Analytic Number Theory

Introduction

These are notes I wrote up from my study of analytic field theory. I give proofs of the Prime Number Theorem and PNT for primes in arithmetic progressions, with error bounds. I paid extra attention to working out the asymptotics in detail.

Please pardon the appearance of Chapter A. It contains some background on characters needed for PNT for arithmetic progressions, is ripped off from a final paper, and is not very well-integrated. Chapter 2 is incomplete, but contains the statement of all theorems on Dirichlet series needed in the proof of PNT. The proofs of the Prime Number Theorems (Chapters 3 and 4) are complete.

This is open-source (and under a creative commons license), so feel free to change it to your liking. The notes and tex files can be found at `http://web.mit.edu/~holden1/www/math/notes.htm`. In particular, if you find mistakes, add the missing sections, or edit the material, please email me at `holden1@mit.edu` and I'll update the document and give you credit on my website. (After all, writing one more section is a lot easier than writing up the whole proof again, right?)

Recent update (7/29/2012): Finally added the section on the Siegel zero 4.5 (an improved version of PNT for arithmetic progressions).

Contents

1 Crash course in complex analysis **1**
- 1 Holomorphic functions . . . 1
- 2 Complex integration . . . 2
- 3 Cauchy's Theorem . . . 3
- 4 Power series and Laurent series . . . 4
  - 4.1 Cauchy's residue formula . . . 5
- 5 Convergence . . . 6
- 6 Series and product developments . . . 6
- 7 Gamma function . . . 8

2 Dirichlet series **11**
- 1 Dirichlet series, convergence . . . 11
- 2 Basic properties . . . 12
- 3 Dirichlet generating functions . . . 14
- 4 Summing coefficients . . . 15

3 Zeta functions and the prime number theorem **17**
- 1 Prime number theorem: Outline . . . 17
  - 1.1 The big picture . . . 18
  - 1.2 Main steps . . . 18
- 2 Riemann zeta function . . . 20
- 3 Zeros of zeta . . . 24
- 4 Prime number theorem: proof . . . 29
- 5 The Riemann hypothesis . . . 34

4 L-functions and Dirichlet's theorem **37**
- 1 Outline . . . 37
- 2 L-functions . . . 38
- 3 Zeros of L . . . 45
- 4 Prime number theorem in arithmetic progressions . . . 49
- 5 Siegel zero . . . 53
  - 5.1 $L'(\beta, \chi)$ is not too large . . . 54
  - 5.2 $L(1, \chi)$ is not too small . . . 55

5.3 Proof of Siegel-Walfisz 58

A Arithmetic over Finite Fields **61**
1 Characters 61
1.1 Dirichlet characters 63
1.2 Characters on finite fields 64
2 Gauss Sums 65
3 Enumerating Solutions 67
4 Applications to Waring's Problem 69
5 Finite calculus 70

Chapter 1

Crash course in complex analysis

complex-analysis Complex analysis is calculus on the complex numbers. The main functions of study are complex differentiable functions.

Reference books: Lang or Ahlfors

1 Holomorphic functions

Definition 1.1.1: Let $U \subseteq \mathbb{C}$ be an open set and $f : U \to \mathbb{C}$ be a function. The **derivative** of f is

$$f'(z) := \lim_{\Delta z \to 0} \frac{f(z + \Delta z) - f(z)}{\Delta z}$$

if it exists. f is **holomorphic** if its derivative exists at every point of U. f is **meromorphic** if it is defined and holomorphic on U except at a discrete set of points.

Write $f(x + iy) = u(x, y) + iv(x, y)$. Note that f being differentiable is a much *stronger* condition than being simply u and v being differentiable, because the limit of f as $\Delta z \to 0$ along the real and complex directions must be equal:

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial x} = \frac{1}{i}\left(\frac{\partial u}{\partial y} + i\frac{\partial v}{\partial y}\right).$$

Thus we get the Cauchy-Riemann criteria: If f is differentiable as a function of (x, y), then f is holomorphic iff

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}.$$

Another way to think about complex differentiability is that holomorphic maps preserve angles (i.e. are *conformal*); we have

$$f(z + re^{i\theta}) - f(z) \approx re^{i\theta} f'(z).$$

Because complex differentiability is such a strong property, holomorphic functions have many nice properties. Hence it is often useful to take functions defined on the reals and

extend them as far as possible on $\mathbb{C}$. Some of the good properties are the following (to be explained in the rest of the chapter); note they are not necessarily true for real differentiable functions!

- A function is holomorphic iff it is analytic (has a power series expansion).
- A sequence of holomorphic functions with good convergence properties converges to a holomorphic function.
- A bounded entire function is constant.
- If two holomorphic functions agree on a set containing a limit point, then they are equal. Thus analytic continuations are unique.
- Bounds on a function give bounds on the derivative. Hence we can "differentiate" asymptotic formulas.
- We can expand holomorphic functions into products or sums depending on their poles and zeros—in much the same way that rational functions can be expanded into partial fractions or factored.

2 Complex integration

We now give two definitions of the integral.

Definition 1.2.1: A **path** is a continuous function $\gamma : [a, b] \to \mathbb{C}$. It is called a **loop** if $\gamma(a) = \gamma(b)$. Let f be a holomorphic function on U and γ be a path in U.

1. If γ is differentiable (except possibly at a finite number of points), define
$$\int_\gamma f(z)\,dz = \int_a^b f(\gamma(t))\gamma'(t)\,dt.$$
2. Define an (indefinite) **integral** of f on a set V to be a function F on V such that $F'(z) = f(z)$. Given holomorphic f, choose points $t_0, \ldots, t_n$ such that there exist open sets $U_j \supseteq f(\gamma([t_{j-1}, t_j]))$ such that f has an integral F_j on U_j. Define
$$\int_\gamma f(z)\,dz = \sum_{k=1}^{n} [F_j(\gamma(t_j)) - F_j(\gamma(t_{j-1})].$$

Note that unlike in the real case, indefinite integrals may not exist globally, for example, $\ln t$ is locally an integral for $\frac{1}{t}$ but cannot be extended holomorphically to $\mathbb{C}\backslash\{0\}$. We need to establish the well-definedness of the second definition.

Theorem 1.2.2 (Cauchy's Theorem, version 1). cauchy1 *Let f be holomorphic on a closed rectangle R, with boundary ∂R. Then (using the first definition),*

$$\int_{\partial R} f = 0.$$

From this one can show that integrals exist locally by defining

$$F(z) = \int_{z_0}^{z} f(s)\, ds$$

where the integral is along horizontal and vertical lines; moreover one gets well-definedness in the second definition.

We can now define the logarithm of a function.

Definition 1.2.3: Let f be a holomorphic function on a simply connected set U (see Definition 1.3.1), with $f(z) \neq 0$ on U. Choose $z_0 \in U$ and a_0 such that $e^{a_0} = z_0$.

$$(\ln f)(z) = \int_{z}^{z_0} \frac{f'}{f}(z)\, dz.$$

Note different definitions of the logarithm will differ by integer multiples of $2\pi i$, and $e^{(\ln f)(z)} = f(z)$. The motivation comes from the fact that one would expect the derivative of $\ln f(z)$ to be $\frac{f'}{f}(z)$. We write $(\ln f)(z)$ to emphasize that this is *not* simply a composite of functions: We could have $f(z_1) = f(z_2)$ but $(\ln f)(z_1) \neq (\ln f)(z_2)$.[1]

We seek a generalization of Theorem 1.2.2 to meromorphic functions and arbitrary paths.

3 Cauchy's Theorem

Definition 1.3.1: homotopic Two paths γ and $\eta : [a, b] \to \mathbb{C}$ are **homotopic** if there exists a continuous map

$$\gamma_s(t) : [0, 1] \times [a, b] \to \mathbb{C}$$

such that $\gamma_0(t) = \gamma(t)$ and $\gamma_1(t) = \eta(t)$.

A subset of $\mathbb{C}$ is simply connected if it is pathwise connected and every loop in $\mathbb{C}$ is homotopic to a point.

Theorem 1.3.2. *Let U be a simply connected open set containing z_0. Every path γ around z_0 in $U\backslash\{z_0\}$ is homotopic to a circle going around z_0 n times for some $n \in \mathbb{Z}$. This n can be calculated by*

$$n = W(\gamma, z_0) := \frac{1}{2\pi i} \int_{\gamma} \frac{f(z)}{z - z_0}\, dz$$

and is called the ***winding number****.*

[1] Consider, for example, the case where $f(z) = z^2$ on $\mathbb{C}\backslash\mathbb{R}_{\leq 0}$, and $z_1 = i$, $z_2 = -i$.

Theorem 1.3.3 (Global Cauchy's formula)**.** *Let U be a simply connected open set and $f : U \to \mathbb{C}$ be holomorphic. Suppose γ is a loop in U. Then*

$$\frac{1}{2\pi i}\int_\gamma \frac{f(z)}{z-z_0}\,dz = W(\gamma, z_0)f(z_0).$$

4 Power series and Laurent series

As complex differentiability is a much stronger condition than differentiability for real functions, holomorphic functions enjoy nicer properties. The most important one is the following.

Definition 1.4.1: A function $f : U \to \mathbb{C}$ is **analytic** at z_0 if it can be written as a power series in a neighborhood around z_0:

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n.$$

If f is given by its power series representation then we must have $a_n = \frac{f^{(n)}(z)}{n!}$.

Theorem 1.4.2. *A function $f : U \to \mathbb{C}$ is analytic iff and only iff it is holomorphic.*

Note this is not true for real functions: for example, $e^{-\frac{1}{x^2}}$ has Taylor expansion equal to 0 at 0, but is not the zero function. This kind of irregularity does not happen for holomorphic functions.

Corollary 1.4.3. *A holomorphic function has infinitely many derivatives.*

The following theorem says that for holomorphic functions, the radius of convergence is "as large as it could possibly be."

Theorem 1.4.4. radius-convergence *Suppose f is holomorphic on a disc $N_r(z_0)$ of radius r around z_0. Then the Taylor series around z_0 converges absolutely to f on $N_r(z_0)$.*

Proof. Estimate coefficients using Cauchy's theorem. Complex Analysis, Lang III.7.3. □

We can generalize power series to allow terms with negative exponents.

Theorem 1.4.5. *Suppose f is defined on an annulus $A = \{z : r < |z - z_0| < R\}$. Let C be the circle of radius $r' \in (r, R)$ around z_0. Then f has a Laurent expansion on A:*

$$f(z) = \sum_{n=-\infty}^{\infty} a_n (z - z_0)^n, \quad a_n = \frac{1}{2\pi i}\int_C \frac{f(z)}{(z-z_0)^{n+1}}\,dz.$$

If f is defined on $\{z : |z - z_0| < R\}$ then

$$f^{(n)}(z_0) = \frac{n!}{2\pi i}\int_C \frac{f(z)}{(z-z_0)^{n+1}}\,dz.$$

The coefficient a_{-1} is called the **residue** of f at z_0:

$$\mathrm{Res}_{z_0}(f) = a_{-1}.$$

The following theorem controls the size of the derivatives of a complex analytic function by its values of the function in a circle. Note that in the real analytic case we can't make such a statement!

Corollary 1.4.6. *cor:cauchy-ineq Suppose f is defined on $\{z : |z - z_0| < R\}$, and let C be a circle of radius $r < R$ around z_0. Then*

$$|f^{(n)}(z)| \leq \frac{n!}{r^n} \max_{z \in C} f(z)$$

and the nth coefficient in the power series expansion satisfies

$$a_n \leq \frac{1}{r^n} \max_{z \in C} f(z).$$

Proof. Simply note that in the integral $\int_C \frac{f(z)}{(z-z_0)^{n+1}}\, dz$, the denominator has constant absolute value r^{n+1}, the numerator is bounded by $\max_{z \in C} f(z)$, and the arc length is $2\pi r$. □

Corollary 1.4.7 (Liouville)**.** *A bounded entire function is constant.*

Proof. We can take $r \to \infty$ in the inequality for $n = 1$ to find that $f'(z) = 0$ everywhere. □

4.1 Cauchy's residue formula

Using residues, we can state the most comprehensive form of Cauchy's formula:

Theorem 1.4.8 (Residue formula)**.** *residue Suppose f is meromorphic on simply connected open U, and γ is a loop in U. Then*

$$\int_\gamma f(s)\, ds = 2\pi i \sum_{z \text{ pole of } f} W(\gamma, z)\, \mathrm{Res}_z(f).$$

One useful application of this is counting zeros and poles of a function f.

Definition 1.4.9: Define the **order** of f at z_0 to be the least integer so that the Laurent expansion of f at z_0 has $a_m \neq 0$:

$$\mathrm{ord}_f(z_0) = m.$$

Note that $\mathrm{ord}_f(z_0) > 0$ signals a zero and $\mathrm{ord}_f(z_0) < 0$ signals a pole.

Corollary 1.4.10. *Suppose f is meromorphic on simply connected open U, and γ is a loop in U. Then*

$$\frac{1}{2\pi i} \int_\gamma \frac{f'}{f}(s)\, ds = \sum_\rho W(\gamma, \rho)\, \mathrm{ord}_f(\rho).$$

Proof. If f has Laurent expansion $a_m(z-z_0)^m + \cdots$ at z_0 then $\frac{f'}{f}$ has Laurent expansion

$$\frac{ma_m(z-z_0)^{m-1} + \cdots}{a_m(z-z_0)^m + \cdots} = m(z-z_0)^{-1} + \cdots$$

□

5 Convergence

Unlike in the real case, holomorphic functions behave nicely under infinite sums and pointwise convergence. This is because by Cauchy's theorem we can write f as an integral, and integrals preserve convergence.

Theorem 1.5.1 (Holomorphic functions converge to holomorphic functions)**.** *Let $\{f_n\}_{n=1}^{\infty}$ be a sequence of holomorphic functions on U.*

1. *Suppose $f_n \to f$ uniformly on compact subsets of U. Then f is holomorphic.*
2. *Suppose $\sum_{n=1}^{\infty} f_n = f$ converges absolutely and uniformly on compact subsets of U. Then f is holomorphic.*

6 Series and product developments

We know that locally, we can write a meromorphic function f as a Laurent series $\sum_{n=-\infty}^{\infty} a_n x^n$. There are two other representations that are useful, depending on what information we have about the function f.

1. If we know the *poles* of f, we can write f as a sum of rational functions

$$f(z) = \sum_{n=1}^{\infty} \left[P_n\left(\frac{1}{z-z_n}\right) - Q_n(z) \right] + g(z).$$

2. If f is entire and we know the *zeros* of f, we can write f as an infinite product

$$f(z) = z^m e^{g(z)} \prod_{n=1}^{\infty} \left(1 - \frac{z}{z_n}\right) e^{P_n\left(\frac{z}{z_n}\right)}.$$

(Think of this as "factoring" f, much like a polynomial can be factored as in the fundamental theorem of algebra.) These representations come about from convergence properties of holomorphic functions—so we can be sure the infinite products converge to holomorphic functions—and by Liouville's theorem—if we engineer a function that is close enough to f then it must be equal to f.

Theorem 1.6.1 (Mittag-Leffler). *Let z_n be a sequence with $\lim_{n\to\infty} |z_n| = \infty$ (or a finite sequence), and P_n polynomials without constant term.*

1. *(Existence) There is a meromorphic function f with poles exactly at z_n, with Laurent expansion $P_n\left(\frac{1}{z-z_n}\right) + \cdots$ at z_n.*
2. *(Uniqueness) All such functions f are in the form*

$$\sum_{n=1}^{\infty} \left(P_n\left(\frac{1}{z-z_n}\right) - Q_n(z) \right) + g(z)$$

where Q_n is a polynomial and $g(z)$ is analytic.

Proof. See Ahlfors [Ahl79, p. 187]. □

Warning: this does not converge for all P_n. Typically we take Q_n to the the first terms of the Laurent expansion of $P_n\left(\frac{1}{z-z_n}\right)$, to ensure cancellation of high-order terms.

Definition 1.6.2: The **order** of an entire function f is the smallest $\alpha \in [0, \infty]$ such that

$$|f(z)| \precsim_\varepsilon e^{|z|^{\alpha+\varepsilon}}$$

for all $\varepsilon > 0$.

Theorem 1.6.3 (Product development). product-development *Let z_n be a sequence with $\lim_{n\to\infty} |z_n| = \infty$. If f is entire with order $\alpha < \infty$ with zeros $z_1, z_2, \ldots$ (with multiplicity, not including 0), then it has a product formula*

$$\text{product-formula} f(z) = z^r e^{g(z)} \prod_{n=1}^{\infty} \left(1 - \frac{z}{z_n}\right) e^{\frac{z}{z_n} + \frac{1}{2}\left(\frac{z}{z_n}\right)^2 + \cdots + \frac{1}{m}\left(\frac{z}{z_m}\right)^m}, \tag{1.1}$$

where

- *$m = \lfloor \alpha \rfloor$,*
- *r is the order of vanishing of f at 0, and*
- *g is a polynomial of degree at most a.*

The product converges uniformly locally. Moreover,

$$\text{num-zeros} |\{k : z_k < R\}| \precsim_\varepsilon R^{\alpha+\varepsilon}. \tag{1.2}$$

Conversely, if $a = \lfloor \alpha \rfloor$ and z_k is a sequence satisfying (1.2), then the RHS of (1.1) defines an entire function of order at most α.

Proof. See Ahlfors [Ahl79, p. 195]. □

Hence the order of a entire function gives an asymptotic bound for the number of zeros.[2]

[2] A function which grows faster is allowed to have more zeros—much like a polynomial with lots of zeros grows fast simply because it has higher degree.

7 Gamma function

To prove basic properties of the zeta function in the next chapter, we need to know the properties of the gamma function.

Definition 1.7.1: Define the **gamma function** by

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

We will begin by analytically continuing the gamma function and giving its basic properties.

Proposition 1.7.2 (Facts about Γ)**:** gamma-facts

1. $\Gamma(s)$ can be analytically continued to a meromorphic function with poles $-n, n \in \mathbb{N}$, with residue $\frac{(-1)^n}{n!}$.
2. $\Gamma(s) = \lim_{n\to\infty} \frac{n^s n!}{s(s+1)\cdots(s+n)}$ when $s \notin -\mathbb{N}$.
3. $\frac{1}{\Gamma(s)} = se^{Cs} \prod_{n=1}^\infty \left(1 + \frac{s}{n}\right) e^{-\frac{s}{n}}$.
4. $\Gamma(s+1) = s\Gamma(s)$ so $\Gamma(n+1) = n!, n \in \mathbb{N}_0$.
5. $\Gamma(s)\Gamma(1-s) = \frac{\pi}{\sin \pi s}$.
6. $\Gamma(s)\Gamma\left(s + \frac{1}{m}\right) \cdots \Gamma\left(s + \frac{m-1}{m}\right) = (2\pi)^{\frac{m-1}{2}} m^{\frac{1}{2}-ms}\Gamma(ms)$. In particular, $\Gamma(s)\Gamma\left(s + \frac{1}{2}\right) = \pi^{\frac{1}{2}} 2^{1-2s}\Gamma(2s)$.

From the product development 1.6.3 we get the following.

