frac{\Lambda(n)}{n^{\sigma}}=-\sum_{(n,q)=1}\frac{\Lambda(n)}{n^{\sigma}}=-\frac{\zeta'}{\zeta}(\sigma)+\sum_{p\mid q}(\log p)\left(\frac{1}{p^{\sigma}}+\frac{1}{p^{2\sigma}}+\cdots\right),$$

and the second of these terms is

$$\ll\log q=O(\log(q(2+|t|))).$$

Thus

$$\mathrm{Re}\left(\frac{1}{\sigma-\beta_1-it_1}\right)+\mathrm{Re}\left(\frac{1}{\sigma-\beta_2-it_2}\right)+O(\log(q(2+|t|)))$$

for all $1\leqslant\sigma<10$. Now let us take $\sigma=1+c\log(q(2+|t|))^{-1}$, with c small enough so that we get a contradiction with

$$(2-\beta_1-\beta_2)c^{-1}\log(q(2+|t|))\leqslant c^{-1}\log(q(2+|t|))+O(\log(q(2+|t|))),$$

since $2-\varepsilon > 1$.

Now if $\beta+it$ is a zero of $L(s,\chi)$ with $1-\beta,\ t \ll \log(q(2+|t|))^{-1}$, and if $t \neq 0$, then (since $\chi = \overline{\chi}$), so it $\beta - it$. So we get two zeros in this region since $\overline{L(s,\chi)} = L(\overline{s},\overline{\chi}) = L(\overline{s},\chi)$. Contradiction. Thus any possible zero in this region satisfies $t = 0$.

If a real zero β occurs with multiplicity $\geqslant 2$, then we apply previous arguments with $\beta_1 = \beta_2 = \beta$. Thus any real zero with $1-\beta \ll 1$ is simple. □

CHAPTER 10

Siegel's Theorem

The possible simple real zeros $\beta \in \mathbb{R}$ with $1-\beta \ll (\log q)^{-1}$ that could occur when $\chi^2 = \chi_0$ according to Lemma 9.3 are called *exceptional* zeros, or sometimes *Siegel* zeros. The possible primitive quadratic χ for which an exceptional zero may occur is called an *exceptional* character. In this chapter, we will show that even though we cannot exclude (at present!) the existence of exceptional zeros, we can at least show that any exceptional zero cannot be too close to 1, quantified in terms of the conductor of χ.

In a sense that will be made more precise later, if $L(s,\chi)$ admits a zero β which is very close to 1, then $L(1,\chi)$ must be small. So we first embark on showing lower bounds for $L(1,\chi)$ in terms of the conductor of χ.

Let χ be primitive quadratic modulo $q > 1$. Recall from Theorem 4.13 that we showed that $L(1,\chi) \neq 0$ for any primitive χ. Dirichlet's theorem could be considered a "qualitative" theorem, and we now make that theorem quantitative, following the same basic ideas of proof. The reader is encouraged to compare the proof of the following proposition to the proof of Theorem 4.13.

Proposition 10.1. *For χ primitive quadratic modulo $q > 1$ we have $L(1,\chi) \gg q^{-1/2}$.*

Remark 10.2. The constant implicit in the $\gg$ notation above could be given a specific numerical value, if we were sufficiently motivated to do so. Such an implicit constant is called "effective". This will be in contrast to the constant appearing in Siegel's theorem, later.

PROOF. Let $r \in \mathscr{A}$ be defined by $r = 1 * \chi$. That is,

$$r : n \mapsto \sum_{d|n} \chi(d).$$

I claim that $r(n) \geqslant 0$ for all n, and $r(n) \geqslant 1$ when n is a perfect square. Indeed: The function r is multiplicative (see Proposition 2.17), so we have

$$r(n) = \prod_{p|n} r(p^{\nu_p(n)}),$$

and

$$r(p^{\nu_p(n)}) = 1 + \chi(p) + \chi(p)^2 + \cdots + \chi(p)^{\nu_p(n)},$$

since χ is completely multiplicative. There are only 3 possibilities for $\chi(p)$, since χ is quadratic, we have $\chi(p) \in \{-1, 0, 1\}$ (see subsubsection 1.1.2, and the definition of a Dirichlet character). We have that $r(p^{\nu_p(n)}) \geqslant 0$ since

