

Analytic Number Theory

Course Notes for UCLA Math 205A

Nickolas Andersen

November 26, 2018

Contents

1	Introduction	4
I	Primes	8
2	Equivalent forms of the Prime Number Theorem	9
2.1	Partial summation	9
2.2	The Chebyshev functions	10
2.3	Approximating the n -th prime	12
3	Approximations of $\vartheta(x)$ and Mertens' theorems	12
3.1	Stirling's formula and Euler summation	12
3.2	The Chebyshev ϑ function	14
3.3	Mertens' estimates	15
4	Basic Properties of the Riemann Zeta Function	16
4.1	The Euler product	17
4.2	The logarithmic derivative of $\zeta(s)$	18
4.3	The analytic continuation of $\zeta(s)$	19
5	The functional equation for $\zeta(s)$	20
5.1	The Poisson summation formula	20
5.2	Modularity of the theta function	21
5.3	The functional equation for $\zeta(s)$	22
6	Outline of the proof of the Prime Number Theorem	23
6.1	The Mellin transform	23
6.2	Outline of the proof of PNT	25
7	The Gamma function	26

8	Stirling's Formula	29
9	Weierstrass factorization of entire functions	32
10	A zero-free region for $\zeta(s)$	35
10.1	Nonvanishing of zeta on the 1-line	36
10.2	The infinite product for $\xi(s)$	37
10.3	The classical zero-free region	39
11	The number of nontrivial zeros below height T	40
11.1	Approximations of ζ'/ζ	41
11.2	The number $N(T)$	42
12	The Prime Number Theorem	44
12.1	The test function $\phi_x(t)$	44
12.2	Contour integration	45
12.3	A quantitative explicit formula	46
12.4	The Prime Number Theorem for $\psi(x)$	48
II	Primes in Arithmetic Progressions	50
13	Dirichlet Characters	51
13.1	The dual group	52
13.2	Orthogonality	53
14	Dirichlet L-functions and Dirichlet's Theorem	55
14.1	Basic properties of $L(s, \chi)$	55
14.2	Dirichlet's Theorem	56
15	The nonvanishing of $L(1, \chi)$	58
15.1	Real characters	59
16	The functional equation of $L(s, \chi)$	60
16.1	Primitive characters	61
16.2	Gauss sums	62
16.3	The functional equation for primitive characters	64
16.3.1	Even characters	65
16.3.2	Odd characters	66
17	Zero-free regions for $L(s, \chi)$	68
17.1	Complex characters	69
17.2	Real characters	69
17.3	The number $N(T, \chi)$	73

18 PNT in arithmetic progressions	73
18.1 The explicit formula for $L(s, \chi)$	73
18.2 The prime number theorem for arithmetic progressions	77
19 Siegel's Theorem	78
19.1 The proof of Siegel's Theorem	80
19.2 Primes in arithmetic progressions revisited	82
20 The Bombieri-Vinogradov Theorem	83
20.1 The large sieve inequality	83
20.2 Bilinear forms with Dirichlet characters	87
20.3 The proof of Bombieri-Vinogradov	89

1 Introduction

Number Theory is, broadly speaking, the study of the integers. Analytic Number Theory is the branch of Number Theory that attempts to solve problems involving the integers using techniques from real and complex analysis. Many of the problems that analytic methods are well-suited for involve the primes (and in this course, problems involving primes will be our main focus).

Theorems in analytic number theory are often about the behavior of some number theoretic quantity on average or when some parameter is very large. For example, the two main quantities we will study in this course are

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

and

$$\pi(x; a \bmod q) = \#\{p \leq x : p \text{ is prime and } p \equiv a \bmod q\}.$$

The famous Prime Number Theorem is a statement about the behavior of $\pi(x)$ as x tends to ∞ ; it does not say anything about the exact value of $\pi(x)$ for any specific x . The Prime Number Theorem states that

$$\pi(x) \sim \frac{x}{\log x} \text{ as } x \rightarrow \infty.$$

This statement requires some explanation. It is read as “ $\pi(x)$ is asymptotically equal to $x/\log x$ as x tends to infinity.” The symbol $\sim$ is defined by:

$$f(x) \sim g(x) \iff \lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} = 1.$$

Sometimes we will write $f(x) \sim g(x)$ as $x \rightarrow 0$ (or some other value) and the definition above should be changed in the obvious way. Usually, if we don’t specify what x tends to, it is assumed that $x \rightarrow \infty$.

There are stronger forms of the Prime Number Theorem that specify a bound on the error between $\pi(x)$ and $x/\log x$, but in order to state them we need to modify the main term $x/\log x$ a bit. For $x \geq 2$, define

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

Integrating by parts, it is not too difficult to show that $\text{Li}(x) \sim x/\log x$. However, $\text{Li}(x)$ is a much better approximation to $\pi(x)$, as we shall see later. In 1958 Vinogradov and Korobov proved that there is a number $c > 0$ for which

$$\pi(x) = \text{Li}(x) + O\left(x \exp\left(-c \frac{\log^{\frac{3}{5}} x}{(\log \log x)^{1/5}}\right)\right). \quad (1.1)$$

(You will likely not fully appreciate the monumental effort that this result represents until much later in the course.) Like the asymptotic $\sim$ notation, the result above requires some explanation.

We say that $f(x) = O(g(x))$ (read “ $f(x)$ is big-Oh of $g(x)$ ”) if $|f(x)| \leq C|g(x)|$ for some constant $C > 0$. So (1.1) means that there exists some constant C for which

$$|\pi(x) - \text{Li}(x)| \leq Cx \exp\left(-c \frac{\log^{\frac{3}{5}} x}{(\log \log x)^{1/5}}\right).$$

Often, determining an explicit value of the constant C is both interesting and quite difficult. Big-O has a sibling called Little-O, which is defined as follows: we say that $f(x) = o(g(x))$ if for every constant $c > 0$ we have $f(x) \leq cg(x)$ for sufficiently large x (here “sufficiently large” can depend on c). Equivalently,

$$f(x) = o(g(x)) \iff \lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} = 0.$$

For example, our first statement of the Prime Number Theorem could be stated as

$$\pi(x) = \frac{x}{\log x} (1 + o(1)).$$

Sometimes, to save a pair of parentheses, we write $f(x) \ll g(x)$ when $f(x) = O(g(x))$. This is read as “ $f(x)$ is less-than-less-than $g(x)$.” The opposite notation $f(x) \gg g(x)$ is also used to denote that $f(x) \geq Cg(x)$ for some constant $C > 0$. When $g(x) \ll f(x) \ll g(x)$, we write $f(x) \asymp g(x)$.

The function $\pi(x)$ is a step-function with infinitely many discontinuities, so how can we expect to use analytic methods to study it? The key to this is the Riemann Zeta function

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}, \quad \text{Re}(s) > 1.$$

defined for complex numbers $s = \sigma + it$ with real part $\sigma > 1$. We will prove several important properties of the zeta function in the coming weeks, but for now I’ll just list a few of them without proof. The zeta function is analytic as a function of s in the region of absolute convergence $\sigma > 1$ and it has the infinite product expansion

$$\zeta(s) = \prod_{p \text{ prime}} (1 - p^{-s})^{-1},$$

also valid for $\sigma > 1$. It follows from this that the logarithmic derivative ζ'/ζ can be written

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

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

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

Here we see the intimate connection between the primes and the Riemann zeta function; this is the main connection that we will exploit throughout this course to study the prime counting function $\pi(x)$.

To study primes in arithmetic progressions and the function $\pi(x; a \bmod q)$ we will use Dirichlet L -functions. These are closely related to the Riemann zeta function and are defined (for $\text{Re}(s) > 1$) by the Dirichlet series

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

where $\chi : \mathbb{Z} \rightarrow \mathbb{C}$ is a “Dirichlet character” which we’ll define later. For example, we can write

$$\sum_{\substack{n=1 \\ n \equiv 1(4)}}^{\infty} \frac{\Lambda(n)}{n^s} = \frac{1}{2} \sum_{n=1}^{\infty} \frac{\Lambda(n)\chi_0(n)}{n^s} + \frac{1}{2} \sum_{n=1}^{\infty} \frac{\Lambda(n)\chi_1(n)}{n^s},$$

where

$$\chi_0(n) = \begin{cases} 1 & \text{if } n \text{ is odd,} \\ 0 & \text{if } n \text{ is even,} \end{cases} \quad \chi_1(n) = \begin{cases} 1 & \text{if } n \equiv 1 \pmod{4}, \\ -1 & \text{if } n \equiv 3 \pmod{4}, \\ 0 & \text{if } n \text{ is even.} \end{cases}$$

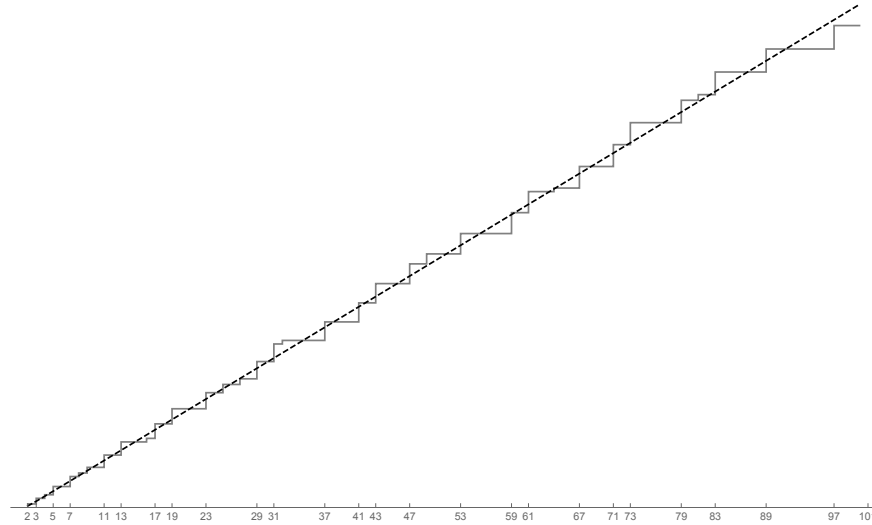


Figure 1: Plots of $\psi(x)$ (solid) and $x - \log 2\pi - \frac{1}{2} \log(1 - x^{-2})$ (dashed)

I will end this introduction with a visual representation of the connection between the zeta function and the primes. This is sometimes called the “music of the primes”. As we will show later, the Prime Number Theorem is equivalent to the statement that

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

Since $\Lambda(n)$ are the coefficients that appear in the logarithmic derivative of ζ , it is easier to use analytic tools to study $\psi(x)$ than it is to study $\pi(x)$ directly. An argument that involves the Residue Theorem from Complex Analysis leads to the “explicit formula”¹

$$\psi(x) = x - \log 2\pi - \sum_{\rho} \frac{x^{\rho}}{\rho}, \quad (1.2)$$

where the sum is over $\rho \in \mathbb{C}$ such that $\zeta(\rho) = 0$. Figure 1 shows a plot of $\psi(x)$ together with an asymptotic approximation. There are two big points to notice about the formula (1.2): 1) it is an equality, not involving any asymptotics, and 2) it involves the zeros of a function which, so far as we know right now, never vanishes. (What do I mean by that? Well, the infinite product expansion for ζ shows that $\zeta(s) \neq 0$ for $\sigma > 1$.) This formula involves the zeros of the *analytic continuation* of the zeta function to $\mathbb{C}$, which we will establish soon. There is a so-called “trivial zero” at each negative even integer: $\zeta(-2n) = 0$, but there are also infinitely many nontrivial zeros somewhere in the region $0 < \sigma < 1$ (the “critical strip”). Note that the contribution from the trivial zeros can be written as

$$\sum_{\rho \text{ trivial}} \frac{x^{\rho}}{\rho} = \frac{1}{2} \log \left(1 - \frac{1}{x^2} \right).$$

It is widely believed that all of the nontrivial zeros lie on the critical line $\sigma = \frac{1}{2}$, but this is a wide open problem known as the Riemann Hypothesis.

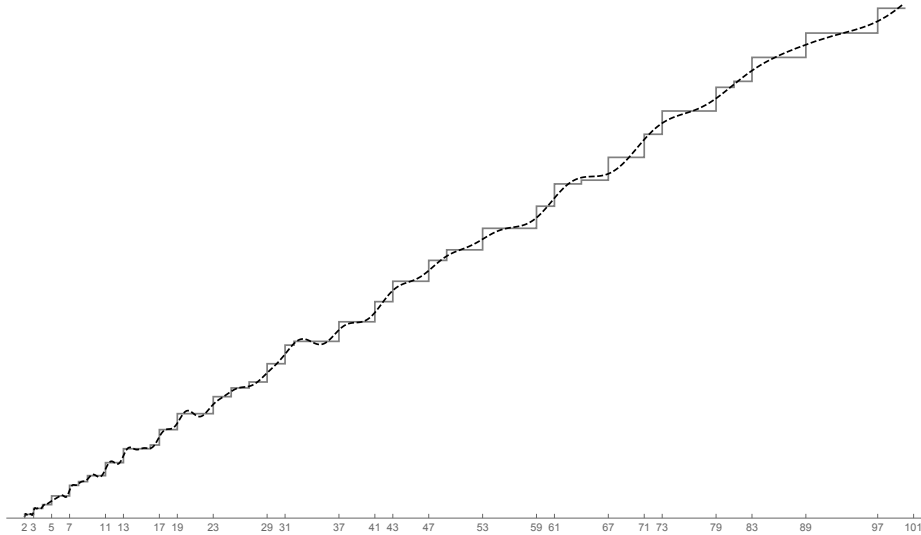


Figure 2: Approximation of $\psi(x)$ using the first 10 nontrivial zeros of $\zeta(s)$.

The explicit formula (1.2) gives a kind of “Fourier expansion” for the function $\psi(x) - x + \log 2\pi + 1/2 \log(1 - 1/x^2)$ in the following sense. It is not too difficult to show that the

¹Technically, this is only correct when x is not equal to a prime power, but we’ll ignore that for now.

zeros of ζ come in pairs ρ and $\bar{\rho}$. If we assume that each nontrivial zeta zero is of the form $\rho = \frac{1}{2} + i\gamma$, then for large γ we have

$$\frac{x^\rho}{\rho} + \frac{x^{\bar{\rho}}}{\bar{\rho}} \approx x^{\frac{1}{2}} \left(\frac{x^{i\gamma} - x^{-i\gamma}}{i\gamma} \right) = \frac{2\sqrt{x} \sin(\gamma \log x)}{\gamma}.$$

So for large x we have

$$\psi(x) \approx x - \log 2\pi - \frac{1}{2} \log \left(1 - \frac{1}{x^2} \right) - 2\sqrt{x} \sum_{\gamma > 0} \frac{\sin(\gamma \log x)}{\gamma}.$$

The sum on the right looks like a Fourier expansion, but involving very complicated frequencies (the ordinates of the zeros of zeta are conjectured to be “random” in some sense). This is theoretically very beautiful, but you can also see it happening concretely. Figures 1–3 show plots of $\psi(x)$ together with increasingly better approximations using the explicit formula.

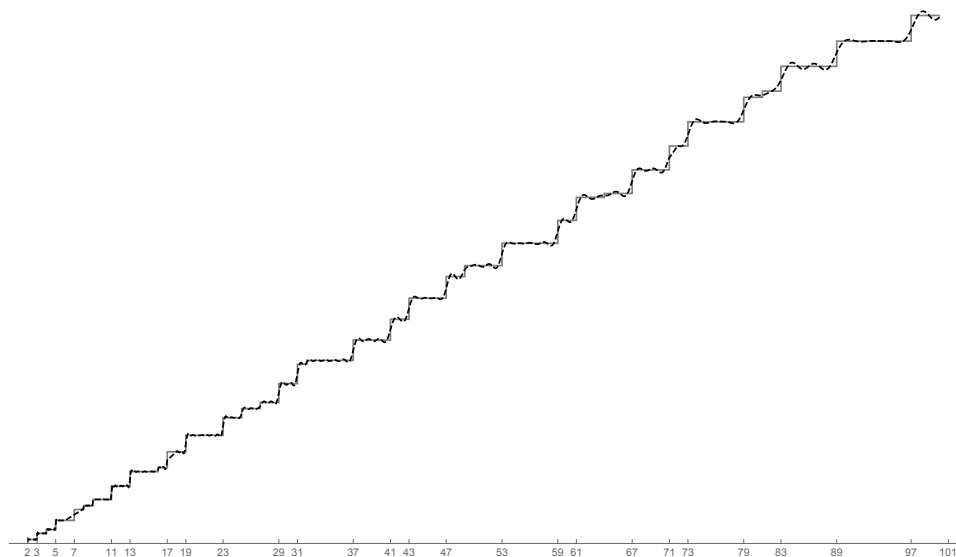


Figure 3: Approximation of $\psi(x)$ using the first 100 nontrivial zeros of $\zeta(s)$.

This course is roughly organized as follows. We’ll start by studying the primes, $\pi(x)$, and the Riemann zeta function in detail. As we encounter theorems/tools we need from Complex Analysis, we’ll prove them (instead of building up all of the tools at the beginning of the course). The second part of the course will focus on primes in arithmetic progressions, $\pi(x; a \bmod q)$, and the Dirichlet L -functions. Some of the ideas in this part will mirror those in the first part of the course, so we may skip a few proofs here and there when the ideas are straightforward. Then, time permitting, we’ll talk about other L -functions (e.g. those associated with elliptic curves and modular forms) and their roles in number theory.

Part I

Primes

2 Equivalent forms of the Prime Number Theorem

In this section we'll prove that two statements are equivalent to the prime number theorem. Here, and throughout the course, I'll use the term "equivalent" to mean that each statement follows from the other in a straightforward manner. Of course, all true statements are logically equivalent but that's not what I mean here.

2.1 Partial summation

We'll first need a tool which I've heard several people call the most useful tool in analytic number theory. It goes by a few names, but here I'll call it "partial summation" (it's also known as Abel summation). It's a discrete analogue of integration by parts.

Theorem 2.1 (Partial Summation). *Suppose that $\{a(n)\}$ is any sequence and define*

$$A(x) := \sum_{1 \leq n \leq x} a(n).$$

If $f(x)$ is continuously differentiable on $[y, x]$ with $0 < y < x$, then

$$\sum_{y < n \leq x} a(n)f(n) = A(x)f(x) - A(y)f(y) - \int_y^x A(t)f'(t) dt.$$

Proof. Let $M = \lfloor y \rfloor$ and $N = \lfloor x \rfloor$. (Here $\lfloor \cdot \rfloor$ is the "floor function" which gives the greatest integer less than or equal to the input.) The idea is that if x and y aren't integers then the sum only picks up the terms from $M + 1$ to N , so we might as well just sum over those. Note that $A(x) = A(N)$ and $A(y) = A(M)$. Writing $a(n) = A(n) - A(n - 1)$, splitting up the sum into two parts, and re-indexing the second sum, we have

$$\sum_{y < n \leq x} a(n)f(n) = \sum_{n=M+1}^N A(n)f(n) - \sum_{n=M}^{N-1} A(n)f(n+1).$$

We'll take out the $n = N$ term in the first sum and the $n = M$ term in the second sum to get that the above expression equals

$$A(N)f(N) - A(M)f(M+1) - \sum_{n=M+1}^{N-1} A(n)[f(n+1) - f(n)].$$

Now we write $f(n+1) - f(n) = \int_n^{n+1} f'(t) dt$ and note that since $A(n)$ is a step function, we have $A(n) = A(t)$ for all $t \in [n, n+1]$. Then the expression above equals

$$A(N)f(N) - A(M)f(M+1) - \sum_{n=M+1}^{N-1} \int_n^{n+1} A(t)f'(t) dt.$$

In the third term, we can write the sum of integrals as a single integral on $[M + 1, N]$. Also, a similar trick as before shows that $A(N)f(N) = A(x)f(x) - \int_N^x A(t)f'(t) dt$, and similarly $-A(M)f(M + 1) = -A(y)f(y) - \int_y^{M+1} A(t)f'(t) dt$. Thus we have

$$\sum_{y < n \leq x} a(n)f(n) = A(x)f(x) - A(y)f(y) - \left(\int_y^{M+1} + \int_{M+1}^N + \int_N^x \right) A(t)f'(t) dt.$$

This completes the proof. $\square$

2.2 The Chebyshev functions

We are now ready for two equivalent forms of the prime number theorem involving the following summatory functions. The von Mangoldt summatory function $\psi(x)$ was defined in the introduction, but I'll repeat it here:

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

The Chebyshev ϑ -function is defined as

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

(Let's agree that in these notes p always denotes a prime, so $\sum_{p \leq x}$ means that the sum is over primes $\leq x$.) Note that the definition for $\psi(x)$ can be written

$$\psi(x) = \sum_{k=1}^{\infty} \sum_{p^k \leq x} \log p = \sum_{k=1}^{\infty} \vartheta(x^{1/k}).$$

The sum is actually finite since $\vartheta(x) = 0$ if $x < 2$. So we can truncate it at $k = \lfloor \log_2(x) \rfloor$ if we desire. In a minute we'll show that $\vartheta(x) \sim x$ is equivalent to the prime number theorem. So morally we should expect that the $k = 1$ term above dominates, since $\vartheta(\sqrt{x}) \sim \sqrt{x}$, etc. Indeed, this is the case: the sums $\psi(x)$ and $\vartheta(x)$ behave quite similarly for large x . Using partial summation we can prove the following.

Proposition 2.2. *As $x \rightarrow \infty$ we have*

$$\pi(x) \sim \frac{x}{\log x} \iff \vartheta(x) \sim x \iff \psi(x) \sim x. \quad (2.1)$$

Proof. We begin by showing that $\psi(x)/x$ and $\vartheta(x)/x$ tend to the same limit (if either tends to a limit). We have

$$0 \leq \psi(x) - \vartheta(x) = \sum_{2 \leq k \leq \log_2 x} \vartheta(x^{1/k}).$$

By the definition of $\vartheta(x)$, we have the crude bound $\vartheta(x) \leq x \log x$, from which it follows that

$$0 \leq \psi(x) - \vartheta(x) \leq \sum_{2 \leq k \leq \log_2 x} x^{1/k} \log(x^{1/k}) \leq \sqrt{x} \log x \sum_{2 \leq k \leq \log_2 x} \frac{1}{k} \ll \sqrt{x} (\log x)^2.$$

We conclude that

$$\left| \frac{\psi(x)}{x} - \frac{\vartheta(x)}{x} \right| \ll \frac{(\log x)^2}{\sqrt{x}},$$

So $\psi(x)/x - \vartheta(x)/x \rightarrow 0$ as $x \rightarrow \infty$. Thus it suffices to prove the first $\iff$ in (2.1).

Let $a(n)$ denote the characteristic function of the primes:

$$a(n) = \begin{cases} 1 & \text{if } n = p \text{ is prime,} \\ 0 & \text{otherwise.} \end{cases}$$

With $f(t) = \log t$, partial summation gives

$$\vartheta(x) = \sum_{1 < n \leq x} a(n) \log n = \pi(x) \log x - \int_2^x \frac{\pi(t)}{t} dt$$

(note that $\pi(t) = 0$ if $t < 2$). Now suppose that $\pi(x) \sim x/\log x$. Then to show that $\vartheta(x) \sim x$ it is enough to prove that

$$\lim_{x \rightarrow \infty} \frac{1}{x} \int_2^x \frac{\pi(t)}{t} dt = 0.$$

By assumption we have $\pi(t) \ll t/\log t$, so

$$\frac{1}{x} \int_2^x \frac{\pi(t)}{t} dt \ll \frac{1}{x} \int_2^x \frac{dt}{\log t} = \frac{1}{x} \int_2^{\sqrt{x}} \frac{dt}{\log t} + \frac{1}{x} \int_{\sqrt{x}}^x \frac{dt}{\log t} \ll \frac{1}{\sqrt{x}} + \frac{1}{\log \sqrt{x}},$$

and the latter expression tends to zero as $x \rightarrow \infty$.

The reverse direction is quite similar. Define

$$b(n) = \begin{cases} \log n & \text{if } n = p \text{ is prime,} \\ 0 & \text{otherwise.} \end{cases}$$

Then by partial summation with $f(t) = 1/\log t$ we have

$$\pi(x) = \sum_{3/2 < n \leq x} b(n) \frac{1}{\log n} = \frac{\vartheta(x)}{\log x} + \int_2^x \frac{\vartheta(t)}{t \log^2 t} dt.$$

If $\vartheta(x) \sim x$ then it suffices to show that

$$\lim_{x \rightarrow \infty} \frac{\log x}{x} \int_2^x \frac{\vartheta(t)}{t \log^2 t} dt = 0.$$

But if $\vartheta(t) \ll t$ then we have

$$\frac{\log x}{x} \int_2^x \frac{\vartheta(t)}{t \log^2 t} dt \ll \frac{\log x}{x} \int_2^x \frac{dt}{\log^2 t}.$$

The latter integral is $\ll x/(\log \sqrt{x})^2$ by the same trick as before, so we are done. $\square$

2.3 Approximating the n -th prime

We can also get an approximation for the n -th prime number p_n if we assume the PNT.

Proposition 2.3. *Let p_n denote the n -th prime number. If $\pi(x) \sim x/\log x$ then*

$$p_n \sim n \log n.$$

Proof. We first show that $\log \pi(x) \sim \log x$. Taking logs in the prime number theorem we have

$$\lim_{x \rightarrow \infty} (\log \pi(x) + \log \log x - \log x) = 0.$$

Dividing through by $\log x$ (and using that $\log x \rightarrow \infty$) we see that

$$\lim_{x \rightarrow \infty} \left(\frac{\log \pi(x)}{\log x} + \frac{\log \log x}{\log x} - 1 \right) = 0,$$

from which it follows that

$$\lim_{x \rightarrow \infty} \frac{\log \pi(x)}{\log x} = 1.$$

Now we notice that $\pi(p_n) = n$. Thus, by the prime number theorem, as $n \rightarrow \infty$ we have

$$p_n \sim \pi(p_n) \log p_n \sim \pi(p_n) \log \pi(p_n) = n \log n,$$

as desired. □

By a similar method, the statement $p_n \sim n \log n$ implies the prime number theorem.

3 Approximations of $\vartheta(x)$ and Mertens' theorems

Using some elementary ideas we can show that $\vartheta(x) \asymp x$ (which shows that the difficulty in proving the prime number theorem is establishing the exact asymptotic, not merely the order of growth). To do this, we'll first need a rough approximation for the factorial function.

3.1 Stirling's formula and Euler summation

I'll refer to this as Stirling's formula, though later we'll prove a more precise version which I'll also call Stirling's formula.

Theorem 3.1 (Stirling's formula). *As $n \rightarrow \infty$ we have*

$$\log(n!) = n \log n - n + \frac{1}{2} \log n + O(1).$$

To prove Stirling's formula, it will be convenient to have another version of partial summation which we can use whenever we're summing a continuously differentiable function. You can prove this directly, but we'll work it out as a corollary to partial summation.

Theorem 3.2 (Euler's summation formula). *If f is continuously differentiable on $[y, x]$ with $0 < y < x$, then*

$$\sum_{y < n \leq x} f(n) = \int_y^x f(t) dt + \int_y^x (t - \lfloor t \rfloor) f'(t) dt + f(x)(\lfloor x \rfloor - x) - f(y)(\lfloor y \rfloor - y).$$

(Note that the last two terms vanish if $x, y \in \mathbb{Z}$.)

Proof. We apply partial summation with $a(n) = 1$ to get

$$\sum_{y < n \leq x} f(n) = f(x)\lfloor x \rfloor - f(y)\lfloor y \rfloor - \int_y^x \lfloor t \rfloor f'(t) dt.$$

This, together with the integration by parts formula

$$\int_y^x t f'(t) dt = x f(x) - y f(y) - \int_y^x f(t) dt$$

yields Euler's summation formula. □

Proof of Proposition 3.1. We begin by writing

$$\log(n!) = \log \prod_{k=1}^n k = \sum_{1 < k \leq n} \log k.$$

Applying Euler summation with $f(k) = \log k$ we find that

$$\sum_{1 < k \leq n} \log k = n \log n - n + 1 + \int_1^n \frac{t - \lfloor t \rfloor}{t} dt.$$

We split the integral into intervals of length one:

$$\int_1^n \frac{t - \lfloor t \rfloor}{t} dt = \sum_{\ell=1}^{n-1} \int_{\ell}^{\ell+1} \frac{t - \ell}{t} dt.$$

Evaluating the inner integral and expanding the log term as a Taylor series (around $\ell = \infty$) we find that

$$\int_{\ell}^{\ell+1} \frac{t - \ell}{t} dt = 1 - \ell \log(1 + \frac{1}{\ell}) = \frac{1}{2\ell} + O\left(\frac{1}{\ell^2}\right).$$

It follows that

$$\int_1^n \frac{t - \lfloor t \rfloor}{t} dt = \frac{1}{2} \sum_{\ell=1}^{n-1} \frac{1}{\ell} + O\left(\sum_{\ell=1}^{n-1} \frac{1}{\ell^2}\right).$$

The latter term is $O(1)$ (the full sum with $n = \infty$ is convergent, so the partial sum is bounded by a fixed constant). Applying Euler summation again with $f(\ell) = 1/\ell$ we find that

$$\sum_{\ell=1}^{n-1} \frac{1}{\ell} = \sum_{\ell=1}^n \frac{1}{\ell} - \frac{1}{n} = \log n - \frac{1}{n} + \int_1^n \frac{t - \lfloor t \rfloor}{t^2} dt = \log n + O(1).$$

Here we used that $|t - \lfloor t \rfloor| \leq 1$ and that $\int_1^{\infty} dt/t^2 < \infty$. The result follows. □

3.2 The Chebyshev ϑ function

We now prove that $\vartheta(x) \asymp x$.

Proposition 3.3. *There exist constants $a, b > 0$ such that $ax \leq \vartheta(x) \leq bx$.*

Proof. We begin with the upper bound. Let $P_k = \prod_{p \leq k} p$ so that $\vartheta(k) = \log P_k$. We prove by induction that $P_k < 4^k$. Certainly this is true for $k = 1$. To check it for P_{k+1} , we split into two cases. If $k + 1$ is even then $P_{k+1} = P_k < 4^k < 4^{k+1}$. Now suppose that $k + 1$ is odd and write $k + 1 = 2m + 1$. Then

$$P_{k+1} = \prod_{p \leq m+1} p \prod_{m+2 \leq p \leq 2m+1} p = P_{m+1} Q_m,$$

say. Note that Q_m divides $\frac{(2m+1)!}{m!(m+1)!} = \binom{2m+1}{m}$ since the primes dividing Q_m are all larger than $m + 1$. By the binomial theorem we have

$$(1+1)^{2m+1} = \sum_{\ell=0}^{2m+1} \binom{2m+1}{\ell} = 2 \sum_{\ell=0}^m \binom{2m+1}{\ell} \geq 2 \binom{2m+1}{m}.$$

It follows that $Q_m \leq \frac{1}{2} \cdot 2^{2m+1} = 4^m$. So by the inductive hypothesis we compute that $P_{k+1} = P_{m+1} Q_m < 4^{m+1} 4^m = 4^{k+1}$, as desired. Thus $\vartheta(n) \leq n \log 4$.

We turn to the lower bound. Let $v_p(n)$ denote the exponent of p in the prime factorization of n . Then for the factorial function we have

$$v_p(n!) = \left\lfloor \frac{n}{p} \right\rfloor + \left\lfloor \frac{n}{p^2} \right\rfloor + \left\lfloor \frac{n}{p^3} \right\rfloor + \dots$$

This follows by counting the number of integers $\leq n$ which are divisible by p , then those which are divisible by p^2 (note we already counted those once), then those divisible by p^3 (we already counted those twice), etc. Therefore

$$n! = \prod_{p \leq n} p^{v_p(n!)} \implies \log n! = \sum_{p \leq n} v_p(n!) \log p = \sum_{p \leq n} \left\lfloor \frac{n}{p} \right\rfloor \log p + A_n,$$

where

$$A_n = \sum_{p \leq n} \sum_{j=2}^{\infty} \left\lfloor \frac{n}{p^j} \right\rfloor \log p \leq \sum_{p \leq n} \sum_{j=2}^{\infty} \frac{n}{p^j} \log p \leq n \sum_p \frac{\log p}{p^2} \cdot \frac{1}{1 - \frac{1}{p}} \ll n,$$

So by Stirling's formula we have

$$\sum_{p \leq n} \left\lfloor \frac{n}{p} \right\rfloor \log p = n \log n + O(n).$$

Since $\vartheta(n) \ll n$ we have

$$\sum_{p \leq n} \left\lfloor \frac{n}{p} \right\rfloor \log p = n \sum_{p \leq n} \frac{\log p}{p} + O\left(\sum_{p \leq n} \log p\right) = n \sum_{p \leq n} \frac{\log p}{p} + O(n),$$

from which it follows that

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

Let $0 < a < 1$; then

$$\sum_{an \leq p \leq n} \frac{\log p}{p} = \log n - \log(an) + O(1) = \log\left(\frac{1}{a}\right) + O(1).$$

So there is a constant $c > 0$ such that

$$\sum_{an \leq p \leq n} \frac{\log p}{p} \geq \log\left(\frac{1}{a}\right) - c.$$

It follows that

$$\frac{\vartheta(n)}{n} \geq \frac{1}{n} \sum_{an \leq p \leq n} \log p \geq a \sum_{an \leq p \leq n} \frac{\log p}{p} \geq a \left(\log\left(\frac{1}{a}\right) - c \right).$$

Choosing $a = e^{-c-1}$, we conclude that $\vartheta(n) \geq e^{-c-1}n$. □

Remark 3.4. This proof illustrates one of the most-used tricks in analytic number theory: introducing an auxiliary variable (in this case a) and choosing its value at the end of the argument once you know what a good choice is.