Theorem 1.7.3 (Product development of Γ)**.** gamma-product-development *We have*

$$\Gamma(s) = \frac{e^{-\gamma s}}{s} \prod_{k=1}^\infty \frac{e^{\frac{s}{k}}}{1 + \frac{s}{k}}.$$

In the region

$$R_\varepsilon = \mathbb{C}\backslash(\{s : \arg(s) \in [\pi - \varepsilon, \pi + \varepsilon]\} \cup \{0\}),$$

i.e. $\mathbb{C}$ with a wedge containing $\mathbb{R}_{\leq 0}$ deleted, we can define the function $(\ln \Gamma)(s)$. By the product formula, it equals

$$(\ln \Gamma)(s) = -\gamma s - \ln s + \sum_{k=1}^\infty \left(\frac{s}{k} - \ln\left(1 + \frac{s}{k}\right)\right).$$

The following asymptotic formulas will be useful.

Theorem 1.7.4 (Stirling's approximation)**.** stirling *Let* $P_1(t) = \{t\} - \frac{1}{2}$*. For* $s \in R_\varepsilon$*,*

$$\begin{aligned}(\ln\Gamma)(s) &= \left(s-\frac{1}{2}\right)\ln s - s + \frac{1}{2}\ln(2\pi) - \int_0^\infty \frac{P_1(t)}{z+t} \\ &= \left(s-\frac{1}{2}\right)\ln s - s + \frac{1}{2}\ln(2\pi) + O_\varepsilon(|s|^{-1}) \\ \frac{\Gamma'(s)}{\Gamma(s)} &= \ln s - \frac{1}{2s} + O_\varepsilon(|s|^{-2}) \\ \Gamma(s) &\sim s^{s-\frac{1}{2}}e^{-s}\sqrt{2\pi}\end{aligned}$$

Chapter 2

Dirichlet series

dirichlet For proofs see [Apo94].

1 Dirichlet series, convergence

Dirichlet series are the "power series of number theory." As such, we will first need to get acquainted with their analytic properties.

Definition 2.1.1: A **Dirichlet series** is a series of the form

$$F(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s}$$

where $f(n)$ is an arithmetical function. Following convention, we let $s = \sigma + it$, with σ, t real.

Let $\{\lambda(n)\}$ be a sequence strictly increasing to ∞. A **general Dirichlet series** with exponents $\{\lambda(n)\}_{n=1}^{\infty}$ is in the form

$$F(s) = \sum_{n=1}^{\infty} f(n)e^{-s\lambda(n)}.$$

An ordinary Dirichlet series has $\lambda(n) = \ln(n)$. [1]

Theorem 2.1.2 (Half-plane of convergence)**.** ***Convergence:*** *If the series $\sum_{n=1}^{\infty} |f(n)e^{-s\lambda(n)}|$ does not converge or diverge for all n, then there exists a real number σ_c, called the* ***abscissa of convergence****, such that $\sum_{n=1}^{\infty} f(n)n^{-s}$*

- *converges locally uniformly for $\sigma > \sigma_c$, but*
- *does not converge for $\sigma < \sigma_c$.*

[1] A further generalization is given by the Laplace-Stieltjes transform, $\int_0^{\infty} e^{-st} d\alpha(t)$, where α is a measure. The "step" part of α gives a Dirichlet while the continuous part gives a Laplace transform.

In fact, if the series diverges for all s with $\sigma < 0$, then

$$\sigma_c = \limsup_{n\to\infty} \frac{\ln |\sum_{k=1}^{n} a(k)|}{\lambda(n)}.$$

Absolute convergence: *If the series $\sum_{n=1}^{\infty} |e^{-s\lambda(n)}|$ does not converge or diverge for all n, then there exists a real number σ_a, called the* ***abscissa of absolute convergence****, such that $\sum_{n=1}^{\infty} f(n)n^{-s}$*

- *converges locally uniformly absolutely for $\sigma > \sigma_a$, but*
- *does not converge absolutely for $\sigma < \sigma_a$.*

In fact, if the series diverges for all s with $\sigma < 0$, then

$$\sigma_a = \limsup_{n\to\infty} \frac{\ln \sum_{k=1}^{n} |a(k)|}{\lambda(n)}.$$

In particular, for ordinary Dirichlet series (that diverge when $\sigma < 0$),

$$\sigma_a = \limsup_{n\to\infty} n^{\sum_{k=1}^{n} |a(k)|}.$$

2 Basic properties

Proposition 2.2.1 (General facts)**:** Let $F(s) = \sum_{n=1}^{\infty} f(n)n^{-s}$.

1. $\lim_{\sigma\to\infty} F(\sigma + it) = f(1)$ uniformly
2. (Uniqueness) If $F(s) = G(s)$ are absolutely convergent for $\sigma > \sigma_a$ and are equal for s in an infinite sequence $\{s_k\}$ with $\sigma_k \to \infty$, then $f(n) = g(n)$.
3. (Non-vanishing in half-plane) Suppose $F(s) \neq 0$ for some s with $\sigma > \sigma_a$. Then there is a half-place $\sigma > c \geq \sigma_a$ in which $F(s)$ is never 0.

Proposition 2.2.2: (Operations on Dirichlet series)oper-on-dir Let $F(s) = \sum_{n=1}^{\infty} f(n)n^{-s}$ and $G(s) = \sum_{n=1}^{\infty} g(n)n^{-s}$. Then

$$F(s)G(s) = \sum_{n=1}^{\infty} \frac{h(n)}{n^s}$$

where

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

Proof. Formally, by grouping together terms where mn is constant,

$$\begin{aligned} F(s)G(s) &= \sum_{m,n\in\mathbb{N}} \frac{f(m)}{n^s}\frac{g(n)}{n^s} \\ &= \sum_{k=1}^{n} \left(\sum_{m,n\in\mathbb{N},\, mn=k} f(m)g(n) \right) \frac{1}{k^s}. \end{aligned}$$

Since the sums for F and G converge absolutely, so does the double sum above, and the rearrangement of terms is valid. □

Theorem 2.2.3 (Euler products)**.** euler-product *Let f be a multiplicative arithmetical function such that $\sum_{n=1}^{\infty} f(n)$ converges absolutely. Then when $\Re s > \sigma_a$,*

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

If f is completely multiplicative,

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

Proposition 2.2.4 (Derivatives)**:** dir-derivative The derivative is

$$F'(s) = -\sum_{n=1}^{\infty} \frac{f(n)\ln n}{n^s}.$$

Theorem 2.2.5 (Landau)**.** landau *Suppose $F(s)$ is a holomorphic function that can be represented in $\sigma > c$ by the Dirichlet series*

$$F(s) = \sum_{n=1}^{\infty} f(n)n^{-s}$$

with $f(n) \geq 0$ for all $n \geq n_0$. If $F(s)$ is analytic in some disc of radius r around $s = c$, then $F(s)$ converges in $\sigma > \sigma - \varepsilon$ for some $\varepsilon > 0$.

Hence, $F(s)$ has a singularity at $s = \sigma_c$.

Proof. We reinterpret in terms of power series and apply Theorem 1.4.4.

Take $a = c + \frac{r}{2}$. Since F is analytic at in $N_r(a) \subseteq N_r(c) \cup \{z : \Re z > c\}$, it equals its Taylor expansion there:

$$F(s) = \sum_{k=1}^{\infty} \frac{F^{(k)}(a)}{k!}(s-a)^k.$$

From Proposition 2.2.4, $F^{(k)}(a) = (-1)^k \sum_{n=1}^{\infty} f(n)(\ln n)^k n^{-s}$. Plugging in and noting that the sum converges absolutely (since $f(n) \geq 0$ for large n), we have, for $s \in N_r(a)$,

$$\begin{aligned} F(s) &= \sum_{k=0}^{\infty} \left[\left(\frac{(-1)^k}{k!} \sum_{n=1}^{\infty} f(n)(\ln n)^k n^{-a} \right) (s-a)^k \right] \\ &= \sum_{n=1}^{\infty} \left[\left(\sum_{k=0}^{\infty} \frac{(s-a)^k (\ln n)^k}{k!} \right) n^{-a} \right] \\ &= \sum_{n=1}^{\infty} f(n) e^{(a-s)\ln n} n^{-a}. \end{aligned}$$

This converges for $c - \varepsilon \in N_r(a)$. But because it has nonnegative real coefficients, this shows $\sigma_c > c - \varepsilon$. □

Proposition 2.2.6 (Logarithms)**:** Assume $f(1) \neq 0$. if $F(s) \neq 0$ for $\sigma > \sigma_0 \geq \sigma_a$, then for $\sigma > \sigma_0$,

$$\ln F(s) = \ln f(1) + \sum_{n=1}^{\infty} \frac{f' * f^{-1}(n)}{\ln n} n^{-s}.$$

Also talk about log diff of Euler product

3 Dirichlet generating functions

Definition 2.3.1: Let $f : \mathbb{N} \to \mathbb{C}$ be an arithmetic function. The **Dirichlet generating function** of f is

$$F(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s}.$$

To get the generating function of $g(n) = \sum_{d|n} f(n)$, by Proposition 2.2.2, we simply multiply by $\zeta(s)$:

$$F(s)\zeta(s) = \left(\sum_n \frac{f(n)}{n^s} \right) \left(\sum_n \frac{1}{n^s} \right) = \sum_n \left(\sum_{d|n} f(d) \right) \frac{1}{n^s}.$$

Note that the inverse of $\zeta(s)$ is

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

Hence by matching coefficients of

$$(F(s)\zeta(s)) \frac{1}{\zeta(s)}$$

we get the Mobius inversion formula.

Table of dgf's here

4 Summing coefficients

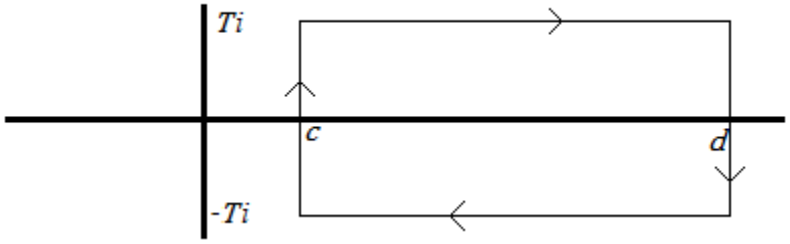
Lemma 2.4.1. *Dir-Mellin* *For $y, c, T > 0$,*[2]

$$\begin{aligned}
\left|\frac{1}{2\pi i}\int_{c-iT}^{c+iT} y^s \frac{ds}{s}\right| &\le y^c \min\left(\frac{1}{\pi T|\ln y|}, \frac{1}{2}\right), && 0 < y < 1\\
\left|\frac{1}{2\pi i}\int_{c-iT}^{c+iT} y^s \frac{ds}{s} - \frac{1}{2}\right| &\le \frac{y^c}{\pi T}, && y = 1\\
\left|\frac{1}{2\pi i}\int_{c-iT}^{c+iT} y^s \frac{ds}{s} - 1\right| &\le y^c \min\left(\frac{1}{\pi T|\ln y|}, 1\right), && y > 1
\end{aligned}$$

Proof. First suppose $y < 1$. Take $d > c$. By Cauchy's theorem, since $\frac{y^s}{s}$ is analytic in the region below, we have

$$\int_{c-iT}^{c+iT} y^s \frac{ds}{s} + \int_{c+iT}^{d+iT} y^s \frac{ds}{s} + \int_{d+iT}^{d-iT} y^s \frac{ds}{s} + \int_{d-iT}^{c-iT} y^s \frac{ds}{s} = 0$$

where the path of integrations are those shown in the picture.



Hence,

$$\begin{aligned}
\left|\int_{c-iT}^{c+iT} y^s \frac{ds}{s}\right| &= \left|\int_{c+iT}^{d+iT} y^s \frac{ds}{s} + \int_{d-iT}^{c-iT} y^s \frac{ds}{s} + \int_{d+iT}^{d-iT} y^s \frac{ds}{s}\right|\\
&\le 2\int_c^d y^\sigma \frac{d\sigma}{T} + \left|\int_{d+iT}^{d-iT} y^s \frac{ds}{s}\right|.
\end{aligned}$$

Note that the last integral goes to 0 as $d \to \infty$, because $|y^s| = |y^d| \to 0$. Hence, taking $d \to \infty$ gives

$$\left|\int_{c-iT}^{c+iT} y^s \frac{ds}{s}\right| \le 2\int_c^\infty \frac{y^\sigma}{T}\, d\sigma = -\frac{2y^c}{T\ln y} = \frac{2y^c}{T|\ln y|}.$$

This gives $\left|\frac{1}{2\pi i}\int_{c-iT}^{c+iT} y^s \frac{ds}{s}\right| \le \frac{y^c}{\pi T}|\ln y|$.

[2]The integral $\frac{1}{2\pi i}\int_{c-\infty i}^{c+\infty i} f(s)\frac{ds}{s}$ is called the *Mellin transform* of f.

By Cauchy's theorem applied to the smaller segment bounded by $\Re s = c$ and the circle with radius $R = \sqrt{c^2 + T^2}$, (picture) we have

$$\left|\int_{c-iT}^{c+iT} y^s \frac{ds}{s}\right| = \left|\int_C y^s \frac{ds}{s}\right|$$
$$\leq \pi R \frac{y^c}{R} = \pi y^c,$$

since $y < 1$ and $\Re s > c$ on the arc. Hence $\left|\frac{1}{2\pi i}\int_{c-iT}^{c+iT} y^s \frac{ds}{s}\right| \leq \frac{y^c}{2}$.

For $y > 1$, take $d < 0$. Note $\frac{y^s}{s}$ is analytic in the region below except for a simple pole at 0 with residue 1 (since $y^s = 1$ when $s = 0$). Hence by Cauchy's Theorem,

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

[INSERT PICCY]

Then

$$\left|\int_{c-iT}^{c+iT} y^s \frac{ds}{s} - 1\right| = \left|\int_{c+iT}^{d+iT} y^s \frac{ds}{s} + \int_{d-iT}^{c-iT} y^s \frac{ds}{s} + \int_{d+iT}^{d-iT} y^s \frac{ds}{s}\right|$$
$$\leq 2\int_d^c y^\sigma \frac{d\sigma}{T} + \left|\int_{d+iT}^{d-iT} y^s \frac{ds}{s}\right|.$$

The last term goes to 0 as $d \to -\infty$, so the same argument applies as in the first part to show $\left|\frac{1}{2\pi i}\int_{c-iT}^{c+iT} y^s \frac{ds}{s} - 1\right| \leq \frac{y^c}{\pi T \ln y}$.

By Cauchy's theorem applied to the larger segment bounded by $\Re s = c$ and the circle with radius $R = \sqrt{c^2 + T^2}$, (picture) we have

$$\int_{c-iT}^{c+iT} y^s \frac{ds}{s} + \int_C y^s \frac{ds}{s} = 2\pi i$$
$$\left|\int_{c-iT}^{c+iT} y^s \frac{ds}{s} - 1\right| \leq \left|\int_C y^s \frac{ds}{s}\right|$$
$$\leq 2\pi R \frac{y^c}{R} = 2\pi y^c,$$

since $y > 1$ and $\Re s < c$ on the arc. Hence $\left|\frac{1}{2\pi i}\int_{c-iT}^{c+iT} y^s \frac{ds}{s}\right| \leq y^c$.

Proof for $y = 1$ omitted. □

Corollary 2.4.2. *sum-coeff-Dir The partial sum of the coefficients of a Dirichlet series is given by*

$$\sum_{n<x} a_n + \frac{a_x}{2}(x \in \mathbb{N}_0) = \frac{1}{2\pi i}\lim_{T\to\infty}\int_{c-iT}^{c+iT} x^s f(s)\frac{ds}{s}.$$

The error from truncating the integral is

$$\left|\left(\sum_{n<x} a_n + \frac{a_x}{2}(x \in \mathbb{N}_0)\right) - \left(\frac{1}{2\pi i}\int_{c-iT}^{c+iT} x^s f(s)\frac{ds}{s}\right)\right| \leq \sum_{n=1}^{\infty}\left(\frac{x}{n}\right)^c a_n \min\left(1, \frac{1}{T\left|\ln\left(\frac{x}{n}\right)\right|}\right).$$

Chapter 3

Zeta functions and the prime number theorem

zeta-l-pnt

1 Prime number theorem: Outline

Definition 3.1.1: Define the prime-counting function

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

Our goal in this chapter is to prove the following famous theorem (in all its error-bounded glory).

Theorem 3.1.2 (Prime number theorem)**.** pnt *There is an effective constant $C > 0$ such that*

$$\pi(x) = li(x) + O(xe^{-C\sqrt{\ln x}})$$

for all $x \geq 1$.

Here $\mathrm{li}(x)$ denotes the **logarithmic integral**

$$\mathrm{li}(x) = \int_2^x \frac{dt}{\ln t}.$$

Note that $\mathrm{li}(x) = \frac{x}{\ln x} + O\left(\frac{x}{(\ln x)^2}\right)$ as $x \to \infty$, since integration by parts gives

$$\begin{aligned} \mathrm{li}(x) &= \int_2^x \frac{dy}{\ln y} + O(1) = \frac{x}{\ln x} + \int_2^x \frac{dy}{(\ln y)^2} + O(1) \\ &= \frac{x}{\ln x} + O\left(\frac{x}{(\ln x)^2}\right). \end{aligned} \tag{3.1}$$

li-ibp

1.1 The big picture

We recommend Andrew Granville's article IV.2 Analytic Number Theory in [GBGL10] for an overview.

How might we guess at the asymptotics for $\pi(x)$? (In particular, why is it closer to $\mathrm{li}(x)$ than $\frac{x}{\ln x}$?) By studying tables of primes up to 3 million, Gauss hypothesized that the density of primes at around x is around $\frac{1}{\ln x}$, and hence that the number of primes up to x would be the integral $\mathrm{li}(x) = \int_2^x \frac{dt}{\ln t}$. Making a table of $\pi(x)$ and the difference $\mathrm{li}(x) - \pi(x)$, we find that the difference is slightly more than on the order of $\sqrt{x}$, so this seems to be a good estimate.

It is a common theme in analytic number theory to make conjectures about the distribution of primes (or other subsets of interest) by assuming they are randomly distributed according to some probability model. Often a simple model works for simple asymptotics up to x, and the model needs to be refined or corrected when dealing with more complicated quantities such as number of primes in a small interval, or spacing between primes.

Model 3.1.3 (Gauss-Cramér model)**:** For $n \geq 3$, let X_n be the random variable such that

$$X_n = 1 \text{ with probability } \frac{1}{\ln n}$$
$$X_n = 0 \text{ with probability } 1 - \frac{1}{\ln n}.$$

Then the sequence X_n behaves similarly to the sequence

$$a_n = 1 \text{ if } n \text{ is prime}$$
$$a_n = 0 \text{ otherwise.}$$

The Gauss-Cramér model exactly predicts $\pi(x) \sim \mathrm{li}(x)$. The model gives more than just the asymptotics of $\pi(x)$, though, it can also be used to think about primes in short intervals $\pi(x+y) - \pi(x)$.

Problem 3.1.4: What are the shortcomings of the Gauss-Cramér model?

1.2 Main steps

The main steps in the proof are as follows.