- If $\chi(p) = 1$, then $r(p^{\nu_p(n)}) \geqslant \nu_p(n) + 1 \geqslant 1$.
- If $\chi(p) = 0$, then $r(p^{\nu_p(n)}) = 1$.
- If $\chi(p) = -1$, then

$$r(p^{\nu_p(n)}) = \begin{cases} 0 & \text{if } \nu_p(n) \text{ is odd} \\ 1 & \text{if } \nu_p(n) \text{ is even.} \end{cases}$$

And observe also that if $v_p(n)$ is even for all $p \mid n$ (i.e. n is a perfect square), then $r(p^{v_p(n)}) \geqslant 1$ for all $p \mid n$, so $r(n) \geqslant 1$ for all perfect squares. So we have proved the claim.

Now let us consider the sum

$$\sum_{n \geqslant 1} r(n) e^{-n/x}.$$

This is to be thought of as a smooth sum of $r(n)$ of length x, i.e. it behaves like $\sum_{n \leqslant x} r(n)$, but with the harmonic analytic difficulties inherent in the sharp cut-off at x suppressed. We can lower bound this sum by the claim that we have just proven. We have

$$\sum_{n \geqslant 1} r(n) e^{-n/x} \geqslant \sum_{m \geqslant 1} r(m^2) e^{-(m/x^{1/2})^2} \geqslant \sum_{m \geqslant 1} e^{-(m/x^{1/2})^2} \gg x^{1/2}, \tag{10.1}$$

by the monotone comparison theorem (Thm. 2.2) in the last step.

On the other hand, we can also evaluate this sum using the theory of Mellin transforms that we have developed. Recall that the Mellin transform of e^{-y} is by definition $\Gamma(s)$, so that by the Mellin inversion theorem we have

$$\sum_{n \geqslant 1} r(n) e^{-n/x} = \frac{1}{2\pi i} \int_{(2)} \left(\sum_{n=1}^{\infty} \frac{r(n)}{n^s} \right) \Gamma(s) x^s \, ds.$$

But since the abscissa of convergence $\sigma_r \leqslant 1$ (by Thm. 3.4), we have

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

when $\mathrm{Re}(s) = 2$. For $T > 1$, let C_T be the contour defined by: $2 - i\infty$ to $2 - iT$, $2 - iT$ to $-1/2 - iT$, $-1/2 - iT$ to $-1/2 + iT$, $-1/2 + iT$ to $2 + iT$, and $2 + iT$ to $2 + i\infty$. Then by the residue theorem and the unique meromorphic continuation of $\zeta(s)L(s, \chi)$ we have

$$\sum_{n \geqslant 1} r(n) e^{-n/x} = xL(1, \chi) + \zeta(0) L(0, \chi) + \frac{1}{2\pi i} \int_{C_T} \zeta(s) L(s, \chi) \Gamma(s) x^s \, ds.$$

I claim that the horizontal integrals

$$\pm \frac{1}{2\pi i} \int_{-1/2 \pm iT}^{2 \pm iT} \zeta(s) L(s, \chi) \Gamma(s) x^s \, ds$$

tend rapidly to 0 as $T \to \infty$. Indeed, we have from exercise 2 of sheet 8 "Stirling's approximation", which says that

$$|\Gamma(\sigma + it)| \sim \sqrt{2\pi} |t|^{\sigma - 1/2} \exp(-\frac{\pi}{2} |t|),$$

as $|t| \to \infty$. We also have by exercise 4 from sheet 6 and the functional equation that for $\sigma \in [-1/2, 2]$ and $|t| \geqslant 1$ we have

$$\zeta(\sigma + it) \ll |t|,$$

and

$$L(\sigma + it, \chi) \ll q|t|. \tag{10.2}$$

By the triangle inequality, and these estimates, the claim is proven.