3.3 Mertens' estimates

Note that in the course of the previous proof we showed the following. This is sometimes called Mertens' first theorem (though, he proved it with an explicit error bound).

Corollary 3.5. *As $x \rightarrow \infty$ we have*

$$\sum_{p \leq x} \frac{\log p}{p} = \log x + O(1). \tag{3.1}$$

This, together with partial summation, gives us Mertens' second theorem.

Theorem 3.6. *As $x \rightarrow \infty$ we have*

$$\sum_{p \leq x} \frac{1}{p} = \log \log x + m + O\left(\frac{1}{\log x}\right),$$

where $m \approx 0.261\,497\,212\,847\,642\,783\,755\,426\,838\,608$ is Mertens' constant.

Proof. Applying partial summation with $f(n) = 1/\log n$ and

$$a(n) = \begin{cases} \frac{\log p}{p} & \text{if } p \text{ is prime,} \\ 0 & \text{otherwise,} \end{cases}$$

we find that

$$\sum_{p \leq x} \frac{1}{p} = \sum_{n \leq x} a(n)f(n) = \frac{1}{\log x} \sum_{p \leq x} \frac{\log p}{p} + \int_2^x \left(\sum_{p \leq t} \frac{\log p}{p} \right) \frac{dt}{t \log^2 t}.$$

The first term above is handled by (3.1). For the second term, let $E(x)$ denote the error term in (3.1); then we know that $|E(x)| \leq C$ for some constant $C > 0$. It follows that the integral above equals

$$\int_2^x \frac{dt}{t \log t} + \int_2^x \frac{E(t)}{t \log^2 t} dt = \log \log x - \log \log 2 + \int_2^\infty \frac{E(t)}{t \log^2 t} dt + O\left(\int_x^\infty \frac{dt}{t \log^2 t}\right).$$

Note that the integral involving $E(t)$ is absolutely convergent, though its value would be difficult to compute exactly. Putting this all together (and evaluating the integral in the big-O), we find that

$$\sum_{p \leq x} \frac{1}{p} = \log \log x + m + O\left(\frac{1}{\log x}\right),$$

where

$$m = 1 - \log \log 2 + \int_2^\infty \frac{E(t)}{t \log^2 t} dt.$$

This is the aforementioned Mertens constant, approximated above. $\square$

Remark 3.7. Mertens' second theorem is a strengthening of Euclid's theorem that there are infinitely many primes. By contrast, there are infinitely many squares, but the sum of their reciprocals converges:

$$\sum_{n=1}^\infty \frac{1}{n^2} = \frac{\pi^2}{6}.$$

So Mertens' second theorem says that, not only are there infinitely many primes, but they appear "more frequently" than the squares (or any sequence that grows like $n^{1+\delta}$ for some fixed $\delta > 0$), whatever that means; this resonates with the statement that $p_n \sim n \log n$.

4 Basic Properties of the Riemann Zeta Function

The most important tool for studying the prime numbers is the Riemann Zeta Function

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

As is customary in analytic number theory, we'll follow Riemann's notation $s = \sigma + it$. To determine the convergence of the series defining $\zeta(s)$, let's compute

$$|n^s| = |e^{s \log n}| = |e^{\sigma \log n}| \cdot |e^{it \log n}| = n^\sigma.$$

We'll use the following standard theorems from complex analysis to determine the analyticity of the zeta function.

Theorem 4.1. *Let f_n be a sequence of functions analytic in a domain D and converging uniformly to f on all compact subsets of D . Then f is analytic in D , and $f' = \lim f'_n$.*

Proof sketch. By Morera's theorem, it is enough to show that $\int_{\Gamma} f = 0$ for every loop Γ in D . But since $f_n \rightarrow f$ uniformly on Γ (a compact subset of D) we have $0 = \int_{\Gamma} f_n \rightarrow \int_{\Gamma} f$. $\square$

Theorem 4.2 (The Weierstrass M -test). *Let $T \subset \mathbb{C}$ and let f_j be a sequence of complex-valued functions on T . Suppose that for each j we have $|f_j(z)| \leq M_j$ for all $z \in T$ and suppose that $\sum_j M_j$ converges. Then the series $\sum_j f_j(z)$ converges uniformly on T .*

Proof sketch. Fix $\epsilon > 0$. By the Cauchy criterion, there is a number $N \geq 1$ such that $\sum_{j=m}^n M_j < \epsilon$ whenever $n \geq m \geq N$. It follows that $\sum f_j(z)$ also satisfies the Cauchy criterion, and therefore converges to some function $F(z)$. The convergence is uniform because

$$\left| F(z) - \sum_{j=0}^m f_j(z) \right| \leq \sum_{j=m+1}^{\infty} M_j \leq \epsilon$$

whenever $m \geq N$, for all $z \in T$ (let $n \rightarrow \infty$ in the Cauchy criterion). $\square$

Applying these theorems to the Riemann zeta function, we see that $\zeta(s)$ is analytic in the domain $D = \{\sigma > 1\}$. Indeed, let $K \subseteq D$ be compact. Then $\sigma \geq 1 + \delta$ for some fixed $\delta > 0$ for all $s \in K$. It follows that $|n^{-s}| = n^{-\sigma} \leq n^{-1-\delta}$ for all $s \in K$. So by Theorem 4.2, $\zeta(s)$ converges uniformly on K ; thus, by Theorem 4.1, $\zeta(s)$ is analytic in D .

4.1 The Euler product

The connection between $\zeta(s)$ and primes comes from the Euler product (for $\sigma > 1$)

$$\zeta(s) = \prod_p (1 - p^{-s})^{-1}. \quad (4.1)$$

To show that the infinite product converges, we apply the following standard theorem from complex analysis.

Theorem 4.3. *If the series $\sum a_n$ converges absolutely, then the infinite product $\prod(1 + a_n)$ converges absolutely, i.e. the product $\prod(1 + |a_n|)$ converges.*

To apply this to $\zeta(s)$, we first expand the factor $(1 - p^{-s})^{-1}$ as a geometric series

$$\frac{1}{1 - p^{-s}} = 1 + \sum_{m=1}^{\infty} p^{-ms}.$$

Thus, to show that the infinite product on the right-hand side of (4.1) converges absolutely we verify the convergence of the series

$$\sum_p \sum_{m=1}^{\infty} |p^{-ms}| = \sum_p \sum_{m=1}^{\infty} p^{-m\sigma} = \sum_p p^{-\sigma} \cdot \frac{1}{1 - p^{-\sigma}} \leq \frac{1}{1 - 2^{-\sigma}} \sum_p \frac{1}{p^{\sigma}} < \frac{1}{1 - 2^{-\sigma}} \sum_n \frac{1}{n^{\sigma}}.$$

Since $\sigma > 1$, this series converges, so the infinite product converges absolutely.

It remains to show that this convergent infinite product equals $\zeta(s)$. Consider the finite product

$$\prod_{p \leq x} \frac{1}{1 - p^{-s}} = \prod_{p \leq x} \left(1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \dots \right) = \sum_{n \in A} \frac{1}{n^s},$$

where A is the set of positive integers all of whose prime factors are $\leq x$ (note that we are using the fundamental theorem of arithmetic here, together with the fact that we can rearrange a finite product of absolutely convergent infinite series). It follows that

$$\left| \sum_{n=1}^{\infty} \frac{1}{n^s} - \prod_{p \leq x} \frac{1}{1 - p^{-s}} \right| \leq \sum_{n > x} \frac{1}{n^\sigma},$$

since all of the leftover terms have at least one prime factor larger than x . As $x \rightarrow \infty$, the sum on the right-hand side goes to zero because $\sum n^{-\sigma}$ is convergent. It follows that

$$\lim_{x \rightarrow \infty} \prod_{p \leq x} \frac{1}{1 - p^{-s}} = \zeta(s).$$

Our first observation from the Euler product is the nonvanishing of $\zeta(s)$ in $\sigma > 1$.

Theorem 4.4. *If $\sigma > 1$ then $\zeta(s) \neq 0$.*

Proof. Let P be a large prime and consider the product

$$(1 - 2^{-s})(1 - 3^{-s}) \cdots (1 - P^{-s})\zeta(s) = 1 + \sum_{n \in B} \frac{1}{n^s},$$

where B is the set of integers whose smallest prime factor is larger than P . By the reverse triangle inequality, we have

$$|(1 - 2^{-s})(1 - 3^{-s}) \cdots (1 - P^{-s})\zeta(s)| \geq 1 - \sum_{n \in B} \frac{1}{n^\sigma} > 1 - \sum_{n > P} \frac{1}{n^\sigma}.$$

If P is large enough, then the latter sum is less than 1, so we have

$$|(1 - 2^{-s})(1 - 3^{-s}) \cdots (1 - P^{-s})\zeta(s)| > 0.$$

Since none of the factors $(1 - p^{-s})$ is zero when $\sigma > 1$, it follows that $|\zeta(s)| > 0$. $\square$

4.2 The logarithmic derivative of $\zeta(s)$

We can now connect $\zeta(s)$ to the von Mangoldt function $\Lambda(n)$.

Theorem 4.5. *For $\sigma > 1$ we have*

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

Note that sometimes we will write $\frac{\zeta'}{\zeta}(s)$ for $\frac{\zeta'(s)}{\zeta(s)}$.

Proof. Taking the logarithm² of the Euler product for $\zeta(s)$, we find that

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

Differentiating, and expanding the geometric series, we find that

$$\frac{\zeta'(s)}{\zeta(s)} = - \sum_p \frac{p^{-s} \log p}{1 - p^{-s}} = - \sum_p \log p \sum_{m=1}^{\infty} \frac{1}{p^{ms}} = - \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s}.$$

The last equality follows from rearranging the series and remembering that $\Lambda(n) = \log p$ if $n = p^m$ and otherwise $\Lambda(n) = 0$. $\square$

4.3 The analytic continuation of $\zeta(s)$

The last property we'll prove in this section is the analytic continuation of $\zeta(s)$ to $\sigma > 0$. Applying Euler summation, we find that for $\sigma > 1$ we have

$$\sum_{2 \leq n \leq x} \frac{1}{n^s} = \int_1^x \frac{dt}{t^s} - s \int_1^x \frac{t - \lfloor t \rfloor}{t^{s+1}} dt = \frac{x^{1-s}}{1-s} - \frac{1}{1-s} - s \int_1^x \frac{t - \lfloor t \rfloor}{t^{s+1}} dt.$$

Since $\sigma > 1$, everything on the right-hand side converges as $x \rightarrow \infty$. Adding in the $n = 1$ term, we find that

$$\zeta(s) = \frac{1}{s-1} + 1 - s \int_1^{\infty} \frac{t - \lfloor t \rfloor}{t^{s+1}} dt.$$

I claim that the integral defines an analytic function in the region $\sigma > 0$. Indeed, we can use Theorem 4.1 again: define

$$f_m(z) = \int_1^m \frac{t - \lfloor t \rfloor}{t^{s+1}} dt,$$

and let $f(z) = \lim_m f_m(z)$. If K is a compact subset of $\{\sigma > 0\}$ then there is some $\delta > 0$ such that $\sigma \geq \delta$ for all $s \in K$. It follows that

$$|f(z) - f_m(z)| \leq \int_m^{\infty} \frac{|t - \lfloor t \rfloor|}{t^{\sigma+1}} dt \leq \int_m^{\infty} \frac{dt}{t^{1+\delta}} = \frac{1}{\delta m^{\delta}},$$

so the sequence f_m converges uniformly to f on K . Therefore the function

$$\frac{1}{s-1} + 1 - s \int_1^{\infty} \frac{t - \lfloor t \rfloor}{t^{s+1}} dt \tag{4.2}$$

²We should be careful about branch cuts, because we're going to take a derivative. However, I'm going to ignore these kinds of issues here; this can be done very carefully, but it's not very enlightening. Alternatively, one could restrict to $t = 0$, i.e. real values of s , so that $\zeta(s)$ is positive, and then there are no branch cut issues. After doing the rest of the computation, the equality for general s follows by analytic continuation.

is meromorphic in $\sigma > 0$ with a simple pole at $s = 1$ with residue 1. We take this to be the *definition* of $\zeta(s)$ in the region $\sigma > 0$; this is reasonable because it agrees with our original definition on the subset $\sigma > 1$ and it is analytic except at $s = 1$. Actually, by the theory of analytic continuation this is the *only* definition of $\zeta(s)$ in the larger region which preserves analyticity, so we can say it is *the* analytic continuation (really, it should be called the meromorphic continuation, but nobody says that).

We collect the results above in a theorem.

Theorem 4.6. *The Riemann zeta function $\zeta(s)$ defined by (4.2) is meromorphic in the region $\sigma > 1$ with a simple pole at $s = 1$ with residue 1 and no other singularities.*

5 The functional equation for $\zeta(s)$

Among the most important properties of the Riemann zeta function is the functional equation, which relates the value of $\zeta(s)$ to the value of $\zeta(1 - s)$. Note that since we have analytically continued $\zeta(s)$ to the region $\operatorname{Re}(s) > 0$, the functional equation provides the glue that extends the definition of $\zeta(s)$ to the whole complex plane.

The functional equation for $\zeta(s)$ follows from the “modularity of a theta function of weight $1/2$.” For me to fully explain that phrase, we would need to take a big detour into the theory of modular forms; we won’t do that here, but you should be aware that something bigger is happening in the background.

5.1 The Poisson summation formula

We require a result from classical Fourier analysis called the Poisson summation formula. While we could state it in more generality, we will assume that we are working with suitably nice functions in order to streamline the exposition; what follows is certainly enough for our purposes.

Recall that if $\int_{\mathbb{R}} |f| < \infty$ then the Fourier transform of f is defined as

$$\hat{f}(y) = \int_{\mathbb{R}} f(x) e(-xy) dx,$$

where we have used the standard shorthand notation

$$e(x) := e^{2\pi i x}.$$

Theorem 5.1. *Suppose that both $f, \hat{f}$ are differentiable on $\mathbb{R}$. If both f and $\hat{f}$ are in $L^1(\mathbb{R})$ and have bounded variation³ then*

$$\sum_{m \in \mathbb{Z}} f(m) = \sum_{n \in \mathbb{Z}} \hat{f}(n) \tag{5.1}$$

and both series converge absolutely.

³The first condition means that $\int_{\mathbb{R}} |g| < \infty$ and the second means that $\int_{\mathbb{R}} |g'| < \infty$, for $g = f, \hat{f}$.

Proof. First note that the absolute convergence of the series in (5.1) follows from an application of Euler summation. Now consider the function

$$F(x) = \sum_{m \in \mathbb{Z}} f(x + m).$$

Since F is periodic of period one, it has a Fourier series expansion

$$F(x) = \sum_{n \in \mathbb{Z}} c(n) e(nx),$$

where the Fourier coefficients are given by

$$c(n) = \int_0^1 F(t) e(-nt) dt = \int_0^1 \sum_{m \in \mathbb{Z}} f(t + m) e(-nt) dt = \int_{\mathbb{R}} f(t) e(-nt) dt = \hat{f}(n).$$

Thus we have

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

and taking $x = 0$ we obtain (5.1). □

5.2 Modularity of the theta function

Here we establish a functional equation for the theta function

$$\theta(z) := \sum_{n \in \mathbb{Z}} \exp(\pi i n^2 z),$$

where z is in the upper half-plane $\mathcal{H} = \{\operatorname{Im}(z) > 0\}$. If we write $z = x + iy$, then $|\exp(\pi i n^2 z)| = \exp(-\pi n^2 y)$. So if $K \subset \mathcal{H}$ is compact, then $y \geq y_0$ for some fixed $y_0 > 0$, and thus the terms of the series rapidly decay as $|n| \rightarrow \infty$; it follows that $\theta(z)$ defines an analytic function in $\mathcal{H}$. The following lemma shows that $\theta(z)$ also satisfies a functional equation relating $\theta(z)$ and $\theta(-1/z)$.

Lemma 5.2. *For all $z \in \mathcal{H}$ we have*

$$\theta(-1/z) = (-iz)^{\frac{1}{2}} \theta(z), \tag{5.2}$$

where the branch of $z^{1/2}$ is determined by $1^{1/2} = 1$.

Proof. Since both sides of (5.2) are analytic on $\mathcal{H}$, it suffices to prove the equality for $z = iy$ with $y > 0$; the result then follows by analytic continuation. To apply Poisson summation, we need to evaluate the integral

$$\int_{-\infty}^{\infty} e^{-\pi x^2 y - 2\pi i n x} dx.$$

Writing

$$-\pi x^2 y - 2\pi i n x = -\pi(x + in/y)^2 y - \pi n^2/y,$$

the integral equals

$$e^{-\pi n^2/y} \int_{-\infty}^{\infty} e^{-\pi(x+in/y)^2 y} dx.$$

Let's think about this integral as a contour integral in the complex plane, and consider x as a complex variable. Note that the integrand decays rapidly as $|\operatorname{Re}(x)| \rightarrow \infty$ as long as $\operatorname{Im} x$ remains constant. So by Cauchy's theorem we can shift the line of integration to the line $x - in/y$ with $x \in \mathbb{R}$. So we find that

$$\int_{-\infty}^{\infty} e^{-\pi x^2 y - 2\pi i n x} dx = e^{-\pi n^2/y} \int_{-\infty}^{\infty} e^{-\pi x^2 y} dx = y^{-1/2} e^{-\pi n^2/y} \int_{-\infty}^{\infty} e^{-\pi x^2} dx,$$

where the last equality comes from a change of variables. The latter integral evaluates to 1, which finishes the proof of the lemma. $\square$

5.3 The functional equation for $\zeta(s)$

We begin with the Gamma function

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

which is analytic in the region $\sigma > 0$. Making the change of variables $t = \pi n^2 x$ we find that

$$\Gamma(s) = \pi^s n^{2s} \int_0^{\infty} e^{-\pi n^2 x} x^s \frac{dx}{x},$$

from which it follows that

$$\Gamma(s/2) \pi^{-s/2} n^{-s} = \int_0^{\infty} e^{-\pi n^2 x} x^{s/2} \frac{dx}{x}.$$

If $\sigma > 1$ then we can sum both sides over positive integers n to obtain

$$\Gamma(s/2) \pi^{-s/2} \zeta(s) = \int_0^{\infty} \theta_1(x) x^{s/2-1} dx,$$

where $\theta_1(x) = \frac{1}{2}(\theta(ix) - 1)$. The exchange of summation and integration is justified by absolute convergence. We will use the transformation property of $\theta(z)$ to manipulate the integral on the right-hand side to obtain an expression that is absolutely convergent for all $s \in \mathbb{C}$ except at $s = 0, 1$. This way we'll be able to analytically extend the definition of $\zeta(s)$ to the entire complex plane.

Let

$$I_1 = \int_0^1 \theta_1(x) x^{s/2-1} dx, \quad I_2 = \int_1^{\infty} \theta_1(x) x^{s/2-1} dx.$$

Making the change of variables $x = 1/u$ in I_1 , we find that

$$I_1 = \int_1^{\infty} \theta_1(1/u) u^{-s/2-1} du.$$

Then the functional equation for $\theta(z)$ gives

$$\theta_1(1/u) = \frac{\theta(i/u) - 1}{2} = \frac{u^{1/2}\theta(iu) - 1}{2} = \frac{2u^{1/2}\theta_1(u) + u^{1/2} - 1}{2}.$$

It follows that

$$I_1 = -\frac{1}{s} + \frac{1}{s-1} + \int_1^\infty \theta_1(u)u^{-s/2-1/2} du.$$

We can estimate $\theta_1(x)$ for large x by using that $n^2 \geq n$, so

$$\sum_{n=1}^\infty e^{-\pi n^2 y} \leq \sum_{n=1}^\infty e^{-\pi n x} = \frac{1}{e^{\pi x} - 1} \ll e^{-\pi x}.$$

Therefore the integral above (also the integral I_2) is absolutely convergent for any $s \in \mathbb{C}$ and defines an entire function of s . Putting this together with the work above, we find that

$$\pi^{-s/2}\Gamma(s/2)\zeta(s) = \frac{1}{s(s-1)} + \int_1^\infty \theta_1(x)(x^{-s/2-1/2} + x^{s/2-1}) dx. \quad (5.3)$$

The right-hand side is meromorphic for all $s \in \mathbb{C}$, with poles only at $s = 0, 1$. This provides the meromorphic continuation of the left-hand side to all of $\mathbb{C}$. Furthermore, the right-hand side is invariant under the transformation $s \leftrightarrow 1-s$. We collect these facts in a theorem.

Theorem 5.3. *Define*

$$\xi(s) := s(s-1)\pi^{-s/2}\Gamma(s/2)\zeta(s).$$

Then $\xi(s)$ is an entire function that satisfies the functional equation

$$\xi(s) = \xi(1-s). \quad (5.4)$$

6 Outline of the proof of the Prime Number Theorem

We will prove a few versions of the prime number theorem in this course, with various error estimates. But each time the general idea will be the same. Since this can get kind of technical, it's good to have a big picture view of what's going on before we zoom in on the details. To start, we'll need to know about Mellin transforms.

6.1 The Mellin transform

Let $\phi(t)$ be a continuous⁴ function on the nonnegative real line which decays rapidly at ∞ . For concreteness, let's say that $\phi(t) \ll t^{-A}$ for any $A > 0$. In practice, we'll usually pick something with compact support (meaning that $\phi(t) = 0$ for $t \geq M$) so the rapid decay is immediate. The Mellin transform of ϕ is

$$\tilde{\phi}(s) := \int_0^\infty \phi(t)t^s \frac{dt}{t},$$

⁴You don't really need continuity here (and the growth conditions can be relaxed), but for our purposes this is enough.

where s is a complex parameter. By the rapid decay assumption, this integral converges absolutely for any $s \in \mathbb{C}$ with $\operatorname{Re}(s) > 0$ and defines an analytic function of s in that region. The real power of the Mellin transform comes from the Mellin inversion formula.

Theorem 6.1 (Mellin inversion). *Suppose that ϕ satisfies the conditions above. Then*

$$\phi(x) = \frac{1}{2\pi i} \int_{(c)} \tilde{\phi}(s) x^{-s} ds. \quad (6.1)$$

The notation $\int_{(c)}$ is shorthand for $\int_{c-i\infty}^{c+i\infty}$.

Proof. If you are familiar with Fourier inversion, this follows quickly from that. The Fourier inversion formula states that

$$f(y) = \int_{\mathbb{R}} \hat{f}(u) e^{2\pi i y u} du, \quad (6.2)$$

where $\hat{f}$ is the Fourier transform

$$\hat{f}(u) = \int_{\mathbb{R}} f(v) e^{-2\pi i u v} dv.$$

Applying this to $f(y) = \phi(e^y)$ we find that

$$\hat{f}(u) = \int_{\mathbb{R}} \phi(e^v) e^{-2\pi i u v} dv = \int_0^\infty \phi(t) t^s \frac{dt}{t} = \tilde{\phi}(s),$$

where $s = -2\pi i u$. Then a simple change of variables in (6.2) yields (6.1).

If you are not familiar with Fourier inversion, I'll sketch a direct proof of (6.1) here (with some extra assumptions that streamline the proof). Let R be a large parameter and consider the integral

$$\frac{1}{2\pi i} \int_{c-iR}^{c+iR} \tilde{\phi}(s) x^{-s} ds = \int_0^\infty \frac{\phi(t)}{t} \left[\frac{1}{2\pi i} \int_{c-iR}^{c+iR} \left(\frac{t}{x} \right)^s ds \right] dt.$$

The decay conditions on ϕ justify switching the order of integration. For $t \neq x$, the integral in brackets evaluates to

$$\frac{1}{2\pi i} \frac{(t/x)^{c+iR} - (t/x)^{c-iR}}{\log(t/x)} = \frac{1}{\pi} \left(\frac{t}{x} \right)^c \frac{\sin(R \log(t/x))}{\log(t/x)}.$$

Changing variables to $t = x e^u$, we find that

$$\frac{1}{2\pi i} \int_{c-iR}^{c+iR} \tilde{\phi}(s) x^{-s} ds = \frac{1}{\pi} \int_{-\infty}^\infty \phi(x e^u) e^{cu} \sin(Ru) \frac{du}{u}.$$

Now suppose that $f : \mathbb{R} \rightarrow \mathbb{R}$ is smooth and that $\int_{\mathbb{R}} |f| < \infty$ and $\int_{\mathbb{R}} |f'| < \infty$. I claim that

$$\lim_{R \rightarrow \infty} \int_{\mathbb{R}} f(y) \sin(Ry) \frac{dy}{y} = \pi f(0).$$

Write

$$\int_{\mathbb{R}} f(y) \sin(Ry) \frac{dy}{y} = f(0) \int_{\mathbb{R}} \sin(Ry) \frac{dy}{y} + \int_{\mathbb{R}} g(y) \sin(Ry) dy,$$

where $g(y) = (f(y) - f(0))/y$ for $y \neq 0$, $g(0) = f'(0)$. Note that g is continuous at $y = 0$; we also have $\int_{\mathbb{R}} |g'| < \infty$. The first integral evaluates to π for any R (this is a famous integral – look up “sinc” function – and can be computed using a clever contour integration). For the second integral, we integrate by parts to see that

$$\int_{\mathbb{R}} g(y) \sin(Ry) dy = \frac{1}{R} \int_{\mathbb{R}} \cos(Ry) g'(y) dy \ll \frac{1}{R} \int_{\mathbb{R}} |g'(y)| dy \ll \frac{1}{R}.$$

So the second integral goes to zero as $R \rightarrow \infty$. This completes the proof. $\square$

6.2 Outline of the proof of PNT

Now suppose that we choose $\phi_x(t)$ to smoothly approximate the characteristic function for the interval $[0, x]$ (i.e. $\phi_x(t) \approx 1$ when $0 \leq t \leq x$ and $\phi_x(t) \approx 0$ otherwise). Then we can approximate $\psi(x)$ by the sum

$$\sum_{n=1}^{\infty} \Lambda(n) \phi_x(n) = \sum_{n=1}^{\infty} \Lambda(n) \cdot \frac{1}{2\pi i} \int_{(c)} \tilde{\phi}_x(s) n^{-s} ds = \frac{1}{2\pi i} \int_{(c)} \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s} \tilde{\phi}_x(s) ds,$$

assuming everything converges. Now we have some choice of where we perform the integral (what value of c to choose). If we can show that ζ'/ζ is somewhat well-behaved for $\sigma \in (0, 1)$ then we can use the residue theorem to move the contour of integration to $c = \delta$ slightly larger than 0. Along the way we pick up terms corresponding to the poles of the integrand

$$-\frac{\zeta'}{\zeta}(s) \tilde{\phi}_x(s).$$

Since ϕ_x is analytic in $\text{Re}(s) > 0$ the poles come from the poles of ζ'/ζ . You may recall the following basic fact from complex analysis.

Lemma 6.2. *If f is meromorphic at s_0 then*

$$\text{Res}(f; s_0) = \begin{cases} m & \text{if } f \text{ has a zero of order } m \text{ at } s_0, \\ -m & \text{if } f \text{ has a pole of order } m \text{ at } s_0, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. Write $f(s) = (s - s_0)^k g(s)$ where g is analytic and nonvanishing at s_0 . Then

$$\frac{f'(s)}{f(s)} = \frac{k(s - s_0)^{k-1} g(s) + (s - s_0)^k g'(s)}{(s - s_0)^k g(s)} = \frac{k}{s - s_0} + \frac{g'(s)}{g(s)}.$$

Since g'/g is analytic at s_0 , the residue of f at s_0 equals k . $\square$

It follows that when we move the contour of integration to $c = \delta$ close to zero, we pick up the following terms:

$$\tilde{\phi}_x(1) - \sum_{\rho} \tilde{\phi}_x(\rho),$$

where the sum is over all ρ such that $\zeta(\rho) = 0$ and $\text{Re}(\rho) > \delta$, counted with multiplicity. The first term comes from the pole of ζ at $s = 1$ (the residue there is -1) and the sum comes from the zeros of ζ . It then remains to estimate the contour integral with the contour (δ) and compute the Mellin transforms (once we've made a choice of ϕ_x). It will also be helpful to approximate how many terms there are in the sum, which we will do.

For the ϕ_x we will eventually choose, it turns out that

$$\tilde{\phi}_x(s) \approx \frac{x^s}{s}.$$

So you can immediately see the main term $\tilde{\phi}_x(s) \approx x$ which will give us $\psi(x) \sim x$. However, this also shows that if there is a zero of $\zeta(s)$ on the line $\sigma = 1$ then the sum over zeros will have a term of size $\gg x$. It turns out that the asymptotic formula $\psi(x) \sim x$ is actually *equivalent* to the statement that $\zeta(1 + it) \neq 0$.

These are the main ideas that will lead us to the “explicit formula” and to the prime number theorem with a quantitative error term.

7 The Gamma function

Here we will prove several important properties of the gamma function. We begin by obtaining the meromorphic continuation of $\Gamma(s)$ to the entire complex plane. Recall that we defined $\Gamma(s)$ by the integral

$$\Gamma(s) = \int_0^\infty e^{-t} t^s \frac{dt}{t}, \quad \text{Re}(s) > 0.$$

Integrating by parts, we find that

$$\Gamma(s + 1) = s\Gamma(s) \tag{7.1}$$

and, more generally, if n is a positive integer then

$$\Gamma(s + n) = (s + n)(s + n - 1) \cdots (s + 1)s\Gamma(s).$$

Since $\Gamma(1) = 1$, this shows that $\Gamma(n + 1) = n!$, so the Gamma function provides an analytic interpolation of the factorial function.

We will use the relation (7.1) to extend the definition of $\Gamma(s)$ to the left of $\text{Re}(s) = 0$. This is done inductively; we'll do the first step, and it should be clear from there how to proceed. Suppose that $\text{Re}(s) > -1$ and that $s \neq 0$. We *define*

$$\Gamma(s) = \frac{\Gamma(s + 1)}{s}$$

in that region. Since $\text{Re}(s + 1) > 0$ this only uses values of $\Gamma(s)$ for which the integral representation is valid. It follows that $\Gamma(s)$ is meromorphic in $\text{Re}(s) > -1$, with a pole only at $s = 0$. The residue there equals $\Gamma(1) = 1$.

We repeat this process now for $\operatorname{Re}(s) > -2$, etc. In this way, we extend the definition of $\Gamma(s)$ to the entire complex plane except for the points $s = 0, -1, -2, \dots$. The Gamma function has a simple pole at each of the nonpositive integers and

$$\operatorname{Res}(\Gamma; -n) = \frac{(-1)^n}{n!}.$$

What does this tell us about the Riemann zeta function? Recall that the function

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

is entire. The factor $(s-1)$ cancels out the pole of $\zeta(s)$ and the factor s cancels out the pole of $\Gamma(s/2)$ at $s = 0$. But all of the other poles of $\Gamma(s/2)$ at the negative even integers need to be cancelled as well; it follows that $\zeta(s)$ has a simple zero at each even negative integer. These zeros are called the “trivial zeros.”

Consider the function $\Gamma(s)\Gamma(1-s)$; it is meromorphic on $\mathbb{C}$ with a simple pole at every integer. Another function that satisfies that property is $1/\sin \pi z$. It is therefore natural to guess that the two functions are related.