1. When we have a Dirichlet series

$$F(s) = \sum_{n=0}^{\infty} a_n n^{-s},$$

we can get estimates for $\sum_{n=0}^{N} a_n$ by "plucking out" those coefficients: The equation

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

gives

$$\frac{1}{2\pi i}\lim_{T\to\infty}\int_{c-iT}^{c+iT} x^s f(s)\frac{ds}{s} = \sum_{n<x} a_n + \frac{a_x}{2}(x\in\mathbb{N}_0).$$

We use the more precise statement giving error bounds (Corollary 1).

We want a Dirichlet series where the sum of the first N terms is related to $\pi(N)$. Let

$$\zeta(s) = \prod_{p \text{ prime}}\frac{1}{1-p^{-s}} = \sum_{n=1}^{\infty}\frac{1}{n^s}.$$

We consider the function

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

We use this function because $\psi(x) := \sum_{n<x}\Lambda(n)$ gives information on $\pi(x)$, and $-\frac{\zeta'}{\zeta}$ continues into a meromorphic function on $\mathbb{C}$ (since ζ does). We now have the estimate

$$\psi(x) = \frac{1}{2\pi i}\int_{c-iT}^{c+iT} -\frac{\zeta'(s)}{\zeta(s)}x^s\frac{ds}{s} + (\text{error}).$$

2. We know ζ has analytic continuation (Theorem 3.2.2). Hence we can move the path of integration to $c<0$. From Cauchy's integral formula, we get extra terms from the horizontal integrals (integrals involving $-\frac{\zeta'}{\zeta}$) and terms $\frac{x^\rho}{\rho}$ from Cauchy's integral theorem from the zeros of $\zeta(s)$. *This is why we care about its zeros!* Zeros with large real part contribute large error terms. We will need the following.

   (a) We apply the product development (Theorem 1) on $\xi(s) = \pi^{-\frac{s}{2}}\zeta(s)\Gamma\left(\frac{s}{2}\right)$ to obtain

   $$\frac{\zeta'(s)}{\zeta(s)} = \sum_{\rho \text{ zero of } \zeta}\left(\frac{1}{s-\rho}+\frac{1}{\rho}\right)+\cdots$$

   (Theorem 3.2.5).

   (b) Using the above equation for $\frac{\zeta'}{\zeta}$, we calculate the asymptotics of $N(T)$, the number of zeros in $\{\sigma+it : (\sigma,t)\in[0,1]\times[-T,T]\}$ (Theorem 3.3.2).

   (c) From (a) to (b) we get a zero-free region for ζ (which includes $\Re s\geq 1$) (Theorems 3.3.1 and 3.3.3).

From the zero-free region we get a bound for $\sum \frac{x^\rho}{\rho}$, as well as the horizontal integrals. If the Riemann hypothesis is true, then we can enlarge our zero-free region to $\Re s > \frac{1}{2}$, which is even better.

3. Finally we use the estimate for $\psi(x)$ to get an estimate for $\pi(x)$ (Lemma 3.4.2).

2 Riemann zeta function

Definition 3.2.1: The **Riemann zeta function** is defined by

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

when $\Re s > 1$. This will be generalized to L-functions $L(s, \chi)$ in Definition 1.

By Theorem 2.2.3 and by unique factorization in $\mathbb{Z}$, we can write

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

By taking the logarithmic derivative, we have

$$-\frac{\zeta'(s)}{\zeta(s)} = \sum_p \frac{d}{ds} \ln(1 - p^{-s}) = \sum_p (\ln p) \frac{p^{-s}}{1 - p^{-s}} = \sum_p \ln p \sum_{k=1}^{\infty} p^{-ks}.$$

Interchanging order of summation gives

$$\text{log-diff-zeta} \quad -\frac{\zeta'(s)}{\zeta(s)} = \sum_{n=1}^{\infty} \Lambda(n) n^{-s}, \quad \Re s > 1, \tag{3.2}$$

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

$$\Lambda(n) = \begin{cases} \ln p, & n = p^r,\ p \text{ prime},\ r \in \mathbb{N}. \\ 0, & \text{else} \end{cases}$$

The most important property of ζ is its analytic continuation and functional equation.

Theorem 3.2.2 (Analytic continuation and functional equation for ζ)**.** *zeta-continues* *$\zeta(s)$ can be analytically continued to a meromorphic function with a simple pole at $s = 0, 1$. It satisfies the functional equation*

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

Letting $\xi(s) = \pi^{-\frac{s}{2}}\zeta(s)\Gamma\left(\frac{s}{2}\right)$, *we have*[1]

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

Moreover, $\zeta(s)$ *has zeros* $-2\mathbb{N}$ *(the trivial zeros); all other zeros are in the critical strip* $0 \le \Re s \le 1$.

To prove this, we first need the transformation law for the theta function; we will show the functional equation for ζ by writing it in terms of θ. As we will prove a more generalized transformation law, we will postpone the proof for θ.

Definition 3.2.3: Define the **theta function** by

$$\theta(u) = \sum_{n\in\mathbb{Z}} e^{-\pi n^2 u}, \quad \Re u > 0.$$

Proposition 3.2.4 (Transformation law for θ)**:** theta-law For all u with $\Re u > 0$,

$$\theta\left(\frac{1}{u}\right) = u^{\frac{1}{2}}\theta(u).$$

This is a special case of Proposition 4.2.4.

Proof of Theorem 3.2.2. We first analytically continue ζ to $\Re s > 0$, show the functional equation is true for $0 < \Re s < 1$, and use it to establish analytic continuation to $\mathbb{C}$.

Note

$$\text{continue-zeta-to-0}\quad \zeta(s) = \frac{1}{s-1} + \sum_{n=1}^{\infty}\left[n^{-s} - \int_n^{n+1} x^{-s}\,dx\right] = \frac{1}{s-1} + \sum_{n=1}^{\infty}\int_n^{n+1}(n^{-s} - x^{-s})\,dx \quad (3.3)$$

Since for $n \le x \le n+1$ we have

$$|n^{-s} - x^{-s}| = \left|\int_n^x sx^{-s-1}\,dx\right| \le |s|n^{-s-1}$$

$$\text{bound-zeta-summands}\quad \left|\int_n^{n+1} n^{-s} - x^{-s}\,dx\right| \le |s|n^{-s-1}, \quad (3.4)$$

the sum (3.3) converges uniformly locally for $\Re s > 0$ and extends ζ to an analytic function for $\Re s > 0$.

We claim that

$$\text{zeta-theta}\quad 2\xi(s) = \int_0^{\infty}(\theta(u) - 1)u^{\frac{s}{2}}\frac{du}{u}, \quad \Re s > 1 \quad (3.5)$$

[1]The factor $\Gamma\left(\frac{s}{2}\right)$ can be thought of as coming from the infinite place—see Chapter **??**.

Indeed, we have

$$\begin{aligned}
\int_0^\infty (\theta(u)-1)u^{\frac{s}{2}}\frac{du}{u} &= \int_0^\infty 2\sum_{n=1}^\infty e^{-\pi n^2 u}u^{\frac{s}{2}}\frac{du}{u} \\
&= 2\sum_{n=1}^\infty \int_0^\infty e^{-\pi n^2 u}u^{\frac{s}{2}}\frac{du}{u} \\
&= 2\sum_{n=1}^\infty \int_0^\infty e^{-u}\left(\frac{u}{\pi n^2}\right)^{\frac{s}{2}}\frac{du}{u} && u \longmapsto \frac{u}{\pi n^2} \\
&= 2\pi^{-\frac{s}{2}}\left(\sum_{n=1}^\infty \frac{1}{n^s}\right)\left(\int_0^\infty e^{-u}u^{\frac{s}{2}}\frac{du}{u}\right) \\
&= 2\pi^{-\frac{s}{2}}\zeta(s)\Gamma\left(\frac{s}{2}\right) = 2\xi(s).
\end{aligned}$$

The theta transformation law 3.2.4 give that for $\Re s > 1$,

$$\begin{aligned}
2\xi(s) &= \int_0^1 (\theta(u)-1)u^{\frac{s}{2}}\frac{du}{u} + \int_1^\infty (\theta(u)-1)u^{\frac{s}{2}}\frac{du}{u} \\
&= \int_1^\infty \left(\theta\left(\frac{1}{u}\right)-1\right)u^{\frac{s}{2}}\frac{du}{u} + \int_1^\infty (\theta(u)-1)u^{\frac{s}{2}}\frac{du}{u} && u \longmapsto \frac{1}{u} \\
&= \int_1^\infty \left(u^{-\frac{1}{2}}\theta\left(\frac{1}{u}\right)-1\right)u^{\frac{1-s}{2}}\frac{du}{u} + \int_1^\infty (u^{\frac{1-s}{2}}-u^{-\frac{s}{2}})\frac{du}{u} + \int_1^\infty (\theta(u)-1)u^{\frac{s}{2}}\frac{du}{u} \\
&= -\frac{2}{s} - \frac{2}{1-s} + \int_1^\infty (\theta(u)-1)u^{\frac{1-s}{2}}\frac{du}{u} + \int_1^\infty (\theta(u)-1)u^{\frac{s}{2}}\frac{du}{u}.
\end{aligned}$$

The last expression converges for all $\Re s > 0$, so in fact equals $2\zeta(s)$ for all $\Re s > 0$ by uniqueness of analytic continuation. Since the last expression is symmetric under $1-s \mapsto s$, the functional equation for ξ follows.

The functional equation for ξ gives

$$\begin{aligned}
\zeta(s) &= \pi^{\frac{s}{2}}\Gamma\left(\frac{s}{2}\right)^{-1}\pi^{-\frac{1-s}{2}}\Gamma\left(\frac{1-s}{2}\right)\zeta(1-s) \\
&= \pi^{s-\frac{1}{2}}\frac{\Gamma\left(\frac{1-s}{2}\right)}{\Gamma\left(\frac{s}{2}\right)}\zeta(1-s) \\
&= \pi^{s-\frac{1}{2}}\Gamma\left(\frac{1-s}{2}\right)\Gamma\left(1-\frac{s}{2}\right)\frac{\sin\left(\frac{\pi s}{2}\right)}{\pi}\zeta(1-s) && \text{by Proposition 1.7.2(5)} \\
&= 2(2\pi)^{s-1}\sin\left(\frac{\pi s}{2}\right)\Gamma(1-s)\zeta(1-s) && \text{by Proposition 1.7.2(6)}
\end{aligned}$$

Finally, the statement about zeros follows from the fact that ζ has no zeros with $\Re s > 1$ (as $\frac{\zeta'}{\zeta}$ is holomorphic there) and the functional equation, noting $\sin\left(\frac{\pi s}{2}\right) = 0$ exactly when s is an even integer, with the zero at $s = 0$ cancelled by the pole at 1 of ζ. □

Theorem 3.2.5 (Product development of ξ). xi-product-development *The function $(s^2-s)\xi(s)$ is entire of order 1, and $\xi(s)$ has the product expansion*

$$\xi(s) = \frac{e^{A+Bs}}{s^2-s} \prod_{\rho \text{ zero of } \zeta} \left(1-\frac{s}{\rho}\right) e^{\frac{s}{\rho}}.$$

Then $\frac{\zeta'}{\zeta}(s)$ has the partial-fraction expansion

$$\frac{\zeta'}{\zeta}(s) = B - \frac{1}{s-1} + \frac{1}{2}\ln(\pi) - \frac{1}{2}\frac{\Gamma'}{\Gamma}\left(\frac{s}{2}+1\right) + \sum_{\rho \text{ nontrivial zero of } \zeta} \left(\frac{1}{s-\rho}+\frac{1}{\rho}\right).$$

From now on, unless otherwise specified, when we say zero of ζ we mean *nontrivial* zero.

Proof. Note $(s^2-s)\xi(s)$ is entire because ξ only has 2 simple poles at $0, 1$. To show it has order 1 we need two inequalities.

Step 1: There is no constant C so that $(s^2-s)\xi(s) \precsim e^{C|s|}$: Indeed, for real s and any constant C, by Stirling's approximation 1.7.4 we have

$$\begin{aligned}(s^2-s)\xi(s) &= (s^2-s)\pi^{-\frac{s}{2}}\Gamma\left(\frac{s}{2}\right)\zeta(s)\\ &\succsim s^{-\frac{1}{2}}\left(\frac{s}{2e\pi}\right)^{\frac{s}{2}} \succsim e^{Cs}.\end{aligned}$$

Step 2: There is a constant C so that $(s^2-s)\xi(s) \precsim e^{C|s|\ln|s|}$: $e^{|s|\ln|s|} \geq 1$ for all s so it suffices to prove this for sufficiently large s. By the integral and sum formulas for Γ and ξ, and the fact that $|x^s| = |x^{\Re s}|$, we have

$$|\xi(\sigma+ti)| \leq \pi^{-\frac{\sigma}{2}}\Gamma\left(\frac{\sigma}{2}\right)\zeta(\sigma), \quad \sigma > 1.$$

By symmetry of ξ is suffices to consider $\sigma \geq \frac{1}{2}$. ("Nudging" $|s|$ in $e^{C|s|\ln|s|}$ by a constant changes it by at most a constant factor.) Consider 2 cases.

1. $\sigma > 2$: Then $\pi^{-\frac{\sigma}{2}} < 1$ and $\zeta(\sigma) < \zeta(2)$ so by Stirling's approximation 1.7.4,

   $$|\xi(\sigma+ti)| \precsim \Gamma\left(\frac{\sigma}{2}\right) = e^{|(\ln\Gamma)(\sigma)|} = e^{\left(\frac{\sigma}{2}-1\right)\ln\frac{\sigma}{2}-\frac{\sigma}{2}+O(1)}$$

   from which the result follows.

2. $\frac{1}{2} \leq \sigma \leq 2$: From (3.4), we have for s bounded away from 1,

   $$\zeta(s) \leq O(1) + |s|\sum_{n=1}^{\infty} n^{-\frac{3}{2}} = O(|s|).$$

   This time $\Gamma\left(\frac{\sigma}{2}\right) = O(1)$ so

   $$|(s^2-s)\xi(s)| \leq \left|s^2\pi^{-\frac{\sigma}{2}}\zeta(s)\Gamma\left(\frac{\sigma}{2}\right)\right| = O(|s|^3) \precsim e^{C|s|\ln|s|}.$$

This shows $(s^2 - s)\xi(s)$ has order 1.

Step 3: By the product development 1.6.3, noting the the zeros of $(s^2-s)\xi$ are the nontrivial zeros of ζ (since Γ has no zeros and trivial zeros of ζ come from the poles of Γ in the definition of ξ), we get

$$(s^2 - s)\xi(s) = e^{A+Bs} \prod_{\rho \text{ zero of } \zeta} \left(1 - \frac{s}{\rho}\right) e^{\frac{s}{\rho}}.$$

Dividing by $s^2 - s$ and log-differentiating gives

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

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

$$\begin{aligned} \frac{\zeta'}{\zeta}(s) &= \frac{1}{2}\ln \pi + \frac{1}{2}\frac{\Gamma'}{\Gamma}\left(\frac{s}{2}\right) + B - \frac{1}{s} - \frac{1}{s-1} + \sum_{\rho} \left(\frac{1}{s-\rho} + \frac{1}{\rho}\right) \\ &= \frac{1}{2}\ln \pi + \frac{1}{2}\frac{\Gamma'}{\Gamma}\left(\frac{s}{2} + 1\right) + B - \frac{1}{s-1} + \sum_{\rho} \left(\frac{1}{s-\rho} + \frac{1}{\rho}\right), \qquad \Gamma(z) = \frac{\Gamma(z+1)}{z}. \end{aligned}$$

□

3 Zeros of zeta

Note that from the function equation, $\zeta(s)$ has simple zeros at $-2\mathbb{N}$. We call these trivial zeros. More importantly for us are the zeros with real part in $[0, 1]$.

Denote by $N(T)$ be the number of zeros of ζ in $\{\sigma + it : (\sigma, t) \in [0, 1] \times [-T, T]\}$, counting multiplicity. We first give asymptotics on the vertical distribution of zeros of ζ (von Mangoldt's formula, Theorem 3.3.2), then give a zero-free region for ζ (Theorem 3.3.3).

Lemma 3.3.1. weak-zeta-zeros *Define $\mathcal{L}(t) = \ln(|t|+2)$. For $s = \sigma + it$ with $\sigma \in [-1, 2]$, we have*[2]

$$\begin{aligned} \text{weak-zeta-zeros-eq1} \frac{\zeta'(s)}{\zeta(s)} &= -\frac{1}{s-1} + \sum_{\rho} \left(\frac{1}{s-\rho} + \frac{1}{\rho}\right) + O(\mathcal{L}) \qquad (3.6) \\ &= -\frac{1}{s-1} + \sum_{|\Im(s-\rho)|<1} \frac{1}{s-\rho} + O(\mathcal{L}). \end{aligned}$$

Moreover, there are $O(\mathcal{L})$ zeros ρ with $|\Im(s-\rho)| < 1$, i.e. the number of zeros with imaginary part in $[t, t+1]$ is $O(\ln t)$, as $t \to \infty$.

Note this gives $N(T) = O(T \ln T)$. The next theorem will give an improvement of this estimate.

[2] Note $\frac{1}{s-1} = O(1)$ when s is bounded away from 1.

Proof. Our strategy is this: at a point where we know $\frac{\zeta'}{\zeta}$ is bounded ($s = 2 + it$), we use Theorem 3.2.5 to get information on how many zeros of ζ can be close to s. Then we use compare $\frac{\zeta'}{\zeta}(\sigma + it)$ with $\frac{\zeta'}{\zeta}(2 + it)$ to get the general estimate.

Step 1: Theorem 3.2.5 gives us

$$\text{zeta2-zero-sum}\quad \frac{\zeta'(s)}{\zeta(s)} = -\frac{1}{s-1} + \underbrace{B + \frac{1}{2}\ln\pi}_{O(1)} - \frac{1}{2}\underbrace{\frac{\Gamma'}{\Gamma}\left(\frac{s}{2}+1\right)}_{(A)} + \underbrace{\sum_\rho \left(\frac{1}{s-\rho}+\frac{1}{\rho}\right)}_{(B)}. \tag{3.7}$$

From Stirling's approximation 1, (A) equals

$$\text{gamma2-estimate}\quad \ln\left|\frac{\sigma}{2} + 1 + i\frac{t}{2}\right| + O(1) = O(\mathcal{L}) \tag{3.8}$$

These two equations show (3.6).

Now suppose $s = 2 + it$. Note that

$$\left|\frac{\zeta'(2+it)}{\zeta(2+it)}\right| = \left|\sum_{n=1}^\infty \Lambda(n) n^{-2-it}\right| \le \left|\sum_{n=1}^\infty (\ln n) n^{-2}\right| < \infty,$$

so the LHS of (3.7) is $O(1)$. Hence (3.7) becomes

$$\text{zeta2-zero-sum2}\quad O(\mathcal{L}) = \sum_\rho \left(\frac{1}{s-\rho}+\frac{1}{\rho}\right). \tag{3.9}$$

We estimate the terms with $|\Im(s-\rho)| < 1$ by a constant to show that there aren't too many of them. From (3.9) and (3.8),

$$\begin{aligned}
O(\mathcal{L}) &= \Re \sum_\rho \left(\frac{1}{2+it-\rho}+\frac{1}{\rho}\right) \\
&\ge \Re \sum_\rho \left(\frac{(2-\Re\rho)-(t-\Im\rho)i}{(2-\Re\rho)^2+(t-\Im\rho)^2}\right) && \text{since } \Re\left(\frac{1}{\rho}\right) > 0 \\
&\ge \sum_\rho \frac{1}{4+(t-\Im\rho)^2} && \text{since } 0 \le \Re\rho \le 1 \\
\text{zero-olnt}\quad &\ge \frac{1}{5}|\{\rho : |\Im(s-\rho)| < 1\}| + \frac{1}{5}\sum_{|\Im(s-\rho)|\ge 1} \frac{1}{(t-\Im\rho)^2}.
\end{aligned} \tag{3.10}$$

This proves the second part of the lemma.