Now letting $T \to \infty$ we find that

$$\sum_{n \geqslant 1} r(n) e^{-n/x} = xL(1, \chi) + \zeta(0) L(0, \chi) + \frac{1}{2\pi i} \int_{(-1/2)} \zeta(s) L(s, \chi) \Gamma(s) x^s \, ds.$$

By (10.2) and the fact that $|x^s| = x^{-1/2}$ for $\mathrm{Re}(s) = -1/2$, we have that

$$\sum_{n \geqslant 1} r(n) e^{-n/x} = xL(1, \chi) + \zeta(0) L(0, \chi) + O(qx^{-1/2}).$$

One can compute that $\zeta(0) = -1/2$ by the functional equation, and also $L(0,\chi) \geqslant 0$ by the functional equation, so

$$\sum_{n\geqslant 1} r(n)e^{-n/x} \leqslant xL(1,\chi) + O(qx^{-1/2}).$$

Then, finally, putting in (10.1), we find

$$x^{1/2} \ll xL(1,\chi) + O(qx^{-1/2}).$$

This equation holds for all $x > 1$, so we may choose x as we please. Taking $x = Cq$, with $C > 0$ sufficiently large in terms of the other implicit constants, we derive a contradiction unless $L(1,\chi) \gg q^{-1/2}$, as was to be shown. □

Theorem 10.3 (Siegel's theorem)**.** *For every $\varepsilon > 0$ there exists $C(\varepsilon) > 0$ such that*

$$L(1,\chi) \geqslant \frac{C(\varepsilon)}{q^{\varepsilon}}$$

for all χ primitive quadratic Dirichlet characters modulo $q > 1$.

Unfortunately, we do *not* know how to give a numerical value for $C(\varepsilon)$ for any value of $\varepsilon < 1/2$. Such a constant is called *ineffective*. They are the bane of our existence, and a major flaw in this theorem. Nonetheless, the (ineffective) bound on exceptional zeros given by Siegel's theorem is nearly as good as the zero free region given by Lemma 9.3.

Proof. Consider $r_2 \in \mathscr{A}$ defined by $r_2 = 1 * \chi_1 * \chi_2 * \chi_1\chi_2$, where χ_1, χ_2 are primitive quadratic Dirichlet characters modulo q_1 and q_2, respectively. Let

$$f(s) = L(s, r_2) = \zeta(s)L(s,\chi_1)L(s,\chi_2)L(s,\chi_1\chi_2).$$

It was shown in exercise 3 of sheet 6 that $r_2(n) \geqslant 0$ for all $n \geqslant 1$. Note also that r_2 is multiplicative, so $r_2(1) = 1$.

The heart of the proof of Siegel's theorem is the following Lemma.

Lemma 10.4. *For every $\varepsilon > 0$ there exists χ_1 modulo q_1 and $1-\varepsilon < \beta < 1$ such that $f(\beta) \leqslant 0$ for all χ_2 modulo q_2.*

Proof.
- Suppose there are no zeros in $[1-\varepsilon, 1)$ for any quadratic χ. Then we choose any χ_1 modulo q_1 and β satisfying $1-\varepsilon < \beta < 1$, and we have $L(\beta,\chi_1)$, $L(\beta,\chi_2)$, $L(\beta,\chi_1\chi_2)$ are all positive, while $\zeta(\beta) < 0$, hence $f(\beta) < 0$.
- Suppose there exists a χ modulo q with a real zero in $[1-\varepsilon, 1)$. Then we choose $\chi_1 = \chi$ modulo $q_1 = q$ and $1-\varepsilon < \beta < 1$ to be this real zero. Then $f(\beta) = 0$ regardless of what χ_2 is, since $L(\beta,\chi_1) = 0$.

□

Given then lemma, the rest of the proof of Siegel's theorem follows the same lines as the proof of Proposition 10.1.