Proposition 7.1. *For $s \in \mathbb{C} \setminus \mathbb{Z}$ we have*

$$\Gamma(s)\Gamma(1-s) = \frac{\pi}{\sin \pi s}. \quad (7.2)$$

Proof. Consider the function $f(s) = \Gamma(s)\Gamma(1-s)\sin \pi s$. The zeros of $\sin \pi s$ cancel out the poles of the Gamma factors, so $f(s)$ is an entire function. By the functional relation (7.1) we have $f(s+1) = f(s)$. Writing $s = x + iy$ we have, in the region $0 < x < 1$, the inequality

$$|\Gamma(x + iy)| \leq \int_0^\infty e^{-t} t^x \frac{dt}{t} = \Gamma(x).$$

It follows that, for $0 < x < 1$, $f(s) \ll |\sin(\pi iy)| \ll e^{\pi y}$. Now, the periodicity of f shows that the function $g(w) = f(\frac{\log w}{2\pi i})$ is well-defined (here, we are using the multi-valued logarithm) and analytic on $\mathbb{C} \setminus \{0\}$. The bound for f translates to $g(w) \ll |w|^{1/2}$ as $|w| \rightarrow \infty$ and $g(w) \ll |w|^{-1/2}$ as $w \rightarrow 0$. But this shows that the singularities of g at both 0 and ∞ are removable, thus g is a constant (by Liouville’s theorem).

To compute the constant:

$$f(1/2) = \Gamma(1/2)^2 = \left(\int_0^\infty e^{-t} t^{-1/2} dt \right)^2 = \left(2 \int_0^\infty e^{-u^2} du \right)^2 = \pi.$$

You can also compute the residue of both sides of (5.3) at $s = 1$, but that also implicitly involves the Gaussian integral. $\square$

Corollary 7.2. $\Gamma(s)$ is nonvanishing on $\mathbb{C}$.

The corollary shows that $1/\Gamma(s)$ is an entire function with simple zeros at the nonpositive integers. We can construct another such function using an infinite product:

$$g(s) = s \prod_{n=1}^{\infty} \left(1 + \frac{s}{n} \right) e^{-s/n}.$$

The product converges absolutely because $(1+z)e^{-z} = 1 + O(|z|^2)$ for z close to zero. To construct a function $G(s)$ which satisfies the recurrence relation $sG(s+1) = G(s)$ (the same relation that $1/\Gamma(s)$ satisfies) we compute

$$\begin{aligned} \frac{g(s+1)}{g(s)} &= \frac{s+1}{s} \lim_{N \rightarrow \infty} \prod_{n=1}^N \frac{1+s/n+1/n}{1+s/n} e^{-1/n} \\ &= \frac{s+1}{s} \lim_{N \rightarrow \infty} \exp\left(-\sum_{n=1}^N \frac{1}{n}\right) \prod_{n=1}^N \frac{s+n+1}{s+n} \\ &= \frac{s+1}{s} \lim_{N \rightarrow \infty} \exp\left(-\sum_{n=1}^N \frac{1}{n}\right) \frac{s+N+1}{s+1} \cdot \frac{N}{N} \\ &= \frac{1}{s} \lim_{N \rightarrow \infty} \frac{s+N+1}{N} \exp\left(\log N - \sum_{n=1}^N \frac{1}{n}\right) \end{aligned}$$

The factor $(s+N+1)/N$ approaches 1 as $N \rightarrow \infty$. It follows that the limit of the argument of the exponential exists (you can also show this using Euler summation), so let's give it a name: Euler's constant,

$$\gamma = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} - \log n\right).$$

So we have $g(s+1) = e^{-\gamma}g(s)/s$. If we define $G(s) = e^{-\gamma s}g(s)$ then $G(s)$ satisfies the relation $sG(s+1) = G(s)$. We will show that $1/\Gamma(s) = G(s)$.

Proposition 7.3. *For all $s \in \mathbb{C}$ we have*

$$\frac{1}{\Gamma(s)} = s e^{-\gamma s} \prod_{n=1}^{\infty} \left(1 + \frac{s}{n}\right) e^{-s/n}. \quad (7.3)$$

Proof. We follow the outline of our proof of Proposition 7.1. Let $f(s) = \Gamma(s)G(s)$, where $G(s)$ is the expression on the right-hand side of (7.3). The poles of $\Gamma(s)$ cancel with the zeros of $G(s)$, so $f(s)$ is entire. Then since $\Gamma(s+1) = s\Gamma(s)$ and $sG(s+1) = G(s)$ it follows that f is periodic: $f(s+1) = f(s)$. Thus the function $F(w) = f(\frac{\log w}{2\pi i})$ is analytic in $\mathbb{C} \setminus \{0\}$. It suffices to bound $|f(s)|$ in $0 < \operatorname{Re}(s) \leq 1$ and evaluate $f(s)$ at a single point. We have

$$f(0) = \lim_{s \rightarrow 0} \Gamma(s)G(s) = \lim_{s \rightarrow 0} s\Gamma(s) = \operatorname{Res}(\Gamma; 0) = 1.$$

We have $|\Gamma(x+iy)| \leq \Gamma(x)$ and

$$\frac{|g(x+iy)|}{|g(x)|} = \prod_{n=0}^{\infty} \frac{|x+n+iy|}{x+n} = \exp\left(\frac{1}{2} \sum_{n=0}^{\infty} \log\left(1 + \frac{y^2}{(x+n)^2}\right)\right).$$

The summand is decreasing as a function of n , so we have (since $x > 0$)

$$\sum_{n=0}^{\infty} \log\left(1 + \frac{y^2}{(x+n)^2}\right) \leq \int_0^{\infty} \log(1 + (y/t)^2) dt = |y| \int_0^{\infty} \log(1 + (1/t)^2) dt.$$

Integrating by parts, the integral equals

$$2|y| \int_0^\infty \frac{dy}{y^2 + 1} = \pi|y|.$$

It follows that $|g(x + iy)| \leq \exp(\frac{\pi}{2}|y|)g(x)$, and thus $|F(w)| \ll |w|^{1/2} + |w|^{-1/2}$. Since this growth rate is sub-linear, we have that $F(w)$ is a constant by Liouville's theorem. $\square$

One consequence of the product formula is one last important property of $\Gamma(s)$. For brevity, we'll skip the proof of this one.

Proposition 7.4 (Duplication formula). *For all $s \in \mathbb{C}$ we have*

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

In the duplication formula, if we multiply both sides by s , apply the functional equation (7.1), and replace s by $-s/2$, we find that

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

This, together with the reflection formula (7.2) with $s \rightarrow -s/2$ gives

$$\frac{\Gamma(\frac{1-s}{2})}{\Gamma(\frac{s}{2})} = \frac{2^s}{\sqrt{\pi}} \sin\left(\frac{\pi s}{2}\right) \Gamma(1 - s). \quad (7.4)$$

The functional equation for $\xi(s)$ states that

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

Using (7.4) we get the asymmetric functional equation for $\zeta(s)$.

Proposition 7.5. *The Riemann zeta function satisfies*

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

8 Stirling's Formula

In this section we prove Stirling's formula, with a few caveats (below).

Theorem 8.1 (Stirling's formula). *Fix $\delta > 0$. As $|s| \rightarrow \infty$ in the sector $|\operatorname{Arg} s| \leq \pi - \delta$ we have*

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

Furthermore, under the same assumptions we have

$$\frac{\Gamma'}{\Gamma}(s) = \log s + O\left(\frac{1}{|s|}\right).$$

We will actually prove a slightly weaker form of Stirling's formula since this is all we really need in this course anyway. Also, throughout this section we will generally exchange orders of summation and integration and limits without comment; every step can be justified, but this is long enough as it is already. We begin with a lemma.

Lemma 8.2. *We have*

$$\gamma = \int_0^\infty \left(\frac{1}{e^x - 1} - \frac{1}{xe^x} \right) dx. \quad (8.1)$$

Proof. First observe that, by expanding the partial geometric series, we have

$$\sum_{k=1}^n \frac{1}{k} = \int_0^1 \frac{1-t^n}{1-t} dt = \int_0^\infty \frac{1-e^{-nx}}{e^x-1} dx. \quad (8.2)$$

The second equality comes from letting $t = e^{-x}$. We have a similar expression for $\log n$:

$$\log n = \int_0^\infty \frac{e^{-x} - e^{-nx}}{x} dx \quad (8.3)$$

To prove this, write the integral as

$$\int_0^\infty \int_1^n e^{-xy} dy dx = \int_1^n \int_0^\infty e^{-xy} dx dy = \int_1^n \frac{dy}{y} = \log n.$$

Putting (8.2) and (8.3) together, we find that

$$\gamma = \lim_{n \rightarrow \infty} \left(\sum_{k=1}^n \frac{1}{k} - \log n \right) = \lim_{n \rightarrow \infty} \int_0^\infty \left(\frac{1-e^{-nx}}{e^x-1} - \frac{e^{-x} - e^{-nx}}{x} \right) dx.$$

Taking the limit on the inside, we obtain (8.1). □

We start the proof of Stirling's formula by first proving a useful integral representation of the logarithmic derivative $\psi(s) = \frac{\Gamma'}{\Gamma}(s)$.

Proposition 8.3. *If $\operatorname{Re}(s) > 0$ then*

$$\psi(s) = \frac{\Gamma'}{\Gamma}(s) = \int_0^\infty \left(\frac{e^{-t}}{t} - \frac{e^{-st}}{1-e^{-t}} \right) dt. \quad (8.4)$$

Proof. The infinite product formula for $\Gamma(s)$ gives

$$\psi(s) = -\gamma - \frac{1}{s} + \sum_{n=1}^\infty \left(\frac{1}{n} - \frac{1}{s+n} \right).$$

We write

$$\frac{1}{s+n} = \int_0^\infty e^{-t(s+n)} dt,$$

which gives

$$\begin{aligned}
\psi(s) &= -\gamma - \int_0^\infty e^{-ts} dt + \lim_{N \rightarrow \infty} \int_0^\infty \sum_{n=1}^N (e^{-nt} - e^{-t(s+n)}) dt \\
&= \int_0^\infty \left(\frac{e^{-t}}{t} - \frac{e^{-t}}{1 - e^{-t}} \right) dt + \lim_{N \rightarrow \infty} \int_0^\infty \left(\frac{e^{-t} - e^{-st} - e^{-(N+1)t} + e^{-(N+s+1)t}}{1 - e^{-t}} \right) dt \\
&= \int_0^\infty \left(\frac{e^{-t}}{t} - \frac{e^{-st}}{1 - e^{-t}} \right) dt - \lim_{N \rightarrow \infty} \int_0^\infty \left(\frac{1 - e^{-st}}{1 - e^{-t}} \right) e^{-(N+1)t} dt.
\end{aligned}$$

We have $\frac{1 - e^{-st}}{1 - e^{-t}} \ll 1$, so the second integral is $\ll 1/N$. The formula (8.4) follows. $\square$

Corollary 8.4. *For $\operatorname{Re}(s) > 0$ we have*

$$\psi(s) = \log s + \int_0^\infty \left(\frac{1}{t} - \frac{1}{1 - e^{-t}} \right) e^{-st} dt = \log s - \frac{1}{2s} + O\left(\frac{1}{|s|^2}\right). \quad (8.5)$$

Proof. We use that

$$\int_0^\infty \frac{e^{-t} - e^{-st}}{t} dt = \log s$$

which we proved near (8.3) above. The remaining integrand is $g(t)e^{-st}$, where

$$g(t) = \frac{1}{t} - \frac{1}{1 - e^{-t}}.$$

To bound the integral, we use integration by parts twice. The function $g(t)$ is analytic at $t = 0$; its Taylor expansion there begins

$$g(t) = -\frac{1}{2} - \frac{t}{12} + O(t^3).$$

It is elementary, but somewhat tedious, to check that $g^{(k)}(t) \ll 1$ on $[0, \infty)$ for $k = 0, 1, 2$. It follows that

$$\begin{aligned}
\int_0^\infty \left(\frac{1}{t} - \frac{1}{1 - e^{-t}} \right) e^{-st} dt &= -\frac{1}{2s} + \frac{1}{s} \int_0^\infty g'(t) e^{-st} dt \\
&= -\frac{1}{2s} - \frac{1}{12s^2} - \frac{1}{s^2} \int_0^\infty g''(t) e^{-st} dt = -\frac{1}{2s} + O\left(\frac{1}{|s|^2}\right).
\end{aligned}$$

Further applications of integration by parts would yield more terms. $\square$

We can now prove a weak version of Stirling's formula with an error of $O(1)$ instead of $O(1/|s|)$. Since we won't ever need that level of accuracy in this course, I'm content to prove this weaker version. I'll also state this for $\operatorname{Re}(s) > 0$; the full version of Stirling's formula in Theorem 8.1 (without the constant term) follows after using the reflection relation (7.2).

Proposition 8.5. *For $\operatorname{Re}(s) > 0$ we have*

$$\log \Gamma(s) = (s - \tfrac{1}{2}) \log s - s + O(1).$$

Proof. Let $E(s)$ denote the error term in (8.5), i.e. $\psi(s) = \log s - 1/(2s) + E(s)$. Since $\psi(s) - \log s + 1/(2s)$ is analytic in $\{\operatorname{Re}(s) > 0\}$, so is $E(s)$. We also have $E(s) \ll 1/|s|^2$. Since $\psi(s) = \frac{d}{ds} \log \Gamma(s)$ and $\log \Gamma(1) = 0$, it follows that

$$\begin{aligned} \log \Gamma(s) &= \int_1^s \psi(w) dw = s \log s - s + 1 - \frac{1}{2} \log s + \int_1^s E(w) dw \\ &= (s - \frac{1}{2}) \log s - s + O(1), \end{aligned}$$

as desired. □

9 Weierstrass factorization of entire functions

Our main aim right now is to use the functional equation (5.4) to derive further properties of the Riemann zeta function. In order to do that, we first need to take a short detour through complex analysis.

The basic idea behind the following theorem is that we want to represent an entire function as a (possibly infinite) product over its zeros in the same way that a polynomial is a finite product over its zeros.

Theorem 9.1 (Weierstrass factorization). *Suppose that f is an entire function such that*

1. *f has a zero of order $K \geq 0$ at 0,*
2. *the other zeros of f are $z_1, z_2, \dots$, counted with multiplicity, and*
3. *there is a constant $\alpha < 2$ such that*

$$|f(z)| \ll \exp(|z|^\alpha) \quad \text{as } |z| \rightarrow \infty.$$

Then there exist numbers A, B such that

$$f(z) = z^K e^{A+Bz} \prod_{k=1}^{\infty} \left(1 - \frac{z}{z_k}\right) e^{z/z_k}$$

for all $z \in \mathbb{C}$. The product is uniformly convergent for z in compact subsets of $\mathbb{C}$.

Before proving this theorem, we first need two lemmas.

Lemma 9.2 (Jensen's inequality). *If f is analytic in a domain containing the disk $|z| \leq R$, if $|f(z)| \leq M$ in this disk, and if $f(0) \neq 0$, then for $r < R$ the number of zeros of f in the disk $|z| \leq r$ is less than or equal to*

$$\frac{\log(M/|f(0)|)}{\log(R/r)}.$$

Proof. Let $z_1, \dots, z_K$ denote the zeros of f in $|z| \leq R$, and define

$$g(z) = f(z) \prod_{k=1}^K \frac{R^2 - z\bar{z}_k}{R(z - z_k)}.$$

Each factor in the product has a pole at z_k and has modulus 1 when $|z| = R$ (this is fun to check). It follows that $g(z)$ is analytic in $|z| \leq R$ with modulus $|g(z)| = |f(z)| \leq M$ on $|z| = R$. By the maximum modulus principle, $|g(0)| \leq M$. Suppose that f has L zeros in the subdisk $|z| \leq r$. Then

$$M \geq |g(0)| = |f(0)| \prod_{k=1}^K \frac{R}{|z_k|} \geq |f(0)| \left(\frac{R}{r}\right)^L.$$

The result follows after taking logs. $\square$

Lemma 9.3. *Suppose that h is analytic in a domain containing $|z| \leq R$, that $h(0) = 0$, and that $\operatorname{Re} h(z) \leq M$ for $|z| \leq R$. Then for all $|z| \leq r < R$ we have*

$$|h(z)| \leq \frac{2Mr}{R-r}. \quad (9.1)$$

Proof. Consider the function

$$\phi(z) = \frac{h(z)}{2M - h(z)},$$

where $M = \sup_{|z| \leq R} \operatorname{Re} h(z)$. Then $\phi(z)$ is analytic in $|z| \leq R$ because the real part of the denominator does not vanish. Furthermore, $\phi(0) = 0$. Write $h = u + iv$; then we have the inequalities

$$-2M + u \leq u \leq 2M - u \implies u^2 \leq (2M - u)^2,$$

and we find that

$$|\phi(z)|^2 = \frac{u^2 + v^2}{(2M - u)^2 + v^2} \leq 1.$$

By the Schwarz lemma⁵ it follows that $|\phi(z)| \leq r/R$. After applying the triangle inequality and rearranging, we obtain (9.1). $\square$

Proof of Theorem 9.1. After replacing $f(z)$ with $f(z)/z^K$, we may assume that $f(0) \neq 0$. Let $N(R)$ be the number of zeros of f in $|z| \leq R$. Then by Jensen's inequality, $N(R) \ll R^\alpha$. It follows that

$$\sum_{R \leq |z_k| \leq 2R} \frac{1}{|z_k|^2} \ll R^{\alpha-2},$$

so by summing over dyadic blocks we find that

$$\sum_{k=1}^{\infty} \frac{1}{|z_k|^2} < \infty.$$

⁵If f is analytic, $f(z) = 0$, and $|f(z)| \leq 1$ in the open unit disk D , then $|f(z)| \leq |z|$ in D .

(Note that we can actually replace the exponent 2 by $\alpha + \epsilon$.) Note that $(1 - z)e^z = 1 + O(|z|^2)$ uniformly for $|z| \leq 1$, so the convergence of the sum above shows that the product

$$g(z) = \prod_{k=1}^{\infty} \left(1 - \frac{z}{z_k}\right) e^{z/z_k}$$

converges uniformly in compact subsets of $\mathbb{C}$. Hence g is an entire function whose zeros match the zeros of $f(z)$. Thus the function

$$h(z) = \frac{f(z)}{f(0)g(z)}$$

is entire and nonvanishing on $\mathbb{C}$, and $h(0) = 1$.

It remains to show that $h(z) = e^{Bz}$. We first need a bound for $|h(z)|$ in $|z| \leq R$. Write the product defining $g(z)$ as

$$P_1(z)P_2(z)P_3(z) = \prod_{|z_k| \leq R/2} \left(1 - \frac{z}{z_k}\right) e^{z/z_k} \prod_{R/2 < |z_k| \leq 3R} \left(1 - \frac{z}{z_k}\right) e^{z/z_k} \prod_{|z_k| > 3R} \left(1 - \frac{z}{z_k}\right) e^{z/z_k}.$$

Suppose that $R \leq |z| \leq 2R$. If $|z_k| \leq R/2$ then $|1 - z/z_k| \geq |z/z_k| - 1 \geq 1$, so

$$|P_1(z)| \geq \prod_{|z_k| \leq R/2} e^{-2R/|z_k|}.$$

Furthermore, we have

$$\sum_{|z_k| \leq R/2} \frac{1}{|z_k|} \ll R^{\alpha-1},$$

so

$$|P_1(z)| \geq \prod_{|z_k| \leq R/2} e^{-2R/|z_k|} = \exp\left(-2R \sum_{|z_k| \leq R/2} \frac{1}{|z_k|}\right) \gg \exp(-c_1 R^\alpha)$$

for some $c_1 > 0$. For the second product, we first observe that $\#\{R/2 < |z_k| \leq 3R\} \ll R^\alpha$. Since $\alpha < 2$, the pigeonhole principle shows that there is some r in $[R, 2R]$ such that $|r - |z_k|| \geq 1/R^2$ for all k . So if $|z| = r$ we have

$$|1 - z/z_k| \geq \frac{|r - |z_k||}{|z_k|} \gg \frac{1}{R^3}$$

for all k in the second product. Therefore when $|z| = r$ we have

$$|P_2(z)| \geq e^{-c_2 R^\alpha \log R}$$

for some $c_2 > 0$. Finally, we have

$$|P_3(z)| \geq e^{-c_3 R^\alpha}$$

for some $c_3 > 0$, when $|z| \leq 2R$. We conclude that for each large R , there is an $r \in [R, 2R]$ such that $|g(z)| \geq \exp(-cR^\alpha \log R)$ when $|z| = r$, for some $c > 0$. By the maximum modulus principle, it follows that

$$\max_{|z| \leq R} |h(z)| \leq e^{cR^\alpha \log R}.$$

Let $j(z) = \log h(z)$ (since h is nonvanishing, there are no branches to consider, so $j(z)$ is analytic). Since $h(0) = 1$ we have $j(0) = 0$. Furthermore $\operatorname{Re} j(z) = \log |h(z)| \leq cR^\alpha \log R$ for all large R . So by Lemma 9.3 we conclude that

$$j(z) \ll R^\alpha \log R.$$

But $\alpha < 2$ so by Liouville's theorem (actually, a simple corollary to Liouville), j must be a polynomial of degree at most 1 with $j(0) = 0$, so $j(z) = Bz$ for some B . $\square$

As a simple example of the Weierstrass factorization theorem, we obtain

$$\sin \pi z = \pi z e^{Bz} \prod_{n \neq 0} \left(1 - \frac{z}{n}\right) e^{z/n} = \pi z e^{Bz} \prod_{n=1}^{\infty} \left(1 - \frac{z^2}{n^2}\right)$$

for some B . Clearly $B \in \mathbb{R}$ since $\sin \pi x$ is real for $x \in \mathbb{R}$. Letting $z = iy$, we obtain

$$\sinh \pi y = \pi y e^{iBy} \prod_{n=1}^{\infty} \left(1 + \frac{y^2}{n^2}\right).$$

Since the left-hand side is real for all y , it follows that $B = 0$. We conclude that

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

As a fun application of this, consider the Taylor expansion of $\sin \pi z$: the coefficient of z^3 is $-\pi^3/6$. On the other hand, expanding out the product on the right-hand side, we find that the coefficient of z^3 equals

$$-\pi \sum_{n=1}^{\infty} \frac{1}{n^2} = -\pi \zeta(2).$$

It follows that $\zeta(2) = \pi^2/6$.

10 A zero-free region for $\zeta(s)$

All we need to know for the simple form of the PNT $\pi(x) \sim x/\log x$ is that $\zeta(s)$ is nonvanishing on the line $\sigma = 1$. However, we can obtain a quantitative form of PNT with an error term if we put in a bit more effort to show that $\zeta(s)$ is nonvanishing a little bit to the left of $\sigma = 1$. In this section we will prove the following.

Theorem 10.1. *There is an absolute constant $c > 0$ such that $\zeta(s) \neq 0$ for*

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

This is the classical zero-free region for the zeta function. Essentially every improvement to the prime number theorem has been a result of proving stronger zero-free regions.

10.1 Nonvanishing of zeta on the 1-line

We begin by proving that $\zeta(s)$ is nonvanishing on the line $\sigma = 1$; while this argument is quite simple, it serves as the motivation for our more complicated proof of Theorem 10.1.

Lemma 10.2. (a) If $\sigma > 1$ then

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

(b) $\zeta(1 + it) \neq 0$ for all $t \in \mathbb{R}$.

Proof. (a) Using the Dirichlet series form of ζ'/ζ , we see that

$$\operatorname{Re}\left(-\frac{\zeta'}{\zeta}(\sigma + it)\right) = \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^{\sigma}} \operatorname{Re}(n^{-it}).$$

We compute $\operatorname{Re}(n^{-it}) = \operatorname{Re}(\exp(it \log n)) = \cos(t \log n)$. It follows that

$$\operatorname{Re}\left(-3\frac{\zeta'}{\zeta}(\sigma) - 4\frac{\zeta'}{\zeta}(\sigma + it) - \frac{\zeta'}{\zeta}(\sigma + 2it)\right) = \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^{\sigma}} (3 + 4\cos(t \log n) + \cos(2t \log n)).$$

The result follows from the simple inequality

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

(b) Suppose that $s = 1 + i\gamma$ is a zero of $\zeta(s)$ of multiplicity m . Then by Lemma 6.2 we have

$$\frac{\zeta'}{\zeta}(1 + \delta + i\gamma) \sim \frac{m}{\delta} \quad \text{as } \delta \rightarrow 0^+.$$

But we have

$$\operatorname{Re} \frac{\zeta'}{\zeta}(1 + \delta + i\gamma) = - \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^{1+\delta}} \cos(\gamma \log n) \leq \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^{1+\delta}} = -\frac{\zeta'}{\zeta}(1 + \delta) \sim \frac{1}{\delta},$$

where we have used Lemma 6.2 again for the pole of $\zeta(s)$ at $s = 1$. It follows that $m = 1$ (if there is a zero, it must be simple). But then the inequality in (a) implies that

$$\frac{\zeta'}{\zeta}(1 + \delta + 2i\gamma) \sim -\frac{1}{\delta},$$

i.e. $\zeta(s)$ has a pole at $1 + 2i\gamma$, contradicting Theorem 4.6. □

10.2 The infinite product for $\xi(s)$

In order to extend our zero-free region to the left of $\sigma = 1$ we require a formula for $\frac{\zeta'}{\zeta}(s)$ involving its zeros. This will follow from the Weierstrass factorization of the entire function

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

We first need to verify the growth condition in Theorem 9.1 for some $\alpha < 2$; we will actually prove the stronger bound

$$\xi(s) \ll \exp(C|s| \log |s|). \quad (10.1)$$

Since $\xi(s) = \xi(1-s)$ it suffices to prove this bound for $\sigma \geq 1/2$. We clearly have the bound $s(s-1)\pi^{-s/2} \ll e^{c_1|s|}$ for some $c_1 > 0$.⁶ By Stirling's formula,

$$\Gamma(\frac{s}{2}) \ll \exp(c_2|s| \log |s|).$$

(Note that we only use Proposition 8.5 since $\sigma \geq 1/2$.) So it suffices to bound $\zeta(s)$; recall the analytic continuation

$$\zeta(s) = \frac{1}{s-1} + 1 - s \int_1^\infty \frac{t - \lfloor t \rfloor}{t^{s+1}} dt$$

which is valid for $\sigma > 0$. For $\sigma \geq 1/2$ the integral is bounded by some absolute constant, so $\zeta(s) \ll |s|$ for large $|s|$. Thus we have proved (10.1).

It will be useful to know that we *cannot* remove the $\log |s|$ factor in the bound (10.1). Indeed, Stirling's formula tells us that $\Gamma(\sigma) \gg \exp(c_3|\sigma| \log |\sigma|)$ as $\sigma \rightarrow \infty$ and by the Dirichlet series definition of $\zeta(s)$ it follows that $\zeta(\sigma) \sim 1$ as $\sigma \rightarrow \infty$.

This fact is useful because it tells us something about the zeros of $\zeta(s)$. Let $\rho_1, \rho_2, \dots$ denote the zeros of $\xi(s)$. Since all of the trivial zeros of $\zeta(s)$ are cancelled by the poles of $\Gamma(s/2)$, and since the factor $s(s-1)$ cancels the pole of $\Gamma(s/2)$ at 0 and the pole of $\zeta(s)$ at 1, the ρ_n are also the (nontrivial) zeros of $\zeta(s)$. By Theorem 9.1 we have

$$\xi(s) = e^{A+Bs} \prod_{\rho} \left(1 - \frac{s}{\rho}\right) e^{s/\rho}.$$

Recall that in the proof of the Weierstrass factorization theorem we showed that $\sum_k 1/|z_k|^{\alpha+\epsilon}$ converges, where z_k are the zeros of f and α is the exponent in the growth condition. It follows that the series

$$\sum_n \frac{1}{|\rho_n|^{1+\epsilon}} \quad (10.2)$$

converges for any $\epsilon > 0$.

Suppose, for the moment, that the series (10.2) converges with $\epsilon = 0$. Then I claim that we have the bound $\xi(s) \ll e^{c_4|s|}$ (which we already showed *doesn't* hold). To show this, we use the inequality

$$|(1-z)e^z| \leq e^{2|z|},$$

⁶We will use c_1, c_2 , etc to denote positive constants without further comment.

which for $|z|$ large is obvious, and for $|z|$ small, follows from the Taylor expansion of $(1-z)e^z$ at $z = 0$. Then

$$\prod_{\rho} \left(1 - \frac{s}{\rho}\right) e^{s/\rho} \ll \prod_{\rho} e^{2|s|/|\rho|} = \exp\left(2|s| \sum_{\rho} \frac{1}{|\rho|}\right) \ll e^{2|s|},$$

so $\xi(s) \ll \exp((B+2)|s|)$. We conclude that the series

$$\sum_{\rho} \frac{1}{|\rho|}$$

diverges, and thus $\zeta(s)$ has infinitely many nontrivial zeros (all in the critical strip $0 \leq \sigma \leq 1$).

We conclude from Theorem 9.1 that there exist constants A, B such that

$$\xi(s) = e^{A+Bs} \prod_{\rho} \left(1 - \frac{s}{\rho}\right) e^{s/\rho} \quad (10.3)$$

To compute A , we have (by the functional equation)

$$e^A = \xi(0) = \xi(1) = \pi^{-1/2} \Gamma\left(\frac{1}{2}\right) \lim_{s \rightarrow 1} (s-1)\zeta(s) = 1.$$

The constant B takes a bit more (tedious but not hard) work to compute; I'll just tell you

$$B = \frac{1}{2} \log 4\pi - \frac{1}{2}\gamma - 1.$$

Logarithmic differentiation of (10.3) leads to the following theorem.

Theorem 10.3. *For all $s \in \mathbb{C}$ we have*

$$\frac{\xi'}{\xi}(s) = \log\left(\frac{2\pi}{e}\right) - \frac{1}{2}\gamma - \frac{1}{s-1} - \frac{\Gamma'}{\Gamma}\left(\frac{s}{2} + 1\right) + \sum_{\rho} \left(\frac{1}{s-\rho} + \frac{1}{\rho}\right).$$

Proof. Logarithmic differentiation of (10.3) gives

$$\frac{\xi'}{\xi}(s) = B + \sum_{\rho} \left(\frac{1}{s-\rho} + \frac{1}{\rho}\right). \quad (10.4)$$

We combine this with the logarithmic derivative of the right-hand side of

$$\xi(s) = 2(s-1)\pi^{-s/2}\Gamma\left(\frac{s}{2} + 1\right)\zeta(s),$$

where we have used that $s\Gamma\left(\frac{s}{2}\right) = 2\Gamma\left(\frac{s}{2} + 1\right)$. □

Remark 10.4. From the definition of $\zeta(s)$ we have the relation $\zeta(s) = \overline{\zeta(\bar{s})}$ for $\sigma > 1$ (which extends to all s by analytic continuation), from which it follows that if ρ is a zero of $\zeta(s)$ then so is $\bar{\rho}$. And if ρ is a nontrivial zero of $\zeta(s)$ then the functional equation $\xi(s) = \xi(1-s)$

shows that $1 - \rho$ and $1 - \bar{\rho}$ are also zeros. So even though $\sum 1/|\rho|$ diverges, the sum $\sum 1/\rho$ is convergent, assuming we sum in a particular order:

$$\lim_{T \rightarrow \infty} \sum_{\substack{\rho = \beta + i\gamma \\ |\gamma| \leq T}} \frac{1}{\rho} = \lim_{T \rightarrow \infty} \sum_{0 \leq \gamma \leq T} \left(\frac{1}{\rho} + \frac{1}{\bar{\rho}} \right) = \lim_{T \rightarrow \infty} \sum_{0 \leq \gamma \leq T} \frac{2\beta}{\beta^2 + \gamma^2} \leq 2 \sum_{\rho} \frac{1}{|\gamma|^2}.$$

Then by (10.4) and the functional equation $\xi(s) = \xi(1-s)$ we have

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

By the pairing up zeros $1 - \rho$ and ρ and setting $s = 0$ we find that

$$B = - \sum_{\rho} \frac{1}{\rho} = -2 \sum_{\substack{\rho = \beta + i\gamma \\ \gamma > 0}} \frac{\beta}{\beta^2 + \gamma^2}, \quad (10.5)$$

where we sum in the order described above. Since $B = -0.023\dots$ it follows (using the fact that we can take $1/2 \leq \beta \leq 1$) that the imaginary part of the lowest nontrivial zero is > 6.5 (actually, more detailed computations show that the first zero is at $\frac{1}{2} + 14.1347\dots i$).