Step 2: Now we consider general $s = \sigma + it$, by comparing it to $2 + it$. We have by (3.7)

and (3.8) that

$$\begin{aligned}
&\frac{\zeta'}{\zeta}(s) - \underbrace{\frac{\zeta'}{\zeta}(2+it)}_{O(1)} \\
&= -\frac{1}{s-1} + O(1) + \underbrace{\frac{1}{2}\left(\ln\left|\frac{\sigma}{2}+1+\frac{t}{2}i\right| - \ln\left|2+\frac{t}{2}i\right|\right)}_{O(1)} + \sum_\rho \left(\frac{1}{s-\rho} - \frac{1}{2+it-\rho}\right) \\
&= -\frac{1}{s-1} + O(1) + \sum_{|\Im(s-\rho)|<1} \frac{1}{s-\rho} - \underbrace{\sum_{|\Im(s-\rho)|<1} \frac{1}{2+it-\rho}}_{O(\mathcal{L})} + \underbrace{\sum_{|\Im(s-\rho)|\geq 1} \frac{(2-\sigma)}{(s-\rho)(2+it-\rho)}}_{O(\mathcal{L})}.
\end{aligned}$$

The first $O(\mathcal{L})$ is because there are at most $O(\mathcal{L})$ terms and each term is at most 1 in absolute value; the second is from

$$\sum_{|\Im(s-\rho)|\geq 1} \frac{2-\sigma}{(s-\rho)(2+it-\rho)} = O\left(\sum_{|\Im(s-\rho)|\geq 1} \frac{1}{\Im(s-\rho)^2}\right) = O(\mathcal{L});$$

the first equality is from $2-\sigma = O(1)$ and $\Im(s-\rho) = \Im(2+it-\rho)$; the second is by (3.10). □

Theorem 3.3.2 (von Mangoldt)**.** zeta-zeros(*) *As $T \to \infty$,*

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

Proof. As ζ has only a countable number of zeros, we may assume T is not the imaginary part of any zero.

Let

$$\mathcal{R} = \{\sigma + it : (s,t) \in [-1,2] \times [-T,T]\}$$

and let C be the boundary of $\mathcal{R}$. (PICTURE) From $\xi(s) = \pi^{-\frac{s}{2}}\zeta(s)\Gamma\left(\frac{s}{2}\right)$, we see that ξ has the same zeros as ζ in this region, and simple poles at 0 and 1. Hence by Cauchy's residue formula 1,

$$\frac{1}{2\pi i}\oint_C \frac{\xi'(s)}{\xi(s)} ds = 2N(T) - 2.$$

Noting that $\xi(\overline{s}) = \overline{\xi(s)}$ and $\xi(s) = \xi(1-s)$, changes of variable show that the integral on each of the sections of C between 2, $\frac{1}{2}+iT$, -1, and $\frac{1}{2}-iT$ are the same.[3] Let C' be the part from 1 to $\frac{1}{2}+iT$. Thus the above equals

$$\begin{aligned}
\frac{2}{\pi i}\int_{C'} \frac{\xi'(s)}{\xi(s)} ds &= \frac{2}{\pi i}\int_{C'} -\frac{\ln\pi}{2} + \frac{\zeta'(s)}{\zeta(s)} + \frac{\left(\Gamma\left(\frac{s}{2}\right)\right)'}{\Gamma\left(\frac{s}{2}\right)} ds && \frac{\left(\prod_{k=1}^n f_k\right)'}{\prod_{k=1}^n f_k} = \sum_{k=1}^n \frac{f_k'}{f_k} \\
&= \frac{2}{\pi}\Im\int_{C'} -\frac{\ln\pi}{2} + \frac{\zeta'(s)}{\zeta(s)} + \frac{\left(\Gamma\left(\frac{s}{2}\right)\right)'}{\Gamma\left(\frac{s}{2}\right)} ds && \text{(expression is real)}.
\end{aligned}$$

We break this up into 3 integrals and estimate each part separately.

[3] We used ξ because its symmetry allows us to do this.

1. $\Im \int_{C'} -\frac{\ln \pi}{2}\,ds = -\frac{T}{2}\ln \pi$.

2. Using the estimate for $\frac{\zeta'}{\zeta}$ in Lemma 3.3.1, we evaluate the second integral. Note that $\ln \zeta$ is defined for $\Re s > 1$ and is uniformly bounded for $\Re s = 2$:

$$(\ln \zeta)(s) = \sum_{p \text{ prime}} \ln(1 - p^{-s})$$
$$|(\ln \zeta)(2 + it)| \le \sum_{p \text{ prime}} 2p^{-2}.$$

   (Just bound ln linearly near 1, or expand in Taylor series.) Note $\ln(x-\rho)$ is well-defined on C' for any ρ. Hence by Theorem 3.3.1,

$$\begin{aligned}
\Im \int_{C'} \frac{\zeta'}{\zeta}(s)\,ds &= (\Im(\ln \zeta)(2 + iT) - \Im(\ln \zeta)(2)) + \int_{2+iT}^{\frac{1}{2}+iT} \frac{\zeta'}{\zeta}(s)\,ds \\
&= O(1) + \int_{2+iT}^{\frac{1}{2}+iT} \Im\left(\sum_{|\Im(s-\rho)|<1} \frac{1}{s-\rho}\right) + O(\ln T)\,ds \\
&= O(\ln T) + \sum_{|\Im(s-\rho)|<1} \Im(\ln(x-\rho))\Big|_{2+Ti}^{\frac{1}{2}+Ti} \\
&\le O(\ln T) + 2\pi O(\ln T)
\end{aligned}$$

   since there are at most $\ln T$ terms in the sum.

3. We estimate the last integral using Stirling's formula 1. (Note that $\ln \Gamma$ is well-defined for $s \in C'$.)

$$\begin{aligned}
\int_{C'} \frac{\left(\Gamma\left(\frac{s}{2}\right)\right)'}{\Gamma\left(\frac{s}{2}\right)} &= \left[\Im(\ln \Gamma)\left(\frac{s}{2}\right)\right]_2^{\frac{1}{2}+Ti} \\
&= \Im(\ln \Gamma)\left(\frac{1}{4} + \frac{T}{2}i\right) \\
&= \Im\left[\left(-\frac{1}{4} + \frac{T}{2}i\right)\ln\left(\frac{1}{4} + \frac{T}{2}i\right) - \left(\frac{1}{4} + \frac{T}{2}i\right) + O(1)\right] \\
&= \frac{T}{2}\ln\left(\frac{T}{2}\right) - \frac{T}{2} + O(1). \qquad \square
\end{aligned}$$

Now put everything together to get

$$\begin{aligned}
N(T) - 2 &= \frac{2}{\pi}\left(-\frac{T}{2}\ln \pi + O(\ln T) + \left(\frac{T}{2}\ln\left(\frac{T}{2}\right) - \frac{T}{2} + O(1)\right)\right) \\
N(T) &= \frac{T}{\pi}\ln\left(\frac{T}{2\pi}\right) - \frac{T}{\pi} + O(\ln T).
\end{aligned}$$

Theorem 3.3.3 (Zero-free region for ζ). zeta-zero-free *There are no zeros of ζ with $\Re s \ge 1$. Moreover, there is a constant $c > 0$ such that for $|t| > 2$, every zero $\sigma + it$ satisfies*

$$\sigma < 1 - \frac{c}{\ln |t|}.$$

PICTURE!

Proof. We already noted ζ has no zero for $\Re s > 1$ (Theorem 3.2.2), so for the first part it suffices to prove that no zero has real part 1.

If ζ had a zero $1+it$, then $\frac{\zeta'}{\zeta}$ would have a pole of positive residue at $1+it$. For $s = \sigma+it$, $\sigma > 1$ we have $-\frac{\zeta'}{\zeta}(s) = \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s}$, so this means that as $\sigma \to 1^+$, many of the important terms would have n^{-it} "close" to -1, to make it blow up in the negative direction. For those terms, we have n^{-2it} "close" to 1. This would force $-\frac{\zeta'}{\zeta}(\sigma + 2ti)$ to have a pole of positive residue at $1 + 2ti$, i.e ζ to have a pole at $1 + 2ti$, contradicting the fact that it is analytic there.

We now make this idea precise. What we want is an inequality between some function of an angle and its double, so that if one is small it forces the other to be large. So we consider

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

This gives

$$0 \le 3 + 4\Re(n^{-it}) + \Re(n^{-2it}).$$

Multiplying by $\Lambda(n)n^{-\sigma}$ and summing, we get

zero-free-zeta-inequality
$$0 \le 3\left(-\frac{\zeta'}{\zeta}(\sigma)\right) + 4\Re\left(-\frac{\zeta'}{\zeta}(\sigma + ti)\right) + \Re\left(-\frac{\zeta'}{\zeta}(\sigma + 2ti)\right), \quad \sigma > 1. \tag{3.11}$$

Letting r be the degree of the zero at $1 + ti$, we have by Lemma 3.3.1

$$0 \le \left(\frac{3}{\sigma - 1} + O(1)\right) - \left(\frac{4r}{\sigma - 1} + O(\mathcal{L})\right) + \Re\left(-\frac{\zeta'}{\zeta}(\sigma + 2ti)\right) \text{ as } \sigma \to 1^+.$$

If $r \ge 1$, then this gives $-\frac{\zeta'}{\zeta}(\sigma + 2ti) \to \infty$ as $\sigma \to 1^+$, contradiction. Hence $r = 0$; $1 + it$ is not a zero.

For the second statement, we have to use the partial fraction decomposition 3.2.5. Suppose $\rho = (1 - \delta) + it$ is a zero. By Lemma 3.3.1, we have

$$-\frac{\zeta'(s)}{\zeta(s)} = O(\ln|t|) - \sum_{\rho}\left(\frac{1}{s - \rho} + \frac{1}{\rho}\right) \le O(\ln|t|) - \frac{1}{s - \rho}.$$

Then

$$-\Re\frac{\zeta'}{\zeta}(\sigma + ti) \le O(\ln|t|) - \frac{1}{\sigma + \delta - 1}$$
$$-\Re\frac{\zeta'}{\zeta}(\sigma + 2ti) \le O(\ln|2t|) = O(\ln|t|).$$

For $\sigma > 1$, plugging this into (3.11) gives

$$0 \le \frac{3}{\sigma - 1} + O(\ln|t|) - \frac{4}{\sigma + \delta - 1}$$
$$\implies \frac{4}{\sigma + \delta - 1} < \frac{3}{\sigma - 1} + C_1 \ln|t|$$

for some C_1. Now take $\sigma = 1 + 4\delta$ to get

$$\frac{4}{5\delta} < \frac{3}{4\delta} + C_1 \ln |t|,$$

giving

$$\delta > \frac{1}{20 C_1 \ln |t|}$$

as needed. □

4 Prime number theorem: proof

Now we gather everything together to prove the prime number theorem. We first show the following.

Theorem 3.4.1 (von Mangoldt's formula). *von-Mangoldt-formula For an integer $x > 2$ and $x \geq T$,*

$$\psi(x) = x - \sum_{|\Im(\rho)|<T} \frac{x^\rho}{\rho} + O\left(\frac{x(\ln x)^2}{T}\right). \tag{3.12}$$

Proof. Step 1: We estimate $\psi(x)$ using Theorem 2.4.2. Suppose x is an integer; the theorem gives

$$\begin{aligned}\left|\psi(x) - \left(\int_{c-iT}^{c+iT} x^s \left(-\frac{\zeta'}{\zeta}(s)\frac{ds}{s}\right)\right)\right| &\leq \Lambda(x) + \sum_{n\geq 1, n\neq x} \left(\frac{x}{n}\right)^c \Lambda(n) \frac{1}{T\left|\ln\left(\frac{x}{n}\right)\right|} \\ &\leq \ln(x) + \sum_{n\geq 1, n\neq x} \left(\frac{x}{n}\right)^c \frac{\ln(n)}{T\left|\ln\left(\frac{x}{n}\right)\right|}.\end{aligned}$$

Take

$$c = 1 + \frac{1}{\ln x}.$$

Note that this makes $x^c = ex = O(x)$. To estimate the sum we split it into several parts.

1. $1 \leq n < \frac{x}{e}$: We have

$$\begin{aligned}\sum_{1\leq n<\frac{x}{e}} \left(\frac{x}{n}\right)^c \frac{\ln n}{T\left|\ln\left(\frac{x}{n}\right)\right|} &\asymp \frac{x \ln x}{T} \sum_{1\leq n<x} \frac{1}{n} \\ &\sim \frac{x(\ln x)^2}{T}.\end{aligned}$$

2. $\frac{x}{e} \le n < ex$: We have

$$\begin{aligned}
\sum_{\frac{x}{e}\le n<ex,\, n\neq x} \left(\frac{x}{n}\right)^c \ln n \frac{1}{T\left|\ln\left(\frac{x}{n}\right)\right|} &\precsim \sum_{\frac{x}{e}\le n<ex,\, n\neq x} \cancel{e^{1+\frac{1}{\ln x}}} \frac{\ln n}{T\left|\ln\left(\frac{x}{n}\right)\right|} \\
&\precsim \frac{1}{T} \sum_{\frac{x}{e}\le n<ex,\, n\neq x} \frac{\ln x}{\left|1-\frac{x}{n}\right|} && \text{using } \ln x \sim x-1 \text{ when } x \approx 1 \\
&\precsim \frac{x\ln x}{T} \sum_{\frac{x}{e}\le n<ex, n\neq x} \frac{1}{|n-x|} \\
&\precsim \frac{x\ln x}{T} \sum_{1\le n<(e-1)x} \frac{1}{n} \\
&\sim \frac{x(\ln x)^2}{T}.
\end{aligned}$$

3. $n \ge ex$: We have

$$\begin{aligned}
\sum_{n\ge ex} \left(\frac{x}{n}\right)^c \frac{\ln n}{T} &< \frac{x}{T}\int_{ex-1}^{\infty} \frac{\ln y}{y^c}\,dy && \frac{\ln y}{y^c} \text{ decreasing for } y > e \\
&= \frac{x}{T}\left[\frac{-y^{-c+1}\ln y}{c-1} - \frac{y^{-c+1}}{(c-1)^2}\right]_{ex-1}^{\infty} \\
&\sim \frac{x(\ln x)^2}{T}.
\end{aligned}$$

Putting everything together gives

$$\text{von-M-1}\quad \left|\psi(x) - \left(\int_{c-iT}^{c+iT} x^s \left(-\frac{\zeta'}{\zeta}(s)\right)\frac{ds}{s}\right)\right| = O\left(\frac{x(\ln x)^2}{T} + \ln x\right). \tag{3.13}$$

Step 2: We move the line of integration to $\Re s = -1$. Assuming that T is not the imaginary part of any root, by Cauchy's residue theorem 1.4.8 PICTURE

$$\int_{c-iT}^{c+iT} \frac{x^s}{s}\frac{\zeta'}{\zeta}(s)\,ds + \underbrace{\int_{c+iT}^{-1+iT} \frac{x^s}{s}\frac{\zeta'}{\zeta}(s)\,ds}_{I_{h,1}} + \underbrace{\int_{-1+iT}^{-1-iT} \frac{x^s}{s}\frac{\zeta'}{\zeta}(s)\,ds}_{I_v} + \underbrace{\int_{-1-iT}^{c-iT} \frac{x^s}{s}\frac{\zeta'}{\zeta}(s)\,ds}_{I_{h,2}} = \frac{\zeta'}{\zeta}(0) - x + \sum_{|\Im\rho|<T} \frac{x^\rho}{\rho}.$$

Here $\frac{x^\rho}{\rho}$ are the resuides at the zeros, $-x$ comes from the pole of ζ at 1, and $\frac{\zeta'}{\zeta}(0)$ comes from the pole of $\frac{1}{s}$. Then

$$\text{von-M-2}\quad \int_{c-iT}^{c+iT} \frac{x^s}{s}\left(-\frac{\zeta'}{\zeta}(s)\right)ds - x = 1 + I_{h,1} + I_{h,2} + I_v - \sum_{\Im\rho<T} \frac{x^\rho}{\rho}. \tag{3.14}$$

We estimate each summand.

1. For the horizontal integrals, we use the estimate 3.3.1 to get

$$\begin{aligned}\left|\frac{\zeta'}{\zeta}(s)\right| &= \left|\sum_{|\Im(s-\rho)|<1} \frac{1}{s-\rho}\right| + O(\ln T), \quad s = \sigma + Ti \\ &\leq \sum_{|\Im(s-\rho)|<1} \frac{1}{\Im(s-\rho)} + O(\ln T).\end{aligned}$$

   We would like to bound $\Im(s-\rho)$ away from 0. To do this, note that there are $O(\ln T)$ roots in with $\Im\rho \in [T, T+1]$ by Lemma 3.3.1. Hence by tweaking T slightly[4], we can assume $|\Im(s-\rho)| > \frac{C}{\ln T}$ for all ρ. Also by Lemma 3.3.1 there are at most $O(\ln T)$ terms in the sum, so the sum is $O((\ln T)^2)$. Integrating gives

$$\begin{aligned}\left|\int_{c\pm Ti}^{-1\pm Ti} \frac{x^s}{s}\frac{\zeta'}{\zeta}(s)ds\right| &= O((\ln T)^2)O\left(\frac{1}{T}\right)\int_c^{-1} |x^s|\, ds \\ &= O\left(\frac{(\ln T)^2}{T}\right)O(x) \\ &= O\left(\frac{x(\ln x)^2}{T}\right).\end{aligned}$$

2. For the vertical integral, we use the same estimate, this time noting that $|s-\rho| > 1$ for every root ρ, since every zero satisfies $\Re\rho > 0$. This gives that $\frac{\zeta'}{\zeta}(s) = O(\ln T)$, and

$$\begin{aligned}\left|\int_{-1+Ti}^{-1-Ti} \frac{x^s}{s}\frac{\zeta'}{\zeta}(s)\, ds\right| &= O(\ln T)\int_{-1-Ti}^{-1+Ti} \frac{x^{-1}}{|s|}\, ds \\ &= O\left(\frac{\ln T}{x}\right)\int_{-T}^{T} \frac{1}{\sqrt{t^2+1}}\, dt \\ &= O\left(\frac{\ln T}{x}\right)\int_1^{T+1} \frac{1}{t}\, dt \\ &= O\left(\frac{(\ln T)^2}{x}\right) = O\left(\frac{x(\ln x)^2}{T}\right).\end{aligned}$$

Equations (3.13) and (3.14) together with the above two estimates give the theorem. □

The final ingredient in the proof of the Prime Number Theorem is the estimate for $\sum_{|\Im(\rho)|<T} \frac{x^\rho}{\rho}$ using the zero-free regions for ζ and the estimate for number of zeros of ζ.