Let

$$\lambda = \operatorname{res}_{s=1} f(s) = L(1,\chi_1)L(1,\chi_2)L(1,\chi_1\chi_2),$$

and χ_1, q_1 and β be as produced by Lemma 10.4. Suppose $x \geqslant 1$. By taking the first term of the series, and by Mellin inversion

$$\frac{1}{e} \leqslant \sum_{n\geqslant 1} \frac{r_2(n)}{n^{\beta}} e^{-n/x} = \frac{1}{2\pi i}\int_{(2)} f(s+\beta)\Gamma(s)x^s\, ds.$$

Using the same estimates for $\zeta, L,$ and Γ as in the proof of Proposition 10.1, we may shift the contour to the left so that we find by the residue theorem

$$e^{-1} \leqslant \lambda x^{1-\beta}\Gamma(1-\beta) + f(\beta) + \frac{1}{2\pi i}\int_{(-\beta)} f(s+\beta)\Gamma(s)x^s\,ds. \tag{10.3}$$

By exercise 4 of sheet 6, and the functional equation for Dirichlet L-functions, we have for χ a primitive character modulo q that

$$L(it,\chi) \ll ((2+|t|)q)^{1/2+\varepsilon},$$

and similarly for ζ, so that we find

$$f(it) \ll (2+|t|)^{2+\varepsilon}(q_1q_2)^{1+\varepsilon}.$$

We use this to bound the integral on the $\mathrm{Re}(s) = -\beta$ line appearing in (10.3). Using the Laurent series expansion for $\Gamma(s)$ near $s=-1$, and Stirling's approximation (exercise 2 sheet 8), we find

$$e^{-1} \leqslant \lambda x^{1-\beta}\Gamma(1-\beta) + f(\beta) + O_\varepsilon\left(\frac{(q_1q_2)^{1+\varepsilon}x^{-\beta}}{1-\beta}\right).$$

By Lemma 10.4, we have that $f(\beta) \leqslant 0$, independently of χ_2. So we have

$$e^{-1} \leqslant \lambda x^{1-\beta}\Gamma(1-\beta) + O_\varepsilon\left(\frac{(q_1q_2)^{1+\varepsilon}x^{-\beta}}{1-\beta}\right).$$

We have that $1-\beta$ is close to 0, where $\Gamma(s)$ has a pole, so $\Gamma(1-\beta) = O(\frac{1}{1-\beta})$. Meanwhile, by Proposition 10.1 we have $\lambda \gg (q_1q_2)^{-1+\varepsilon}$, so that we have

$$1 \ll \lambda\frac{x^{1-\beta}}{1-\beta} \quad \text{for any} \quad (q_1q_2)^{2+\varepsilon} \ll x. \tag{10.4}$$

We have also that

$$L(1,\chi) \ll \log q,$$

by choosing $s=1$ and $X=q$ in (4.12), and estimating the left hand side trivially (by the Monotone comparison theorem 2.5, e.g.). Thus

$$\lambda \ll L(1,\chi_2)(\log q_1)(\log q_1q_2). \tag{10.5}$$

Putting (10.5) together with (10.4) and choosing $x = (q_1q_2)^{2+\varepsilon}$, we find that

$$L(1,\chi_2) \gg \frac{1}{(\log q_1)(\log q_1q_2)}(q_1q_2)^{-(2+\varepsilon)(1-\beta)}(1-\beta).$$

We only care about the q_2 dependence here, however. That is to say, there is some function $C(q_1,\beta)$ of q_1 and β so that

$$L(1,\chi_2) \geqslant C(q_1,\beta)q_2^{-(2+\varepsilon)(1-\beta)}(\log q_2)^{-1}$$

for all $q_2 > 1$. But in the beginning of the proof, we chose q_1 and β in terms of $\varepsilon > 0$ only. So in fact this quantity $C(q_1,\beta)$ only depends on $\varepsilon > 0$. So, we have

$$L(1,\chi_2) \geqslant C(\varepsilon)q_2^{-(2+\varepsilon)(1-\beta)}(\log q_2)^{-1}.$$

Finally, we have that

$$(2+\varepsilon)(1-\beta) < 3\varepsilon,$$

and note that $\log q_2 \leqslant \varepsilon^{-1} q_2^{\varepsilon}$ for all $\varepsilon > 0$, so that we have shown

$$L(1, \chi_2) \geqslant \frac{\varepsilon C(\varepsilon)}{q_2^{4\varepsilon}},$$

which finishes the proof of Siegel's theorem. □

Finally, we relate Siegel's theorem back to exceptional zeros.