10.3 The classical zero-free region

We begin with the inequality

$$\operatorname{Re} \left(-3 \frac{\zeta'}{\zeta}(\sigma) - 4 \frac{\zeta'}{\zeta}(\sigma + it) - \frac{\zeta'}{\zeta}(\sigma + 2it) \right) \geq 0 \quad (10.6)$$

and we estimate each piece more carefully than we did before. The pole of $\zeta(s)$ at $s = 1$ shows that for the first term we have

$$-3 \frac{\zeta'}{\zeta}(\sigma) - \frac{3}{\sigma - 1} \leq c_1$$

uniformly for $1 < \sigma \leq 2$. For the other two terms, it is clear that if $\zeta(s)$ has a zero to the left of $\sigma = 1$, but close to that line, then the behavior of ζ'/ζ is greatly influenced by that zero even to the right of $\sigma = 1$. We can make this explicit by looking at

$$-\frac{\zeta'}{\zeta}(s) = -\log\left(\frac{2\pi}{e}\right) + \frac{1}{2}\gamma + \frac{1}{s-1} + \frac{\Gamma'}{\Gamma}\left(\frac{s}{2} + 1\right) - \sum_{\rho} \left(\frac{1}{s-\rho} + \frac{1}{\rho} \right).$$

Since there are no zeros below $t = 2$, let us restrict ourselves to the region $t \geq 2$ and $1 \leq \sigma \leq 2$. By Stirling's formula for $\psi(s)$, we have

$$\operatorname{Re} \left\{ -\log\left(\frac{2\pi}{e}\right) + \frac{1}{2}\gamma + \frac{1}{s-1} + \frac{\Gamma'}{\Gamma}\left(\frac{s}{2} + 1\right) \right\} \leq c_2 \log t,$$

hence

$$-\operatorname{Re} \frac{\zeta'}{\zeta}(s) \leq -\sum_{\rho} \operatorname{Re} \left(\frac{1}{s-\rho} + \frac{1}{\rho} \right) + c_2 \log t. \quad (10.7)$$

For $\rho = \beta + i\gamma$ we have

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

It follows that the sum over ρ is positive (since $\sigma > \beta$), so the corresponding term in (10.7) is only making the right-hand side smaller. So for $s = \sigma + 2it$ we have

$$-\operatorname{Re} \frac{\zeta'}{\zeta}(\sigma + 2it) \leq c_3 \log t.$$

When $s = \sigma + it$, we suppose that $t = \gamma$ for some zero $\rho = \beta + i\gamma$ and we drop all but the term corresponding to that zero:

$$-4 \operatorname{Re} \frac{\zeta'}{\zeta}(\sigma + it) + \frac{4}{\sigma - \beta} \leq c_4 \log t.$$

Inserting these inequalities into (10.6), we find that

$$0 \leq \frac{3}{\sigma - 1} - \frac{4}{\sigma - \beta} + c_5 \log t.$$

Now suppose that $\sigma = 1 + \delta / \log t$, where δ is a positive constant. Then, solving for β above,

$$\beta \leq 1 - \frac{c}{\log t}, \quad \text{where } c = \frac{4}{3/\delta + c_5} - \delta.$$

We want $c > 0$, so we need to choose δ appropriately (note that this only works because $4 > 3$; this is essential!). We can choose any δ in $(0, 1/c_5)$, but the optimal choice is

$$\delta = \frac{2\sqrt{3} - 3}{c_5}, \quad \text{so that } c = \frac{7 - 4\sqrt{3}}{c_5}.$$

This proves Theorem 10.1.

11 The number of nontrivial zeros below height T

Our aim in this section is to count the number of nontrivial zeros of $\zeta(s)$ inside the critical strip up to height T . More precisely, we will estimate the number

$$N(T) = \#\{\rho = \beta + i\gamma : \zeta(\rho) = 0 \text{ and } \beta \in [0, 1], \gamma \in [0, T]\}.$$

To accomplish this, we recall the argument principle from complex analysis: if $f(z)$ is meromorphic in the interior of a simple closed contour Γ and analytic and nonzero on Γ then

$$\frac{1}{2\pi i} \int_{\Gamma} \frac{f'(z)}{f(z)} dz = Z - P,$$

where Z is the number of zeros of f inside Γ and P is the number of poles (both counted with multiplicity). We apply the argument principle to the entire function $\xi(s)$ since the nontrivial zeros of $\zeta(s)$ are exactly the zeros of $\xi(s)$. Thus

$$N(T) = \frac{1}{2\pi i} \int_C \frac{\xi'}{\xi}(s) ds,$$

where C is the boundary of the rectangle in the definition of $N(T)$. This integral formula is called the argument principle because the integral can also be written as the change in the argument of the function on the contour, which can be seen by writing $\frac{\xi'}{\xi}(s) = \frac{d}{ds} \log \xi(s)$. Thus, we can reinterpret $N(T)$ as

$$N(T) = \frac{1}{2\pi} \Delta_C \arg(\xi(s)).$$

Before estimating $N(T)$ we start with a few basic properties of the zeta zeros.

11.1 Approximations of ζ'/ζ

Using the ideas from the last section we can prove useful bounds for how many zeros of $\zeta(s)$ are close together and also an approximation for $\zeta'(s)/\zeta(s)$.

Lemma 11.1. 1. The number of nontrivial zeros of $\zeta(s)$ of the form $\rho = \beta + i\gamma$ with $|\gamma - t| \leq 1$ is $O(\log t)$.

2. Uniformly for $t \geq 2$ and $-1 \leq \sigma \leq 2$ we have

$$\frac{\zeta'}{\zeta}(s) = \sum_{\substack{\rho=\beta+i\gamma \\ |\gamma-t|\leq 1}} \frac{1}{s-\rho} + O(\log t).$$

Proof. (1) Suppose that $s = 2 + it$ with $t \geq 1$. Arguing as above, we have

$$\sum_{\rho} \frac{1}{s-\rho} = \frac{\zeta'}{\zeta}(s) + \log |s| + O(1) \tag{11.1}$$

since, by (10.5), the sum $\sum 1/\rho$ is $O(1)$. For $\sigma = 2$ we have

$$\frac{\zeta'}{\zeta}(2 + it) \ll \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^2} \ll \sum_{n=1}^{\infty} \frac{1}{n^{3/2}} \ll 1.$$

It follows that

$$\operatorname{Re} \left(\sum_{\rho} \frac{1}{2 + it - \rho} \right) \ll \log t$$

But for $\rho = \beta + i\gamma$ with $0 \leq \beta \leq 1$ we have

$$\operatorname{Re} \left(\frac{1}{2 + it - \rho} \right) = \frac{2 - \beta}{(2 - \beta)^2 + (t - \gamma)^2} \geq \frac{1}{4 + (t - \gamma)^2},$$

so

$$\sum_{|t-\gamma|\leq 1} 1 \ll \sum_{|t-\gamma|\leq 1} \frac{1}{4+(t-\gamma)^2} \leq \sum_{\rho} \frac{1}{4+(t-\gamma)^2} \leq \operatorname{Re} \left(\sum_{\rho} \frac{1}{2+it-\rho} \right) \ll \log t. \quad (11.2)$$

(2) Subtract (11.1) with $s = 2 + it$ from (11.1) with $s = \sigma + it$, where $-1 \leq \sigma \leq 2$ and $\sigma + it$ is not one of the zeros. We have

$$\left| \frac{1}{s-\rho} - \frac{1}{2+it-\rho} \right| = \frac{2-\sigma}{|s-\rho| \cdot |2+it-\rho|} \leq \frac{2-\sigma}{(t-\gamma)^2} \ll \frac{1}{4+(t-\gamma)^2}.$$

It follows that

$$\begin{aligned} \frac{\zeta'}{\zeta}(s) - \sum_{|t-\gamma|\leq 1} \frac{1}{s-\rho} &\ll \sum_{|t-\gamma|\leq 1} \frac{1}{|2+it-\rho|} + \sum_{|t-\gamma|\geq 1} \left| \frac{1}{s-\rho} - \frac{1}{2+it-\rho} \right| + \log t \\ &\ll \sum_{|t-\gamma|\leq 1} 1 + \sum_{|t-\gamma|\geq 1} \frac{1}{4+(t-\gamma)^2} + \log t. \end{aligned}$$

The bound $\ll \log t$ follows from (11.2) applied to each sum. $\square$

11.2 The number $N(T)$

We showed in the previous section that $\xi(s)$ has no zeros on the line $\sigma = 1$; by the functional equation it follows that $\xi(s)$ also has no zeros on the line $\sigma = 0$. We also showed that there are no zeros in $|t| \leq 6$, so $\xi(s)$ is not zero on the real line. If we choose T such that $\zeta(\sigma + iT) \neq 0$ for $\sigma \in [0, 1]$ then there are no zeros on the contour C .

We can argue by symmetry to make things a bit simpler. Since $\xi(s) = \xi(1-s) = \overline{\xi(1-\bar{s})}$ we see that $\xi(\sigma + it) = \xi(1-\sigma + it)$. It follows that the change of argument as we move along the right half of the contour ($\sigma \geq 1/2$) is the same as when we move on the left half of the contour. Furthermore, $\xi(s)$ is real on the real line (and therefore it's positive since it has no real zeros), so there is no change in argument along the real line. Therefore, $N(T) = \frac{1}{\pi} \Delta_L \arg(\xi(s))$, where L is the upside-down L-shaped contour from 1 to $1 + iT$, then from $1 + iT$ to $1/2 + iT$.

To avoid the pole of $\zeta(s)$ at $s = 1$ it is more convenient to widen the region defining $N(T)$ to $-1 \leq \sigma \leq 2$. Since $\xi(s)$ has no extra zeros in this larger region, all our arguments above go through exactly the same, i.e. we can consider instead the contour L from 2 to $1/2 + iT$. By definition we have

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

On L the arguments of both s and $s-1$ change from 0 to $\pi/2 + O(1/T)$ because

$$\arg\left(\pm \frac{1}{2} + iT\right) = \frac{\pi}{2} \mp \arctan\left(\frac{1}{2T}\right) = \frac{\pi}{2} + O\left(\frac{1}{T}\right).$$

The next term is

$$\arg(\pi^{-s/2}) = \arg(e^{-\frac{1}{2}s \log \pi}) = \arg(e^{-\frac{1}{2}it \log \pi}) = -\frac{1}{2}t \log \pi,$$

so the change in argument for that term is $-\frac{1}{2}T \log \pi$. For the gamma term we can use Stirling's formula since s is always in the sector $|\arg s| \leq \pi/2$. Hence

$$\begin{aligned}\Delta_L \arg(\Gamma(\tfrac{s}{2})) &= \operatorname{Im} \log \Gamma(\tfrac{1}{4} + \tfrac{1}{2}iT) - \operatorname{Im} \log \Gamma(1) \\ &= \operatorname{Im} \left\{ \left(-\tfrac{1}{4} + \tfrac{1}{2}iT \right) \log \left(\tfrac{1}{4} + \tfrac{1}{2}iT \right) - \left(\tfrac{1}{4} + \tfrac{1}{2}iT \right) + \tfrac{1}{2} \log 2\pi + O\left(\tfrac{1}{T}\right) \right\}.\end{aligned}$$

Using the Taylor expansion of $\log(1+x)$ we find that

$$\begin{aligned}\log\left(\tfrac{1}{4} + \tfrac{1}{2}iT\right) &= \frac{\pi i}{2} - \log 2 + \log T + \log\left(1 - \frac{2i}{T}\right) \\ &= \frac{\pi i}{2} + \log \frac{T}{2} + O\left(\frac{1}{T}\right).\end{aligned}$$

Thus

$$\Delta_L \arg(\Gamma(\tfrac{s}{2})) = \frac{1}{2}T \log \frac{T}{2} - \frac{1}{2}T - \frac{\pi}{8} + O\left(\frac{1}{T}\right),$$

and

$$N(T) = \frac{T}{2\pi} \log\left(\frac{T}{2\pi e}\right) + \frac{7}{8} + S(T) + O\left(\frac{1}{T}\right),$$

where $S(T) = \frac{1}{\pi} \Delta_L \arg(\zeta(s))$.

Along the line $\sigma = 2$ we have

$$\arg(\zeta(2+it)) = \operatorname{Im} \log(\zeta(2+it)) = -\operatorname{Im} \sum_{n=1}^{\infty} \frac{\log n}{n^2} n^{-it} = \sum_{n=1}^{\infty} \frac{\log n}{n^2} \sin(t \log n) \ll 1.$$

It follows that

$$\begin{aligned}\Delta_L \arg(\zeta(s)) &= \arg(\zeta(\tfrac{1}{2} + iT)) - \arg(\zeta(2 + iT)) + O(1) \\ &= - \int_{1/2+iT}^{2+iT} \operatorname{Im} \left(\frac{\zeta'}{\zeta}(s) \right) ds + O(1).\end{aligned}$$

Now we can appeal to Lemma 11.1(2) to get

$$\begin{aligned}\Delta_L \arg(\zeta(s)) &= - \sum_{|\gamma-T| \leq 1} \int_{1/2+iT}^{2+iT} \operatorname{Im} \left(\frac{1}{s-\rho} \right) ds + O(\log T) \\ &= - \sum_{|\gamma-T| \leq 1} \left(\arg(2+iT-\rho) - \arg(\tfrac{1}{2}+iT-\rho) \right) + O(\log T).\end{aligned}$$

The difference in argument between $(2-\beta) + i(T-\gamma)$ and $(\frac{1}{2}-\beta) + i(T-\gamma)$ is at most π , and since there are $\ll \log T$ terms in the sum, we find that $S(T) \ll \log T$, which yields

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

As a corollary, we estimate the sum

$$\sum_{|\gamma| \leq T} \frac{1}{|\rho|}.$$

We already showed that as $T \rightarrow \infty$ the sum above is infinite, but we can now measure the order of growth. By partial summation,

$$\sum_{|\gamma| \leq T} \frac{1}{|\rho|} = \frac{2N(T)}{T} + 2 \int_1^T \frac{N(t)}{t^2} dt = \frac{1}{\pi} \int_1^T \frac{\log t}{t} dt + O(\log T) = \frac{\log^2 T}{2\pi} + O(\log T). \quad (11.3)$$

12 The Prime Number Theorem

We are finally ready to prove the prime number theorem. We begin by choosing an explicit test function, then we prove the explicit formula and finally obtain a quantitative error bound.

12.1 The test function $\phi_x(t)$

For our proof of the prime number theorem, we will use the test function

$$\phi_x(t) = \begin{cases} 1 & \text{if } 0 \leq t \leq x, \\ 1 + \frac{x}{y} - \frac{1}{y}t & \text{if } x \leq t \leq x + y, \\ 0 & \text{if } t \geq x + y. \end{cases}$$

Here y is a parameter satisfying $1 \leq y \leq x$ which we will choose later. Note that $\phi_x(t)$ is continuous and piecewise linear on $[0, \infty)$ and that it satisfies $0 \leq \phi_x(t) \leq 1$. In the interval $[x, x + y] \subseteq [x, 2x]$, the function $\Lambda(n)$ is approximately $\log x$ (when it is nonzero), so the price we pay for using $\phi_x(t)$ instead of a sharp cutoff function is

$$\sum_{n \leq x} \Lambda(n) - \sum_{n \leq x} \Lambda(n) \phi_x(n) \ll \sum_{x \leq n \leq x+y} \Lambda(n) \ll y \log x. \quad (12.1)$$

We now compute the Mellin transform of $\phi_x(t)$.

Lemma 12.1. *With $\phi_x(t)$ as above, we have*

$$\begin{aligned} \tilde{\phi}_x(s) &= \frac{x^{s+1}}{ys(s+1)} \left(\left(1 + \frac{y}{x}\right)^{s+1} - 1 \right) \\ &= \frac{x^s}{s} + yx^{s-1}w_s, \end{aligned} \quad (12.2)$$

where

$$w_s = \frac{(1 + \frac{y}{x})^{s+1} - 1 - (s+1)\frac{y}{x}}{s(s+1)(\frac{y}{x})^2}.$$

Furthermore, if $1 \leq y \leq x$ and $0 \leq \operatorname{Re}(s) \leq 1$ then

$$|w_s| \leq \frac{1}{2}.$$

Proof. The expressions for $\tilde{\phi}_x(s)$ are simple computations, starting from the definition

$$\tilde{\phi}_x(s) = \int_0^\infty \phi_x(t) t^s \frac{dt}{t} = \int_0^x t^{s-1} dt + \int_x^y \left(1 + \frac{x}{y} - \frac{t}{y}\right) t^{s-1} dt.$$

To prove the bound for w_s , we recall Taylor's theorem (for suitable twice-differentiable functions) which says that

$$f(u) = f(0) + f'(0)u + \int_0^u f''(t)(u-t) dt.$$

Applying this to the function $f(u) = (1+u)^{s+1}$ we find that

$$(1+u)^{s+1} = 1 + (s+1)u + s(s+1) \int_0^u (1+t)^{s-1}(u-t) dt.$$

Thus, for $y \in [0, 1]$ and $\operatorname{Re}(s) \in [0, 1]$ we have

$$|w_s| = \left(\frac{x}{y}\right)^2 \left| \int_0^{y/x} (1+t)^{s-1} \left(\frac{y}{x} - t\right) dt \right| \leq \left(\frac{x}{y}\right)^2 \int_0^{y/x} \left(\frac{y}{x} - t\right) dt = \frac{1}{2},$$

as desired. □

12.2 Contour integration

Let T be a large parameter such that $|\sigma + iT - \rho| \gg (\log T)^{-1}$ for all nontrivial zeros ρ . This is possible by Lemma 11.1(1) and the pigeonhole principle. By the Mellin inversion formula we have

$$\sum_{n=1}^{\infty} \Lambda(n) \phi_x(n) = -\frac{1}{2\pi i} \int_{(2)} \frac{\zeta'}{\zeta}(s) \tilde{\phi}_x(s) ds = -\frac{1}{2\pi i} \int_{2-iT}^{2+iT} \frac{\zeta'}{\zeta}(s) \tilde{\phi}_x(s) ds + R_x(T),$$

where (using (18.2))

$$R_x(T) \ll x^3 y^{-1} \frac{\zeta'}{\zeta}(2) \int_T^\infty \frac{dt}{t^2} \ll \frac{x^3}{yT}.$$

We now apply the residue theorem to the integrand $-\frac{\zeta'}{\zeta}(s) \tilde{\phi}_x(s)$ over the rectangular contour with corners $2 - iT$, $2 + iT$, $-N + iT$, and $-N - iT$, where $-N$ is a large negative odd integer (to avoid the poles of ζ'/ζ). In that box the integrand has poles coming from (1) the pole of $\zeta(s)$ at $s = 1$, (2) the nontrivial zeros of $\zeta(s)$ with $|\gamma| \leq T$, (3) the pole of $\tilde{\phi}_x(s)$ at $s = 0$, and (4) the trivial zeros of $\zeta(s)$ at negative integers $-2k \geq -N$. Therefore

$$\begin{aligned} -\frac{1}{2\pi i} \int_{2-iT}^{2+iT} \frac{\zeta'}{\zeta}(s) \tilde{\phi}_x(s) ds &= \frac{1}{2\pi i} \left\{ -\int_{-N-iT}^{-N+iT} - \int_{-N+iT}^{2+iT} + \int_{-N-iT}^{2-iT} \right\} \frac{\zeta'}{\zeta}(s) \tilde{\phi}_x(s) ds \\ &\quad + \tilde{\phi}_x(1) - \sum_{|\gamma| \leq T} \tilde{\phi}_x(\rho) - \frac{\zeta'}{\zeta}(0) - \sum_{k=1}^{\lfloor N/2 \rfloor} \tilde{\phi}_x(-2k). \end{aligned}$$

Our aim is to show that all of the integrals on the right-hand side go to zero as $N, T \rightarrow \infty$. To simplify things a bit, let's assume that $N \asymp T$, so that they are going to ∞ at the same rate. The upshot is that $\log |s| \asymp \log N \asymp \log T$ for s on the three lines above.

On the horizontal segments, Lemma 11.1 provides a bound for ζ'/ζ for $-1 \leq \sigma \leq 2$:

$$\frac{\zeta'}{\zeta}(\sigma + iT) \ll \sum_{|T-\gamma| \leq 1} \frac{1}{|T-\gamma|} + \log T \ll \log^2 T$$

since we chose T such that $|\gamma - T| \gg (\log T)^{-1}$ and there are $\ll \log T$ terms in the sum. So we need a bound for ζ'/ζ in the region $\sigma \leq -1$. By our choice of N , the distance from the rectangular contour to a trivial zero of $\zeta(s)$ is always at least 1, so we can restrict ourselves to the set $L = \{s : \sigma \leq -1 \text{ and } |s + 2k| \geq 1/2 \text{ for all } k \geq 1\}$. By the asymmetrical functional equation Proposition 7.5 we have

$$\frac{\zeta'}{\zeta}(s) = \log 2\pi + \frac{\pi}{2} \cot\left(\frac{\pi s}{2}\right) - \frac{\Gamma'}{\Gamma}(1-s) - \frac{\zeta'}{\zeta}(1-s).$$

(Using this we can compute $\frac{\zeta'}{\zeta}(0) = \log 2\pi$.) In the region L we have $\operatorname{Re}(1-s) \geq 2$, and we are avoiding the poles of $\cot(\pi s/2)$ by a distance of at least $1/2$. So everything is $O(1)$ except the gamma factor. By Stirling's formula it follows that

$$\frac{\zeta'}{\zeta}(s) \ll \log |s|$$

for $s \in L$. We bound the horizontal segments by

$$\left\{ \int_{-N+iT}^{2+iT} + \int_{-N-iT}^{2-iT} \right\} \frac{\zeta'}{\zeta}(s) \tilde{\phi}_x(s) ds \ll \frac{\log^2 T}{yT^2} \int_{-\infty}^2 x^{\sigma+1} d\sigma \ll \frac{x^3 \log^2 T}{yT^2 \log x}.$$

For the vertical segment we have

$$\int_{-N-iT}^{-N+iT} \frac{\zeta'}{\zeta}(s) \tilde{\phi}_x(s) ds \ll \frac{T \log N}{yN^2 x^{N-1}}.$$

So all three integrals go to zero as $T \asymp N \rightarrow \infty$. Thus we obtain our first form of the explicit formula.

Proposition 12.2. *With $\phi_x(t)$ as above, we have*

$$\sum_{n=1}^{\infty} \Lambda(n) \phi_x(n) = \tilde{\phi}_x(1) - \sum_{\rho} \tilde{\phi}_x(\rho) - \log 2\pi - \sum_{k=1}^{\infty} \tilde{\phi}_x(-2k).$$

12.3 A quantitative explicit formula

Let $1 \leq T \leq x$. Using the second expression for $\tilde{\phi}_x(s)$, namely

$$\tilde{\phi}_x(s) = \frac{x^s}{s} + yx^{s-1}w_s, \quad \text{where } |w_s| \leq \frac{1}{2} \text{ if } 0 \leq \sigma \leq 1,$$

we find that

$$\sum_{\rho} \tilde{\phi}_x(\rho) = \sum_{|\gamma| \leq T} \frac{x^{\rho}}{\rho} + y \sum_{|\gamma| \leq T} x^{\rho-1} w_{\rho} + \sum_{|\gamma| > T} \tilde{\phi}_x(\rho).$$

To estimate the second sum, fix a function $\beta_0(T)$ such that

$$\max\{\beta : \zeta(\beta + i\gamma) = 0 \text{ and } \beta \in [0, 1], \gamma \in [-T, T]\} \leq \beta_0(T)$$

for all $T \geq 1$. Then our zero-free region tells us that we can take $\beta_0(T) = 1 - c/\log T$, where c is the constant from Theorem 10.1. Under the Riemann Hypothesis, we can take $\beta_0(T) = 1/2$. Since $\operatorname{Re}(\rho) \in [0, 1]$ we have

$$\sum_{|\gamma| \leq T} |x^{\rho-1} w_{\rho}| \leq x^{\beta_0(T)-1} N(T) \ll x^{\beta_0(T)-1} T \log T.$$

For the third sum we use the first expression for $\tilde{\phi}_x(s)$:

$$\tilde{\phi}_x(s) = \frac{x^{s+1}}{ys(s+1)} \left(\left(1 + \frac{y}{x}\right)^{s+1} - 1 \right).$$

Since $\operatorname{Re} \rho \leq 1$ and $y \leq x$ we have

$$\left| \left(1 + \frac{y}{x}\right)^{\rho+1} - 1 \right| \leq 2^{\operatorname{Re} \rho + 1} + 1 \ll 1.$$

Now fix a function $\beta_1(T)$ such that

$$\sup\{\beta : \zeta(\beta + i\gamma) = 0 \text{ and } \beta \in [0, 1], |\gamma| \geq T\} \leq \beta_1(T)$$

for all $T \geq 1$. Our zero free region doesn't help here, so we have to take $\beta_1(T) = 1$. Under RH we can take $\beta_1(T) = \beta_0(T) = 1/2$. Then

$$\sum_{|\gamma| > T} \tilde{\phi}_x(\rho) \ll \frac{x^{\beta_1(T)+1}}{y} \sum_{|\gamma| > T} \frac{1}{|\rho(\rho+1)|} \ll \frac{x^{\beta_1(T)+1}}{y} \sum_{|\gamma| > T} \frac{1}{|\gamma|^2}.$$

By partial summation, we have

$$\sum_{T < \gamma < Z} \frac{1}{\gamma^2} = \frac{N(Z)}{Z^2} - \frac{N(T)}{T^2} + 2 \int_T^Z \frac{N(t)}{t^3} dt,$$

and using that $N(t) \ll t \log t$, this implies that

$$\sum_{\gamma > T} \frac{1}{\gamma^2} \ll \frac{\log T}{T} + \int_T^{\infty} \frac{\log t}{t^2} dt \ll \frac{\log T}{T}.$$

Collecting all of the above, we conclude that

$$\sum_{\rho} \tilde{\phi}_x(\rho) = \sum_{|\gamma| \leq T} \frac{x^{\rho}}{\rho} + O\left(yx^{\beta_0(T)-1} T \log T + \frac{x^{\beta_1(T)+1} \log T}{yT}\right). \quad (12.3)$$

We can now pick y to balance the two error terms above, the optimal choice being

$$y = \frac{1}{T} x^{1+\frac{1}{2}(\beta_1(T)-\beta_0(T))}. \quad (12.4)$$

Using our zero-free region we have $\beta_1(T) - \beta_0(T) = c(\log T)^{-1}$, so to ensure that $y \leq x$ we need to impose the restriction

$$(\log T)^2 \geq \frac{1}{2} c \log x.$$

As we'll see soon, the optimal choice of T is very close to this bound.

When we combine Proposition 12.2 with (18.1), (12.3), and (12.4) we arrive at the second form of the explicit formula.

Theorem 12.3. *Let $\beta_0(T)$ and $\beta_1(T)$ be functions chosen as above for all $T \geq 1$. Then for $x \geq 2$ and T satisfying*

$$x^{\frac{1}{2}(\beta_1(T)-\beta_0(T))} \leq T \leq x$$

we have

$$\sum_{n \leq x} \Lambda(n) = x - \sum_{|\gamma| \leq T} \frac{x^\rho}{\rho} + O\left(x^{\frac{1}{2}(\beta_0(T)+\beta_1(T))} \log T + \frac{\log x}{T} x^{1+\frac{1}{2}(\beta_1(T)-\beta_0(T))}\right).$$

12.4 The Prime Number Theorem for $\psi(x)$

We'll prove the prime number theorem for $\psi(x)$ as a corollary to the explicit formula.

Theorem 12.4 (PNT for $\psi(x)$). *There exists a constant $\alpha > 0$ such that for all $x \geq 2$,*

$$\psi(x) = \sum_{n \leq x} \Lambda(n) = x + O\left(x \exp\left(-\alpha \sqrt{\log x}\right)\right).$$

Proof. By (11.3) we have

$$\sum_{|\gamma| \leq T} \frac{x^\rho}{\rho} \ll x^{\beta_0(T)} \sum_{|\gamma| \leq T} \frac{1}{|\rho|} \ll x^{\beta_0(T)} \log^2 T.$$

We choose β_0 and β_1 corresponding to our zero-free region, namely

$$\beta_0(T) = 1 - \frac{c}{\log T}, \quad \beta_1(T) = 1.$$

Then the sum and the error term in Theorem 12.3 together are (using that $T \leq x$)

$$\begin{aligned} &\ll x^{1-\frac{c}{\log T}} \log^2 T + x^{1-\frac{c}{2\log T}} \log T + \frac{\log x}{T} x^{1+\frac{c}{2\log T}} \\ &= x \log x \left\{ \log x \exp\left(\frac{-c \log x}{\log T}\right) + \exp\left(\frac{-c \log x}{2\log T}\right) + \frac{1}{T} \exp\left(\frac{c \log x}{2\log T}\right) \right\}. \end{aligned}$$

We can balance the second and third terms by choosing

$$T = \exp(\sqrt{c \log x}).$$

(For some perspective, $\log^A x \ll \exp(\sqrt{c \log x}) \ll x^\epsilon$ for any $A > 0$ and $\epsilon > 0$.) This yields a total error term of

$$x \exp(-\sqrt{c \log x}) \log^2 x \ll x \exp(-\alpha \sqrt{\log x})$$

for some $\alpha > 0$. □

As you might expect, we get a much better result by assuming RH.

Theorem 12.5. *Assume the Riemann Hypothesis. Then for every $x \geq 2$ we have*

$$\psi(x) = x + O(\sqrt{x} \log^2 x).$$

Proof. Assuming RH we can take $\beta_0(T) = \beta_1(T) = 1/2$. Then the error terms add up to

$$\sqrt{x} \log^2 T + \frac{x \log x}{T}.$$

Choosing any $T \geq \sqrt{x}/\log x$, we obtain the theorem. □

The really amazing thing is that the converse of this theorem is true; i.e. a good bound for the error term in the prime number theorem yields the Riemann hypothesis.

Theorem 12.6. *Suppose that for some $\alpha < 1$ we have*

$$\psi(x) = x + O(x^\alpha).$$

Then all of the zeros $\rho = \beta + i\gamma$ of $\zeta(s)$ satisfy $\beta \leq \alpha$.

Proof. Let $E(x)$ denote the error term: $\psi(x) = x + E(x)$. By partial summation we have (for $\sigma > 1$)

$$-\frac{\zeta'}{\zeta}(s) = \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s} = s \int_1^{\infty} \frac{\psi(t)}{t^{s+1}} dt = s \int_1^{\infty} \frac{dt}{t^s} + s \int_1^{\infty} \frac{E(t)}{t^{s+1}} dt.$$

The first term evaluates to $s/(s-1)$; not surprising since this corresponds to the pole of $\zeta(s)$ at $s = 1$. The remaining integral is analytic in the region $\sigma > \alpha$ since $E(t) \ll t^\alpha$. It follows that the only pole of ζ'/ζ to the right of $\sigma = \alpha$ is at $s = 1$, and therefore $\zeta(s)$ has no zeros in that region (otherwise, ζ'/ζ would have another pole). □

One amusing consequence of the previous theorem is that a bound of the form $\psi(x) = x + O(x^{\theta+\epsilon})$ for any $\epsilon > 0$ is automatically self-improving: it follows that the zeros are all to the left of $\sigma = \theta$, and by the same argument as in the proof of Theorem 12.5, we obtain

$$\psi(x) = x + O(x^\theta \log^2 x).$$

So the term x^ϵ is automatically replaced with the much more precise $\log^2 x$.

Part II

Primes in Arithmetic Progressions

We turn our attention to the function

$$\pi(x; a \bmod q) := \#\{p \leq x : p \equiv a \bmod q\}.$$

Unlike in the case of $\pi(x)$, it is not at all obvious that $\pi(x; a \bmod q)$ tends to ∞ as $x \rightarrow \infty$. Actually, in some cases it's not true at all. For example, there are no primes congruent to $0 \bmod 4$ and there is only one prime congruent to $2 \bmod 4$. In general, the number of primes $p \equiv a \bmod q$ is finite whenever $\gcd(a, q) = 1$. With that “local” obstruction out of the way, we will prove the following result.

Theorem 12.7 (Dirichlet's Theorem). *If $\gcd(a, q) = 1$ then*

$$\lim_{x \rightarrow \infty} \pi(x; a \bmod q) = \infty.$$

Actually, we will prove this as a corollary to a stronger result which is an analogue of Mertens' estimate for $\sum 1/p$. What do we expect is true? Consider the case $q = 4$. Then there are two choices for a , namely $a = 1$ and $a = 3$. There isn't any immediate reason to suspect that primes are more likely to be in one progression than the other, so we might guess that “half” of the primes satisfy $p \equiv 1 \bmod 4$, for some suitable definition of “half.” Indeed, this is the case. For each q , define Euler's totient function by

$$\varphi(q) = \#\{0 \leq a \leq q - 1 : \gcd(a, q) = 1\}.$$

We often shorten notation to $(a, q) := \gcd(a, q)$. We will prove the following.