Proof of Theorem 3.1.2. First, note there can only be a finite number of zeros of ζ with $|\Im(\rho)| < 2$, so $\sum_{|\Im(\rho)|<2} \frac{x^\rho}{\rho} = O(x^r)$ for some fixed $r < 1$.[5] We estimate $\sum_{2\leq|\Im(\rho)|<T} \frac{x^\rho}{\rho}$ in two steps.

[4] Changing T by a constant does not change the error term of (3.12); moreover the change in the LHS sum is $O\left(\frac{x}{T}\ln T\right) = O\left(\frac{x(\ln x)^2}{T}\right)$.

[5] In fact, there are zero such zeros.

1. By Theorem 3.3.3, there is c such that for ρ with $2 \le |\Im(\rho)| < T$,

$$|x^\rho| = x^{\Re\rho} \le x^{1-\frac{c}{\ln T}} = xe^{-\frac{c\ln x}{\ln T}}.$$

2. Using $N(T) = O(T\ln T)$ (Theorem 3.3.2 or the weaker remark after Lemma 3.3.1),

$$\begin{aligned}
\sum_{2\le|\Im(\rho)|<T} \frac{1}{|\rho|} &\le \sum_{2\le|\Im(\rho)|<T} \frac{1}{\Im(\rho)} \\
&\le \int_2^T \frac{dN(t)}{t} && \text{(Riemann-Steltjes integral)} \\
&= \frac{N(T)}{T} - \frac{N(2)}{2} + \int_2^T \frac{N(t)}{t^2}\,dt && \text{integration by parts} \\
&= O(\ln T) + \int_2^T O\left(\frac{\ln t}{t}\right) dt \\
&\overset{\text{pnt-step2}}{=} O(\ln T) + O((\ln T)^2) = O((\ln T)^2). && (3.15)
\end{aligned}$$

Putting these two estimates together,

$$\begin{aligned}
\left|\sum_{|\Im(\rho)|<T} \frac{x^\rho}{\rho}\right| &\le O(x^r) + \max_{2\le|\Im(\rho)|<T}(|x^\rho|) \sum_{2\le|\Im(\rho)|<t} \frac{1}{|\rho|} \\
&\le O(x^r) + O\left(xe^{-\frac{c\ln x}{\ln T}}(\ln T)^2\right). \text{xrhorho} && (3.16)
\end{aligned}$$

Combining with Theorem 3.4.1, and setting $T = e^{\sqrt{\ln x}}$ (so that $xe^{-\frac{\ln x}{\ln T}} = \frac{x}{T}$), we get

$$\begin{aligned}
|\psi(x) - x| &= O\left(x^r + xe^{-\frac{c\ln x}{\ln T}}(\ln T)^2 + \frac{x(\ln x)^2}{T}\right) \\
&= O\left(x^r + xe^{-c\sqrt{\ln x}}\ln x + x(\ln x)^2 e^{-\sqrt{\ln x}}\right) \\
&= O(xe^{-C\sqrt{\ln x}}),
\end{aligned}$$

for some $C > 0$. This shows

$$\text{psi-asymptotic}\quad \psi(x) = x + O(xe^{-C\sqrt{\ln x}}). \qquad (3.17)$$

Finally, we extract the asymptotics of π from the following.

Lemma 3.4.2. *partial-sum-pi* *We have the following estimates:*

$$\begin{aligned}
\pi(x) &= \frac{\psi(x)}{\ln x} + \int_2^x \psi(y)\frac{dy}{y(\ln y)^2} + O(x^{\frac12}), \\
\psi(x) &= \pi(x)\ln x - \int_2^x \frac{\pi(y)}{y}\,dy + O(x^{\frac12}\ln x).
\end{aligned}$$

Proof. Define

$$\gamma(n) = \begin{cases} 1, & n \text{ prime}, \\ 0, & n \text{ not prime}, \end{cases} \qquad \Lambda_1(n) = \begin{cases} \ln n, & n \text{ prime}, \\ 0, & n \text{ not prime}, \end{cases}$$

and

$$\psi_1(x) = \sum_{n \le x} \Lambda_1(n).$$

First note

$$\begin{aligned} |\psi(x) - \psi_1(x)| &= \sum_{2 \le r \le \log_2(x)} \sum_{p \mid p^r \le x} \ln p \\ &\le \sum_{2 \le r \le \log_2(x)} x^{\frac{1}{r}} \ln x \\ &= O(x^{\frac{1}{2}} \ln x + x^{\frac{1}{3}} (\ln x)^2) = O(x^{\frac{1}{2}} \ln x). \end{aligned} \tag{3.18}$$

psi-psi1-estimate

Part 1: By partial summation 1.5.1 with $u = \Lambda_1$, $U = \psi_1$, and $v = \frac{1}{\ln x}$,

$$\begin{aligned} \pi(x) &= \sum_{n \le x} \gamma(n) \\ &= \sum_{n \le x} \Lambda_1(n) \frac{1}{\ln n} \\ &= \frac{\psi_1(x)}{\ln x} + \int_2^x \psi_1(t) \frac{dt}{t(\ln t)^2} \\ &= \frac{\psi(x)}{\ln x} + O(x^{\frac{1}{2}}) + \int_2^x \psi(t) \frac{dt}{t(\ln t)^2} + \int_2^x O(t^{-\frac{1}{2}})\, dt \qquad \text{by (3.18)} \\ &= \frac{\psi(x)}{\ln x} + \int_2^x \psi(t) \frac{dt}{t(\ln t)^2} + O(x^{\frac{1}{2}}). \end{aligned}$$

Part 2: By partial summation,

$$\begin{aligned} \psi_1(x) &= \sum_{n \le x} \gamma(n) \ln(n) \\ &= \pi(x) \ln x - \int_2^x \frac{\pi(t)}{t}\, dt. \end{aligned}$$

Combining with (3.18) gives the result. □

Putting (3.17) into Lemma 3.4.2,

$$\begin{aligned} \pi(x) &= \frac{x}{\ln x} + O\left(\frac{x e^{-C\sqrt{\ln x}}}{\ln x}\right) + \int_2^x \left(\frac{1}{(\ln y)^2} + O\left(\frac{e^{-C\sqrt{\ln y}}}{(\ln y)^2}\right)\right) dy + O(x^{\frac{1}{2}}) \\ &= \mathrm{li}(x) + O(x e^{-C\sqrt{\ln x}}). \qquad \text{by (3.17)} \end{aligned}$$

□

5 The Riemann hypothesis

The following conjecture is worth one million dollars:

Conjecture 3.5.1 (Riemann hypothesis)**.** *All nontrivial zeros s of $\zeta(s)$ satisfy $\Re s = \frac{1}{2}$.*

Note that for no $\varepsilon > 0$ has it been proved that all zeros satisfy $\Re s < 1 - \varepsilon$. Our zero-free region, sadly, has a boundary approaching real part 1 as $t \to \infty$.

One reason that the Riemann hypothesis is important is that it gives a strong error bound in the prime number theorem (as well as many other theorems of analytic number theory).

Theorem 3.5.2. *Suppose $\frac{1}{2} \le \theta < 1$. The following are equivalent.*

1. *$\zeta(s)$ has no zeros with $\Re s > \theta$.*
2. *$\pi(x) = \operatorname{li}(x) + O(x^\theta \ln x)$.*
3. *$\pi(x) = \operatorname{li}(x) + O(x^{\theta+\varepsilon})$ for every $\varepsilon > 0$, where the constant depends on ε.*

In particular, the Riemann hypothesis is equivalent to $\pi(x) = \operatorname{li}(x) + O(x^{\frac{1}{2}} \ln x)$.

Proof. (1) $\implies$ (2): Suppose $\zeta(s)$ has no zeros with $\Re s > \theta$. Then using the estimate in (3.15), we have

$$\begin{aligned}\sum_{|\Im(\rho)|<T} \frac{x^\rho}{\rho} &\le \max_\rho |x^\rho| \sum_{|\Im(\rho)|<T} \frac{1}{|\rho|} \\ &\le x^\theta (\ln T)^2.\end{aligned}$$

Now take $T = x$ to find that

$$\begin{aligned}|\psi(x) - x| &= O\left(x^\theta (\ln T)^2 + \frac{x(\ln x)^2}{T}\right) \\ &= O(x^\theta (\ln x)^2).\end{aligned}$$

Then using Lemma 3.4.2 and (3.1),

$$\begin{aligned}\pi(x) &= \frac{\psi(x)}{\ln x} + \int_2^x \psi(y) \frac{dy}{y(\ln y)^2} + O(y^{\frac{1}{2}}) \\ &= \operatorname{li}(x) + O\left(\frac{x^\theta (\ln x)^2}{\ln x}\right) + \int_2^x O\left(\frac{x^{\frac{1}{2}-1}(\ln x)^2}{(\ln x)^2}\right) dx \\ &= \operatorname{li}(x) + O(x^\theta \ln x).\end{aligned}$$

(2) $\implies$ (3): Item 2 is stronger than item 3.

(3) $\implies$ (1): Going the other way in Lemma 3.4.2,

$$\begin{aligned}
\psi(x) &= \pi(x)\ln x - \int_2^x \frac{\pi(y)}{y}\,dy + O(x^{\frac{1}{2}}\ln x) \\
&= \left(\frac{x}{\ln x} + \int_2^x \frac{dy}{(\ln y)^2} + O(x^{\theta+\varepsilon})\right)\ln x - \int_2^x \left(\frac{1}{\ln x} + \frac{1}{y}\int_2^y \frac{dt}{(\ln t)^2} + \frac{O(y^{\theta+\varepsilon})}{y}\right) dy + O(x^{\frac{1}{2}}\ln x) \\
&= x + O(x^{\theta+\varepsilon'}) - \underbrace{\int_2^x \frac{dy}{\ln y} + \int_2^x \frac{dy}{(\ln y)^2}\ln x - \int_2^x \left(\int_2^y \frac{dt}{(\ln t)^2}\cdot\frac{1}{y}\right) dy}_{0}
\end{aligned}$$

for any $\varepsilon' > \varepsilon$. Note the integrals above sum to 0 by integration by parts ($u = \ln y$, $dv = \frac{dy}{(\ln y)^2}$).

By partial summation, for $\sigma > 1$,

$$\begin{aligned}
-\frac{\zeta'}{\zeta}(s) &= \sum_n \Lambda(n)n^{-s} \\
&= -\int_1^\infty \psi(n)sn^{-s-1}\,ds \\
&= \frac{s}{s-1} + s\int_1^\infty \underbrace{(\psi(x)-x)}_{O(x^{\theta+\varepsilon'})} x^{-s-1}\,dx.
\end{aligned}$$

The last integral converges whenever $\sigma > \theta + \varepsilon'$, so $\frac{\zeta'}{\zeta}$ has analytic continuation to $\sigma > \theta$. This means ζ has no zeros for $\sigma > \theta$. $\square$

Chapter 4

L-functions and Dirichlet's theorem

l-func-dirichlet

1 Outline

Our goal in this chapter is to study the asymptotics of

$$\pi(x, a \bmod N) = |\{p \leq x : p \text{ prime}, p \equiv a \pmod N\}|$$

where a is relatively prime to N. We define $\psi(x, a \bmod N) = \sum_{n \leq x,\, n \equiv a \,(\mathrm{mod}\ N)} \Lambda(n)$.

To study the distribution of primes in the arithmetic progression $n \equiv a \pmod N$, we study the asymptotics of $\psi(x, a \bmod N)$. However, this does not come from a Dirichlet series that we can easily estimate and that has nice multiplicative properties, like $\psi(x)$ comes from $\zeta(x) = \prod_p \frac{1}{1-p^{-s}}$ (after logarithmic differentiation and extracting coefficients).

The solution is to write $\psi(x, a \bmod N)$ in terms of Dirichlet series whose coefficients are multiplicative. For example, when considering primes $p \equiv 1 \pmod 4$, we consider

$$L(s, \chi_1) = \frac{1}{1^s} + \frac{1}{3^s} + \frac{1}{5^s} + \frac{1}{7^s} + \frac{1}{9^s} \cdots = \prod_p \frac{1}{1 - p^{-s}}.$$

$$L(s, \chi_2) = \frac{1}{1^s} - \frac{1}{3^s} + \frac{1}{5^s} - \frac{1}{7^s} + \frac{1}{9^s} \cdots = \prod_{p \equiv 1 \,(\mathrm{mod}\ 4)} \frac{1}{1 - p^{-s}} \prod_{p \equiv 3 \,(\mathrm{mod}\ 4)} \frac{1}{1 + p^{-s}}$$

The multiplicative structure is from the fact that the coefficients come from group homomorphisms $(\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{C}$, i.e. Dirichlet characters (see Definition 1.1.8).

Logarithmic differentiation gives

$$-\frac{L'}{L}(s, \chi_1) = \frac{\Lambda(1)}{1^s} + \frac{\Lambda(3)}{3^s} + \frac{\Lambda(5)}{5^s} + \frac{\Lambda(7)}{7^s} + \frac{\Lambda(9)}{9^s} \cdots$$

$$-\frac{L'}{L}(s, \chi_2) = \frac{\Lambda(1)}{1^s} - \frac{\Lambda(3)}{3^s} + \frac{\Lambda(5)}{5^s} - \frac{\Lambda(7)}{7^s} + \frac{\Lambda(9)}{9^s} \cdots$$

$$\frac{1}{2}\left(-\frac{L'}{L}(s, \chi_1) - \frac{L'}{L}(s, \chi_2)\right) = \frac{\Lambda(1)}{1^s} + \frac{\Lambda(5)}{5^s} + \frac{\Lambda(9)}{9^s} \cdots$$

Taking the partial sum of coefficients of the last Dirichlet series gives the desired result. In general, we can always estimate $\psi(x, a \bmod N)$ using an average of these *L-functions*.

The main steps in the proof are the same, except with ζ replaced by L and an extra recombination step at the end using character theory. The main steps are the following.

1. Functional equation and analytic continuation for L, Theorem 4.2.5.
2. Product development, Theorem 4.2.6.
3. Estimates on $\frac{L'}{L}$ and asymptotics on number of zeros $N(T, \chi)$, Lemma 4.3.1.
4. Zero-free region for L, Theorem 4.3.3.
5. von Mangoldt's formula 4.4.1.

If we conly cared about bounds for a fixed modulus N, then that's all there is to it.

However, to obtain error bounds independent of N, we need a zero free region independent of N (Theorem 4.3.3). While in Theorem 3.3.3 we had the luxury of restricting to large $|t|$, here we have to work with small $|t|$, and our resulting region may miss an "exceptional" zero. We show there is at most 1 exception (Theorem 4.4.2) and prove a version of the Prime Number Theorem for arithmetic progressions (Theorem 4.4.4). Later we prove a stronger but ineffective bound on the "exceptional zero" (Theorem 4.5.4) and obtain improved asymptotics (Theorem 4.5.1).

2 L-functions

Definition 4.2.1: Let χ be a Dirichlet character. Define the L function

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

By multiplicativity of χ, L has a product expansion

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

Only the factors with $p \nmid N$ contribute. Note that if χ is of level N and $\chi = \chi_1\chi_2$ with χ_1 primitive of level N_1, then

$$\text{in-terms-of-primitive} L(s, \chi) = L(s, \chi_1) \prod_{p|N,\, p\nmid N_1} (1 - \chi(p)p^{-s}). \tag{4.1}$$

Thus for convenience we can often just prove results about primitive characters.

By logarithmic differentiation we have

$$\frac{L'}{L}(s, \chi) = -\sum_p \frac{(\ln p)\chi(p)p^{-s}}{1 - p^{-s}} = -\sum_{n=1}^{\infty} \frac{\chi(n)\Lambda(n)}{n^s}.$$

Theorem 4.2.2 (Generalized Poisson summation). *gen-ps Let g be a function $\mathbb{Z}/N\mathbb{Z} \to \mathbb{R}$, and suppose f is a $\mathcal{C}^2$ function satisfying*

$$|f(x)|, |\hat{f}(x)| \leq C(1+|x|)^{-1-\delta}$$

for some $C, \delta > 0$. Then

$$\sum_{m\in\mathbb{Z}} f\left(\frac{m}{N}\right) g(m) = \sum_{n\in\mathbb{Z}} \hat{f}(n)\hat{\bar{g}}(n).$$

In particular, if χ is a primitive multiplicative character modulo N, then

$$\sum_{m\in\mathbb{Z}} \chi(m) f\left(\frac{m}{N}\right) = G(\chi, \chi_1^+) \sum_{n\in\mathbb{Z}} \overline{\chi}(-n)\hat{f}(n).$$

where $\chi_j^+(k) := e^{\frac{2\pi ijk}{N}}$.

Here $\hat{f}(n)$ denotes the Fourier transform

$$\hat{f}(y) = \int_{-\infty}^{\infty} f(x)e^{-2\pi ixy}\,dx$$

and $\hat{g}(n)$ denotes the finite Fourier transform

$$\hat{g}(n) = \sum_{m \,(\mathrm{mod}\ N)} g(m)e^{-\frac{2\pi im}{N}}.$$

Proof. Consider the function

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

Note this sum converges absolutely to a continuous function by the given conditions. Since $F(x)$ has period 1 and is continuous, we can expand it in Fourier series:

$$F(x) = \sum_{n=0}^{\infty} a_n e^{2\pi inx},$$

$$a_n = \int_0^1 F(x)e^{-2\pi inx}\,dx = \int_0^1 \sum_{m\in\mathbb{Z}} f(x+m)e^{-2\pi inx}\,dx = \int_{-\infty}^{\infty} f(x)e^{-2\pi inx}\,dx = \hat{f}(n).$$

Plugging in $x = \frac{a}{N}$ gives

$$F\left(\frac{a}{N}\right) = \sum_{n\in\mathbb{Z}} \hat{f}(n)e^{2\pi in\left(\frac{a}{N}\right)}.$$

Now we calculate

$$\begin{aligned}
\sum_{m\in\mathbb{Z}} f\left(\frac{m}{N}\right) g(m) &= \sum_{a \,(\mathrm{mod}\ N)} g(a) F\left(\frac{a}{N}\right) \\
&= \sum_{a \,(\mathrm{mod}\ N)} g(a) \sum_{n\in\mathbb{Z}} \hat{f}(n)e^{2\pi in\left(\frac{a}{N}\right)} \\
&= \sum_{n\in\mathbb{Z}} \hat{f}(n) \sum_{a \,(\mathrm{mod}\ N)} g(a)e^{2\pi in\left(\frac{a}{N}\right)} \\
&= \sum_{n\in\mathbb{Z}} \hat{f}(n)\hat{\bar{g}}(n).
\end{aligned}$$

For the second part, note that

$$\begin{aligned}\sum_{m\in\mathbb{Z}} \chi(m) f\left(\frac{m}{N}\right) &= \sum_{n\in\mathbb{Z}} \hat{\overline{\chi}}(n)\hat{f}(m) \\ &= \sum_{n\in\mathbb{Z}} G(\chi,\chi_1^+)\overline{\chi(n)}\hat{f}(n).\end{aligned}$$

□

We apply Poisson summation to derive a transformation law for generalized theta functions.