Corollary 10.5. *For any $\varepsilon > 0$ there exists $c(\varepsilon) > 0$ such that, if χ is a primitive quadratic character modulo $q > 1$, then $L(s, \chi) \neq 0$ for all*

$$s > 1 - c(\varepsilon) q^{-\varepsilon}.$$

Proof. Recall from exercise 4 of sheet 6 that

$$L'(\sigma, \chi) \ll (\log q)^2,$$

for

$$1 - (\log q)^{-1} \ll \sigma \leqslant 1. \tag{10.6}$$

By Lemma 9.3, any zero of $L(s, \chi)$ for q large enough will lie in the interval (10.6), so that we have

$$L(1, \chi) = L(1, \chi) - L(\beta, \chi) \ll (\log q)^2 (1 - \beta) \ll (\log q)^2 c(\varepsilon) q^{-\varepsilon},$$

by assumption. This contradicts Siegel's Theorem 10.3, if we replace ε by $\frac{1}{2}\varepsilon$ and take q sufficiently large. □

It must be emphasized that the implicit constants in Siegel's theorem and its corollary are ineffective. To give them a numerical value, we would either have to guarantee that there are no exceptional zeros (which we do not know how to do), or to produce the numerical values for the exceptional zeros (which presumably do not exist).

For reference, if the generalized Riemann hypothesis holds for $L(s, \chi)$, then we know that $L(1, \chi) \gg (\log\log q)^{-1}$. If we assume that exceptional zeros do not exists, i.e. that Lemma 9.3 holds with no exceptions, then we can conclude that $L(1, \chi) \gg (\log q)^{-3}$, we know that $L(1, \chi) \gg q^{-\varepsilon}$, but the implicit constant is ineffective (Siegel's theorem), and that $L(1, \chi) \gg q^{-1/2}$ with an effective constant (Proposition 10.1). In the 1980s, a major advance was made by Gross and Zagier, in which they were able to improve this to

$$L(1, \chi) \gg q^{-1/2} \log q \prod_{\substack{p \mid q \\ p \neq q}} \left(1 - \frac{2\sqrt{p}}{p+1}\right),$$

using deep results from the theory of modular forms and elliptic curves (with an effective constant). There is obviously a huge gap between what is true and what we can actually prove.

CHAPTER 11

The Prime Number Theorem in Arithmetic Progressions

We apply all of the preceding results of the course to obtain approximations to

$$\psi(X;q,a)=\sum_{n\equiv a\,(\mathrm{mod}\,q)}\Lambda(n)f(n/X),$$

for $f\in\mathscr{C}_c^\infty(\mathbb{R}_{>0})$. The version of this with the condition $n\leqslant X$ instead of the smooth weight $f(n/X)$ follows from the same machinery as exercises 1 and 2 of sheet 12. By the orthogonality relations (4.11), we have if $(a,q)=1$, and $q>1$

$$\psi(X;q,a)=\frac{1}{\varphi(q)}\sum_{\chi\,(\mathrm{mod}\,q)}\overline{\chi}(a)\psi(X,\chi), \tag{11.1}$$

where

$$\psi(X,\chi)=\sum_{n\geqslant 1}\Lambda(n)\chi(n)f(n/X). \tag{11.2}$$

The contribution of the trivial character provides the main term. We have

$$\begin{aligned}\psi(X,\chi_0)=\sum_{\substack{n\geqslant 1\\(n,q)=1}}\Lambda(n)f(n/X)&=\sum_{n\geqslant 1}\Lambda(n)f(n/X)-\sum_{\substack{n\geqslant 1\\(n,q)>1}}\Lambda(n)f(n/X)\\&=\sum_{n\geqslant 1}\Lambda(n)f(n/X)+O_f((\log X)(\log q)).\end{aligned} \tag{11.3}$$