Theorem 12.8. *If $\gcd(a, q) = 1$ then*

$$\sum_{\substack{p \leq x \\ p \equiv a(q)}} \frac{1}{p} = \frac{1}{\varphi(q)} \log \log x + O_q(1).$$

The notation $O_q(1)$ means that the implied constant is allowed to depend on q . So the result above is interesting if q is fixed. If we want to know what happens as we vary q and x , we will have to work much harder (and things get quite interesting).

The key to proving both of these theorems is a careful study of the Dirichlet L -functions

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

which Dirichlet invented for the purpose of proving his theorem.

13 Dirichlet Characters

A Dirichlet character modulo q is a function $\chi : \mathbb{Z} \rightarrow \mathbb{C}$ satisfying

- i) periodicity modulo q : $\chi(n + q) = \chi(n)$ for all $n \in \mathbb{Z}$,
- ii) complete multiplicativity: $\chi(1) = 1$ and $\chi(mn) = \chi(m)\chi(n)$ for all $m, n \in \mathbb{Z}$,
- iii) $\chi(n) = 0$ whenever $\gcd(n, q) > 1$.

The character

$$\chi_0(n) = \begin{cases} 1 & \text{if } \gcd(n, q) = 1, \\ 0 & \text{otherwise,} \end{cases}$$

is called the principal (or trivial) character modulo q .

Examples:

1. If $q = 1$ there is only one character,

$$\chi_0(n) = 1 \quad \text{for all } n.$$

2. If $q = 2$ there is also only one character,

$$\chi_0(n) = \begin{cases} 1 & \text{if } n \text{ is odd,} \\ 0 & \text{if } n \text{ is even.} \end{cases}$$

3. If $q = 3$, there are two characters; the trivial character

$$\chi_0(n) = \begin{cases} 1 & \text{if } n \equiv 0 \pmod{3}, \\ 0 & \text{otherwise,} \end{cases}$$

and the nontrivial character

$$\chi_1(n) = \begin{cases} 1 & \text{if } n \equiv 1 \pmod{3}, \\ -1 & \text{if } n \equiv 2 \pmod{3}, \\ 0 & \text{otherwise.} \end{cases}$$

Since every n with $(n, q) = 1$ has an inverse $\bar{n}$ such that $n\bar{n} \equiv 1 \pmod{q}$, condition (iii) can be strengthened to “ $\chi(n) = 0$ if and only if $\gcd(n, q) > 1$.” Thus χ descends to a group homomorphism

$$\chi : (\mathbb{Z}/q\mathbb{Z})^\times \rightarrow \mathbb{C}^\times.$$

The elements of $(\mathbb{Z}/q\mathbb{Z})^\times$ are residue classes $[n]$, where $n \in \mathbb{Z}$ is a representative. For $[n] \in (\mathbb{Z}/q\mathbb{Z})^\times$, we have $[n]^{\varphi(q)} = [1]$ since $\phi(q)$ is the order of $(\mathbb{Z}/q\mathbb{Z})^\times$. From the complete multiplicativity, it follows that

$$\chi(n)^{\varphi(q)} = \chi(n^{\varphi(q)}) = \chi(1) = 1,$$

so χ takes values in the $\varphi(q)$ -th roots of unity $\mu_{\varphi(q)}$, i.e.

$$\chi : (\mathbb{Z}/q\mathbb{Z})^\times \rightarrow \mu_{\varphi(q)} := \left\{ e\left(\frac{m}{\varphi(q)}\right) : m \bmod \varphi(q) \right\}.$$

13.1 The dual group

The characters themselves form a group, called the dual group, via the multiplication

$$(\chi_1\chi_2)(n) := \chi_1(n)\chi_2(n).$$

The identity element of the group is χ_0 and the inverse of an element χ is given by

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

We often write $\chi \bmod q$ to denote the set of characters modulo q (e.g. in a summation).

Lemma 13.1. *There are exactly $\varphi(q)$ characters modulo q . Furthermore, for every a with $(a, q) = 1$ and $a \not\equiv 1 \pmod{q}$ there is a character χ with $\chi(a) \neq 1$.*

Proof. By the structure theorem for finite abelian groups, we have

$$G := (\mathbb{Z}/q\mathbb{Z})^\times \cong G_1 \times G_2 \times \cdots \times G_r,$$

where each G_k is a cyclic group of order m_k and $m_1 m_2 \cdots m_r = \varphi(q) =: m$. For each k , fix a generator $[g_k]$ of G_k . So if $[n] \in (\mathbb{Z}/q\mathbb{Z})^\times$, there are unique exponents a_k with $0 \leq a_k \leq m_k - 1$ such that $[n] = g_1^{a_1} \cdots g_r^{a_r}$. We can think of G_k as $\mathbb{Z}/m_k\mathbb{Z}$ via the isomorphism

$$g_k^{a_k} \mapsto a_k \pmod{m_k}.$$

For each k , choose a primitive m_k -th root of unity ζ_k , for example $\zeta_k = e(1/m_k)$.

Let $\widehat{G} = \text{Hom}(G, \mu_m)$ denote the character group of G , i.e. the group of homomorphisms from G to the roots of unity μ_m . Define a map $\phi : G \rightarrow \widehat{G}$ via

$$\phi(g_1^{a_1} \cdots g_r^{a_r}) = \left((g_1^{b_1}, \dots, g_r^{b_r}) \mapsto \zeta_1^{a_1 b_1} \cdots \zeta_r^{a_r b_r} \right).$$

You can check that $\phi(gh) = \phi(g)\phi(h)$ so ϕ is a homomorphism.

To construct the inverse of ϕ , let $\varepsilon_k : \mu_{m_k} \rightarrow \mathbb{Z}/m_k\mathbb{Z}$ be the unique isomorphism that sends the primitive m_k -th root of unity ζ_k to $1 + m_k\mathbb{Z}$. For $\eta \in \widehat{G}$, define

$$\psi(\eta) = g_1^{\varepsilon_1(\eta(g_1))} \cdots g_r^{\varepsilon_r(\eta(g_r))}.$$

For $\eta_1, \eta_2 \in \widehat{G}$, we have $\varepsilon_k(\eta_1 \eta_2(g_k)) = \varepsilon_k(\eta_1(g_k) \eta_2(g_k)) = \varepsilon_k(\eta_1(g_k)) + \varepsilon_k(\eta_2(g_k))$, from which it follows that ψ is a homomorphism. We compute

$$\psi(\phi(g_1^{a_1} \cdots g_r^{a_r})) = g_1^{\varepsilon_1(\zeta_1^{a_1})} \cdots g_r^{\varepsilon_r(\zeta_r^{a_r})} = g_1^{a_1} \cdots g_r^{a_r},$$

so ψ and ϕ are inverses. It follows that $G \cong \widehat{G}$, and our construction of the isomorphism shows that there is a character χ with $\chi(a) \neq 1$ associated to each $a \not\equiv 1 \pmod{q}$. $\square$

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
χ_0	0	1	1	0	1	0	0	1	1	0	0	1	0	1	1
χ_1	0	1	i	0	-1	0	0	i	$-i$	0	0	1	0	$-i$	-1
χ_2	0	1	-1	0	1	0	0	-1	-1	0	0	1	0	-1	1
χ_3	0	1	$-i$	0	-1	0	0	$-i$	i	0	0	1	0	i	-1
χ_4	0	1	-1	0	1	0	0	1	-1	0	0	-1	0	1	-1
χ_5	0	1	$-i$	0	-1	0	0	i	i	0	0	-1	0	$-i$	1
χ_6	0	1	1	0	1	0	0	-1	1	0	0	-1	0	-1	-1
χ_7	0	1	i	0	-1	0	0	$-i$	$-i$	0	0	-1	0	i	1

Figure 4: Dirichlet characters modulo 15

13.2 Orthogonality

Arguably the most important property of Dirichlet characters is orthogonality. We begin with a lemma regarding complete sums of characters. We will often use the shorthand “ $a \bmod q$ ” or just “ $a(q)$ ” to denote any complete residue system modulo q (when the choice of representatives does not matter). This can always be replaced by the set $\{0, 1, \dots, q-1\}$ if that is more convenient. We will also write $\chi \bmod q$ to denote the set of all Dirichlet characters modulo q .

Lemma 13.2. (i) *If χ is a Dirichlet character modulo q then*

$$\sum_{a \bmod q} \chi(a) = \begin{cases} \varphi(q) & \text{if } \chi = \chi_0, \\ 0 & \text{otherwise.} \end{cases}$$

(ii) *If a is an integer then*

$$\sum_{\chi \bmod q} \chi(a) = \begin{cases} \varphi(q) & \text{if } a \equiv 1 \pmod{q}, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. (i) If $\chi = \chi_0$ then the result is clear. Suppose that $\chi \neq \chi_0$. Then there is an integer b such that $(b, q) = 1$ and $\chi(b) \neq 1$. Let S denote the sum on the left-hand side. Then

$$\chi(b)S = \sum_{a \bmod q} \chi(b)\chi(a) = \sum_{a \bmod q} \chi(ab) = \sum_{c \bmod q} \chi(c) = S$$

because as a runs through a set of representatives mod q , so does ab (observe that a is invertible mod q). But the equation

$$\chi(b)S = S \iff (\chi(b) - 1)S = 0$$

implies that $S = 0$ since $\chi(b) \neq 1$.

(ii) The only nontrivial case is when $(a, q) = 1$. In that case there is a character ψ such that $\psi(a) \neq 1$. By a similar argument as above,

$$\psi(a)S = \sum_{\chi \bmod q} \chi(a)\psi(a) = \sum_{\chi \bmod q} (\chi\psi)(a) = \sum_{\eta \bmod q} \eta(a) = S,$$

which implies that $S = 0$. □

As an immediate corollary, we can estimate the summatory function of χ .

Lemma 13.3. (i) If $\chi \neq \chi_0$ is a Dirichlet character modulo q then

$$\left| \sum_{n \leq x} \chi(n) \right| \leq \varphi(q).$$

(ii) If $\chi = \chi_0 \pmod{q}$ then

$$\left| \sum_{n \leq x} \chi(n) - \frac{\varphi(q)}{q} x \right| \leq 2\varphi(q).$$

Proof. In either case we have

$$\sum_{n \leq x} \chi(n) = \left\lfloor \frac{x}{q} \right\rfloor \sum_{a=1}^q \chi(a) + R(x), \quad (13.1)$$

where

$$|R(x)| \leq \sum_{a \pmod{q}} |\chi(a)| \leq \varphi(q).$$

In the first case, the sum in (13.1) is zero; in the second case we have

$$\left| \left\lfloor \frac{x}{q} \right\rfloor \sum_{a=1}^q \chi(a) - \frac{x}{q} \sum_{a=1}^q \chi(a) \right| \leq \sum_{a \pmod{q}} |\chi(a)| \leq \varphi(q),$$

as desired. $\square$

The next proposition will provide us with a useful way to encode the condition $p \equiv a \pmod{q}$ in the proof of Dirichlet's theorem.

Proposition 13.4. (i) If χ_1 and χ_2 are two Dirichlet characters modulo q then

$$\sum_{a \pmod{q}} \chi_1(a) \overline{\chi_2}(a) = \begin{cases} \varphi(q) & \text{if } \chi_1 = \chi_2, \\ 0 & \text{otherwise.} \end{cases}$$

(ii) If a_1 and a_2 are integers then

$$\sum_{\chi \pmod{q}} \chi(a_1) \overline{\chi}(a_2) = \begin{cases} \varphi(q) & \text{if } a_1 \equiv a_2 \pmod{q}, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. (i) We apply (i) of Lemma 13.2 with $\chi = \chi_1 \overline{\chi_2}$ and observe that $\chi_1 = \chi_2$ if and only if $\chi_1 \overline{\chi_2} = \chi_0$.

(ii) The only nontrivial case is $(a_1, q) = (a_2, q) = 1$. In that case, a_2 has an inverse modulo q , say $\bar{a}_2$; i.e. $a_2 \bar{a}_2 \equiv 1 \pmod{q}$. Then

$$1 = \chi(a_2 \bar{a}_2) = \chi(a_2) \chi(\bar{a}_2),$$

so $\chi(\bar{a}_2) = \overline{\chi}(a_2)$. Hence $\chi(a_1) \overline{\chi}(a_2) = \chi(a_1) \chi(\bar{a}_2) = \chi(a_1 \bar{a}_2)$. We now apply (i) of Lemma 13.2 with $a = a_1 \bar{a}_2$, observing that the condition $a_1 \bar{a}_2 \equiv 1 \pmod{q}$ is equivalent to $a_1 \equiv a_2 \pmod{q}$. $\square$

14 Dirichlet L -functions and Dirichlet's Theorem

In analogy with the Riemann zeta function, for each Dirichlet character χ we define the Dirichlet L -function

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

Since $|\chi(n)| \leq 1$ for all n , we see (by the same proof as in the $\zeta(s)$ case) that $L(s, \chi)$ defines an analytic function in the region $\sigma > 1$. However, for $\chi \neq \chi_0$ it turns out that the Dirichlet series for $L(s, \chi)$ is conditionally convergent in the larger region $\sigma > 0$. Furthermore, the analytic continuation of $L(s, \chi)$ to $\operatorname{Re}(s) > 0$ only has a pole if $\chi = \chi_0$.

14.1 Basic properties of $L(s, \chi)$

Theorem 14.1. *Suppose that χ is a Dirichlet character modulo q .*

- (i) *if $\chi \neq \chi_0$ then $L(s, \chi)$ is analytic in $\sigma > 0$.*
- (ii) *If $\chi = \chi_0$ then $L(s, \chi)$ has a meromorphic continuation to $\sigma > 0$ with only a simple pole at $s = 1$ with residue $\varphi(q)/q$.*

Proof. (i) By partial summation we have

$$\sum_{n \leq x} \frac{\chi(n)}{n^s} = \frac{1}{x^s} \sum_{n \leq x} \chi(n) + s \int_1^x \left(\sum_{n \leq t} \chi(n) \right) \frac{dt}{t^{s+1}}.$$

By Lemma 13.3 the partial sums $\sum_{n \leq x} \chi(n)$ are $O(1)$, so everything above is convergent for $\sigma > 0$ as $x \rightarrow \infty$. It follows that

$$\sum_{n \leq x} \frac{\chi(n)}{n^s} = s \int_1^{\infty} \left(\sum_{n \leq t} \chi(n) \right) \frac{dt}{t^{s+1}},$$

which is easily seen to be analytic in $\sigma > 0$.

(ii) If $\chi = \chi_0$, let $f(n) = \chi_0(n) - \varphi(q)/q$. Then by Lemma 13.3 we have

$$\sum_{n \leq x} f(n) = \sum_{n \leq x} \chi_0(n) - \frac{\varphi(q)}{q} \sum_{n \leq x} 1 \ll 1,$$

so by the same argument as in (i), the Dirichlet series

$$\sum_{n=1}^{\infty} \frac{f(n)}{n^s} = L(s, \chi_0) - \frac{\varphi(q)}{q} \zeta(s)$$

is analytic in $\sigma > 1$. It follows that $L(s, \chi_0)$ is meromorphic in $\sigma > 1$ with a simple pole at $s = 1$ with residue $\varphi(q)/q$. $\square$

In the region of absolute convergence, each $L(s, \chi)$ has an Euler product of the form

$$L(s, \chi) = \prod_p \left(1 - \frac{\chi(p)}{p^s} \right)^{-1}. \quad (14.1)$$

The convergence is established in the same way as it was for $\zeta(s)$, so we only need to show that the computation works formally. Expanding each term in the product as a geometric series we find that

$$\prod_p \left(1 - \frac{\chi(p)}{p^s} \right)^{-1} = \prod_p \left(1 + \frac{\chi(p)}{p^s} + \frac{\chi(p)^2}{p^{2s}} + \cdots \right).$$

The complete multiplicativity of χ gives $\chi(p)^k = \chi(p^k)$, so the k -th term in the sum above looks like $(\chi(p^k)/p^k)^s$. Expanding out the product and rearranging terms yields

$$\sum_{n=1}^{\infty} \frac{1}{n^s} \prod_{p^a \parallel n} \chi(p^a) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s},$$

again using the complete multiplicativity of χ . Here the notation $p^a \parallel n$ means that p^a is the highest power of p dividing n (so while $|$ means “divides,” the symbol $\parallel$ means “exactly divides”). In the case of the principal character χ_0 , we see from the Euler product that $L(s, \chi_0)$ is quite closely related to $\zeta(s)$:

$$\begin{aligned} L(s, \chi_0) &= \prod_p \left(1 - \frac{\chi_0(p)}{p^s} \right)^{-1} = \prod_{p \nmid q} \left(1 - \frac{1}{p^s} \right)^{-1} \\ &= \prod_{p \mid q} \left(1 - \frac{1}{p^s} \right) \prod_p \left(1 - \frac{1}{p^s} \right)^{-1} = \prod_{p \mid q} \left(1 - \frac{1}{p^s} \right) \zeta(s). \end{aligned}$$

Note that the last product is finite. Actually, since the residue of $L(s, \chi_0)$ at $s = 1$ equals $\varphi(q)/q$, we have

$$\prod_{p \mid q} \left(1 - \frac{1}{p} \right) = \frac{\varphi(q)}{q}.$$

Although this is somewhat out of place, I can't resist stating an immediate corollary of this fact. This elementary fact can be proven in an elementary way, but since it falls right out of the product formula above, we might as well state it here.

Lemma 14.2. *The function $\varphi(q)$ is multiplicative; that is, we have $\varphi(mn) = \varphi(m)\varphi(n)$ whenever $(m, n) = 1$. Furthermore, on prime powers we have $\varphi(p^m) = p^m - p^{m-1}$.*

14.2 Dirichlet's Theorem

Our goal is to obtain an asymptotic formula for the sum

$$\sum_{\substack{p \leq x \\ p \equiv a(q)}} \frac{1}{p}.$$

We will approximate the partial sum above by the Dirichlet series

$$\sum_{p \equiv a(q)} \frac{1}{p^s}$$

evaluated at the special point

$$s = \sigma_x = 1 + \frac{1}{\log x}.$$

Note that $x^{\sigma_x} = ex$. By partial summation

$$\begin{aligned} \sum_{\substack{p \equiv a(q) \\ p > x}} \frac{1}{p^{\sigma_x}} &= -\frac{\pi(x)}{x^{\sigma_x}} + \sigma_x \int_x^\infty \frac{\pi(t)}{t^{\sigma_x+1}} dx \ll \frac{1}{\log x} + \int_x^\infty \frac{dt}{t^{\sigma_x} \log t} \\ &\ll \frac{1}{\log x} + \frac{x^{1-\sigma_x}}{(\sigma_x - 1) \log x} \ll 1. \end{aligned}$$

Note that here we have used the much-weaker-than-PNT statement $\pi(x) \ll x/\log x$, which follows from $\vartheta(x) \ll x$ which we proved in an elementary way. Next,

$$\sum_{\substack{p \leq x \\ p \equiv a(q)}} \frac{1}{p} - \sum_{\substack{p \leq x \\ p \equiv a(q)}} \frac{1}{p^{\sigma_x}} = \sum_{\substack{p \leq x \\ p \equiv a(q)}} \frac{1}{p} (1 - p^{1-\sigma_x}).$$

Observe that $p^{\sigma_x-1} = \exp\{-\frac{\log p}{\log x}\}$. Expanding the Taylor series for $1 - e^{-z}$, we find that

$$1 - p^{1-\sigma_x} = 1 - e^{-\frac{\log p}{\log x}} \ll \frac{\log p}{\log x}.$$

Hence

$$\sum_{\substack{p \leq x \\ p \equiv a(q)}} \frac{1}{p} - \sum_{\substack{p \leq x \\ p \equiv a(q)}} \frac{1}{p^{\sigma_x}} \ll \frac{1}{\log x} \sum_{\substack{p \leq x \\ p \equiv a(q)}} \frac{\log p}{p} \ll 1$$

by Mertens' first theorem. Therefore

$$\sum_{\substack{p \leq x \\ p \equiv a(q)}} \frac{1}{p} = \sum_{p \equiv a(q)} \frac{1}{p^{\sigma_x}} + O(1). \quad (14.2)$$

Now, back to $L(s, \chi)$. Taking logs in (14.1), we find that

$$\log L(s, \chi) = - \sum_p \log \left(1 - \frac{\chi(p)}{p^s} \right) = \sum_p \sum_{m=1}^\infty \frac{\chi(p)}{mp^{ms}} = \sum_p \frac{\chi(p)}{p^s} + O(1),$$

where, as in the $\zeta(s)$ case, we have used that

$$\sum_p \sum_{m=2}^\infty \frac{\chi(p)}{mp^{ms}} \ll \sum_p \sum_{m=2}^\infty \frac{1}{mp^m} \ll \sum_p \frac{1}{p^2} \ll 1.$$

We can detect the condition $p \equiv a \pmod{q}$ using orthogonality of Dirichlet characters:

$$\begin{aligned} \sum_{p \equiv a(q)} \frac{1}{p^s} &= \frac{1}{\varphi(q)} \sum_p \frac{1}{p^s} \sum_{\chi \bmod q} \chi(p) \bar{\chi}(a) = \frac{1}{\varphi(q)} \sum_{\chi \bmod q} \bar{\chi}(a) \sum_p \frac{\chi(p)}{p^s} \\ &= \frac{1}{\varphi(q)} \sum_{\chi \bmod q} \bar{\chi}(a) \log L(s, \chi) + O_q(1). \end{aligned}$$

We expect the main term to come from the principal character, since $L(s, \chi_0)$ has a pole at $s = 1$. However, we run into trouble if any of the other $L(s, \chi)$ is zero at $s = 1$ since this would also make a large contribution to the sum. We will prove the following theorem in the next section. It represents, by a large margin, the most difficult part of the proof of Dirichlet's theorem.

Theorem 14.3. *If $\chi \neq \chi_0$ then $L(1, \chi) \neq 0$.*

As $\sigma \rightarrow 1^+$ along the real line, we have $L(\sigma, \chi_0) = \frac{\varphi(q)}{q}(\sigma - 1)^{-1} + O(1)$ because of the pole at $s = 1$. From that term we get (remember $(a, q) = 1$)

$$\begin{aligned} \bar{\chi}_0(a) \log L(\sigma, \chi_0) &= \log \left(\frac{\varphi(q)}{q} \cdot \frac{1}{(\sigma - 1)} + O(1) \right) \\ &= \log \left(\frac{\varphi(q)}{q} \cdot \frac{1}{(\sigma - 1)} \right) + \frac{1}{\varphi(q)} \log(1 + O_q(\sigma - 1)) \\ &= \log \left(\frac{1}{\sigma - 1} \right) + O_q(1). \end{aligned}$$

Each of the nonprincipal characters contribute $O_q(1)$ since each $L(s, \chi)$ is continuous and nonzero at $s = 1$. It follows that

$$\sum_{p \equiv a(q)} \frac{1}{p^\sigma} = \frac{1}{\varphi(q)} \log \left(\frac{1}{\sigma - 1} \right) + O_q(1).$$

Putting this together with (14.2) and recalling $\sigma_x = 1 + 1/\log x$, we get Dirichlet's Theorem.

Theorem 14.4. *If $(a, q) = 1$ then*

$$\sum_{\substack{p \leq x \\ p \equiv a(q)}} \frac{1}{p} = \frac{1}{\varphi(q)} \log \log x + O_q(1).$$

15 The nonvanishing of $L(1, \chi)$

It remains to prove Theorem 14.3, that $L(1, \chi) \neq 0$. We will prove this in two cases depending on whether χ is complex or real. A real character is one that takes only real values; since these values must also be roots of unity or zero, the image of a real character is a subset of $\{-1, 0, 1\}$. A complex character is any character that is not real.

An important property of complex characters χ is that $\chi \neq \bar{\chi}$ since there is some $a \in \mathbb{Z}$ for which $\chi(a) \notin \mathbb{R}$. Since $L(s, \chi) = \overline{L(\bar{s}, \bar{\chi})}$, it follows that if $L(1, \chi) = 0$ for some complex character χ , then $L(1, \bar{\chi}) = 0$ as well. Then the product

$$P(s) = \prod_{\chi} L(s, \chi)$$

has a zero of order ≥ 1 at $s = 1$ (simple pole from $L(s, \chi_0)$ and double zero from $L(s, \chi)$ and $L(s, \bar{\chi})$). This contradicts the following lemma, and thus Theorem 14.3 is true for complex characters.

Lemma 15.1. *With $P(s)$ as above, we have*

$$P(\sigma) \geq 1 \quad \text{for } \sigma > 1.$$

Proof. In the region $\sigma > 1$ we can use the Dirichlet series:

$$\begin{aligned} \log P(\sigma) &= \sum_{\chi \bmod q} \log L(\sigma, \chi) = - \sum_{\chi \bmod q} \sum_p \log \left(1 - \frac{\chi(p)}{p^\sigma} \right) \\ &= \sum_{\chi \bmod q} \sum_p \sum_{m=1}^{\infty} \frac{\chi(p)^m}{p^{m\sigma}} = \sum_p \sum_{m=1}^{\infty} \frac{1}{p^{m\sigma}} \sum_{\chi \bmod q} \chi(p^m) \\ &= \varphi(q) \sum_{m=1}^{\infty} \sum_{\substack{p \\ p^m \equiv 1(q)}} \frac{1}{p^{m\sigma}}. \end{aligned}$$

Everything in this sum is nonnegative, so $\log P(\sigma) \geq 0$. □

15.1 Real characters

Suppose that χ is a real nonprincipal character such that $L(1, \chi) = 0$. Then the function $L(s, \chi)L(s, \chi_0)$ is analytic in the region $\sigma > 0$ (the zero of $L(1, \chi)$ cancels with the pole of $L(1, \chi_0)$). Thus the function

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

is analytic in the region $\sigma > 1/2$. Furthermore, we have $F(1/2) = 0$ since $L(2s, \chi_0)$ has a pole at $s = 1/2$.

In $\sigma > 1$, we can use the Euler products of $L(s, \chi)$ and $L(s, \chi_0)$ to get

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

since if $\chi(p) = -1$ we have

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

Expanding the geometric series, we find that

$$\frac{1 + p^{-s}}{1 - p^{-s}} = 1 + 2 \sum_{m=1}^{\infty} p^{-ms},$$

from which it follows that, still for $\sigma > 1$,

$$\prod_{\chi(p)=1} \frac{1 + p^{-s}}{1 - p^{-s}} = \sum_{n=1}^{\infty} \frac{a(n)}{n^s}$$

for some coefficients $a(n) \geq 0$, with $a(1) = 1$.

Now consider the Taylor expansion of $F(s)$ centered at $s = 2$. Since F is analytic in $\sigma > 1/2$, the radius of convergence of this Taylor expansion is at least $3/2$. We have

$$F(s) = \sum_{m=0}^{\infty} \frac{F^{(m)}(2)}{m!} (s - 2)^m.$$

The key here is that this series is valid for a portion of a disc which passes through the strip $1/2 < \sigma \leq 1$ where the Dirichlet series is not, a priori, valid. However the coefficients $F^{(m)}(2)$ can be computed from the Dirichlet series:

$$F^{(m)}(2) = \sum_{n=1}^{\infty} \frac{a(n)}{n^2} (-1)^m (\log n)^m = (-1)^m b(m),$$

say, where $b(m) \geq 0$. Note that $b(0) \geq 1$ since $b(0) = F(2) \geq a(1) \geq 1$. It follows that

$$F(s) = \sum_{m=0}^{\infty} \frac{b(m)}{m!} (2 - s)^m$$

for $|2 - s| < 3/2$. So if $\sigma \in (1/2, 2)$ we have

$$F(\sigma) \geq b(0) \geq 1$$

which contradicts the assertion above that $F(1/2) = 0$. This completes the proof of Theorem 14.3. $\square$

16 The functional equation of $L(s, \chi)$

We would like to prove a result stronger than Dirichlet's theorem, namely the prime number theorem for primes in arithmetic progressions. The outline of the proof is the same as in the case of the Riemann zeta function; i.e. we need to prove a functional equation and a zero-free region. Some of this is exactly the same as in the $\zeta(s)$ case, but some of it is quite different, often in very interesting ways. Here we will prove the functional equation for $L(s, \chi)$ when χ is a primitive character.

16.1 Primitive characters

Suppose that χ is a Dirichlet character of modulus q . Then we can construct another character χ' modulo qr by defining

$$\chi'(n) = \begin{cases} \chi(n) & \text{if } (n, qr) = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Note that if r is divisible by a prime p which does not divide q then $\chi \neq \chi'$; for example $\chi'(p) = 0$ while $\chi(p) \neq 0$. In this way we can construct infinitely many Dirichlet characters from the original character χ . But these new characters aren't very interesting because they essentially contain all the same information as χ . So we would like to distinguish these new characters and essentially forget about them.

Suppose that χ and χ' are characters of moduli q and q' , respectively, such that $q' \mid q$. If $\chi(n) = \chi'(n)$ for all n with $(n, q) = 1$ then we say that χ is *induced by* χ' . A *primitive character* is one that is not induced by any character (other than itself).

The principal character modulo q is only primitive if $q = 1$. Also, all nonprincipal characters of prime modulus are automatically primitive.

Lemma 16.1. *A Dirichlet character χ modulo q is induced by some χ' modulo q' if and only if $\chi(a) = \chi(b)$ whenever $(a, q) = (b, q) = 1$ and $a \equiv b \pmod{q'}$. When this holds, the character χ' is uniquely determined.*

Proof. ($\implies$) If χ is induced by χ' and $a \equiv b \pmod{q'}$ with $(a, q) = (b, q) = 1$ then

$$\chi(a) = \chi'(a) = \chi'(b) = \chi(b).$$

($\impliedby$) Suppose that $\chi(a) = \chi(b)$ whenever $(a, q) = (b, q) = 1$ and $a \equiv b \pmod{q'}$. We need to construct a character χ' modulo q which induces χ . It suffices to define $\chi'(n)$ for each n such that $(n, q') = 1$ and $(n, q) > 1$ since $\chi'(m)$ is defined as $\chi(m)$ whenever $(m, q) = 1$. For each such n let $n' = n + q'k$, where k is the product of all the primes dividing q/q' , excluding the primes dividing n . I claim that $(n', q) = 1$. Indeed, if $p \mid q'$ then since $(n, q') = 1$ we have $p \nmid n'$, and each prime $p \mid (q/q')$ divides one of n and k , but not both. We define $\chi'(n) = \chi(n')$. By our assumption on χ , this χ' is completely multiplicative and periodic modulo q' . It is also uniquely determined: if χ is induced by another character χ'' modulo q' then clearly $\chi''(n) = \chi'(n)$ for all n with $(n, q) = 1$. For the remaining n , let n'' be any integer with $n'' \equiv n \pmod{q'}$ and such that $(n'', q) = 1$. (Such an integer exists since n' is an example of one.) Then

$$\chi''(n) = \chi''(n'') = \chi(n'') = \chi'(n'') = \chi'(n).$$

So $\chi'' = \chi'$. □

Theorem 16.2. *Each Dirichlet character χ modulo q is induced by a primitive character χ' that is uniquely determined by χ . Furthermore,*

$$\chi(n) = \chi'(n)\chi_0(n) \tag{16.1}$$

for all integers n , where χ_0 is the principal character modulo q .

Proof. Among all divisors of q , let d be the smallest such that χ is induced by a character χ' of modulus d . By the previous lemma, this χ' is uniquely determined among characters of modulus d . Clearly (16.1) holds because (1) if $(n, q) = 1$ then $\chi(n) = \chi'(n)$ and $\chi_0(n) = 1$, and (2) if $(n, q) > 1$ then $\chi(n) = 0$ and $\chi_0(n) = 0$.

Suppose, by way of contradiction, that χ' is not primitive. Then there is a character χ'' of modulus $< d$ which induces χ . Suppose that $(n, q) = 1$. Then $(n, d) = 1$ also, so

$$\chi(n) = \chi'(n) = \chi''(n)$$

so χ is also induced by χ'' , contradicting the minimality of d . $\square$

We define the *conductor* of a character χ as the modulus of the unique primitive character that induces it (which exists by the previous theorem).