Definition 4.2.3: Let χ be a multiplicative character modulo N. Define

$$\begin{aligned}\theta_\chi(u) &= \sum_{n\in\mathbb{Z}} \chi(n) e^{-\pi n^2 u} \\ \vartheta_\chi(u) &= \sum_{n\in\mathbb{Z}} \chi(n) n e^{-\pi n^2 u}.\end{aligned}$$

Note we need to work with $\vartheta_\chi(u)$ when χ is odd, since in this case $\theta_\chi(u) = 0$ and we cannot express $L(s,\chi)$ in terms of θ_χ.

Proposition 4.2.4 (Transformation law for θ_χ)**:** theta-transforms Suppose χ is primitive. Then

$$\begin{aligned}\theta_\chi(u) &= \frac{G(\chi,\chi_1^+)}{N\sqrt{u}}\theta_{\overline{\chi}}\left(\frac{1}{N^2u}\right) \\ \vartheta_\chi(u) &= -\frac{G(\chi,\chi^+)i}{N^2u^{\frac{3}{2}}}\vartheta_{\overline{\chi}}\left(\frac{1}{N^2u}\right).\end{aligned}$$

Proof. Note the Fourier transform of $e^{-\pi x^2}$ is itself; moreover, if $f(x) = g(ax)$ then $\hat{f}(y) = \hat{g}\left(\frac{y}{a}\right)$. Hence

$$\mathcal{F}(e^{-\pi u(Nx)^2}) = \frac{1}{N\sqrt{u}} e^{-\frac{\pi y^2}{uN^2}}.$$

By the Poisson summation formula 4.2.2,

$$\begin{aligned}\theta_\chi(u) &= \sum_{n\in\mathbb{Z}} \chi(n) e^{-\pi n^2 u} \\ &= \frac{G(\chi,\chi_1^+)}{N\sqrt{u}} \sum_{n\in\mathbb{Z}} \overline{\chi}(-n) e^{-\frac{\pi n^2}{uN^2}} \\ &= \frac{G(\chi,\chi_1^+)}{N\sqrt{u}}\theta_{\overline{\chi}}\left(\frac{1}{N^2u}\right).\end{aligned}$$

For the second part, note first that $\widehat{f'}(y) = 2\pi i x\hat{f}(y)$. Hence

$$\mathcal{F}(Nxe^{-\pi u(Nx)^2}) = \left(-\frac{1}{2\pi uN}\right)\mathcal{F}\left(\frac{d}{dx}(xe^{-\pi u(Nx)^2})\right) = -\frac{1}{\cancel{2\pi}uN}\cdot\cancel{2\pi i}y\frac{1}{N\sqrt{u}}e^{-\frac{\pi y^2}{uN^2}} = -\frac{i}{N^2u^{\frac{3}{2}}}e^{-\frac{\pi y^2}{uN^2}}.$$

Then by Poisson summation,

$$\begin{aligned}\vartheta_\chi(u) &= \sum_{n\in\mathbb{Z}} \chi(n) n e^{-\pi n^2 u} \\ &= -\frac{G(\chi,\chi_1^+)i}{N^2 u^{\frac{3}{2}}} \sum_{n\in\mathbb{Z}} \overline{\chi}(-n) n e^{-\frac{\pi n^2}{uN^2}} \\ &= -\frac{G(\chi,\chi_1^+)i}{N^2 u^{\frac{3}{2}}} \vartheta_{\overline{\chi}}\left(\frac{1}{N^2 u}\right). \end{aligned}$$

□

From this we get the functional equation for the L-function. The proof is similar to that of Theorem 3.2.2.

Theorem 4.2.5 (Analytic continuation and functional equation for L-functions)**.** l-continues *Let χ be any character modulo N. Then $L(s,\chi)$ has an meromorphic continuation to $\mathbb{C}$. If χ is principal then $L(s,\chi)$ has a single pole at 1, and if χ is nonprincipal then $L(s,\chi)$ is entire.*

Now supose χ is primitive. Defining

$$\xi(s,\chi) := \left(\frac{\pi}{N}\right)^{-\frac{s+a}{2}} \Gamma\left(\frac{s+a}{2}\right) L(s,\chi),$$

where

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

we have

$$\xi(s,\chi) := \frac{G(\chi,\chi_1^+)}{i^a\sqrt{q}} \xi(1-s,\overline{\chi}).$$

Moreover, for any χ, $L(s,\chi)$ has zeros at $-2\mathbb{N}+a$ (the trivial zeros) and all other zeros are in the critical strip $0 \le \Re s \le 1$.

Note that for χ nonprincipal, partial cancellation in the Dirichlet series removes the pole at $s = 1$.

Proof. Note that it suffices to prove all statements for χ primitive, in light of (4.1). If χ is principal, the result follows from the result for ζ, so suppose χ is nonprincipal. Use partial summation 1.5.1 to find that for for $s > 1$,

$$\text{bound-L-summands} L(s,\chi) = \int_1^\infty S(x) s x^{-s-1}\, dx \tag{4.2}$$

where $S(x) = \sum_{n\le x} \chi(n)$. (We use the fact that $\lim_{N\to\infty} S(N)N^{-s} = 0$ when $s > 1$.) Since $\chi(1) + \cdots + \chi(N) = 0$ by Corollary 1.1.7, $\chi(1) + \cdots + \chi(n) \le N$. Then for $\Re s > 0$, the above integral converges absolutely, extending $L(s,\chi)$ holomorphically to $\Re s > 0$.

Case 1: Suppose $\chi(-1) = 1$; then $\chi(-n) = \chi(n)$. We calculate

$$\int_0^\infty \theta_\chi(u) u^{\frac{s}{2}} \frac{du}{u}$$

in two different ways.[1] When $0 < \Re s < 1$,

$$\begin{aligned}
\int_0^\infty \theta_\chi(u) u^{\frac{s}{2}} \frac{du}{u} &= \int_0^\infty \sum_{n\in\mathbb{Z}} \chi(n) e^{-\pi n^2 u} u^{\frac{s}{2}} \frac{du}{u} \\
&= 2\sum_{n=1}^\infty \int_0^\infty \chi(n) e^{-\pi n^2 u} u^{\frac{s}{2}} \frac{du}{u} && \chi(-n) = \chi(n),\ \chi(0) = 0 \\
&= 2\sum_{n=1}^\infty \int_0^\infty \chi(n) e^{-u} \left(\frac{u}{\pi n^2}\right)^{\frac{s}{2}} \frac{du}{u} && u \longleftrightarrow \frac{u}{\pi n^2} \\
&= 2\pi^{-\frac{s}{2}} \left(\sum_{n=1}^\infty \frac{\chi(n)}{n^s}\right) \left(\int_0^\infty e^{-u} u^{\frac{s}{2}} \frac{du}{u}\right) \\
&= 2\pi^{-\frac{s}{2}} L(s,\chi) \Gamma\left(\frac{s}{2}\right).
\end{aligned}$$

Now using the transformation law 4.2.4,

$$\begin{aligned}
\int_0^\infty \theta_\chi(u) u^{\frac{s}{2}} \frac{du}{u} &= \int_0^\infty \frac{G(\chi,\chi_1^+)}{N\sqrt{u}} \theta_{\overline{\chi}}\left(\frac{1}{N^2 u}\right) u^{\frac{s}{2}} \frac{du}{u} \\
&= \frac{G(\chi,\chi_1^+)}{N} \int_0^\infty \theta_{\overline{\chi}}\left(\frac{1}{N^2 u}\right) u^{\frac{s}{2}-\frac{1}{2}} \frac{du}{u} \\
&= \frac{2G(\chi,\chi_1^+)}{N} \sum_{n=1}^\infty \int_0^\infty \overline{\chi}(n) e^{-\frac{\pi n^2}{uN^2}} u^{\frac{s}{2}-\frac{1}{2}} \frac{du}{u} \\
&= \frac{2G(\chi,\chi_1^+)}{N} \sum_{n=1}^\infty \int_0^\infty \overline{\chi}(n) e^{-u} \left(\frac{\pi n^2}{uN^2}\right)^{\frac{s}{2}-\frac{1}{2}} \frac{du}{u} && u \longleftrightarrow \frac{\pi n^2}{uN^2} \\
&= \frac{2G(\chi,\chi_1^+)\pi^{\frac{s}{2}-\frac{1}{2}}}{N^s} \sum_{n=1}^\infty \frac{\overline{\chi}(n)}{n^{(1-s)}} \int_0^\infty e^{-u} u^{\frac{1-s}{2}} \frac{du}{u} \\
&= \frac{2G(\chi,\chi_1^+)\pi^{\frac{s}{2}-\frac{1}{2}}}{N^s} L(1-s,\overline{\chi}) \Gamma\left(\frac{1-s}{2}\right).
\end{aligned}$$

Equating these two calculations gives the result.
Case 2: Suppose $\chi(-1) = -1$. We work with ϑ_χ instead of θ_χ. To compensate for the extra factor of n in ϑ_χ, we need an extra factor of $u^{\frac{1}{2}}$. We calculate

$$\int_0^\infty \vartheta_\chi(u) u^{\frac{s+1}{2}} \frac{du}{u}$$

[1] Unlike in Theorem 3.2.2, there is no "−1" since $\chi(0) = 0$.

in two different ways. First,

$$\begin{aligned}
\int_0^\infty \theta_\chi(u) u^{\frac{s+1}{2}} \frac{du}{u} &= \int_0^\infty \sum_{n\in\mathbb{Z}} \chi(n) n e^{-\pi n^2 u} u^{\frac{s+1}{2}} \frac{du}{u} \\
&= 2\sum_{n=1}^\infty \int_0^\infty \chi(n) n e^{-\pi n^2 u} u^{\frac{s+1}{2}} \frac{du}{u} && -n\chi(-n) = n\chi(n),\ \chi(0) = 0 \\
&= 2\sum_{n=1}^\infty \chi(n) n \int_0^\infty e^{-u} \left(\frac{u}{\pi n^2}\right)^{\frac{s+1}{2}} \frac{du}{u} && u \longmapsto \frac{u}{\pi n^2} \\
&= 2\pi^{-\frac{s+1}{2}} \sum_{n=1}^\infty \frac{\chi(n)}{n^s} \int_0^\infty e^{-u} u^{\frac{s+1}{2}} \frac{du}{u} \\
&= 2\pi^{-\frac{s+1}{2}} L(s,\chi)\Gamma\left(\frac{s+1}{2}\right).
\end{aligned}$$

Now using the transformation law 4.2.4,

$$\begin{aligned}
\int_0^\infty \theta_\chi(u) u^{\frac{s+1}{2}} \frac{du}{u} &= \int_0^\infty -\frac{G(\chi,\chi^+) i y}{N^2 u} \theta_{\overline{\chi}}\left(\frac{1}{N^2 u}\right) u^{\frac{s+1}{2}} \frac{du}{u} \\
&= -\frac{G(\chi,\chi^+) i}{N^2} \int_0^\infty \theta_{\overline{\chi}}\left(\frac{1}{N^2 u}\right) u^{\frac{s}{2}-1} \frac{du}{u} \\
&= -\frac{2G(\chi,\chi^+) i}{N^2} \sum_{n=1}^\infty n\overline{\chi}(n) \int_0^\infty e^{-\frac{\pi n^2}{uN^2}} u^{\frac{s}{2}-1} \frac{du}{u} \\
&= -\frac{2G(\chi,\chi^+) i}{N^2} \sum_{n=1}^\infty \int_0^\infty \overline{\chi}(n) n e^{-u} \left(\frac{\pi n^2}{uN^2}\right)^{\frac{s}{2}-1} \frac{du}{u} && u \longmapsto \frac{\pi n^2}{uN^2} \\
&= -\frac{2G(\chi,\chi^+) i \pi^{\frac{s}{2}-1}}{N^2 N^{n-2}} \sum_{n=1}^\infty \frac{\overline{\chi}(n)}{n^{1-s}} \int_0^\infty e^{-u} u^{1-\frac{s}{2}} \frac{du}{u} \\
&= -\frac{2G(\chi,\chi^+) i \pi^{\frac{s}{2}-1}}{N^s} L(1-s,\overline{\chi})\Gamma\left(1-\frac{s}{2}\right). && \square
\end{aligned}$$

Again matching the two calculations gives the result.

From Proposition 1.7.2(5), Γ has no zeros, so we find that $L(s,\chi)$ is defined whenever $L(s,\overline{\chi})$ is defined; this L is entire. The description of the zeros of L follow from the functional equation and the fact that Γ has poles at $-\mathbb{N}_0$.

Theorem 4.2.6 (Product development of $\xi(s,\chi)$)**.** *xi-chi-product-development* *Suppose χ is primitive of level $N > 1$. The function $\xi(s,\chi)$ is entire of order 1 and has the product expansion*

$$\xi(s,\chi) = \xi(0,\chi) e^{Bs} \prod_{\rho \text{ zero of } \xi(s,\chi)} \left(1-\frac{s}{\rho}\right) e^{\frac{s}{\rho}}.$$

Then $\frac{L'}{L}(s,\chi)$ has the partial-fraction expansion

$$\frac{L'}{L}(s,\chi) = B + \frac{1}{2}\ln\left(\frac{N}{\pi}\right) - \frac{1}{2}\frac{\Gamma'}{\Gamma}\left(\frac{s+a}{2}\right) + \sum_{\rho \text{ nontrivial zero of } \zeta} \left(\frac{1}{s-\rho} + \frac{1}{\rho}\right).$$

From now on, we only talk about nontrivial zeros of ζ.

Proof. We proceed as in Theorem 3.2.5. The argument is the same, the only major differences being that $\xi(s,\chi)$ has no poles at $s = 0, 1$, and the slight difference in definition of $\zeta(s,\chi)$ in terms of $L(s,\chi)$, versus the definition of $\xi(s)$ in terms of $\zeta(s)$. (Namely, we have $s+a$ instead of s, and an extra $N^{-\frac{s+a}{2}}$. For completeness we give the proof.

To show it has order 1 we need two inequalities.

Step 1: There is no constant C so that $\xi(s,\chi) \precsim e^{C|s|}$: Indeed, for real s and any constant C' we have

$$\begin{aligned}\xi(s) &= \left(\frac{\pi}{N}\right)^{-\frac{s+a}{2}} \Gamma\left(\frac{s+a}{2}\right) L(s,\chi) \\ &\succsim s^{-\frac{1}{2}} \left(\frac{(s+a)N}{2e\pi}\right)^{\frac{s+a}{2}} \succsim e^{C's}.\end{aligned}$$

Step 2: There is a constant C so that $\xi(s,\chi) \precsim e^{C|s|\ln|s|}$: $e^{|s|\ln|s|} \geq 1$ for all s so it suffices to prove this for sufficiently large s. By the integral and sum formulas for Γ and ξ, and the fact that $|x^s| = |x^{\Re s}|$, we have

$$|\xi(\sigma+ti,\chi)| \leq \left(\frac{\pi}{N}\right)^{-\frac{\sigma+a}{2}} \Gamma\left(\frac{\sigma+a}{2}\right) L(\sigma,\chi), \quad \sigma > 1.$$

By symmetry of ξ is suffices to consider $\sigma \geq \frac{1}{2}$. (We have $\xi(s,\chi) = \frac{G(\chi,\chi^+)}{i^a\sqrt{q}}\xi(1-s,\overline{\chi})$, and the multiplier has absolute value 1.) Consider 2 cases.

1. $\sigma > 2$: Then $\pi^{-\frac{\sigma+a}{2}} < 1$ and $L(\sigma,\chi) < \zeta(2)$ so we have by Stirling's approximation 1.7.4 that

   $$|\xi(\sigma+ti,\chi)| \precsim \left|N^{\frac{\sigma+a}{2}}\Gamma\left(\frac{\sigma+ti+a}{2}\right)\right| = N^{\frac{\sigma+a}{2}} e^{|(\ln\Gamma)(\sigma+a)|} = N^{\frac{\sigma+a}{2}} e^{\left(\frac{\sigma+a-1}{2}\right)\ln\frac{\sigma+a}{2} - \frac{\sigma+a}{2} + O(1)}$$

   from which the result follows.

2. $\frac{1}{2} \leq \sigma \leq 2$: For s bounded away from 1, from (4.2),

   $$L(s,\chi) = O(|s|).$$

   This time $\Gamma\left(\frac{\sigma+a}{2}\right) = O(1)$ so

   $$|L(s,\chi)| \leq \left|\left(\frac{\pi}{N}\right)^{-\frac{\sigma+a}{2}} L(s,\chi)\Gamma\left(\frac{\sigma+a}{2}\right)\right| = O(|s|) \precsim e^{C|s|\ln|s|}.$$

This shows $\xi(s)$ has order 1.

Step 3: By the product development 1.6.3, noting the the zeros of $\xi(s,\chi)$ are the nontrivial zeros of $L(s,\chi)$, we get

$$\xi(s,\chi) = \xi(0,\chi)e^{Bs} \prod_{\rho \text{ zero of } L(s,\chi)} \left(1 - \frac{s}{\rho}\right) e^{\frac{s}{\rho}}.$$

Logarithmic differentiation gives

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

Since $L(s,\chi) = \left(\frac{\pi}{N}\right)^{\frac{s+a}{2}} \Gamma\left(\frac{s+a}{2}\right)^{-1} \xi(s,\chi)$, we get

$$\frac{L'}{L}(s,\chi) = \frac{1}{2}\ln\left(\frac{\pi}{N}\right) - \frac{1}{2}\frac{\Gamma'}{\Gamma}\left(\frac{s+a}{2}\right) + B + \sum_{\rho} \left(\frac{1}{s-\rho} + \frac{1}{\rho}\right). \qquad \square$$

3 Zeros of L

Lemma 4.3.1. weak-L-zeros *Define $\mathcal{L} = \ln N(|t|+2)$. Let χ be a primitive character of level N. For $s = \sigma + it$ with $\sigma \in [-1,2]$, we have*

$$\begin{aligned}\frac{L'}{L}(s,\chi) &= \sum_{\rho} \left(\frac{1}{s-\rho} + \frac{1}{\rho}\right) + O(\mathcal{L}) \\ &= \sum_{|\Im(s-\rho)|<1} \frac{1}{s-\rho} + O(\mathcal{L}).\end{aligned}$$

Moreover, there are $O(\ln|Nt|)$ zeros ρ with $|\Im(s-\rho)| < 1$, i.e. the number of zeros with imaginary part in $[t,t+1]$ is $O(\ln Nt)$, as $t \to \infty$.

Note this gives $N(T) = O(T\ln(NT))$.
We follow the proof of Theorem 3.3.1.

Proof. The case $N = 1$ follows from there so we assume $N > 1$.
Step 1: Theorem 4.2.6 gives us

$$\frac{L'}{L}(s,\chi) = \underbrace{B + \frac{1}{2}\ln\left(\frac{N}{\pi}\right)}_{O(1+\ln N)} - \frac{1}{2}\underbrace{\frac{\Gamma'}{\Gamma}\left(\frac{s+a}{2}\right)}_{(A)} + \underbrace{\sum_{\rho}\left(\frac{1}{s-\rho} + \frac{1}{\rho}\right)}_{(B)}. \tag{4.3}$$

From Stirling's approximation 1.7.4, (A) equals

$$\ln\left|\frac{\sigma+a}{2} + \frac{t}{2}i\right| + O(1) = O(\mathcal{L}). \tag{4.4}$$

Now suppose $s = 2 + it$. Note that

$$\left|\frac{L'}{L}(s,\chi)\right| = \left|\sum_{n=1}^{\infty} \chi(n)\Lambda(n)n^{-2-it}\right| \le \left|\sum_{n=1}^{\infty}(\ln n)n^{-2}\right| < \infty,$$

so the LHS of (4.3) is $O(1)$. Hence (4.3) becomes

$$\text{L2-zero-sum2} O(\mathcal{L}) = \sum_{\rho}\left(\frac{1}{s-\rho} + \frac{1}{\rho}\right). \tag{4.5}$$

Now finish the same way as in Theorem 1 to conclude the first step.