By the prime number theorem (Corollary 8.11), we have

$$\sum_{n\geqslant 1}\Lambda(n)f(\frac{n}{X})=X\int_{\mathbb{R}}f(x)\,dx+O_f\left(X\exp(-C\sqrt{\log X})\right),$$

so that

$$\begin{aligned}\psi(X;q,a)=\frac{X}{\varphi(q)}\int_{\mathbb{R}}f(x)\,dx+\frac{1}{\varphi(q)}\sum_{\chi\neq\chi_0}\overline{\chi}(a)\psi(X,\chi)\\+O_f\left(\frac{1}{\varphi(q)}\left(X\exp(-C\sqrt{\log X})+(\log X)(\log q)\right)\right).\end{aligned} \tag{11.4}$$

So to prove the prime number theorem in arithmetic progressions, it suffices to give a bound for each of the the $\psi(X,\chi)$, $\chi\neq\chi_0$.

Proposition 11.1. *Suppose that* $q\ll\exp(\frac{1}{2}\sqrt{\log X})$.
If $\chi\neq\chi_0$ *is not an exceptional character, then there exists a constant* $c>0$ *so that we have*

$$\psi(X,\chi)\ll_f X\exp(-c\sqrt{\log X}),$$

uniformly in χ.

If $\chi \neq \chi_0$ is an exceptional character with unique real execptional zero β, then there exists a constant $c > 0$ so that we have

$$\psi(X, \chi) = -\widetilde{f}(\beta) X^{\beta} + O_f(X \exp(-c\sqrt{\log X})),$$

uniformly in χ.

Proof. As in the proof of the (smooth) prime number theorem (Corollary 8.11), we choose X sufficiently large so that $f(nX) = 0$ for all $n \geqslant 1$. Then under this assumption by the explicit formula Theorem 8.9, we have for $\chi \neq \chi_0$ that

$$\psi(X, \chi) = \frac{1}{2\pi i} \int_{(1/2)} \left(\frac{L_\infty'(s, \chi)}{L_\infty(s, \chi)} + \frac{L_\infty'(1-s, \overline{\chi})}{L_\infty(1-s, \overline{\chi})} \right) \widetilde{f}(s) X^s \, ds - \sum_{\substack{L(\rho,\chi)=0 \\ \mathrm{Re}(\rho) \in [0,1]}} \widetilde{f}(\rho) X^{\rho}. \tag{11.5}$$

The first term here is $O_f(X^{1/2})$, since $|X^s| = X^{1/2}$ for $\mathrm{Re}(s) = 1/2$.

Let us adopt the convention that a sum with a prime $\sum_{\rho}'$ over the critical zeros ρ of an L-function means that we exclude any possible exceptional zero. For these sums of non-exceptional zeros, we have by Lemma 9.3 that

$$\sum_{\substack{L(\rho,\chi)=0 \\ \mathrm{Re}(\rho) \in [0,1]}}{}' \widetilde{f}(\rho) X^{\rho} \leqslant X \sum_{\substack{L(\rho,\chi)=0 \\ \mathrm{Re}(\rho) \in [0,1]}}{}' |\widetilde{f}(\rho)| \exp\left(-\frac{c \log X}{\log(q(2+|\gamma|))}\right),$$

where we have written (as usual) $\rho = \beta + i\gamma$.

As in the proof of the prime number theorem (Corollary 8.11), we split the sum over ρ at $\log(q(2+|\gamma|)) = \sqrt{\log X}$. For the smaller ρ we have

$$\begin{aligned} \sum_{\substack{L(\rho,\chi)=0 \\ \log(q(2+|\gamma|)) \leqslant \sqrt{\log X}}}{}' |\widetilde{f}(\rho)| \exp\left(-\frac{c \log X}{\log(q(2+|\gamma|))}\right) &\leqslant \sum_{\substack{L(\rho,\chi)=0 \\ \log(q(2+|\gamma|)) \leqslant \sqrt{\log X}}}{}' |\widetilde{f}(\rho)| \exp(-c\sqrt{\log X}) \\ &\leqslant \exp(-c\sqrt{\log X}) \sum_{L(\rho,\chi)=0}{}' |\widetilde{f}(\rho)| \ll \exp(-c\sqrt{\log X}), \end{aligned}$$

by the zero counting Theorem 7.5.