How are the L -functions associated to induced characters related? Suppose that χ of modulus q has conductor d . Then there is a primitive character χ' of modulus d such that $\chi = \chi'\chi_0$. Hence

$$\begin{aligned} L(s, \chi) &= \prod_p \left(1 - \frac{\chi(p)}{p^s}\right)^{-1} = \prod_p \left(1 - \frac{\chi'(p)\chi_0(p)}{p^s}\right)^{-1} \\ &= \prod_{p \nmid q} \left(1 - \frac{\chi'(p)}{p^s}\right)^{-1} = \prod_{p \nmid q} \left(1 - \frac{\chi'(p)}{p^s}\right) \prod_p \left(1 - \frac{\chi'(p)}{p^s}\right)^{-1} \\ &= \prod_{p \nmid q} \left(1 - \frac{\chi'(p)}{p^s}\right) L(s, \chi'). \end{aligned} \tag{16.2}$$

So, up to a well-understood finite product, a general Dirichlet L -function is essentially equal to the L -function associated to the primitive character inducing it. So we will not really lose anything by restricting ourselves to primitive characters from here on out.

16.2 Gauss sums

Suppose, throughout this subsection, that χ is a primitive character modulo q (thus, the conductor of χ is also q). In the functional equation for $L(s, \chi)$ we will encounter the Gauss sums, which are defined as

$$\tau(\chi) = \sum_{a \bmod q} \chi(a) e(a/q).$$

Recall the notation $e(x) = e^{2\pi i x}$. There are also more general Gauss sums defined as

$$G(n, \chi) = \sum_{a \bmod q} \chi(a) e(an/q).$$

These are related in the following way.

Lemma 16.3. *If $\chi \bmod q$ is a primitive Dirichlet character then for all $n \in \mathbb{Z}$ we have*

$$G(n, \chi) = \bar{\chi}(n) \tau(\chi).$$

Proof. Suppose first that $(n, q) = 1$. Then n has an inverse $\bar{n} \bmod q$ such that $n\bar{n} \equiv 1 \pmod{q}$. So as a runs through a complete set of residues modulo q , so does an (the map $a \mapsto an$ has an inverse, namely multiplication by $\bar{n}$). Thus

$$\tau(\chi) = \sum_{a \bmod q} \chi(a)e(a/q) = \sum_{a \bmod q} \chi(an)e(an/q) = \chi(n)G(n, \chi)$$

by the complete multiplicativity of χ .

It remains to show that $G(n, \chi) = 0$ when $(n, q) > 1$. Let $d = (n, q)$ and write $q = dq'$ and $n = dn'$. Then

$$G(n, \chi) = \sum_{b \bmod q'} \sum_{\substack{a \bmod q \\ a \equiv b \bmod q'}} \chi(a)e(adn'/dq') = \sum_{b \bmod q'} e(an'/q') \sum_{\substack{a \bmod q \\ a \equiv b \bmod q'}} \chi(a).$$

I claim that the inner sum always vanishes. If $(b, q') > 1$ then $\chi(a) = 0$ so the sum equals zero. Suppose that $(b, q') = 1$ and let $\bar{b}$ denote the inverse of $b \bmod q'$. Since χ is primitive, there exists a number $m \equiv b \pmod{q'}$ such that $\chi(m) \neq \chi(b)$ (otherwise χ would contradict Lemma 16.1). Replacing a by $am\bar{b}$ we find (since $m\bar{b} \equiv 1 \pmod{q'}$) that

$$\sum_{\substack{a \bmod q \\ a \equiv b \bmod q'}} \chi(a) = \chi(m)\bar{\chi}(b) \sum_{\substack{a \bmod q \\ a \equiv b \bmod q'}} \chi(a).$$

Since $\chi(m)\bar{\chi}(b) \neq 1$ the sum must equal zero. $\square$

The exact evaluation of the Gauss sums $\tau(\chi)$ is an interesting problem (and one that Gauss spent a lot of time on) but for our purposes we only need to know that $|\tau(\chi)| = \sqrt{q}$.

Lemma 16.4. *If $\chi \bmod q$ is a primitive Dirichlet character then*

$$|\tau(\chi)|^2 = q.$$

Proof. Expanding the left-hand side we find that

$$|\tau(\chi)|^2 = \tau(\chi)\overline{\tau(\chi)} = \tau(\chi) \sum_{a \bmod q} \bar{\chi}(a)e(-a/q).$$

By the previous lemma, this equals

$$\begin{aligned} \sum_{a \bmod q} G(a, \chi)e(-a/q) &= \sum_{a \bmod q} \sum_{b \bmod q} \chi(b)e(ab/q)e(-a/q) \\ &= \sum_{b \bmod q} \chi(b) \sum_{a \bmod q} e(a(b-1)/q). \end{aligned}$$

By expanding the finite geometric sum, we find that

$$\sum_{a \bmod q} e(am/q) = \begin{cases} q & \text{if } m \equiv 0 \pmod{q}, \\ \frac{e(m) - 1}{e(m/q) - 1} = 0 & \text{if } m \not\equiv 0 \pmod{q}. \end{cases}$$

We conclude that

$$|\tau(\chi)|^2 = q \sum_{\substack{b \bmod q \\ b \equiv 1 \bmod q}} \chi(b) = q$$

because there is only one term in the sum, namely $\chi(1) = 1$. $\square$

One nice consequence of the previous result is that $\tau(\chi) \neq 0$ when χ is primitive, so we can write

$$\chi(n) = \frac{1}{\tau(\bar{\chi})} \sum_{a \bmod q} \bar{\chi}(a) e(an/q), \quad (16.3)$$

which gives us a useful way of writing the multiplicative character χ in terms of additive characters $e(an/q)$.

16.3 The functional equation for primitive characters

Similar to we did with the Riemann zeta function, we begin with a version of Poisson summation, this time one which takes into account sums over residue classes modulo q .

Theorem 16.5. *(Same assumptions as in Theorem 5.1) Suppose that both f and $\hat{f}$ are in $L^1(\mathbb{R})$ and have bounded variation. If $u \in \mathbb{R}_+$ and $v \in \mathbb{R}$ then*

$$\sum_{m \in \mathbb{Z}} f(um + v) = \frac{1}{u} \sum_{n \in \mathbb{Z}} \hat{f}\left(\frac{n}{u}\right) e\left(\frac{nv}{u}\right).$$

Proof. Let $g(m) = f(um + v)$. Then a change of variables shows that

$$\hat{g}(n) = \frac{1}{u} \hat{f}\left(\frac{n}{u}\right) e\left(\frac{nv}{u}\right).$$

Applying Poisson summation to g proves the theorem. $\square$

Using this version of Poisson summation we can prove yet another version of Poisson summation which is “twisted” by the character χ .

Theorem 16.6. *With the same assumptions as in the previous theorem, and with χ a primitive Dirichlet character modulo q , we have*

$$\sum_{m \in \mathbb{Z}} \chi(m) f(m) = \frac{\tau(\chi)}{q} \sum_{n \in \mathbb{Z}} \bar{\chi}(n) \hat{f}\left(\frac{n}{q}\right).$$

Proof. Starting on the right-hand side, we apply (16.3) to the character $\bar{\chi}$ to get

$$\frac{\tau(\chi)}{q} \sum_{n \in \mathbb{Z}} \bar{\chi}(n) \hat{f}\left(\frac{n}{q}\right) = \sum_{a \bmod q} \chi(a) \frac{1}{q} \sum_{n \in \mathbb{Z}} \hat{f}\left(\frac{n}{q}\right) e\left(\frac{an}{q}\right) = \sum_{a \bmod q} \chi(a) \sum_{k \in \mathbb{Z}} f(qk + a)$$

by the previous theorem. Now reindex by $m = qk + a$ (this covers all integers as a runs mod q) and observe that $\chi(a) = \chi(m)$ because $m \equiv a \pmod{q}$. $\square$

We will use this to prove transformation laws for theta functions twisted by Dirichlet characters. The most obvious thing to do is define

$$\theta_\chi(z) = \sum_{n \in \mathbb{Z}} \chi(n) e^{\pi i n^2 z}.$$

Suppose, however, that $\chi(-1) = -1$ (observe that $\chi(-1) = \pm 1$ since $\chi^2(-1) = 1$). Then the n -th term of the series cancels with the $-n$ -th term, and we get $\theta_\chi(z) \equiv 0$. So the above definition will only work for $\chi(-1) = 1$; we have to do something different in the other case.

Characters satisfying $\chi(-1) = 1$ are called *even*, and those with $\chi(-1) = -1$ are *odd*.

16.3.1 Even characters

Suppose that $\chi(-1) = 1$. As before, we only need the transformation law for $z = iy$, so we'll restrict to that case. Recall that the Fourier transform of $f(m) = e^{-\pi m^2 y}$ is

$$\hat{f}(n) = y^{-1/2} e^{-\pi n^2 / y}.$$

It follows that

$$\theta_\chi(iy) = \sum_{m \in \mathbb{Z}} \chi(m) e^{-\pi i m^2 y} = y^{-1/2} \frac{\tau(\chi)}{q} \sum_{n \in \mathbb{Z}} \bar{\chi}(n) e^{-\pi n^2 / q^2 y} = \frac{\tau(\chi)}{\sqrt{q^2 y}} \theta_{\bar{\chi}}(i/q^2 y).$$

From this we can prove the functional equation of $L(s, \chi)$ for even characters χ . Since the $q = 1$ case is just $\zeta(s)$, we restrict ourselves to $q > 1$.

Theorem 16.7. *Suppose that χ is an even primitive character of conductor $q > 1$ and define*

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

Then $\xi(s, \chi)$ extends to an entire function of s and satisfies the functional equation

$$\varepsilon(\chi) \xi(1 - s, \bar{\chi}) = \xi(s, \chi),$$

where $\varepsilon(\chi) = \tau(\chi)/\sqrt{q}$ is a constant of absolute value 1 called the root number.

Proof. The proof proceeds as in the case of the Riemann zeta function (and is actually simpler, which is a reflection of the fact that $L(s, \chi)$ is analytic at $s = 1$ when χ is primitive and $q > 1$). Starting with the integral representation of $\Gamma(s/2)$, we have

$$\chi(n) (q/\pi)^{s/2} \Gamma(s/2) n^{-s} = \int_0^\infty \chi(n) t^{s/2-1} e^{-\pi n^2 t/q} dt.$$

Summing over n , and using that $\chi(-n) = \chi(n)$ and $\chi(0) = 0$, we get

$$\xi(s, \chi) = (q/\pi)^{s/2} \Gamma(s/2) L(s, \chi) = \frac{1}{2} \int_0^\infty \theta_\chi(it/q) t^{s/2-1} dt.$$

As before, we split the integral at $t = 1$ and use the transformation law on the first piece:

$$\frac{1}{2} \int_0^1 \theta_\chi(it/q) t^{s/2-1} dt = \frac{\tau(\chi)}{2\sqrt{q}} \int_0^1 \theta_{\bar{\chi}}(i/qt) t^{s/2-3/2} dt = \frac{\tau(\chi)}{2\sqrt{q}} \int_1^\infty \theta_{\bar{\chi}}(it/q) t^{-1/2-s/2} dt.$$

It follows that

$$\xi(s, \chi) = \frac{1}{2} \int_1^\infty \left(\theta_\chi(it/q) t^{s/2-1} + \frac{\tau(\chi)}{\sqrt{q}} \theta_{\bar{\chi}}(it/q) t^{-s/2-1/2} \right) dt.$$

The rapid decay of $\theta_\chi(it)$ shows that $\xi(s, \chi)$ defined in this way is an entire function of $\mathbb{C}$. Furthermore, replacing s by $1 - s$ and χ by $\bar{\chi}$, we find that

$$\begin{aligned} \xi(1-s, \bar{\chi}) &= \frac{1}{2} \int_1^\infty \left(\theta_{\bar{\chi}}(it/q) t^{-s/2-1/2} + \frac{\tau(\bar{\chi})}{\sqrt{q}} \theta_\chi(it/q) t^{s/2-1} \right) dt \\ &= \frac{1}{2} \frac{\tau(\bar{\chi})}{\sqrt{q}} \int_1^\infty \left(\frac{\sqrt{q}}{\tau(\bar{\chi})} \theta_{\bar{\chi}}(it/q) t^{-s/2-1/2} + \theta_\chi(it/q) t^{s/2-1} \right) dt \\ &= \overline{\varepsilon(\chi)} \xi(s, \chi) \end{aligned}$$

because $\tau(\bar{\chi}) = \overline{\tau(\chi)}$ (why?), so $\tau(\chi)\tau(\bar{\chi}) = q$. $\square$

From the fact that $\xi(s, \chi)$ is entire it follows that $L(s, \chi)$ has trivial zeros at the negative even integers and that $L(0, \chi) = 0$, corresponding to the poles of $\Gamma(s/2)$. The remaining zeros of $L(s, \chi)$ are the nontrivial zeros inside the critical strip.

16.3.2 Odd characters

In the case $\chi(-1) = -1$, our previous definition of $\theta_\chi(z)$ vanishes identically. We can fix this by defining

$$\theta_\chi(z) := \sum_{n \in \mathbb{Z}} n \chi(n) e^{\pi i n^2 z}.$$

when χ is odd. It is customary to make these definitions uniform by introducing the constant

$$\mathfrak{a} = \mathfrak{a}(\chi) := \begin{cases} 0 & \text{if } \chi \text{ is even,} \\ 1 & \text{if } \chi \text{ is odd.} \end{cases}$$

Then in both cases

$$\theta_\chi(z) = \sum_{n \in \mathbb{Z}} n^{\mathfrak{a}} \chi(n) e^{\pi i n^2 z}.$$

It turns out that $\theta_\chi(z)$ satisfies a similar transformation law for odd characters, but we need a slight modification of our previous arguments. In Section 5.2 we proved that

$$\int_{-\infty}^\infty e^{-\pi x^2 y - 2\pi i n x} dx = y^{-1/2} e^{-\pi n^2 / y}.$$

Replacing x by $x + \alpha$, where $\alpha \in \mathbb{R}$, we find that

$$\int_{-\infty}^\infty e^{-\pi(x+\alpha)^2 y - 2\pi i n x} dx = y^{-1/2} e^{-\pi n^2 / y + 2\pi i n \alpha}.$$

We differentiate with respect to α (this is justified by the rapid decay of the integrand and the dominated convergence theorem applied to the differentiated integral). This yields

$$-2\pi y \int_{-\infty}^{\infty} (x + \alpha) e^{-\pi(x+\alpha)^2 y - 2\pi i \alpha x} dx = 2\pi i n y^{-1/2} e^{-\pi n^2 / y + 2\pi i n \alpha}.$$

We will apply this with $\alpha = 0$. By Theorem 16.6 with $f(n) = n e^{-\pi n^2 y}$ we have

$$\theta_\chi(iy) = \sum_{m \in \mathbb{Z}} m \chi(m) e^{-\pi m^2 y} = -iy^{-3/2} \frac{\tau(\chi)}{q^2} \sum_{n \in \mathbb{Z}} n \bar{\chi}(n) e^{-\pi n^2 / q^2 y} = -iy^{-3/2} \frac{\tau(\chi)}{q^2} \theta_{\bar{\chi}}(i/q^2 y).$$

This leads us to the following theorem.

Theorem 16.8. *Suppose that χ is an odd primitive character of conductor q and define*

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

Then $\xi(s, \chi)$ extends to an entire function of s and satisfies the functional equation

$$\varepsilon(\chi) \xi(1-s, \bar{\chi}) = \xi(s, \chi),$$

where in this case the root number is $\varepsilon(\chi) = -i\tau(\chi)/\sqrt{q}$.

Proof. We proceed as before, starting instead with

$$\left(\frac{q}{\pi}\right)^{\frac{1}{2}(s+1)} \Gamma\left(\frac{1}{2}(s+1)\right) n^{-s} = \int_0^\infty n e^{-\pi n^2 t / q} t^{s/2-1/2} dt.$$

After multiplying by $\chi(n)$ and summing over positive integers n , we obtain

$$\xi(s, \chi) = \left(\frac{q}{\pi}\right)^{\frac{1}{2}(s+1)} \Gamma\left(\frac{1}{2}(s+1)\right) L(s, \chi) = \frac{1}{2} \int_0^\infty \theta_\chi(it/q) t^{s/2-1/2} dt.$$

We split the integral at $t = 1$ and use the transformation law for θ_χ on the first integral:

$$\int_0^1 \theta_\chi(it/q) t^{s/2-1/2} dt = -i \frac{\tau(\chi)}{\sqrt{q}} \int_0^1 \theta_{\bar{\chi}}(i/qt) t^{s/2-2} dt = -i \frac{\tau(\chi)}{\sqrt{q}} \int_1^\infty \theta_{\bar{\chi}}(it/q) t^{-s/2} dt.$$

It follows that

$$\xi(s, \chi) = \frac{1}{2} \int_1^\infty \left(\theta_\chi(it/q) t^{s/2-1/2} - i \frac{\tau(\chi)}{\sqrt{q}} \theta_{\bar{\chi}}(it/q) t^{-s/2} \right) dt.$$

On replacing s by $1-s$ and χ by $\bar{\chi}$ we find that

$$\xi(s, \chi) = -i \frac{\tau(\chi)}{\sqrt{q}} \xi(1-s, \chi)$$

because for odd characters we have $\tau(\bar{\chi}) = -\overline{\tau(\chi)}$ (why?). □

The shift in the gamma factor implies that the trivial zeros of $L(s, \chi)$ are at the negative odd integers when χ is odd.

Remark 16.9. We can put both functional equations together using the number $\mathfrak{a}$ defined above. Let

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

Then $\xi(s, \chi) = \varepsilon(\chi) \xi(1-s, \bar{\chi})$, where $\varepsilon(\chi) = (-i)^{\mathfrak{a}} \tau(\chi) / \sqrt{q}$.

17 Zero-free regions for $L(s, \chi)$

From here on out it is straightforward to mimic the results we obtained for $\zeta(s)$ and apply them to $L(s, \chi)$ when χ is a *fixed* character. But that's not very interesting or useful, since in most cases we would like to allow the modulus q to vary (maybe restricted with respect to other parameters). For example, we'd like to know how the counting function $\pi(x; a \bmod q)$ for arithmetic progressions behaves as we move x and q . As we shall quickly see, obtaining this uniformity in q is much easier when χ is a complex character than it is when χ is real. In fact, for real characters there is a major obstruction to the zero-free region which is not present for complex characters.

As in the case of $\zeta(s)$, we begin with an inequality for the real part of $\frac{L'}{L}(s, \chi)$ that ultimately comes from the same cosine inequality as before. For any χ we have

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

For each n (and fixed t), let θ_n be the angle for which $\chi(n) e^{-it \log n} = e^{i\theta_n}$. Then the real part of $\frac{L'}{L}(\sigma + it, \chi)$ includes the factor $\cos(\theta_n)$, and the real part of $\frac{L'}{L}(\sigma + 2it, \chi^2)$ includes the factor $\cos(2\theta_n)$. It follows that

$$-3 \operatorname{Re} \frac{L'}{L}(\sigma, \chi_0) - 4 \operatorname{Re} \frac{L'}{L}(\sigma + it, \chi) - \operatorname{Re} \frac{L'}{L}(\sigma + 2it, \chi^2) \geq 0,$$

by the same cosine inequality as before. The obstruction is ultimately due to the fact that when χ is a real character, χ^2 is the principal character χ_0 , and the third term above has a pole at $\sigma + 2it = 1$.

As in the case of $\xi(s)$, the entire function $\xi(s, \chi)$ grows like $e^{c|s| \log |s|}$ as $s \rightarrow \infty$, and so has a Weierstrass factorization of the form

$$\xi(s, \chi) = e^{A(\chi) + B(\chi)s} \prod_{\rho} \left(1 - \frac{s}{\rho}\right) e^{s/\rho}$$

for some constants $A(\chi)$ and $B(\chi)$. As before, we can write $B(\chi)$ in terms of the zeros:

$$\operatorname{Re} B(\chi) = - \sum_{\rho} \operatorname{Re} \frac{1}{\rho},$$

where the real part shows up because the functional equation relates $\xi(s, \chi)$ and $\xi(1-s, \bar{\chi})$. Here, and throughout, we will use the same notation $\rho = \beta + i\gamma$ for the nontrivial zeros of $L(s, \chi)$ even though, of course, they are different from the nontrivial zeros of $\zeta(s)$. From the factorization above, and the definition of $\xi(s, \chi)$, we find that

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

When we take real parts, the terms corresponding to $B(\chi)$ and $\sum 1/\rho$ drop out.

17.1 Complex characters

In the case where χ is a complex character, the proof of the zero-free region follows that of $\zeta(s)$ almost exactly, with the following simple modifications. First, we can't ignore the factor $\log q$ in the expansion of $-\frac{L'}{L}(s, \chi)$ above, since we want some uniformity in q . This has the effect of replacing every instance of $\log t$ with $\log q + \log(t+2) = \log(q(t+2))$. (Note that we need the wiggle room $t+2$ in the log because we don't know that $L(s, \chi)$ is nonzero for small t , like we did for $\zeta(s)$). Second, even if χ is primitive, it might not be the case that χ^2 is primitive. However, this isn't much of a problem: suppose that ψ is the primitive character that induces χ^2 ; then (16.2) shows that

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

The error term is small because for $\sigma > 1$

$$\left| \sum_{p|q} \frac{\psi(p) \log p}{p^s - \psi(p)} \right| \leq \sum_{p|q} \log p \leq \log q,$$

and this error gets absorbed in the factor $\log(q(t+2))$. Following the argument we gave for $\zeta(s)$, we obtain the following theorem.

Theorem 17.1. *If χ is a complex primitive character, then there exists a constant $c > 0$ such that $L(s, \chi)$ is nonvanishing in the region*

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

17.2 Real characters

When χ is real we need to be more careful. We start with the main inequality

$$-3 \operatorname{Re} \frac{L'}{L}(\sigma, \chi_0) - 4 \operatorname{Re} \frac{L'}{L}(\sigma + it, \chi) - \operatorname{Re} \frac{L'}{L}(\sigma + 2it, \chi^2) \geq 0.$$

The first term is essentially the same as the corresponding term for $\zeta(s)$, since

$$-\operatorname{Re} \frac{L'}{L}(\sigma, \chi_0) = \sum_{n=1}^{\infty} \frac{\chi_0(n) \Lambda(n)}{n^\sigma} \leq -\frac{\zeta'}{\zeta}(\sigma) < \frac{1}{\sigma-1} + c_1$$

for some $c_1 > 0$, uniformly in $1 < \sigma \leq 2$. For the second term, we use

$$-\operatorname{Re} \frac{L'}{L}(s, \chi) = \frac{1}{2} \log\left(\frac{q}{\pi}\right) + \frac{1}{2} \operatorname{Re} \frac{\Gamma'}{\Gamma}(s/2 + \mathfrak{a}/2) - \sum_{\rho} \operatorname{Re} \left(\frac{1}{s - \rho} \right). \quad (17.1)$$

The log factor and the gamma factor (by Stirling's formula) together are at most $\log(q(t+2))$. As before, we have

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

so we can drop as many of the terms in the ρ -sum as we want to get an upper bound. Suppose that $t = \gamma$ for some $\rho = \beta + i\gamma$. We drop all of the terms in the ρ -sum except for the term corresponding to that ρ to get

$$-\operatorname{Re} \frac{L'}{L}(\sigma + it, \chi) < -\frac{1}{\sigma - \beta} + c_2 \log(q(t+2)).$$

The third term requires some extra care. Since χ is real, the character χ^2 is the principal character χ_0 . Thus χ^2 is induced by the (only) character mod 1, i.e. the character corresponding to $\zeta(s)$. It follows from (16.2) that

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

as we obtained above. Thus it suffices to deal with $-\operatorname{Re} \frac{\zeta'}{\zeta}(s)$, which satisfies

$$-\operatorname{Re} \frac{\zeta'}{\zeta}(s) < \operatorname{Re} \left(\frac{1}{s-1} \right) + c_3 \log(t+2).$$

(When we did this argument before, we ignored the $1/(s-1)$ term because we could restrict to the region $t \geq 2$, but here we can't do that, and this term $1/(s-1)$ is now significant when t is small.) It follows that

$$-\operatorname{Re} \frac{L'}{L}(\sigma + 2it, \chi^2) < \operatorname{Re} \left(\frac{1}{\sigma - 1 + 2it} \right) + c_4 \log(q(t+2)).$$

Putting this together with the inequalities for $\frac{L'}{L}(\sigma)$ and $\frac{L'}{L}(\sigma + it)$ and the main cosine inequality, we find that

$$\frac{4}{\sigma - \beta} < \frac{3}{\sigma - 1} + \operatorname{Re} \left(\frac{1}{\sigma - 1 + 2it} \right) + c_5 \log(q(t+2)),$$

where we recall that we have fixed $t = \gamma$ for some $\rho = \beta + i\gamma$.

For simplicity, we write $\mathcal{L} = \log(q(t+2))$. Fix $\delta > 0$. First, we restrict to the zeros which satisfy $|\gamma| \geq \delta/\mathcal{L}$ (note that if $|\gamma| \geq \delta/\log q$, this condition is satisfied). If we set $\sigma = 1 + \delta/\mathcal{L}$ we find that

$$\frac{4}{\sigma - \beta} < \frac{3\mathcal{L}}{\delta} + \frac{\mathcal{L}}{5\delta} + c_5\mathcal{L} = \frac{16\mathcal{L} + 5\delta c_5\mathcal{L}}{5\delta}.$$

Here we used that $\operatorname{Re}(\frac{1}{\delta/\mathcal{L} + 2it}) = \frac{\delta/\mathcal{L}}{\delta^2/\mathcal{L}^2 + 4t^2} \leq \frac{\delta/\mathcal{L}}{5\delta^2/\mathcal{L}^2} = \frac{\mathcal{L}}{5\delta}$. Solving this inequality for β ,

$$\beta < \sigma - \frac{20\delta}{16\mathcal{L} + 5\delta c_5\mathcal{L}} = 1 - \frac{4 - 5\delta c_5}{16 + 5\delta c_5} \cdot \frac{\delta}{\mathcal{L}}.$$

We want the factor multiplying $\delta/\mathcal{L}$ to be positive, so we need to choose $\delta < \frac{4}{5c_5}$. To get a clean result (these constants don't matter, but it looks nice) let's choose $\delta < \frac{2}{15c_5}$. Then

$$\beta < 1 - \frac{\delta}{5\mathcal{L}}.$$

So there is a constant c_6 (we can take $c_6 = \frac{2}{15c_5}$) such that for any fixed $0 < \delta < c_6$, the function $L(s, \chi)$ is nonvanishing for

$$\sigma \geq 1 - \frac{\delta}{5 \log(q(|t| + 2))} \quad \text{and} \quad |t| \geq \frac{\delta}{\log q}.$$

Now we need to investigate possible zeros near the real line. Suppose that $|\gamma| < \delta / \log q$. The equation (17.1) gives the inequality

$$-\frac{L'}{L}(\sigma, \chi) < -\sum_{\rho} \frac{1}{\sigma - \rho} + c_7 \log q.$$

Note that the sum is real because the zeros come in complex conjugate pairs when χ is real (since $\chi = \bar{\chi}$). If there is a zero $\rho = \beta + i\gamma$ with $\gamma \neq 0$ (or a double real zero at ρ), then the two corresponding terms in the sum are

$$\frac{1}{\sigma - (\beta + i\gamma)} + \frac{1}{\sigma - (\beta - i\gamma)} = \frac{2(\sigma - \beta)}{(\sigma - \beta)^2 + \gamma^2},$$

from which we get (dropping all but these terms)

$$-\frac{L'}{L}(\sigma, \chi) < -\frac{2(\sigma - \beta)}{(\sigma - \beta)^2 + \gamma^2} + c_7 \log q.$$

We have a crude bound for the left-hand side:

$$-\frac{L'}{L}(\sigma, \chi) = \sum_{n=1}^{\infty} \frac{\chi(n)\Lambda(n)}{n^{\sigma}} \geq -\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^{\sigma}} = \frac{\zeta'}{\zeta}(\sigma) > -\frac{1}{\sigma - 1} - c_8,$$

from which it follows that

$$-\frac{1}{\sigma - 1} < -\frac{2(\sigma - \beta)}{(\sigma - \beta)^2 + \gamma^2} + c_9 \log q. \tag{17.2}$$

Now set $\sigma = 1 + 2\delta / \log q$; then

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

Then the inequality (17.2) implies

$$-\frac{\log q}{2\delta} < -\frac{8}{5(\sigma - \beta)} + c_9 \log q \implies \beta < 1 - \frac{3/5 - 2\delta c_9}{(c_9 + 1/2\delta) \log q}.$$

For any $\delta < \frac{1}{30c_9}$ we have $\frac{3/5 - 2\delta c_9}{(c_9 + 1/2\delta)} > \delta$, so the inequality above implies $\beta < 1 - \delta / \log q$. We conclude that $L(s, \chi)$ has at most one simple real zero in the region

$$\sigma > 1 - \frac{\delta}{\log q} \quad \text{and} \quad |t| < \frac{\delta}{\log q}.$$

(We didn't cover the case where $L(s, \chi)$ has two real zeros but it is mostly the same as our argument, so I'll omit it here.)

We can combine the results for real and complex characters into one statement.

Theorem 17.2. Suppose that χ is a nonprincipal character modulo $q > 1$. Then there exists a positive constant $c > 0$ such that in the region

$$\sigma \geq \begin{cases} 1 - \frac{c}{\log q |t|} & \text{if } |t| \geq 1, \\ 1 - \frac{c}{\log q} & \text{if } |t| < 1, \end{cases}$$

the function $L(s, \chi)$ is nonvanishing if χ is complex, and has at most one simple zero on the real line if χ is real.

Remark 17.3. The single exceptional real zero (conjectured not to exist for any χ) is often called the *Siegel zero*.

Remark 17.4. We can upgrade our results from primitive characters to nonprincipal characters using the fact that if χ is induced by the primitive character ψ then

$$L(s, \chi) = \prod_{p|q} \left(1 - \frac{\psi(p)}{p^s}\right) L(s, \psi),$$

which holds for all $s \in \mathbb{C}$ by analytic continuation. This shows that the zeros of $L(s, \chi)$ are exactly the zeros of $L(s, \psi)$, together with (well-understood) zeros on the line $\sigma = 0$. For principal characters, $L(s, \chi_0)$ is a finite Euler product multiplied by $\zeta(s)$, so we can apply the results for $\zeta(s)$ directly.

If we follow the same sort of argument as before for the quantity

$$-\frac{\zeta'}{\zeta}(\sigma) - \frac{L'}{L}(\sigma, \chi_1) - \frac{L'}{L}(\sigma, \chi_2) - \frac{L'}{L}(\sigma, \chi_1 \chi_2),$$

where χ_1 and χ_2 are distinct real primitive characters modulo q_1 and q_2 , respectively, then we can show that

$$\frac{1}{\sigma - \beta_1} + \frac{1}{\sigma - \beta_2} < \frac{1}{\sigma - 1} + A \log q_1 q_2,$$

where β_1 and β_2 are the corresponding real zeros. Set $\sigma = 1 + \delta / \log(q_1 q_2)$ for some small fixed $\delta > 0$. If we have

$$\beta_1, \beta_2 > 1 - \frac{\delta'}{\log q_1 q_2} \tag{17.3}$$

for some $\delta' > \delta$ then we find that

$$\log q_1 q_2 \cdot \frac{2}{\delta' - \delta} < \frac{1}{\sigma - \beta_1} + \frac{1}{\sigma - \beta_2} < \frac{\log q_1 q_2}{\delta} + A \log q_1 q_2,$$

so

$$\delta' > \delta + \frac{2\delta}{A\delta + 1}.$$

Thus if we choose δ' carefully, we see that (17.3) cannot hold.

Theorem 17.5. *If χ_1 and χ_2 are distinct primitive real characters modulo q_1 and q_2 , respectively (we can have $q_1 = q_2$) such that the corresponding L -functions have real zeros β_1 and β_2 , then there exists an absolute positive constant $c > 0$ such that*

$$\min(\beta_1, \beta_2) < 1 - \frac{c}{\log q_1 q_2}.$$

Corollary 17.6. *For each q , at most one of the real nonprincipal characters modulo q has a Siegel zero.*

17.3 The number $N(T, \chi)$

For the prime number theorem in arithmetic progressions, we will also need a formula for the number $N(T, \chi)$ of zeros of $L(s, \chi)$ in the critical strip with imaginary part in $[-T, T]$. Note that we don't necessarily have symmetry over the real line, so we can't restrict to $[0, T]$. The proof goes almost exactly as in the case of $\zeta(s)$ with the main difference being that we keep track of the dependence on q . So I am content to just state the result.