Step 2: Now we consider general $s = \sigma + it$, by comparing it to $2 + it$. We have by (4.3) and (4.4) that

$$\frac{L'}{L}(s,\chi) - \underbrace{\frac{L'}{L}(2+it)}_{O(1)} = O(1) + \sum_{|\Im(s-\rho)|<1}\frac{1}{s-\rho} + \underbrace{\sum_{|\Im(s-\rho)|<1}\frac{1}{2+it-\rho}}_{O(\mathcal{L})} + \underbrace{\sum_{|\Im(s-\rho)|\ge 1}\frac{(2-\sigma)+it}{(s-\rho)(2+it-\rho)}}_{O(\mathcal{L})}.$$

Finish as in Theorem 3.3.1, the only difference being that $\ln|t|$ is replaced by $\ln|Nt|$. □

Theorem 4.3.2 (von Mangoldt). *L-zeros(*) As $T \to \infty$,*

$$N(T,\chi) = \frac{T}{\pi}\ln\left(\frac{NT}{2\pi}\right) - \frac{T}{\pi} + O(\ln NT).$$

where the constant is independent of N.

Proof. The proof is similar to Theorem 1. We'll only need the weaker estimate $N(T,\chi) = O(T\ln NT)$ so we omit the proof. □

Theorem 4.3.3 (Zero-free region for L). *L-zero-free There exists a constant $c > 0$, independent of χ and N, such that the following holds for all primitive χ of level N.*

1. *If χ is nonreal, and $s = \sigma + it$ is a zero of $L(s,\chi)$, then*

$$\text{L-zero-bound}\sigma < 1 - \frac{c}{\mathcal{L}}. \tag{4.6}$$

2. *If χ is real, then with at most 1 exception (counting multiplicity), all zeros satisfy (4.6). If it exists, the exceptional zero is real.*

Unlike in Theorem 3.3.3, we have to worry about small $|t|$. Fortunately, $L(s,\chi)$ has no pole at $s = 1$ to screw us up. Things are not so easy, however.

Proof. We may assume $N \geq 2$.

As in Theorem 3.3.3, we have $0 \leq 3 + 4\cos\theta + \cos 2\theta$, so

$$0 \leq 3 + 4\Re(\chi(n)n^{-it}) + \Re(\chi(n)^2 n^{-2it}).$$

Multiplying by $\Lambda(n)n^{-\sigma}$ and summing, we get

zero-free-L-inequality
$$0 \leq 3\left(-\frac{L'}{L}(\sigma, \chi_0)\right) + 4\Re\left(-\frac{L'}{L}(\sigma + ti, \chi)\right) + \Re\left(-\frac{L'}{L}(\sigma + 2ti, \chi^2)\right), \quad \sigma > 1. \tag{4.7}$$

Suppose $1 < \sigma < 2$ and $\rho = (1 - \delta) + ti$ is zero. First we have

zfli1
$$-\frac{L'}{L}(\sigma, \chi_0) = -\frac{\zeta'}{\zeta}(\sigma, \chi_0) - \sum_{p|N} \frac{(\ln p)p^{-s}}{1 - p^{-s}} = \frac{1}{\sigma - 1} + O(\ln N). \tag{4.8}$$

Next, we use the partial fraction decomposition 4.2.6. By Theorem 4.3.1 we have

zfli2
$$\Re\left(-\frac{L'}{L}(s, \chi)\right) \leq O(\mathcal{L}) - \sum_{\rho} \Re\left(\frac{1}{s - \rho}\right). \tag{4.9}$$

1. Suppose χ^2 is not principal, i.e. χ is not real. Now (4.9) gives

$$\Re\left(-\frac{L'}{L}(\sigma + ti, \chi)\right) \leq O(\mathcal{L}) - \frac{1}{\sigma + \delta - 1}. \tag{4.10}$$

Also by Theorem 4.3.1

zero-free-L-inequality
$$\Re\left(-\frac{L'}{L}(\sigma + 2ti, \chi^2)\right) \leq O(\mathcal{L}(2t)) = O(\mathcal{L}). \tag{4.11}$$

The remainder of this case follows the lines of Theorem 3.3.3.

2. If χ^2 is principal, then we have

$$-\frac{L'}{L}(\sigma + 2ti, \chi^2) = -\frac{\zeta'}{\zeta}(\sigma + 2it) + \sum_{p|N} \ln p \cdot \underbrace{\frac{p^{-(\sigma+2ti)}}{1 - p^{-(\sigma+2ti)}}}_{O(1) \text{ when } \sigma \geq 1}$$

L-zero-free-eq0
$$\Re\left(-\frac{L'}{L}(\sigma + 2ti, \chi^2)\right) \leq O(\ln(|t| + 2)) + \Re\left(\frac{1}{(\sigma + 2ti) - 1}\right) + O(\ln N), \tag{4.12}$$

the last inequality following from Lemma 3.3.1.

Putting (4.8), (4.9), and (4.12) into (4.11) give

L-zero-free-eq1
$$0 \leq \left(\frac{3}{\sigma - 1} + O(\mathcal{L})\right) + \left(-4\sum_{\rho} \Re\left(\frac{1}{\sigma + ti - \rho}\right) + O(\mathcal{L})\right) + \left(\Re\left(\frac{1}{\sigma + 2ti - 1}\right) + O(\mathcal{L})\right) \tag{4.13}$$

Fix $C' > 0$; when $s = \sigma + it$ and $|t| \geq \frac{C'}{\ln N}$ then $\frac{1}{\sigma+2ti-1} = O(\ln N)$ so (4.11) holds and we proceed as in item 1.

Hence we consider $t < \frac{C'}{\ln N}$. We use a different approach. Note

$$-\frac{L'}{L}(\sigma, \chi_0) \geq \frac{L'}{L}(\sigma, \chi) \quad \text{when } \sigma \geq 1$$

because the coefficients their coefficients are $\Lambda(n) \geq -\chi(n)\Lambda(n)$ (and they are real)[2]. Putting in (4.8) and (4.9) give

$$\frac{1}{\sigma - 1} \geq \sum_{\rho} \Re\left(\frac{1}{\sigma - \rho}\right) + O(\ln N). \tag{4.14}$$

Let $\sigma = 1 + \frac{2\delta}{\ln N}$; we estimate the sum in terms of the real parts of $\sigma - \rho$. For any zero ρ we have

$$|\Im\rho| \leq \frac{\delta}{\ln N} = \frac{1}{2}\frac{2\delta}{\ln N} \leq \Re(\sigma - \rho)$$

L-zero-free-eq2
$$|\sigma - \rho|^2 = [\Im(\sigma - \rho)]^2 + [\Re(\sigma - \rho)]^2 \tag{4.15}$$

$$\leq \left(\frac{1}{4} + 1\right) \Re(\sigma - \rho)^2 = \frac{5}{4}\Re(\sigma - \rho)^2. \tag{4.16}$$

Hence (4.14) gives, for some constant A,

$$\begin{aligned}\left(A + \frac{1}{2\delta}\right) \ln N = \frac{1}{1 - \sigma} + A \ln N &\geq \sum_{|\Im(\rho)| < \frac{\delta}{\ln N}} \Re\left(\frac{1}{\sigma - \rho}\right) \\ &= \sum_{|\Im(\rho)| < \frac{\delta}{\ln N}} \frac{\Re(\sigma - \rho)}{|\sigma - \rho|^2} \\ &\geq \sum_{|\Im(\rho)| < \frac{\delta}{\ln N}} \frac{4}{5} \sum_{\rho} \frac{1}{1 + \frac{2\delta}{\ln N} - \Re(\rho)} \qquad \text{by (4.15).}\end{aligned}$$

If $\Re(\rho) > 1 - \frac{c}{\ln N}$ then it contributes $\frac{4}{5}\frac{\ln N}{2\delta + c}$ to the RHS sum. If there are two zeros (counting multiplicity), then

$$\frac{8}{5}\frac{1}{2\delta + c} \leq A + \frac{1}{2\delta}.$$

This would be a contradiction if

$$c < \frac{2\delta(3 - 10A\delta)}{5(2\delta A + 1)}.$$

Now choose δ small enough and c so that it works for case 1 and satisfies the above inequality.

Finally, note $\zeta(\overline{s}, \chi) = \overline{\zeta(s, \chi)}$ for real characters, so if s is an (exceptional) zero so is $\overline{s}$. Since there is at most one exceptional zero, it can only be real. □

[2] Alternatively, put in $t = 0$ in (4.11).

4 Prime number theorem in arithmetic progressions

Theorem 4.4.1 (von Mangoldt's formula). *L-von-Mangoldt-formula For integer $x > 2$, $x \geq T$, and χ primitive of level $N > 1$,*

$$\psi(x,\chi) = -\sum_{|\Im(\rho)|<T} \frac{x^\rho}{\rho} + O\left(\frac{x[(\ln x)^2 + (\ln NT)^2]}{T}\right).$$

If χ has associated primitive character χ_1, then for $x \geq 1$,

$$|\psi(x,\chi) - \psi(x,\chi_1)| = O(\ln N \ln x).$$

Note that unlike in Theorem 3.4.1, we have $\psi(x,\chi) \approx 0$ as opposed to $\psi(x) \approx x$. Remember this is expected because the average of values for a nontrivial character is 0, so there is cancellation in the sum. Moreover, there is no pole at $s = 1$ for L as there was in ζ, so the application of Cauchy's Theorem in Step 2 will not give the x term.

Proof. Step 1: We estimate $\psi(x)$ using Theorem 2.4.2. Suppose x is an integer; the theorem gives

$$\begin{aligned}\left|\psi(x,\chi) - \left(\int_{c-iT}^{c+iT} x^s \left(-\frac{L'}{L}(s,\chi)\right)\frac{ds}{s}\right)\right| &\leq \Lambda(x) + \sum_{n\geq 1,\, n\neq x} \left(\frac{x}{n}\right)^c \chi(n)\Lambda(n)\frac{1}{T\left|\ln\left(\frac{x}{n}\right)\right|} \\ &\leq \ln(x) + \sum_{n\geq 1,\, n\neq x}^{\infty} \left(\frac{x}{n}\right)^c \frac{\ln(n)}{T\left|\ln\left(\frac{x}{n}\right)\right|}.\end{aligned}$$

The difference is $O\left(\frac{x(\ln x)^2}{T}\right)$ exactly as in (3.13).

Step 2: We move the line of integration to $\Re s = -1$. Assuming that T is not the imaginary part of any root, by Cauchy's theorem

$$\int_{c-iT}^{c+iT} \frac{x^s}{s}\frac{L'}{L}(s,\chi)\,ds + \underbrace{\int_{c+iT}^{-1+iT} \frac{x^s}{s}\frac{L'}{L}(s,\chi)\,ds}_{I_{h,1}} + \underbrace{\int_{-1+iT}^{-1-iT} \frac{x^s}{s}\frac{L'}{L}(s,\chi)\,ds}_{I_v} + \underbrace{\int_{-1-iT}^{c-iT} \frac{x^s}{s}\frac{L'}{L}(s,\chi)\,ds}_{I_{h,2}}$$
$$= \frac{L'}{L}(0,\chi) - \sum_{|\Im\rho|<T} \frac{x^\rho}{\rho}. \quad (4.17)$$

so

$$\int_{c-iT}^{c+iT} \frac{x^s}{s}\left(-\frac{L'}{L}(s)\right)ds = I_{h,1} + I_{h,2} + I_v + \frac{L'}{L}(0,\chi) - \sum_{\Im\rho<T} \frac{x^\rho}{\rho}. \quad (4.18)$$

We estimate each summand.

1. For the horizontal integrals, we use the estimate 4.3.1 to get

$$\begin{aligned}\left|\frac{\zeta'}{\zeta}(s)\right| &= \left|\sum_{|\Im(s-\rho)|<1} \frac{1}{s-\rho}\right| + O(\ln NT), \quad s = \sigma + Ti \\ &\le \sum_{|\Im(s-\rho)|<1} \frac{1}{\Im(s-\rho)} + O(\ln NT).\end{aligned}$$

   We would like to bound $\Im(s-\rho)$ away from 0. To do this, note that for $|T| > 2$ large there are $O(\ln NT)$ roots in with $\Im\rho \in \pm[T, T+1]$ by Lemma 4.3.1. Hence by tweaking T slightly we can assume $|\Im(s-\rho)| > \frac{C}{\ln|NT|}$. Also by Lemma 4.3.1 there are at most $O(\ln NT)$ terms in the sum, so the sum is $O((\ln NT)^2)$. Integrating gives

$$\begin{aligned}\left|\int_{c\pm Ti}^{-1\pm Ti} \frac{x^s}{s}\frac{L'}{L}(s,\chi)\,ds\right| &= O((\ln NT)^2)O\left(\frac{1}{T}\right)\int_c^{-1} |x^s|\,ds \\ &= O\left(\frac{x(\ln NT)^2}{T}\right).\end{aligned}$$

2. For the vertical integral, we use the same estimate, this time noting that $|s-\rho| > 1$ for every nontrivial zero ρ, since $\Re\rho > 0$. This gives that $\frac{\zeta'}{\zeta}(s) = O(\ln NT)$ and

$$\begin{aligned}\int_{-1+Ti}^{-1-Ti} \frac{x^s}{s}\frac{L'}{L}(s,\chi)\,ds &= O(\ln NT)\int_{-1-Ti}^{-1+Ti} \frac{x^{-1}}{|s|}\,ds \\ &= O\left(\frac{\ln(NT)\ln(T)}{x}\right) = O\left(\frac{x(\ln NT)^2}{T}\right).\end{aligned}$$

3. Note by Lemma 4.3.1 that $\frac{L'}{L}(0,\chi) = O(\mathcal{L}) = O(\ln(N+1))$.

Step 1 and (4.18) together with the above estimates give the first part of the theorem.

For the second part, note that

$$\begin{aligned}\psi(x,\chi_1) - \psi(x,\chi) &= \sum_{1\le n\le x} (\chi_1(n) - \chi(n))\Lambda(n)n^{-s} \\ &\le \sum_{1\le n\le x,\, n=p^r,\, p|N} \Lambda(n) \\ &\le \sum_{p|N} \left\lfloor \frac{\ln x}{\ln p} \right\rfloor \ln p \\ &\le \sum_{p|N} \ln x \ln p = \ln x \ln N.\end{aligned}$$

□

Theorem 4.4.2. *only-1-char* *There is a constant $c > 0$ such that for any distinct real χ_1 and χ_2 to moduli N_1 and N_2, at most one of $L(s,\chi_1)$ and $L(s,\chi_2)$ has a zero $\beta > 1 - \frac{c}{\ln(N_1N_2)}$.*

Corollary 4.4.3. *only-1-char-2* *There is a constant $c > 0$ such that the following holds: Fix a level N. There is at most 1 character χ of level N such that $L(s,\chi)$ has a zero with $\sigma \ge 1 - \frac{c}{\mathcal{L}}$.*

Proof of Theorem 4.4.2. The product $\chi_1\chi_2$ is a character with modulus N_1N_2. By Theorem 4.3.1, $-\frac{L'}{L}(\sigma,\chi) < O(\ln N_1N_2)$ for $1<\sigma<2$. Let

$$F(s) = \zeta(s)L(s,\chi_1)L(s,\chi_2)L(s,\chi_1\chi_2).$$

Then by logarithmic differentiation,

$$\begin{aligned} -\frac{F'}{F}(s) &= -\frac{\zeta'}{\zeta}(s) - \frac{L'}{L}(s,\chi_1) - \frac{L'}{L}(s,\chi_2) - \frac{L'}{L}(s,\chi_1\chi_2) \\ &= \sum_{n=1}^{\infty}(1+\chi_1(n)+\chi_2(n)+\chi_1(n)\chi_2(n))\Lambda(n)n^{-s} \\ &= \sum_{n=1}^{\infty}(1+\chi_1(n))(1+\chi_2(n))\Lambda(n)n^{-s} \geq 0 \end{aligned} \tag{4.19}$$

onechar-eq (4.20)

since the coefficients are nonnegative.

Suppose β_1,β_2 are exceptional zeros of $L(s,\chi_1), L(s,\chi_2)$; then putting Lemma 3.3.1 into (4.20) gives

$$O(\ln N_1N_2) + \frac{1}{\sigma-1} - \frac{1}{\sigma-\beta_1} - \frac{1}{\sigma-\beta_2} \geq 0.$$

Let $\delta = \min(1-\beta_1, 1-\beta_2)$. Take $\sigma = 1+2\delta$ to get $\frac{1}{6\delta} \leq O(\ln N_1N_2)$, i.e. $\delta \gtrsim \ln N_1N_2$ with constant independent of N_1, N_2, i.e. there is an appropriate choice of constant so that χ_1,χ_2 are not both exceptional for level N_1N_2. □

Proof of Corollary 5.2. Fix a primitive character χ of level N. Suppose χ' is of level N, whose corresponding primitive characters has level N'. Then the theorem gives c such that at most one of $L(s,\chi')$ and $L(s,\chi)$ has a zero $\beta > 1 - \frac{c}{\ln N'N} \geq 1 - \frac{c}{\ln N}$. □

Theorem 4.4.4 (Prime number theorem in arithmetical progressions)**.** pntap *Let $C>0$ and suppose $x > e^{C(\ln N)^2}$. If there is no exceptional zero for level N, there exists $C'>0$ such that*

$$\pi(x, a \bmod N) = (1+O(e^{-C'\sqrt{\ln x}}))\frac{\operatorname{li}(x)}{\varphi(N)}.$$

If there is an exceptional zero β of level N with associated character χ,

$$\pi(x, a \bmod N) = \frac{1}{\varphi(N)}(\operatorname{li}(x) - \chi(a)\operatorname{li}(x^\beta) + O(xe^{-C'\sqrt{\ln x}})).$$

Proof. We have by column orthogonality 1.1.6 that

psi-mod-chi $$\psi(\chi, a \bmod N) = \sum_{n\leq x,\, n\equiv a \,(\mathrm{mod}\ N)} \chi(n)\Lambda(N) = \sum_{n\leq x}\frac{1}{\varphi(N)}\sum_{\chi\in(\mathbb{Z}/N\mathbb{Z})^\times}\overline{\chi}(a)\chi(n)\Lambda(n) = \frac{1}{\varphi(N)}\sum_{\chi}\overline{\chi}(a)\psi(x$$ (4.21)

Letting χ_1 be the primitive character associated to χ, by Theorem 4.4.1 we have

$$\psi(x,\chi) = \begin{cases} -\sum_{\rho \text{ zero of } \psi(x,\chi_1)} \frac{x^\rho}{\rho} + O\left(\frac{x[(\ln x)^2 + (\ln NT)^2]}{T} + \ln N \ln x\right), & \chi \text{ nontrivial} \\ \psi(x) + O(\ln N \ln x), & \chi \text{ trivial.} \end{cases} \tag{4.22}$$

We estimate $\sum_{\rho \text{ nonexceptional zero of } \psi(x,\chi_1)} \frac{x^\rho}{\rho}$ in two steps.[3] Assume $T \geq 2$.