For the larger ρ, since $|\widetilde{f}(\beta + i\gamma)| \ll (1+|\gamma|)^{-2}$, we have

$$\begin{aligned} \sum_{\substack{L(\rho,\chi)=0 \\ \log(q(2+|\gamma|)) > \sqrt{\log X}}}{}' |\widetilde{f}(\rho)| \exp\left(-\frac{c \log X}{\log(q(2+|\gamma|))}\right) &\ll \sum_{\substack{L(\rho,\chi)=0 \\ \log(q(2+|\gamma|)) > \sqrt{\log X}}}{}' |\widetilde{f}(\beta + i\gamma)| \\ &\ll \sum_{\substack{L(\rho,\chi)=0 \\ \log(q(2+|\gamma|)) > \sqrt{\log X}}}{}' \frac{q^{1/2} \exp(-\frac{1}{2}\sqrt{\log X})}{(1+|\gamma|)^{3/2}} \ll \exp(-\frac{1}{4}\sqrt{\log X}), \end{aligned}$$

by the zero counting Theorem 7.5 and since $q^{1/2} \ll \exp(\frac{1}{4}\sqrt{\log X})$. Putting these estimates together, and noting that $X^{1/2}$ is much smaller than $X \exp(-c\sqrt{\log X})$ for large X, we conclude the proof of the proposition. □

It follows from (11.4) and Proposition 11.1 that if there are no exceptional characters modulo q, then

$$\psi(X; q, a) = \frac{X}{\varphi(q)} \int_{\mathbb{R}} f(x) \, dx + O_f(X \exp(-c\sqrt{\log X})), \tag{11.6}$$

for some effective constant $c > 0$, assuming $q \ll \exp(\frac{1}{4}\sqrt{\log X})$. In fact, if this bound on q does *not* hold, then (11.6) holds anyway by applying trivial bounds to (11.4), so the only assumptions on q are $q > 1$, $(a,q) = 1$, $x \geqslant 2$, and that there are no exceptional characters modulo q.

If there is an exceptional character χ modulo q, with exceptional zero β, then we have by the same reasoning that for all $(a,q) = 1$ that

$$\psi(X;q,a) = \frac{X}{\varphi(q)}\int_{\mathbb{R}} f(x)\,dx - \frac{1}{\varphi(q)}\sum_{\chi \text{ exceptional}} \overline{\chi(a)}\widetilde{f}(\beta)X^{\beta} + O_f(X\exp(-c\sqrt{\log X})). \tag{11.7}$$

All of the implicit constants in (11.6) and (11.7) are effective.

Theorem 11.2 (PNT in AP, or the Siegel-Walfisz theorem)**.** *Let $c > 0$ be the (effective) constant appearing in* (11.6) *and* (11.7)*. Then for all $A \geqslant 0$, $q \ll (\log X)^A$, $(a,q) = 1$ and $X \geqslant 2$ we have*

$$\psi(X;q,a) = \frac{X}{\varphi(q)}\int_{\mathbb{R}} f(x)\,dx + O_{f,A}(X\exp(-c\sqrt{\log X})).$$

The implicit constant (which depends on A) is ineffective.

Remark 11.3. The ineffectiveness of the implicit constant comes from Siegel's theorem 10.3 and derives from the fact that we cannot rule out the existence of exceptional characters. Indeed, note that the constants in (11.6) and (11.7) are effective.

Proof. By Corollary 10.5 we have for any exceptional zero that

$$x^{\beta} = x\exp(-(1-\beta)\log x) \leqslant x\exp(-C(\varepsilon)q^{-\varepsilon}\log x).$$

Since $q \ll (\log x)^A$, we have

$$x^{\beta} \leqslant x\exp(-C(\varepsilon)(\log x)^{1-A\varepsilon}).$$

Let $\varepsilon = 1/(3A)$, say. Then

$$x^{\beta} \ll_A x\exp(-c\sqrt{\log x}),$$

where c is the constant appearing in (11.7). Thus the contribution from exceptional zeros to (11.7) is subsumed into the error term. Putting together (11.6) and (11.7), we establish the theorem. □