Theorem 17.7. *If χ is a primitive character modulo q then for $T \geq 2$ we have*

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

The factor $1/2$ on the left-hand side compensates for the rectangle being twice the size of the rectangle defining $N(T)$. Also, if we want to extend this result to nonprimitive characters, we need to count the zeros on the line $\sigma = 0$ (and make the defining rectangle R slightly larger to compensate for this), and this changes the result to $\frac{1}{2}N_R(T, \chi) = \frac{T}{2\pi} \log \frac{T}{2\pi} + O(T \log q)$.

18 PNT in arithmetic progressions

Our aim in this section is to prove the prime number theorem for the function $\pi(x; a \bmod q)$ where $(a, q) = 1$. We might expect that there are roughly the same number of primes in each residue class $a \bmod q$ as long as a and q are coprime. Since there are $\varphi(q)$ such a , we might expect that

$$\pi(x; a \bmod q) \sim \frac{1}{\varphi(q)} \cdot \frac{x}{\log x}.$$

Indeed, this is the case. The difficulty comes in obtaining an error term that is meaningful in as large a range of q as possible.

18.1 The explicit formula for $L(s, \chi)$

We begin by proving an analogue of the explicit formula, this time relating the zeros of $L(s, \chi)$ to the summatory function of the von Mangoldt function $\Lambda(n)$, twisted by χ , i.e.

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

We employ the same test function as before, namely

$$\phi_x(t) = \begin{cases} 1 & \text{if } 0 \leq t \leq x, \\ 1 + \frac{x}{y} - \frac{1}{y}t & \text{if } x \leq t \leq x + y, \\ 0 & \text{if } t \geq x + y. \end{cases}$$

As before, the price we pay for using $\phi_x(t)$ instead of a sharp cutoff function is

$$\sum_{n \leq x} \chi(n) \Lambda(n) - \sum_{n \leq x} \chi(n) \Lambda(n) \phi_x(n) \ll \sum_{x \leq n \leq x+y} \Lambda(n) \ll y \log x. \quad (18.1)$$

Recall that the Mellin transform of $\phi_x(t)$ is given by

$$\begin{aligned} \tilde{\phi}_x(s) &= \frac{x^{s+1}}{ys(s+1)} \left(\left(1 + \frac{y}{x}\right)^{s+1} - 1 \right) \\ &= \frac{x^s}{s} + yx^{s-1}w_s, \end{aligned} \quad (18.2)$$

where, if $1 \leq y \leq x$ and $0 \leq \operatorname{Re}(s) \leq 1$, we have

$$|w_s| \leq \frac{1}{2}.$$

When χ is odd, the proof of the explicit formula goes through exactly as in the case of $\zeta(s)$, where we note that $L(s, \chi)$ is entire, and the trivial zeros are at negative odd integers instead of negative even integers.

Theorem 18.1. *If χ is a primitive odd character modulo q then*

$$\sum_{n=1}^{\infty} \chi(n) \Lambda(n) \phi_x(n) = - \sum_{\rho} \tilde{\phi}_x(\rho) - \frac{L'}{L}(0, \chi) - \sum_{k=1}^{\infty} \tilde{\phi}_x(1-2k).$$

When χ is even, there is one further complication coming from the fact that $L(s, \chi)$ vanishes at $s = 0$, so the function $-\tilde{\phi}_x(s) \frac{L'}{L}(s, \chi)$ has a double pole at $s = 0$. This makes the residue computation slightly more complicated, but otherwise doesn't change much. Near $s = 0$ we have the Laurent expansion

$$\frac{L'}{L}(s, \chi) = \frac{1}{s} + b(\chi) + O(s),$$

where $b(\chi)$ is some constant that we will compute later. For $\tilde{\phi}_x(s)$ it's a bit more complicated. We start with

$$\frac{x^s}{s} = \frac{1}{s} + \log x + O(s).$$

That's not so bad, but now we have to compute the Laurent expansion of w_s around $s = 0$. This is tedious but not difficult, and I'll skip the details:

$$w_s = -\frac{x}{y} + \frac{(1 + \frac{y}{x}) \log(1 + \frac{y}{x})}{(\frac{y}{x})^2} + O(s).$$

It follows that near $s = 0$ we have

$$\tilde{\phi}_x(s) = \frac{1}{s} + \log x - 1 + \left(1 + \frac{x}{y}\right) \log\left(1 + \frac{y}{x}\right) + O(s).$$

Note that $\frac{x}{y} \log\left(1 + \frac{y}{x}\right) \ll 1$, which can be seen by expanding $\log(1 + u)$ around $u = 0$. Thus

$$\text{Res}(0; -\phi_x(s) \frac{L'}{L}(s, \chi)) = -b(\chi) - \left(\log x - 1 + \left(1 + \frac{x}{y}\right) \log\left(1 + \frac{y}{x}\right)\right).$$

The rest of the computation in this case goes through as before, and we get the following theorem.

Theorem 18.2. *If χ is a primitive even character modulo q then*

$$\sum_{n=1}^{\infty} \chi(n) \Lambda(n) \phi_x(n) = - \sum_{\rho} \tilde{\phi}_x(\rho) - b(\chi) - \left(\log x - 1 + \left(1 + \frac{x}{y}\right) \log\left(1 + \frac{y}{x}\right)\right) - \sum_{k=1}^{\infty} \tilde{\phi}_x(-2k).$$

Using the infinite product formula for $\xi(s, \chi)$, we find that

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

It follows that

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

If χ is odd, then $\mathfrak{a} = 1$, and the function $\frac{\Gamma'}{\Gamma}((s + 1)/2)$ is analytic at $s = 0$. Therefore, in that case, we have

$$\frac{L'}{L}(0, \chi) = - \sum_{\rho} \left(\frac{1}{\rho} + \frac{1}{2 - \rho}\right) + O(1).$$

If χ is even, then $\mathfrak{a} = 0$, and we have the Laurent series expansion

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

Since we defined $b(\chi)$ as the constant term in the Laurent expansion of $\frac{L'}{L}(s, \chi)$ around $s = 0$, it follows that

$$b(\chi) = - \sum_{\rho} \left(\frac{1}{\rho} + \frac{1}{2 - \rho}\right) + O(1).$$

So in both cases, we need to estimate the sum over zeros that appears above. To ease notation, let $b(\chi) = \frac{L'}{L}(0, \chi)$ when χ is odd.

For $|\gamma| \geq 1$ we have

$$\sum_{|\gamma| \geq 1} \left(\frac{1}{\rho} + \frac{1}{2 - \rho}\right) = \sum_{|\gamma| \geq 1} \frac{2}{\rho(2 - \rho)} \ll \sum_{|\gamma| \geq 1} \frac{1}{|\rho|^2} \ll \log q.$$

The last estimate follows from partial summation, using the estimate of $N(T, \chi)$ in the previous section. When $|\gamma| \leq 1$ we have $|2 - \rho| \gg 1$, so

$$\sum_{|\gamma| \leq 1} \frac{1}{2 - \rho} \ll \sum_{|\gamma| \leq 1} 1 \ll N(1, \chi) \ll \log q.$$

It follows that

$$b(\chi) = - \sum_{|\gamma| < 1} \frac{1}{\rho} + O(\log q).$$

Among these zeros, the only ones that might give us trouble are the Siegel zeros (if there is a real zero $\beta_S > 1/2$ then there will also be another $1 - \beta_S < 1/2$). Such a zero only (possibly) occurs when χ is real; it is simple, and satisfies

$$\beta_S > 1 - \frac{c}{\log q}.$$

Here c is a constant which we may assume is $< 1/4$ (since making c smaller makes the inequality above worse) so that $\beta_S > 3/4$ (note that $q \geq 3$ since χ is primitive). The sum $\sum' \rho^{-1}$ over the remaining zeros with $|\gamma| < 1$ and $\beta < 1 - c/\log q$ (so also $1 - \beta < 1 - c/\log q$) can be absorbed into a slightly larger error term, since for these zeros $|\rho^{-1}| \ll \log q$ and there are at most $\log q$ such zeros. Furthermore, the term β_S^{-1} can be absorbed into the error term since $\beta_S > 3/4$. Thus

$$b(\chi) = - \frac{1}{1 - \beta_S} + O(\log^2 q).$$

We treat the term $\sum_{\rho} \tilde{\phi}_x(\rho)$ as we did before, except that we need to treat the terms corresponding to β_S and $1 - \beta_S$ separately. For β_S we have

$$\tilde{\phi}_x(\beta_S) = \frac{x^{\beta_S}}{\beta_S} + O(y)$$

and for $1 - \beta_S$ we have

$$\tilde{\phi}_x(1 - \beta_S) = \frac{x^{1 - \beta_S}}{1 - \beta_S} + O(y).$$

When we combine the terms $\tilde{\phi}_x(1 - \beta_S)$ and $b(\chi)$ we will get a term

$$\frac{x^{1 - \beta_S} - 1}{1 - \beta_S} = x^{\sigma_S} \log x$$

for some $\sigma_S \in (0, 1 - \beta_S)$, by the mean value theorem. The latter expression is $\ll x^{1/4} \log x$.

Omitting the details, which at this point are nearly identical to the $\zeta(s)$ case, we obtain the following theorem.

Theorem 18.3. *Define $\beta_0(T, \chi)$ and $\beta_1(T, \chi)$ analogously to $\beta_j(T)$ for $T \geq 1$. Suppose that χ is a nonprincipal character of modulus q . Suppose that $x \geq 2$ and that T satisfies*

$$x^{\frac{1}{2}(\beta_1(T, \chi) - \beta_0(T, \chi))} \leq T \leq x.$$

If no Siegel zero exists for χ , then

$$\sum_{n \leq x} \chi(n) \Lambda(n) = - \sum_{|\gamma| \leq T} \frac{x^\rho}{\rho} + O\left(x^{\frac{1}{2}(\beta_0(T, \chi) + \beta_1(T, \chi))} \log^2 qx + \frac{\log^2 qx}{T} x^{1 + \frac{1}{2}(\beta_1(T, \chi) - \beta_0(T, \chi))}\right).$$

If a Siegel zero β_S exists, then the quantity

$$-\frac{x^{\beta_S}}{\beta_S} + O(x^{1/2} \log x)$$

should be added to the right-hand side, and the terms corresponding to the zeros β_S and $1 - \beta_S$ should be deleted from the sum.

In order to upgrade the discussion above the theorem statement to the case where χ is nonprincipal, but maybe not primitive, suppose that χ is induced by χ' of modulus q' . Then

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

and this is significantly smaller than the error term in the theorem.

18.2 The prime number theorem for arithmetic progressions

We can now estimate the quantity

$$\psi(x; a \bmod q) = \sum_{\substack{n \leq x \\ n \equiv a(q)}} \Lambda(n) = \frac{1}{\varphi(q)} \sum_{\chi(q)} \bar{\chi}(a) \psi(x, \chi),$$

where

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

The main term comes from the principal character χ_0 ; for this character we have

$$|\psi(x, \chi_0) - \psi(x)| \leq \sum_{\substack{n \leq x \\ (n, q) = 1}} \Lambda(n) \ll (\log x)(\log q),$$

and the original prime number theorem tells us that

$$\psi(x) = x + O\left(x \exp(-c_1 \sqrt{\log x})\right).$$

If we follow the same procedure as before with the final estimate for $\psi(x, \chi)$ in the previous section, we find that for $\chi \neq \chi_0$ we have

$$\psi(x, \chi) = -\frac{x^{\beta_S}}{\beta_S} E(\chi) + O\left(x \exp(-c_2 \sqrt{\log x})\right),$$

where β_S is the (possible) Siegel zero of $L(s, \chi_1)$, where χ_1 is the (possible) single exceptional real nonprincipal character modulo q , and $E(\chi) = 0$ unless $\chi = \chi_1$, then $E(\chi_1) = 1$. In order for this estimate to be valid, we need that $q \ll \exp(C\sqrt{\log x})$ for some fixed $C > 0$.

We need an upper bound for β_S in order to estimate the term x^{β_S} . A weak lower bound can be obtained using the class number formula for quadratic fields, namely

$$\beta_S < 1 - \frac{c_3}{q^{1/2} \log^2 q}. \quad (18.3)$$

We won't prove this here, since we'll prove something much stronger (but somehow still unsatisfactory since it's ineffective) in the next section. The inequality (18.3) yields

$$\frac{x^{\beta_S}}{\beta_S} \ll x \exp\left(-c_3 \frac{\log x}{q^{1/2} \log^2 q}\right).$$

In order for this to be on the same order of magnitude as the other error terms we need

$$\frac{\log x}{q^{1/2} \log^2 q} \gg \sqrt{\log x},$$

which is true if

$$q \ll (\log x)^{1-\delta}.$$

for some fixed $\delta > 0$. This uniformity in q is quite weak, but we will upgrade it in the next section. We obtain the following theorem.

Theorem 18.4. *Fix $\delta > 0$ and suppose that $q \ll (\log x)^{1-\delta}$. Then there exists a constant $c > 0$ such that*

$$\psi(x; a \bmod q) = \frac{x}{\varphi(q)} + O\left(x \exp(-c\sqrt{\log x})\right).$$

The Generalized Riemann Hypothesis (GRH) is the conjecture that all of the Dirichlet L -functions $L(s, \chi)$ have all of their nontrivial zeros on the line $\sigma = 1/2$. (There is also an “Extended Riemann Hypothesis” and a “Grand Riemann Hypothesis” associated to other L -functions, but I don't think that there is a consensus as to what their precise scope is. Most people just say “GRH” when referring to the Riemann Hypothesis for L -functions other than $\zeta(s)$, and, if it matters, they specify which L -functions they need a Riemann Hypothesis for.) As in the case of the ordinary prime number theorem, we can do much better if we assume GRH.

Theorem 18.5. *Suppose that $q \leq x$. Assuming GRH, we have*

$$\psi(x; a \bmod q) = \frac{x}{\varphi(q)} + O(\sqrt{x} \log^2 x).$$

19 Siegel's Theorem

The key knowledge that we needed in Dirichlet's Theorem was the nonvanishing of $L(1, \chi)$ and, in the last section, we saw that a (potential) real zero of $L(s, \chi)$ near $s = 1$ has a serious impact on the error term in the prime number theorem for arithmetic progressions. Siegel's Theorem is a much more powerful statement regarding $L(1, \chi)$, namely a pretty strong lower bound in terms of q .

Theorem 19.1 (Siegel's Theorem). *Fix $\varepsilon > 0$. Then there exists a constant $C_1(\varepsilon) > 0$ such that for each real primitive character of modulus q we have*

$$L(1, \chi) > C_1(\varepsilon)q^{-\varepsilon}. \quad (19.1)$$

I have followed Davenport in writing out the constant $C_1(\varepsilon)$ instead of just writing $\gg_\varepsilon$ in order to emphasize the ineffective nature of the result. That is, we don't know, for any fixed ε , an explicit value of $C_1(\varepsilon)$, nor do we have any idea how to compute such a constant. The proof of the theorem simply doesn't provide a way to compute $C_1(\varepsilon)$. Our main application of Siegel's Theorem is an upgraded version of the bound for the location of a (potential) Siegel zero as a function of q .

Theorem 19.2. *Fix $\varepsilon > 0$. Then there exists a constant $C_2(\varepsilon) > 0$ such that for each real primitive character of modulus q , the function $L(\sigma, \chi)$ is nonvanishing for*

$$\sigma > 1 - C_2(\varepsilon)q^{-\varepsilon}.$$

Proof. Suppose that β is a zero of $L(s, \chi)$ satisfying $\beta > 1 - C_2(\varepsilon)q^{-\varepsilon}$. Then by the mean value theorem,

$$L(1, \chi) = L(1, \chi) - L(\beta, \chi) = (1 - \beta)L'(\sigma, \chi) < C_2(\varepsilon)q^{-\varepsilon}L'(\sigma, \chi)$$

for some $\sigma \in (\beta, 1)$. Suppose that $\sigma > 1 - 1/\log q$. For $n \leq q$ we have $n^{-\sigma} \leq n^{-1}e^{\log n/\log q} \leq en^{-1}$. It follows that

$$\begin{aligned} |L'(\sigma, \chi)| &\leq \sum_{n=1}^q \frac{e|\chi(n) \log n|}{n} + \left| \sum_{n=q+1}^{\infty} \frac{\chi(n) \log n}{n^\sigma} \right| \\ &\leq e \sum_{n=1}^q \frac{\log n}{n} - \int_q^\infty \left| \sum_{q < n \leq t} \chi(n) \right| \left(\frac{\log t}{t^\sigma} \right)' dt \\ &\ll \log^2 q - \max_N \left| \sum_{q < n \leq N} \chi(n) \right| \int_q^\infty \left(\frac{\log t}{t^\sigma} \right)' dt \\ &\ll \log^2 q + \varphi(q) \frac{\log q}{q^{1-1/\log q}} \ll \log^2 q. \end{aligned}$$

Note that $\log t/t^\sigma$ is decreasing, hence the minus signs. Thus

$$L(1, \chi) \ll q^{-\varepsilon} \log^2 q,$$

which contradicts Theorem 19.1 if we replace ε by $\varepsilon/2$ in (19.1). □

Note that a similar argument gives the inequality

$$|L(\sigma, \chi)| \ll \log q \quad \text{for } 1 - \frac{1}{\log q} \leq \sigma \leq 1, \quad (19.2)$$

which will be useful later.

19.1 The proof of Siegel's Theorem

Proof of Theorem 19.1. Let χ_1 and χ_2 be real primitive characters of distinct moduli q_1 and q_2 , respectively. The character $\chi_1\chi_2$ is a nonprincipal character of modulus q_1q_2 , though it may not be primitive. Let

$$F(s) = \zeta(s)L(s, \chi_1)L(s, \chi_2)L(s, \chi_1\chi_2).$$

Then $F(s)$ is analytic in $\mathbb{C}$ except for a simple pole at $s = 1$ of residue

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

By considering the Euler products, we find that for $\sigma > 1$

$$\begin{aligned} \log F(s) &= \sum_p \sum_{m=1}^{\infty} \frac{1}{mp^{ms}} (1 + \chi_1(p^m) + \chi_2(p^m) + \chi_1\chi_2(p^m)) \\ &= \sum_p \sum_{m=1}^{\infty} \frac{1}{mp^{ms}} (1 + \chi_1(p^m))(1 + \chi_2(p^m)) = \sum_p \sum_{m=1}^{\infty} \frac{c(p^m)}{mp^{ms}}, \end{aligned}$$

where $c(p^m) \geq 0$. Exponentiating, we obtain

$$F(s) = \prod_p \prod_{m=1}^{\infty} \left(1 + \sum_{k=1}^{\infty} \frac{1}{k!} \frac{c(p^m)^k}{m^k p^{kms}} \right) = \sum_{n=1}^{\infty} \frac{a(n)}{n^s},$$

where $a(1) = 1$ and $a(n) \geq 0$ for $n \geq 2$. Now we follow the same argument as in the proof of $L(1, \chi) \neq 0$. Since $F(s) - \lambda(s-1)^{-1}$ is analytic in $\sigma > 0$, it has a Taylor expansion at $s = 2$ which is valid for at least $|s-2| < 2$. We have

$$F(s) - \frac{\lambda}{s-1} = \sum_{m=0}^{\infty} \frac{F^{(m)}(2)}{m!} (s-2)^m - \frac{\lambda}{1-(2-s)} = \sum_{m=0}^{\infty} (b_m - \lambda)(2-s)^m,$$

where $b_0 \geq 1$ and $b_m \geq 0$ for all $m \geq 2$.

By partial summation, we have, for any nonprincipal character,

$$L(s, \chi) = s \int_1^{\infty} \left(\sum_{n \leq t} \chi(n) \right) \frac{dt}{t^{s+1}}$$

for $\sigma > 0$. Since $|\sum_{n \leq t} \chi(n)| \leq q$ we have, for $\sigma \geq 1/2$,

$$|L(s, \chi)| \leq q|s| \int_1^{\infty} \frac{dt}{t^{\sigma+1}} \leq 2q|s|.$$

Thus, on the circle $|s-2| = 3/2$, we have the bounds

$$\zeta(s) \ll 1, \quad L(s, \chi_1) \ll q_1, \quad L(s, \chi_2) \ll q_2, \quad L(s, \chi_1\chi_2) \ll q_1q_2.$$

Hence $F(s) \ll q_1^2 q_2^2$, and the same holds for $\lambda(s-1)^{-1}$ since $\lambda = L(1, \chi_1)L(1, \chi_2)L(1, \chi_1\chi_2)$. Now, by the residue theorem we have

$$b_m - \lambda = \frac{1}{2\pi i} \int_{|s-2|=3/2} \frac{F(s) - \lambda(s-1)^{-1}}{(2-s)^{m+1}} ds,$$

from which it follows that

$$|b_m - \lambda| \ll q_1^2 q_2^2 \int_{|s-2|=3/2} \frac{|ds|}{|2-s|^{m+1}} \ll q_1^2 q_2^2 \left(\frac{2}{3}\right)^m.$$

So for $\sigma \in [\frac{7}{8}, 1)$ we have

$$\sum_{m=M}^{\infty} |b_m - \lambda| |2-\sigma|^m \ll q_1^2 q_2^2 \sum_{m=M}^{\infty} \left(\frac{2}{3}\right)^m \left(\frac{9}{8}\right)^m \ll q_1^2 q_2^2 e^{-M/4}.$$

Here we have used that $3/4 < e^{-1/4}$. Let c_1 denote the implied constant on the right-hand side above and choose M such that

$$\frac{1}{2} e^{-1/4} \leq c_1 q_1^2 q_2^2 e^{-M/4} < \frac{1}{2}.$$

Then

$$\begin{aligned} F(\sigma) - \frac{\lambda}{\sigma-1} &\geq \sum_{m=0}^{M-1} (b_m - \lambda)(2-\sigma)^m - c_1 q_1^2 q_2^2 e^{-M/4} \\ &> \frac{1}{2} - \lambda \sum_{m=0}^{M-1} (2-\sigma)^m. \end{aligned}$$

Summing the finite geometric series, we conclude that

$$F(\sigma) > \frac{1}{2} - \frac{\lambda}{1-\sigma} (2-\sigma)^M.$$

For our choice of M we have

$$M \leq 4 \log(2e^{1/4} c_1 q_1^2 q_2^2) \leq 8 \log(q_1 q_2) + c_2$$

from which it follows that, for $\sigma \in [\frac{7}{8}, 1)$, we have

$$\begin{aligned} (2-\sigma)^M &= \exp(M \log(2-\sigma)) \leq \exp(M(1-\sigma)) \\ &\leq \exp(8 \log(q_1 q_2)(1-\sigma) + c_2(1-\sigma)) \\ &\leq c_3 (q_1 q_2)^{8(1-\sigma)}. \end{aligned}$$

Therefore we have the inequality

$$F(\sigma) \geq \frac{1}{2} - \frac{c_3 \lambda}{1-\sigma} (q_1 q_2)^{8(1-\sigma)} \quad \text{for } \frac{7}{8} \leq \sigma < 1. \quad (19.3)$$

We choose χ_1 depending on the value of ε , as follows. If there exists a real primitive character χ such that $L(\beta, \chi) = 0$ for some $\beta > 1 - \frac{1}{16}\varepsilon$ then we set $\chi_1 = \chi$ and $\beta_1 = \beta$. Then no matter what χ_2 is, we have $F(\beta_1) = 0$. If no such χ exists, then we fix χ_1 to be any real primitive character, and pick any β_1 in $(1 - \frac{1}{16}\varepsilon, 1)$. I claim that in this case we have $F(\beta_1) < 0$. Indeed, for $0 < \sigma < 1$ we have

$$\zeta(\sigma) = \frac{\sigma}{\sigma - 1} - \sigma \int_1^\infty \frac{t - \lfloor t \rfloor}{t^{\sigma+1}} dt < 0.$$

Then since $L(1, \chi) > 0$ for χ real (the Euler product is strictly positive for $\sigma > 1$ and $L(1, \chi) \neq 0$) and since $L(\sigma, \chi) \neq 0$ for $\sigma \geq \beta_1$, we have $L(\beta_1, \chi_1)L(\beta_1, \chi_2)L(\beta_1, \chi_1\chi_2) > 0$. Therefore, in each case we have $F(\beta_1) \leq 0$. From (19.3) we get the inequality

$$\lambda \gg (1 - \beta_1)(q_1 q_2)^{-8(1-\beta_1)}.$$

Now let χ_2 be any real primitive character of modulus $q_2 > q_1$. By (19.2) we have

$$\lambda \ll \log q_1 \log(q_1 q_2) L(1, \chi_2),$$

from which we get

$$L(1, \chi_2) \gg \frac{(1 - \beta_1)}{\log q_1 \log(q_1 q_2) (q_1 q_2)^{8(1-\beta_1)}} > \frac{C(\varepsilon)}{q_2^{8(1-\beta_1)} \log q_2},$$

where $C(\varepsilon)$ depends only on q_1 , and thus only on ε . Note that $8(1 - \beta_1) < \frac{1}{2}\varepsilon$. Since $\log q_2 < q_2^{\frac{1}{2}\varepsilon}$ for sufficiently large q_2 , we conclude that

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

for sufficiently large q_2 , and this is sufficient to prove Siegel's Theorem. $\square$

19.2 Primes in arithmetic progressions revisited

Recall the estimate

$$\psi(x, \chi) = -\frac{x^{\beta_1}}{\beta_1} E(\chi) + O\left(x \exp(-c_1 \sqrt{\log x})\right)$$

for nonprincipal χ modulo q , where $q \ll \exp(C\sqrt{\log x})$ for some $C > 0$. The main term only appears if β_1 is the Siegel zero associated to the (at most one) exceptional real character modulo q . Using Siegel's Theorem, we can get a better estimate for x^{β_1} . Using that

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

we find that

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

In order for this to be a meaningful error term, we need to impose the restriction

$$q \leq (\log x)^A$$

for some fixed large $A > 0$. (Note that this range is much larger than the range of q we had previously.) Then, taking $\varepsilon = \frac{1}{2A}$, we get $q^\varepsilon \leq \sqrt{\log x}$, so

$$x^{\beta_1} < x \exp\left(-C\left(\frac{1}{2A}\right)\sqrt{\log x}\right).$$

This yields the following improvement on PNT in arithmetic progressions. This is often called the Siegel-Walfisz theorem.

Theorem 19.3 (Siegel-Walfisz). *Fix $A > 0$ and suppose that $q \ll (\log x)^A$. Then there exists a constant $C(A) > 0$ such that*

$$\psi(x; a \bmod q) = \frac{x}{\varphi(q)} + O\left(x \exp(-C(A)\sqrt{\log x})\right).$$

20 The Bombieri-Vinogradov Theorem

From now until the end of the course we will focus on the error term *on average* in the prime number theorem for arithmetic progressions. The celebrated Bombieri-Vinogradov Theorem is the estimate

$$\sum_{q \leq Q} \max_{y \leq x} \max_{\substack{a \bmod q \\ (a,q)=1}} \left| \psi(y; a \bmod q) - \frac{y}{\varphi(q)} \right| \ll Q\sqrt{x}(\log x)^5$$

where $\sqrt{x}(\log x)^{-A} \leq Q \leq \sqrt{x}$. To see the strength of this estimate, consider that GRH gives the bound $\ll \sqrt{x} \log^2 x$ for each term in the sum, so that (up to log factors) the estimate above is essentially of the same strength as what GRH yields. On the other hand, if we estimate the left-hand side trivially: there are at most $\frac{y}{q} + 1$ integers $n \leq y$ such that $n \equiv a \pmod{q}$, from which it follows that $\max_{y \leq x} \psi(y; a \bmod q) \ll \frac{1}{q} x \log x$ for $q \leq x^{1/2}$. Similarly, one can show that $\varphi(q) \gg q/\log q$, from which it follows that the term $y/\varphi(q)$ is also at most $\frac{1}{q} x \log x$ for $q \leq x^{1/2}$ and $y \leq x$. So

$$\sum_{q \leq Q} \max_{y \leq x} \max_{\substack{a \bmod q \\ (a,q)=1}} \left| \psi(y; a \bmod q) - \frac{y}{\varphi(q)} \right| \ll x(\log x)^2$$

if $Q \leq x$. Note that

$$Q\sqrt{x}(\log x)^5 \ll x(\log x)^2 \iff Q \ll \sqrt{x}(\log x)^{-3},$$

so the Bombieri-Vinogradov estimate is only nontrivial in the small range $\sqrt{x}(\log x)^{-A} \ll Q \ll \sqrt{x}(\log x)^{-3}$. However, for applications, this turns out to be enough.

20.1 The large sieve inequality

A key step in the proof of the Bombieri-Vinogradov Theorem is a large sieve inequality for Dirichlet characters. There are many inequalities in analytic number theory that have the name “large sieve” and this is merely one of them. It’s also a bit of a misnomer, because it

really has nothing to do with sieve theory; the name traces back to an idea of Linnik which has since been completely transformed into nothing resembling a sieve. Such inequalities can often be used to prove sieve-type results, however. The result we need is the following theorem.

Theorem 20.1 (The Large Sieve Inequality for Dirichlet Characters). *For any $Q \geq 1$ and any complex numbers a_n we have*

$$\sum_{q \leq Q} \frac{q}{\varphi(q)} \sum_{\chi(q)^*} \left| \sum_{n=M+1}^{M+N} a_n \chi(n) \right|^2 \leq (\pi N + Q^2) \sum_{n=M+1}^{M+N} |a_n|^2,$$

where the notation $\chi(q)^*$ indicates that only primitive characters appear in the sum.

The constants appearing on the right-hand side of the inequality don't matter for our purposes, and with a bit more effort we could improve them. But we won't do that here.

We begin with a theorem which is also itself called the Large Sieve. Some notation will be helpful in what follows. Let us fix the integers M , N , and the complex numbers a_n throughout this subsection. For any real number α , let

$$S(\alpha) = \sum_{n=M+1}^{M+N} a_n e(n\alpha).$$

Let $\|\alpha\|$ denote the distance from α to the nearest integer, so that $\|\alpha_1 - \alpha_2\|$ is the distance from α_1 to α_2 in $\mathbb{R}/\mathbb{Z}$.

Theorem 20.2. *Let M , N , and a_n be as above, and suppose that $\alpha_1, \dots, \alpha_R$ are real numbers which are distinct modulo 1 such that $\|\alpha_r - \alpha_s\| \geq \delta$ for $r \neq s$. Then*

$$\sum_{r=1}^R |S(\alpha_r)|^2 \leq \left(\pi N + \frac{1}{\delta} \right) \sum_{n=M+1}^{M+N} |a_n|^2.$$

Before proving Theorem 20.2, we first need an elementary lemma.

Lemma 20.3. *If $g : [a - h, a + h] \rightarrow \mathbb{C}$ is differentiable then*

$$|g(a)| \leq \frac{1}{2h} \int_{a-h}^{a+h} |g(t)| dt + \frac{1}{2} \int_{a-h}^{a+h} |g'(t)| dt.$$

Proof. Define

$$\rho(t) = \begin{cases} t - a + h & \text{if } a - h \leq t \leq a, \\ t - a - h & \text{if } a < t \leq a + h. \end{cases}$$

Then integration by parts on the intervals $[a - h, a]$ and $[a, a + h]$ gives

$$\int_{a-h}^{a+h} \rho(t) g'(t) dt = 2hg(a) - \int_{a-h}^{a+h} g(t) dt.$$

The lemma follows after using that $|\rho(t)| \leq h$. □

Proof of Theorem 20.2. We begin by noting that the value of M is irrelevant: if we set

$$T_K(\alpha) = \sum_{n=K+1}^{K+N} a_{M-K+n} e(n\alpha) = \sum_{n=M+1}^{M+N} a_n e((n+K-M)\alpha) = e((K-M)\alpha) S(\alpha)$$

then $|T_K(\alpha)| = |S(\alpha)|$. So we only need to prove the theorem for the sum

$$S_K(\alpha) = \sum_{n=-K}^K a_n e(n\alpha)$$

in the form

$$\sum_{r=1}^R |S_K(\alpha_r)|^2 \leq \left(2\pi K + \frac{1}{\delta}\right) \sum_{n=-K}^K |a_n|^2.$$

Indeed, if N is odd then we let K be such that $N = 2K+1$, and if N is even we let $K = N/2$ and let $a_{-K} = 0$.

We apply the lemma above with $g(\alpha) = S_K(\alpha)^2$ and $h = \frac{1}{2}\delta$ to get

$$\sum_{r=1}^R |S(\alpha_r)|^2 \leq \sum_{r=1}^R \int_{\alpha_r - \frac{1}{2}\delta}^{\alpha_r + \frac{1}{2}\delta} \left(\frac{1}{\delta} |S_K(\alpha)|^2 + |S_K(\alpha) S'_K(\alpha)| \right) d\alpha.$$

Since the intervals of integration are non-overlapping and the integrand is nonnegative, we have (note that we can always reduce α_r modulo 1 without affecting $S(\alpha)$)

$$\sum_{r=1}^R |S(\alpha_r)|^2 \leq \frac{1}{\delta} \int_0^1 |S_K(\alpha)|^2 d\alpha + \int_0^1 |S_K(\alpha) S'_K(\alpha)| d\alpha.$$

For the first integral, we expand out $|S_K(\alpha)|^2 = S_K(\alpha) \overline{S_K(\alpha)}$ to get

$$\int_0^1 |S_K(\alpha)|^2 d\alpha = \sum_{n=-K}^K \sum_{m=-K}^K a_n \overline{a_m} \int_0^1 e(\alpha(n-m)) d\alpha = \sum_{n=-K}^K \sum_{\substack{m=-K \\ m=n}}^K a_n \overline{a_m} = \sum_{n=-K}^K |a_n|^2.$$

For the second integral, we apply Cauchy-Schwarz to get

$$\begin{aligned} \int_0^1 |S_K(\alpha) S'_K(\alpha)| d\alpha &\leq \left(\int_0^1 |S_K(\alpha)|^2 d\alpha \right)^{1/2} \left(\int_0^1 |S'_K(\alpha)|^2 d\alpha \right)^{1/2} \\ &= \left(\sum_{n=-K}^K |a_n|^2 \right)^{1/2} \left(\sum_{n=-K}^K |2\pi i n a_n|^2 \right)^{1/2} \\ &\leq 2\pi K \sum_{n=-K}^K |a_n|^2. \end{aligned}$$

The large sieve inequality follows. □

In our application to Dirichlet characters, we will take the numbers α_r to be the Farey fractions

$$\mathcal{F}_Q = \left\{ \frac{a}{q} : 1 \leq a \leq q, (a, q) = 1, q \leq Q \right\}.$$

For any two distinct Farey fractions $\frac{a}{q}$ and $\frac{a'}{q'}$ we have

$$\left\| \frac{a}{q} - \frac{a'}{q'} \right\| = \left\| \frac{aq' - a'q}{qq'} \right\| \geq \frac{1}{qq'} \geq \frac{1}{Q^2}.$$

So the large sieve in this case (with $\delta = Q^{-2}$) reads

$$\sum_{q \leq Q} \sum_{\substack{a \bmod q \\ (a, q) = 1}} \left| S\left(\frac{a}{q}\right) \right|^2 = \sum_{\substack{\frac{a}{q} \in \mathcal{F}_Q}} \left| S\left(\frac{a}{q}\right) \right|^2 \leq (\pi N + Q^2) \sum_{n=M+1}^{M+N} |a_n|^2. \quad (20.1)$$

We are now ready to prove Theorem 20.1.

Proof of Theorem 20.1. Recall that for primitive characters χ we have the expansion into additive characters

$$\chi(n) = \frac{1}{\tau(\bar{\chi})} \sum_{a \bmod q} \bar{\chi}(a) e(an/q).$$

From this and $|\tau(\chi)|^2 = q$ we have

$$\sum_{\chi(q)^*} \left| \sum_{n=M+1}^{M+N} a_n \chi(n) \right|^2 = \frac{1}{q} \sum_{\chi(q)^*} \left| \sum_{a \bmod q} \bar{\chi}(a) S\left(\frac{a}{q}\right) \right|^2 \leq \frac{1}{q} \sum_{\chi(q)} \left| \sum_{a \bmod q} \bar{\chi}(a) S\left(\frac{a}{q}\right) \right|^2.$$

In the last inequality we have forgotten the condition that χ be primitive.⁷ Expanding the right-hand side, we find that

$$\begin{aligned} \frac{1}{q} \sum_{\chi(q)} \left| \sum_{a \bmod q} \bar{\chi}(a) S\left(\frac{a}{q}\right) \right|^2 &= \frac{1}{q} \sum_{\chi(q)} \sum_{a \bmod q} \bar{\chi}(a) S\left(\frac{a}{q}\right) \sum_{b \bmod q} \chi(b) S\left(\frac{-b}{q}\right) \\ &= \frac{1}{q} \sum_{a \bmod q} \sum_{b \bmod q} S\left(\frac{a}{q}\right) S\left(\frac{-b}{q}\right) \sum_{\chi(q)} \bar{\chi}(a) \chi(b) \\ &= \frac{\varphi(q)}{q} \sum_{\substack{a \bmod q \\ (a, q) = 1}} \left| S\left(\frac{a}{q}\right) \right|^2. \end{aligned}$$

The theorem now follows from (20.1). □

⁷This is sometimes called “boosting by positivity” and is frequently used to study a subset of a more “complete” family of objects. In this case the complete family is all the Dirichlet characters modulo q ; since we have the orthogonality relations this family is easier to deal with.

20.2 Bilinear forms with Dirichlet characters

In our proof of Bombieri-Vinogradov, the following estimate will be essential.

Proposition 20.4. *For any complex numbers a_m and b_n we have*

$$\sum_{q \leq Q} \frac{q}{\varphi(q)} \sum_{\chi(q)^*} \max_{u \leq MN} \left| \sum_{\substack{m \leq M \\ n \leq N \\ mn \leq u}} a_m b_n \chi(mn) \right| \\ \ll (M + Q^2)^{1/2} (N + Q^2)^{1/2} \left(\sum_{m \leq M} |a_m|^2 \right)^{1/2} \left(\sum_{n \leq N} |b_n|^2 \right)^{1/2} \log(2MN).$$

We begin with a straightforward lemma that allows us to approximate the condition $mn \leq u$ in the inner sum above.

Lemma 20.5. *For $\alpha, T > 0$ and all $x \neq \pm\alpha$ we have*

$$\begin{cases} 1 & \text{if } |x| < \alpha \\ 0 & \text{if } |x| > \alpha \end{cases} = \frac{1}{\pi} \int_{-T}^T e^{ixy} \frac{\sin \alpha y}{y} dy + O\left(\frac{1}{T|\alpha - |x||}\right).$$

Proof. Recall the integral

$$\int_{\mathbb{R}} \frac{\sin x}{x} dx = \pi$$

from the section on Mellin inversion. Using that the integral of an odd function over $\mathbb{R}$ vanishes, we have

$$\frac{1}{\pi} \int_{\mathbb{R}} e^{ixy} \frac{\sin \alpha y}{y} dy = \frac{1}{\pi} \int_{\mathbb{R}} \cos(xy) \frac{\sin \alpha y}{y} dy = \frac{1}{2\pi} \int_{\mathbb{R}} (\sin((\alpha + x)y) + \sin((\alpha - x)y)) \frac{dy}{y},$$

using a trig identity. After the obvious changes of variable, each piece of the integral evaluates to $\pm\pi$ depending on the sign of $\alpha + x$ and $\alpha - x$. So the expression above equals

$$\frac{\operatorname{sgn}(\alpha + x) + \operatorname{sgn}(\alpha - x)}{2} = \delta(x).$$

To estimate the tail of the integral, we use integration by parts to get

$$\int_T^\infty \frac{\sin \lambda x}{x} dx = \frac{\cos \lambda T}{\lambda T} + \int_T^\infty \frac{\cos(\lambda x)}{x^2} dx \ll \frac{1}{\lambda T}.$$

It follows that

$$\int_T^\infty \sin((\alpha \pm x)y) \frac{dy}{y} \ll \frac{1}{T} \max\left(\frac{1}{|\alpha - x|}, \frac{1}{|\alpha + x|}\right) \ll \frac{1}{T|\alpha - |x||}$$

as desired. □

Proof of Proposition 20.4. We first note that the condition $mn \leq u$ only depends on the value of $\lfloor u \rfloor$, so we can choose a convenient sequence of u , say $u = k + 1/2$ for $k \leq MN$. We set $\alpha = \log u$ and $x = -\log mn$ in the lemma to get

$$\begin{cases} 1 & \text{if } mn < u \\ 0 & \text{if } mn > u \end{cases} = \frac{1}{\pi} \int_{-T}^T (mn)^{-iy} \frac{\sin(y \log u)}{y} dy + O\left(T^{-1} \left| \log\left(\frac{mn}{u}\right) \right|^{-1}\right).$$

It follows that

$$\begin{aligned} & \sum_{m \leq M} \sum_{\substack{n \leq N \\ mn \leq u}} a_m b_n \chi(mn) \\ &= \frac{1}{\pi} \int_{-T}^T A(y, \chi) B(y, \chi) \frac{\sin(y \log u)}{y} dy + O\left(T^{-1} \sum_{m \leq M} \sum_{n \leq N} |a_m b_n| \left| \log\left(\frac{mn}{u}\right) \right|^{-1}\right), \end{aligned}$$

where

$$A(y, \chi) = \sum_{m \leq M} a_m \chi(m) m^{-iy}, \quad B(y, \chi) = \sum_{n \leq N} b_n \chi(n) n^{-iy}.$$

We first deal with the error term. Note that the closest mn ever gets to u is $mn = u \pm 1/2$. So we have

$$\left| \log\left(\frac{mn}{u}\right) \right| \gg \min\left(1, \left| \frac{mn}{u} - 1 \right| \right) \geq \min\left(1, \left| \frac{u \pm 1/2}{u} - 1 \right| \right) \gg \frac{1}{u} \geq \frac{1}{MN},$$

from which we get (by Cauchy-Schwarz)

$$T^{-1} \sum_{m \leq M} \sum_{n \leq N} |a_m b_n| \left| \log\left(\frac{mn}{u}\right) \right|^{-1} \ll \frac{(MN)^{3/2}}{T} \left(\sum_{m \leq M} |a_m|^2 \right)^{1/2} \left(\sum_{n \leq N} |b_n|^2 \right)^{1/2}.$$

Note that the latter expression does not depend on χ or q , so since

$$\sum_{q \leq Q} \frac{q}{\varphi(q)} \sum_{\chi(q)^*} 1 \leq \sum_{q \leq Q} q \ll Q^2,$$

the total contribution from this error term is

$$\ll Q^2 \frac{(MN)^{3/2}}{T} \left(\sum_{m \leq M} |a_m|^2 \right)^{1/2} \left(\sum_{n \leq N} |b_n|^2 \right)^{1/2}.$$

To deal with the integral, we note that

$$\sin(y \log u) \ll \min(1, |y \log u|) \ll \min(1, |y| \log(2MN)),$$

from which we get

$$\max_{u \leq MN} \left| \int_{-T}^T A(y, \chi) B(y, \chi) \frac{\sin(y \log u)}{y} dy \right| \ll \int_{-T}^T |A(y, \chi) B(y, \chi)| \min\left(\frac{1}{|y|}, \log(2MN)\right) dy.$$

By the Large Sieve for Dirichlet characters, together with Cauchy-Schwarz, we get

$$\begin{aligned}
& \sum_{q \leq Q} \frac{q}{\varphi(q)} \sum_{\chi(q)^*} |A(y, \chi) B(y, \chi)| \\
& \leq \left(\sum_{q \leq Q} \frac{q}{\varphi(q)} \sum_{\chi(q)^*} \left| \sum_{m \leq M} a_m \chi(m) m^{-iy} \right|^2 \right)^{1/2} \left(\sum_{q \leq Q} \frac{q}{\varphi(q)} \sum_{\chi(q)^*} \left| \sum_{n \leq N} b_n \chi(n) n^{-iy} \right|^2 \right)^{1/2} \\
& \ll (M + Q^2)^{1/2} (N + Q^2)^{1/2} \left(\sum_{m \leq M} |a_m|^2 \right)^{1/2} \left(\sum_{n \leq N} |b_n|^2 \right)^{1/2}.
\end{aligned}$$

The remaining integral is

$$\int_{-T}^T \min\left(\frac{1}{|y|}, \log(2MN)\right) dy \ll \int_0^1 \log(2MN) dy + \int_1^T \frac{dy}{y} \ll \log(2MN) + \log T.$$

Setting $T = (MN)^{3/2}$ we see that all of the estimates together are at most

$$(M + Q^2)^{1/2} (N + Q^2)^{1/2} \left(\sum_{m \leq M} |a_m|^2 \right)^{1/2} \left(\sum_{n \leq N} |b_n|^2 \right)^{1/2} \log(2MN),$$

which is what we wanted. □

20.3 The proof of Bombieri-Vinogradov

For convenience, we restate the theorem here.

Theorem 20.6 (Bombieri-Vinogradov). *For $\sqrt{x}(\log x)^{-A} \leq Q \leq \sqrt{x}$ we have*

$$\sum_{q \leq Q} \max_{y \leq x} \max_{\substack{a \bmod q \\ (a, q) = 1}} \left| \psi(y; a \bmod q) - \frac{y}{\varphi(q)} \right| \ll Q \sqrt{x} (\log x)^5. \quad (20.2)$$

Recall that

$$\psi(y; a \bmod q) = \frac{1}{\varphi(q)} \sum_{\chi(q)} \bar{\chi}(a) \psi(y, \chi),$$

where

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

If we let

$$\psi'(y, \chi) = \begin{cases} \psi(y, \chi) & \text{if } \chi \neq \chi_0, \\ \psi(y, \chi) - y & \text{if } \chi = \chi_0, \end{cases}$$

then we can write

$$\psi(y; a \bmod q) - \frac{y}{\varphi(q)} = \frac{1}{\varphi(q)} \sum_{\chi} \bar{\chi}(a) \psi'(y, \chi).$$

It follows that

$$\max_{a(q)} \left| \psi(y; a \bmod q) - \frac{y}{\varphi(q)} \right| \leq \frac{1}{\varphi(q)} \sum_{\chi} |\psi'(y, \chi)|.$$

For each χ , let χ' denote the unique primitive character inducing χ . Then (recall that)

$$|\psi'(y, \chi) - \psi'(y, \chi')| \leq \sum_{p|q} \sum_{p^m \leq y} \log p \ll (\log y)(\log q) \ll (\log qy)^2.$$

Thus

$$\max_{y \leq x} \max_{a(q)} \left| \psi(y; a \bmod q) - \frac{y}{\varphi(q)} \right| \ll (\log qx)^2 + \frac{1}{\varphi(q)} \sum_{\chi} \max_{y \leq x} |\psi'(y, \chi')|.$$

Each primitive character χ modulo q induces characters to moduli kq for $k \geq 1$; thus the left-hand side of (20.2) is

$$\ll Q(\log Qx)^2 + \sum_{q \leq Q} \sum_{\chi(q)^*} \max_{y \leq x} |\psi'(y, \chi)| \sum_{k \leq Q/q} \frac{1}{\varphi(kq)}.$$

The first term is much smaller than the desired upper bound, so we will ignore it from now on. For the second term, we use

$$\varphi(kq) = kq \prod_{p|kq} (1 - p^{-1}) \geq kq \prod_{p|k} (1 - p^{-1}) \prod_{p|q} (1 - p^{-1}) = \varphi(k) \varphi(q),$$

to get the upper bound (using $Q \leq x$)

$$\sum_{k \leq Q/q} \frac{1}{\varphi(kq)} \leq \frac{1}{\varphi(q)} \sum_{k \leq x} \frac{1}{\varphi(k)} \leq \frac{1}{\varphi(q)} \prod_{p \leq x} \sum_{k=0}^{\infty} \frac{1}{\varphi(p^k)}.$$

For the inner sum,

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

from which we get

$$\begin{aligned} \prod_{p \leq x} \sum_{k=0}^{\infty} \frac{1}{\varphi(p^k)} &= \exp \left(\sum_{p \leq x} \log \left(1 + \frac{p}{(p-1)^2} \right) \right) \leq \exp \left(\sum_{p \leq x} \left(\frac{1}{p} + O\left(\frac{1}{p^2}\right) \right) \right) \\ &\ll \exp(\log \log x) = \log x \end{aligned}$$

by Mertens' second theorem. Thus we have shown that

$$\sum_{q \leq Q} \max_{y \leq x} \max_{\substack{a \bmod q \\ (a, q)=1}} \left| \psi(y; a \bmod q) - \frac{y}{\varphi(q)} \right| \ll (\log x) \sum_{q \leq Q} \frac{1}{\varphi(q)} \sum_{\chi(q)^*} \max_{y \leq x} |\psi'(y, \chi)|. \quad (20.3)$$

Note that we have essentially traded the complete sum over all $\chi \bmod q$ for the restricted sum over primitive characters, at the cost of one factor of $\log x$.

We split the sum on q into two ranges. In the initial range $q \leq (\log x)^4$ we use the Siegel-Walfisz theorem $|\psi'(x, \chi)| \ll x \exp(-c\sqrt{\log x})$ to obtain

$$\sum_{q \leq (\log x)^4} \frac{1}{\varphi(q)} \sum_{\chi(q)^*} \max_{y \leq x} |\psi'(y, \chi)| \ll x \exp(-c\sqrt{\log x}) (\log x)^4 \ll x (\log x)^{4-A} \ll x^{1/2} Q (\log x)^4 \quad (20.4)$$

for $Q \gg \sqrt{x} (\log x)^{-A}$. So it suffices to prove the analogous upper bound in the second range $(\log x)^4 \leq q \leq Q$. We first prove the following upper bound.

Proposition 20.7. *For all $x \geq 2$ and $Q \geq 1$ we have*

$$\sum_{q \leq Q} \frac{q}{\varphi(q)} \sum_{\chi(q)^*} \max_{y \leq x} |\psi'(y, \chi)| \ll (x + x^{5/6} Q + x^{1/2} Q^2) (\log Q x)^4.$$

Proof. Suppose first that $x < Q^2$. Then Proposition 20.4 with $M = 1$, $a_1 = 1$, $N = x$, and $b_n = \Lambda(n)$ gives

$$\sum_{q \leq Q} \frac{q}{\varphi(q)} \sum_{\chi(q)^*} \max_{u \leq x} \left| \sum_{n \leq u} \Lambda(n) \chi(n) \right| \ll Q(x + Q^2)^{1/2} \left(\sum_{n \leq x} \Lambda(n)^2 \right)^{1/2} \log x \ll Q^2 x^{1/2} \log^2 x,$$

using the crude bound $\Lambda(n) \ll \log x$.

Now suppose that $Q^2 \leq x$. We will apply the following idea of Vaughan. For any functions $F(s)$ and $G(s)$ we have the formal identity

$$-\frac{\zeta'}{\zeta}(s) = F(s) - \zeta(s)F(s)G(s) - \zeta'(s)G(s) + \left(-\frac{\zeta'}{\zeta}(s) - F(s) \right) (1 - \zeta(s)G(s)).$$

We apply this identity with the Dirichlet polynomials

$$F(s) = \sum_{m \leq U} \frac{\Lambda(m)}{m^s}, \quad G(s) = \sum_{d \leq V} \frac{\mu(d)}{d^s},$$

where $\mu(d)$ is the Mobius function (the Dirichlet coefficients of $1/\zeta(s)$). By comparing Dirichlet coefficients on each side of the identity, we find that

$$\Lambda(n) = a_1(n) + a_2(n) + a_3(n) + a_4(n),$$

where

$$\begin{aligned} a_1(n) &= \begin{cases} \Lambda(n) & \text{if } n \leq U, \\ 0 & \text{otherwise,} \end{cases} \\ a_2(n) &= - \sum_{\substack{md|n \\ m \leq U \\ d \leq V}} \Lambda(m) \mu(d), \\ a_3(n) &= \sum_{\substack{d|n \\ d \leq V}} \mu(d) \log(n/d), \\ a_4(n) &= - \sum_{\substack{mk=n \\ m > U \\ k > 1}} \Lambda(m) \sum_{\substack{d|k \\ d \leq V}} \mu(d). \end{aligned}$$

It follows that

$$\psi(y, \chi) = \sum_{n \leq y} \chi(n) \Lambda(n) = \sum_{1 \leq j \leq 4} \sum_{n \leq y} \chi(n) a_j(n) =: \sum_{1 \leq j \leq 4} S_j(y, \chi, U, V).$$

So our goal is now to bound the sums

$$\sum_{q \leq Q} \frac{q}{\varphi(q)} \sum_{\chi(q)^*} \max_{y \leq x} |S_j(y, \chi, U, V)|$$

under the conditions $Q \leq \sqrt{x}$ and $UV \leq x$.

($j = 1$) This is the simplest one.

$$\sum_{q \leq Q} \frac{q}{\varphi(q)} \sum_{\chi(q)^*} |S_1(y, \chi, U, V)| = \sum_{q \leq Q} \frac{q}{\varphi(q)} \sum_{\chi(q)^*} \left| \sum_{n \leq \min(y, U)} \chi(n) \Lambda(n) \right| \leq \sum_{n \leq U} \Lambda(n) \ll UQ^2$$

by the (weak version of the) prime number theorem.

($j = 4$) We begin by writing

$$S_4 = - \sum_{n \leq y} \chi(n) \sum_{\substack{mk=n \\ m > U \\ k > 1}} \Lambda(m) \sum_{\substack{d|k \\ d \leq V}} \mu(d) = - \sum_{U < m \leq y/V} \Lambda(m) \sum_{V < k \leq y/m} \chi(mk) \sum_{\substack{d|k \\ d \leq V}} \mu(d).$$

The lower bound on k (and thus the upper bound on m) comes from the fact that $\sum_{d|k, d \leq V} \mu(d) = 0$ if $k > 1$ and $k \leq V$ (if $k \leq V$ then all of the divisors of k are automatically $\leq V$). We restrict to dyadic blocks $m \in [M, 2M]$ and write the contribution from S_4 in that range in the form of the bilinear estimate to get

$$\begin{aligned} & \sum_{q \leq Q} \frac{q}{\varphi(q)} \sum_{\chi(q)^*} \max_{y \leq x} \left| \sum_{\substack{U < m \leq y/V \\ M \leq m \leq 2M}} \Lambda(m) \sum_{V < k \leq y/m} \chi(mk) \sum_{\substack{d|k \\ d \leq V}} \mu(d) \right| \\ & \ll (Q^2 + M)^{1/2} (Q^2 + x/M)^{1/2} \left(\sum_{m=M}^{2M} \Lambda(m)^2 \right)^{1/2} \left(\sum_{k \leq x/M} d(k)^2 \right)^{1/2} \log x. \end{aligned}$$

In order to get this, we take

$$\begin{aligned} a_m &= \begin{cases} \Lambda(m) & \text{if } m \in [M, 2M] \cap [U, y/V] \\ 0 & \text{otherwise} \end{cases} \\ b_n &= \begin{cases} \sum_{\substack{d|k \\ d \leq V}} \mu(d) & \text{if } k > V, \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

We have the bound

$$\sum_{M \leq m \leq 2M} \Lambda(m)^2 \leq \log(2M) \sum_{M \leq m \leq 2M} \Lambda(m) \ll M \log M$$

from the prime number theorem and the bound

$$\sum_{k \leq z} d(k)^2 \ll z \log z$$

which you can prove using the fact that $d(k) = \sum_{d|k} f(d)$, where f is the multiplicative function defined by $f(p^r) = 2r + 1$. For the sake of brevity I will skip the details here. It follows that

$$\begin{aligned} \sum_{q \leq Q} \frac{q}{\varphi(q)} \sum_{\chi(q)^*} \max_{y \leq x} \left| \sum_{\substack{U < m \leq y/V \\ M \leq m \leq 2M}} \Lambda(m) \sum_{V < k \leq y/m} \chi(mk) \sum_{\substack{d|k \\ d \leq V}} \mu(d) \right| \\ \ll (Q^2 x^{1/2} + QxM^{-1/2} + Qx^{1/2}M^{1/2} + x)(\log x)^3. \end{aligned}$$

Summing over the dyadic intervals, we find that

$$\begin{aligned} \sum_{q \leq Q} \frac{q}{\varphi(q)} \sum_{\chi(q)^*} \max_{y \leq x} |S_4(y, \chi, U, V)| \\ \leq \sum_{\substack{j=0 \\ M=2^j U}}^{\log_2(x/UV)} \sum_{q \leq Q} \frac{q}{\varphi(q)} \sum_{\chi(q)^*} \max_{y \leq x} \left| \sum_{\substack{U < m \leq y/V \\ M \leq m \leq 2M}} \Lambda(m) \sum_{V < k \leq y/m} \chi(mk) \sum_{\substack{d|k \\ d \leq V}} \mu(d) \right| \\ \ll \sum_{\substack{j=0 \\ M=2^j U}}^{\log_2(x/UV)} (Q^2 x^{1/2} + QxM^{-1/2} + Qx^{1/2}M^{1/2} + x)(\log x)^3 \\ \ll (Q^2 x^{1/2} + QxU^{-1/2} + QxV^{-1/2} + x)(\log x)^4. \end{aligned}$$

($j = 2$) We write the sum S_2 as

$$\begin{aligned} -S_2 &= \sum_{n \leq y} \sum_{\substack{md|n \\ m \leq U \\ d \leq V}} \chi(n) \Lambda(m) \mu(d) = \sum_{\substack{m \leq U \\ d \leq V}} \Lambda(m) \mu(d) \sum_{r \leq y/md} \chi(rmd) \\ &= \sum_{t \leq UV} \sum_{\substack{md=t \\ m \leq U \\ d \leq V}} \Lambda(m) \mu(d) \sum_{r \leq y/t} \chi(rt) = S'_2 + S''_2, \end{aligned}$$

where S'_2 comprises the terms $t \leq U$ and S''_2 comprises the terms $U < t \leq UV$. For S''_2 , we proceed as in ($j = 4$), considering dyadic intervals $t \in [T, 2T]$ and applying the

bilinear estimate with

$$a_t = \begin{cases} \sum_{\substack{md=t \\ m \leq U \\ d \leq V}} \Lambda(m)\mu(d) & \text{if } t \in [T, 2T] \cap [U, UV], \\ 0 & \text{otherwise,} \end{cases}$$

$$b_r = 1 \quad \text{for } 1 \leq r \leq x/T.$$

We obtain

$$\sum_{q \leq Q} \frac{q}{\varphi(q)} \sum_{\chi(q)^*} \max_{y \leq x} \left| \sum_{T \leq t \leq 2T} \sum_{1 \leq r \leq x/T} a_t b_r \chi(rt) \right|$$

$$\ll (Q^2 + T)^{1/2} \left(Q^2 + \frac{x}{T} \right)^{1/2} \left(\sum_{T \leq t \leq 2T} |a_t|^2 \right)^{1/2} \left(\frac{x}{T} \right)^{1/2} \log x.$$

For the a_t sum we have

$$|a_t| \leq \sum_{m|t} \Lambda(m) = \log t,$$

from which it follows that

$$\left(\sum_{T \leq t \leq 2T} |a_t|^2 \right)^{1/2} \leq \left(\sum_{T \leq t \leq 2T} \log^2 t \right)^{1/2} \ll \sqrt{T} \log T.$$

Summing the dyadic intervals as before, we find that

$$\sum_{q \leq Q} \frac{q}{\varphi(q)} \sum_{\chi(q)^*} \max_{y \leq x} |S_2''(y, \chi, U, V)| \ll (Q^2 x^{1/2} + QxU^{-1/2} + Qx^{1/2}U^{1/2}V^{1/2} + x)(\log x)^3.$$

For the sum S_2' we estimate a_t as before to get

$$|S_2'| \ll \sum_{t \leq U} |a_t| \left| \sum_{r \leq y/t} \chi(r) \chi(t) \right| \ll \log U \sum_{t \leq U} \left| \sum_{r \leq y/t} \chi(r) \right|.$$

In the case $q = 1$, we have

$$\log U \sum_{t \leq U} \left| \sum_{r \leq y/t} \chi(r) \right| \leq y \log U \sum_{t \leq U} \frac{1}{t} \ll y(\log U)^2,$$

and when $q > 1$ we apply the Polya-Vinogradov inequality to get

$$\log U \sum_{t \leq U} \left| \sum_{r \leq y/t} \chi(r) \right| \ll q^{1/2} U (\log qU)^2.$$

Thus, using that $Q \leq x^{1/2}$ and that $U \leq UV \leq x$ we get

$$\sum_{q \leq Q} \frac{q}{\varphi(q)} \sum_{\chi(q)^*} \max_{y \leq x} |S_2'(y, \chi, U, V)| \ll (x + Q^{5/2}U)(\log x)^2.$$

($j = 3$) We use the same sorts of ideas for S_3 :

$$S_3 = \sum_{n \leq y} \chi(n) \sum_{\substack{hd=n \\ d \leq V}} \mu(d) \log h = \sum_{d \leq V} \mu(d) \sum_{h \leq y/d} \chi(hd) \log h.$$

By partial summation and Polya-Vinogradov we have (for $q > 1$)

$$\sum_{h \leq y/d} \chi(h) \log h = \log(y/d) \sum_{h \leq y/d} \chi(h) - \int_1^{y/d} \left(\sum_{h \leq t} \chi(t) \right) \frac{dt}{t} \ll q^{1/2} (\log qy)^2,$$

while for $q = 1$ we have

$$\sum_{h \leq y/d} \chi_0(h) \log h \ll (\log y)y/d.$$

Together these yield

$$\sum_{d \leq V} (q^{1/2} (\log qy)^2 + (\log y)y/d) \ll q^{1/2} V (\log qy)^2 + y (\log yV)^2.$$

We end up with the same bound as we got with S'_2 :

$$\sum_{q \leq Q} \frac{q}{\varphi(q)} \sum_{\chi(q)^*} \max_{y \leq x} |S_3(y, \chi, U, V)| \ll (x + Q^{5/2}U) (\log x)^2.$$

We put all four cases together to get

$$\begin{aligned} & \sum_{q \leq Q} \frac{q}{\varphi(q)} \sum_{\chi(q)^*} \max_{y \leq x} |\psi'(y, \chi)| \\ & \ll (Q^2 x^{1/2} + x + QxU^{-1/2} + QxV^{-1/2} + Qx^{1/2}U^{1/2}V^{1/2} + Q^{5/2}U + Q^{5/2}V) (\log x)^4. \end{aligned}$$

If $x^{1/3} \leq Q \leq x^{1/2}$ then we take $U = V = x^{2/3}Q^{-1}$ to get

$$\sum_{q \leq Q} \frac{q}{\varphi(q)} \sum_{\chi(q)^*} \max_{y \leq x} |\psi'(y, \chi)| \ll (Q^2 x^{1/2} + Q^{3/2} x^{2/3} + x^{7/6}) (\log x)^4 \ll Q^2 x^{1/2} (\log Qx)^4.$$

If $Q \leq x^{1/3}$ then we take $U = V = x^{1/3}$ to get

$$\begin{aligned} & \sum_{q \leq Q} \frac{q}{\varphi(q)} \sum_{\chi(q)^*} \max_{y \leq x} |\psi'(y, \chi)| \ll (Q^2 x^{1/2} + x + Qx^{5/6} + Q^{5/2} x^{1/3}) (\log Qx)^4 \\ & \ll (Q^2 x^{1/2} + x + Qx^{5/6}) (\log Qx)^4. \end{aligned}$$

In each case this is less than the desired bound, so we win. $\square$

Recall that we need to show that

$$\sum_{(\log x)^A \leq q \leq x^{1/2}} \frac{1}{\varphi(q)} \sum_{\chi(q)^*} \max_{y \leq x} |\psi'(y, \chi)| \ll x^{1/2} Q (\log x)^4.$$

To this end, consider the dyadic sum

$$\sum_{U \leq q \leq 2U} \frac{1}{\varphi(q)} \sum_{\chi(q)^*} \max_{y \leq x} |\psi'(y, \chi)| \ll \frac{1}{U} \sum_{U \leq q \leq 2U} \frac{q}{\varphi(q)} \sum_{\chi(q)^*} \max_{y \leq x} |\psi'(y, \chi)|.$$

By the proposition with $Q = 2U$, the latter expression is

$$\ll \left(\frac{x}{U} + x^{5/6} + x^{1/2} U \right) (\log Ux)^4.$$

Summing the dyadic pieces in the interval $[(\log x)^A, Q]$ the total error is

$$\ll (x(\log x)^{-A} + x^{5/6} + x^{1/2} Q) (\log x)^4 \ll Q x^{1/2} (\log x)^4$$

since $Q \in [x^{1/2}(\log x)^{-A}, x^{1/2}]$. Putting this together with (20.3) and (20.4) we finish the proof of Bombieri-Vinogradov. $\square$