1. By Theorem 4.3.3, there is a constant c such that for all $|\Im(\rho)| < T$,

$$|x^\rho| = x^{\Re\rho} \leq x^{1-\frac{c}{\ln NT}} = xe^{-\frac{c\ln x}{\ln NT}}$$

2. Note the zero free region in Theorem 4.3.3 means there is a constant d_0, independent of N, χ, so that for all nonexceptional roots ρ, $|\rho| \geq d_0$. Hence using $N(T) = O(T \ln NT)$ (Lemma 3.3.1 or Theorem 4.3.2),

$$\begin{aligned} \sum_{|\Im(\rho)|<T} \frac{1}{|\rho|} &\leq \sum_{|\Im(\rho)|<T} \frac{1}{\max(\Im(\rho), d_0)} \\ &\leq \int_0^T \frac{dN(t)}{\max(t, d_0)} && \text{(Riemann-Steltjes integral)} \\ &= \frac{N(T)}{\max(T, d_0)} + \int_{d_0}^T \frac{N(t)}{t^2}\, dt && \text{integration by parts} \\ &= O(\ln NT) + \int_{d_0}^T O\left(\frac{\ln Nt}{t}\right) dt \\ &= O(\ln NT) + O((\ln NT)^2) = O((\ln NT)^2). \end{aligned}$$

Putting these two estimates together,

$$\begin{aligned} \left| \sum_{|\Im(\rho)|<T,\, \rho \text{ nonexceptional}} \frac{x^\rho}{\rho} \right| &\leq \max_{|\Im(\rho)|<T}(|x^\rho|) \sum_{|\Im(\rho)|<t} \frac{1}{|\rho|} \\ &\leq O\left(e^{-\frac{c\ln x}{\ln NT}} (\ln NT)^2\right). \end{aligned}$$

Combining with Theorem 4.4.1, setting $T = e^{\sqrt{\ln x}}$, and using $N < e^{C\sqrt{\ln x}}$ we get

$$\begin{aligned} |\psi(x,\chi) - x| &= O\left(xe^{-\frac{c\ln x}{\ln NT}} (\ln NT)^2 + \frac{x[(\ln x)^2 + (\ln NT)^2]}{T} + \frac{x(\ln T)^2}{T}\right) - \frac{x^\beta}{\beta} - \frac{x^{1-\beta}}{1-\beta} \\ &= O\left(xe^{-\frac{c\sqrt{\ln x}}{C+1}} (C+1)^2 \ln x + xe^{-\sqrt{\ln x}}((\ln x)^2 + (C+1)^2 \ln x) + C(\ln x)^{\frac{3}{2}}\right) - \frac{x^\beta}{\beta} \\ &= O(xe^{-C_1\sqrt{\ln x}}) - \frac{x^\beta}{\beta} \end{aligned} \tag{4.23}$$

[3] Here "nonexceptional" means with respect to level N.

for some $C_1 > 0$ independent of N, χ, where the implied constant is independent of N, χ.

For the trivial character, (4.22) and (3.17) give

psi-chi-asymptotic2 $$\psi(x,\chi) = x + O(xe^{-C_2\sqrt{\ln x}} + \ln x \ln T) = x + O(xe^{-C_2\sqrt{\ln x}}) \tag{4.24}$$

Using (4.21), (4.23), and (4.24), we get

$$\psi(\chi, a \bmod N) = \frac{1}{\varphi(N)}\left(x - \frac{\chi(a)x^\beta}{\beta} + O(xe^{-C_3\sqrt{\ln x}})\right)$$

where the grayed-out portion appears only if there is an exceptional zero. (Note this can happen for at most 1 character by Lemma 4.4.2.) It remains to transfer the asymptotics of ψ to that for π.

The same argument as in Lemma 3.4.2 shows that

$$\pi(x, a \bmod N) = \frac{\psi(x, a \bmod N)}{\ln x} + \int_2^x \psi(y)\frac{dy}{y(\ln y)^2} + O(x^{\frac{1}{2}}),$$

giving the estimate for π. □

5 Siegel zero

sec:siegel-zero In this section we obtain bounds on the exceptional zero to get a better error bound for prime number theorem on arithmetic progressions. We proceed in 2 steps.

1. Show that $L'(\beta, \chi)$ is small for β close to 1.
2. Bound $L(1, \chi)$ away from 0.

From this, we get that $L(\beta, \chi)$ cannot be 0 for β too close to 1.

Then we will be able to show the following improved form of Theorem 4.4.4.

Theorem 4.5.1 (Siegel-Walfisz)**.** *Given any C there exists a constant C' depending only on C so that*

$$\pi(x, a \bmod N) = \frac{\mathrm{li}(x)}{\varphi(N)} + O(xe^{-C'(\ln x)^{\frac{1}{2}}})$$

whenever

$$N \le (\ln x)^C.$$

Unfortunately, this bound is *ineffective*; the proof does not give a way to compute a suitable value of C'.

Of course, if the Riemann hypothesis were true then it would solve all our problems.

Theorem 4.5.2. *If the Extended Riemann hypothesis holds (all nontrivial zeros of $L(s, \chi)$ satisfy $\Re s = \frac{1}{2}$), then*

$$\pi(x, a \bmod N) = \frac{\mathrm{li}(x)}{\varphi(N)} + O(x^{\frac{1}{2}}(\ln x)^2)$$

for $x > N^2$, where the constant is independent of N.

5.1 $L'(\beta, \chi)$ is not too large

Lemma 4.5.3. *lem:L'-not-large There exists an absolute constant C such that*

$$|L'(\sigma, \chi)| < C(\ln N)^2$$

for any nontrivial Dirichlet character χ modulo N and any σ with $1 - \frac{1}{\ln N} \le \sigma \le 1$.

Proof. Because $L(\sigma, \chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^\sigma}$, by Proposition 2.2.4 we can simply differentiate term-by-term to get

$$L'(\sigma, \chi) = -\sum_{n=1}^{\infty} \frac{\chi(n) \ln n}{n^\sigma}.$$

Now we bound this sum by breaking it up into two parts.

First note that for $n \le N$, we have

$$1 - \sigma \le \frac{1}{\ln N} \le \frac{1}{\ln n}.$$

Hence

$$\frac{1}{n^\sigma} = \frac{1}{n} n^{1-\sigma} \le \frac{1}{n} n^{\frac{1}{\ln n}} = \frac{e}{n}. \tag{4.25}$$

eq:siegel-zero-1

Step 1: We bound the sum from $n = 1$ to N. By (4.25),

$$\left| \sum_{n=1}^{N} \frac{\chi(n) \ln n}{n^\sigma} \right| \le \sum_{n=1}^{N} \frac{e \ln n}{n} < C_1 (\ln N)^2 \tag{4.26}$$

eq:siegel-zero-2

for some C_1. The last step follows from estimating using the integral $\int_1^N \frac{\ln x}{x}\, dx = \frac{1}{2}(\ln N)^2$.

Step 2: Now we consider the sum from $N+1$ to ∞. Let $U(n) := \sum_{m=L+1}^{n} \chi(m)$ and $v(n) = \frac{\ln n}{n^\sigma}$. By partial summation 1.5.1, we have

$$\sum_{n=N+1}^{\infty} \frac{\chi(n) \ln n}{n^\sigma} = \lim_{L \to \infty} \left[-U(L)v(L) + \sum_{n=N+1}^{L} U(n-1)(v(n) - v(n-1)) \right].$$

Since $v(n)$ decreases to 0 and $|U(n)| \le N$ (as $\sum_{n=k}^{k+N-1} \chi(n) = 0$ for any k), the first term goes to 0 and we get the bound

$$\left| \sum_{n=N+1}^{\infty} \frac{\chi(n) \ln n}{n^\sigma} \right| \le N v(N) = N \frac{\ln N}{N^\sigma} \le N (\ln N) \frac{e}{N} = e \ln N. \tag{4.27}$$

eq:siegel-zero-3

where in the last step we used (4.25).

Adding (4.26) and (4.27) together gives the desired bound. □

5.2 $L(1,\chi)$ is not too small

Theorem 4.5.4 (Siegel's inequality). except-zero *For each $\varepsilon > 0$ there exists $C_\varepsilon > 0$ such that*

$$L(1,\chi) > C_\varepsilon N^{-\varepsilon}$$

for all real Dirichlet characters χ modulo N.

Thus there exists $C'_\varepsilon > 0$ such that any real zero β of $L(s,\chi)$ satisfies $1-\beta > C'_\varepsilon N^{-\varepsilon}$.

First we prove the following lemma.

Lemma 4.5.5. lem:l1chi-not-small *Let χ_1 and χ_2 be real primitive characters with modulus N_1 and N_2, let*

$$F(s) := \zeta(s)L(s,\chi_1)L(s,\chi_2)L(s,\chi_1\chi_2),$$

and let

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

Then the following inequality holds:

$$F(s) > \frac{1}{2} - \frac{C\lambda}{1-s}(N_1N_2)^{8(1-s)}, \qquad \frac{7}{8} < s < 1.$$

Note the technique of getting information about a L-function of a *single* character by looking at $F(s)$—a function defined using *two* characters—is a lot like what we did in showing Corollary only-1-char-2 using Theorem 4.4.2. We'll comment more later on why we looked at $F(s)$.[4]

Proof. The main idea is to expand $F(s)$ in power series and bound its coefficients (equivalently, bound the derivatives of $F(s)$) using the inequality from Cauchy's formula, Corollary 1.4.6.

We have

$$\begin{aligned}
\ln F(s) &= \ln\zeta(s) + \ln L(s,\chi_1) + \ln L(s,\chi_2) + \ln L(s,\chi_1\chi_2) \\
&= \sum_p \left(\ln\frac{1}{1-p^{-s}} + \ln\frac{1}{1-\chi_1(p)p^{-s}} + \ln\frac{1}{1-\chi_2(p)p^{-s}} + \ln\frac{1}{1-\chi_1(p)\chi_2(p)p^{-s}} \right) \\
&= \sum_p \sum_{m=1}^{\infty} \left(\frac{1}{m}p^{-ms} + \frac{1}{m}\chi_1(p^m)p^{-ms} + \frac{1}{m}\chi_2(p^m)p^{-ms} + \frac{1}{m}\chi_1(p^m)\chi_2(p^m)p^{-ms} \right) \\
&= \sum_p \sum_{m=1}^{\infty} \frac{1}{m}(1+\chi_1(p^m))(1+\chi_2(p^m))p^{-ms}.
\end{aligned}$$

[4] A deeper reason why we often look at $F(s)$ is that it is the zeta function of a *biquadratic field*. Thus we can prove nice facts about $F(s)$ by combining algebraic and analytic theory. We'll give proofs that don't require this knowledge.

This means $\ln F(s)$ is a Dirichlet series with all coefficients positive. Because the power series of e^x has positive coefficients, this means that $F(s)$ also has all coefficients positive. We're allowed to substitute any absolutely convergent series into a power series. (Is this right?) Suppose $F(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s}$.

Now we expand $F(s)$ in Taylor series at $s = 2$. (We can't do it at $s = 1$ because $F(s)$ has a pole there.) We have

$$F(s) = \sum_{m=0}^{\infty} a_m (2-s)^m, \qquad a_m = (-1)^m \frac{F^{(m)}(2)}{m!}.$$

We calculate the coefficients using 2.2.4 and get

$$a_m = \sum_{n=1}^{\infty} \frac{f(n)(\ln n)^m}{n^2} \geq 0.$$

In particular, for $m = 1$ we have $a_m \geq 1$ since $f(1) \geq 1$. It's 4.

Because we know $F(s)$ has a pole of residue $\lambda = L(1,\chi_1)L(1,\chi_2)L(1,\chi_1\chi_2)$, we consider the function

$$F(s) - \frac{\lambda}{s-1} = F(s) - \frac{\lambda}{1-(2-s)} = \sum_{m=0}^{\infty} (a_m - \lambda)(2-s)^m.$$

Let Ω be the circle of radius $\frac{3}{2}$ (not its interior) centered at 2. Then for any χ of modulus N, $|L(s,\chi)| \leq C_1 N$ for some C_1, for all s in a bounded region away from 0 because by (4.2)

$$|L(s,\chi)| = \left| \int_1^{\infty} S(x) s x^{-s-1}\, dx \right| \leq N \int_1^{\infty} |s x^{-s-1}|\, dx, \qquad S(x) = \sum_{n \leq x} \chi(n).$$

Therefore,

$$|F(s)| \leq (C_1 N_1)(C_1 N_2)(C_1 N_1 N_2) = C_2 (N_1 N_2)^2, \qquad C_2 = C_1^4 \tag{4.28}$$

and for $s \in \Omega$,

$$\left| \left(\frac{\lambda}{s-1} \right) \right| \leq 2L(1,\chi_1)L(1,\chi_2)L(1,\chi_1\chi_2) \leq 2C_2 (N_1 N_2)^2. \tag{4.29}$$

Now we use the inequality from Cauchy's formula, Corollary 1.4.6, to get

$$|a_m - \lambda| \leq \frac{1}{\left(\frac{3}{2}\right)^m} \max_{z \in \Omega} F(s) \leq C_3 N_1^2 N_2^2 \left(\frac{2}{3} \right)^m.$$

To bound $F(s) - \frac{\lambda}{s-1} = \sum_{m=0}^{\infty} (b_m - \lambda)(2-s)^m$ when $\frac{7}{8} < s < 1$, we first bound the sum from some M (to be determined) to ∞.

Firstly,

$$\begin{aligned}
\sum_{m=M}^{\infty} |a_m - \lambda|(2-s)^m &\leq \sum_{m=M}^{\infty} C_3 N_1^2 N_2^2 \left|\frac{2}{3}(2-s)\right|^m \\
&\leq \sum_{m=M}^{\infty} C_3 N_1^2 N_2^2 \left(\frac{3}{4}\right)^m, && \frac{7}{8} < s < 1 \\
&\leq C_4 N_1^2 N_2^2 \left(\frac{3}{4}\right)^M \\
&\leq C_4 N_1^2 N_2^2 e^{-M/4}, && e^{-1/4} \approx 0.78.
\end{aligned}$$

We choose M so that $C_4 N_1^2 N_2^2 e^{-\frac{M}{4}} \in \left[\frac{1}{2}e^{-\frac{1}{4}}, \frac{1}{2}\right]$. Note the lower bound rearranges to $M \leq 8 \ln N_1 N_2 + C_5$. Then because the coefficients a_m are all nonnegative, we can drop some of them in the inequality to get

$$\begin{aligned}
F(s) - \frac{\lambda}{s-1} &\geq 1 - \lambda \sum_{m=0}^{M-1} (2-s)^m - C_4 N_1^2 N_2^2 e^{-M/4} \\
&> 1 - \frac{\lambda}{1-s}[(2-s)^M - 1] - \frac{1}{2}, && C_4 N_1^2 N_2^2 e^{-\frac{M}{4}} \leq \frac{1}{2} \\
\implies F(s) &> \frac{1}{2} - \frac{\lambda}{1-s}(2-s)^M \\
&\geq \frac{1}{2} - \frac{\lambda}{1-s} e^{M(1-s)}, && e^x \leq 1 + x \\
&> \frac{1}{2} - \frac{C_6 \lambda}{1-s}(N_1 N_2)^{8(1-s)}, && M \leq 8 \ln N_1 N_2 + C_5.
\end{aligned}$$

This finishes the proof of the lemma. □

Proof of Theorem 4.5.4. Fix $\varepsilon > 0$. We want to choose χ_1 so that $0 \geq F(s)$. Consider two cases.

1. For some χ, $L(s, \chi)$ has a real zero in the range $\left(1 - \frac{1}{16}\varepsilon, 1\right)$. Then choose χ_1 to be this character and β_1 to be this zero. We then have $F(\beta_1) = 0$.

2. Else, let χ_1 be any primitive character and $\beta_1 \in \left(1 - \frac{1}{16}\varepsilon, 1\right)$. Note the following:

   - In this case there are no zeros for any L-function in $\left(1 - \frac{1}{16}\varepsilon, 1\right)$, so they all have the same sign as their value at 1. The value at 1 is nonnegative (in fact, positive) because the product expansion gives that the L-function is positive for $\sigma > 1$.
   - $\zeta(s) < 0$ for $0 < s < 1$, and

   Thus $F(\beta_1) < 0$.

In either case $F(\beta_1) \leq 0$, and the choice of β_1 depends only on ε. From Lemma 4.5.5, we now get the inequality

$$\begin{aligned} 0 &\leq \frac{1}{2} - \frac{C\lambda}{1-\beta_1}(N_1N_2)^{8(1-\beta_1)}. \\ \lambda &> C_{\varepsilon,1}(N_1N_2)^{-8(1-\beta_1)} \end{aligned}$$

for some $C_{\varepsilon,1}$ depending only on ε. Now we also have an upper bound for λ:

$$\begin{aligned} \lambda &= L(1,\chi_1)L(1,\chi_2)L(1,\chi_1\chi_2) \\ &< (C_1 \ln N_1)L(1,\chi_2)(C_1 \ln N_1N_2). \end{aligned}$$

Now suppose that $N_2 \geq N_1$. Combining the two inequalities and noting that $\ln N_1$ is a constant depending only on ε and is less than $\ln N_2$, we have

$$\begin{aligned} L(1,\chi_2) &> C_{\varepsilon,2}N_2^{-8(1-\beta_1)}(\ln N_2)^{-1} \\ &> C_{\varepsilon,2}N_2^{-\frac{\varepsilon}{2}}(\ln N_2)^{-1} \\ &> C_{\varepsilon,3}N^{-\varepsilon}. \end{aligned}$$

By choosing the constant to be smaller, we may ensure that this bound also works for $N_2 < N_1$.

Finally, combining Lemma 4.5.3 and the bound $L(1,\chi) > C_\varepsilon N^{-\varepsilon}$ immediately gives the fact that any real zero of $L(s,\chi)$ must satisfy $\beta < 1 - C'_\varepsilon N^{-\varepsilon}$. □

Note that it was essential to work with $F(s)$ rather than $G(s) = \zeta(s)L(s,\chi)$: Something like Lemma 4.5.5 would go through, but if we used $G(s)$ then $G(s)$ may have a zero close to $s = 1$ so we don't know the region where $G(s)$ is nonpositive, and we may have to take $s = \beta_1$ arbitrarily close to 1. This kills the proof because of the term $\frac{1}{1-s}$. When we work with $F(s)$, the case where there is a zero close to 1 is dealt with nicely.

5.3 Proof of Siegel-Walfisz

Proof of Theorem 4.5.1. Suppose there is an exceptional zero β. By Siegel's inequality 4.5.4, for any $\varepsilon > 0$ we have

$$\beta - 1 < -C_\varepsilon N^{-\varepsilon}.$$

The prime number theorem in arithmetic progressions 4.4.4 gives

$$\pi(x, a \bmod N) = \frac{1}{\varphi(N)}(\mathrm{li}(x) - \chi(a)\,\mathrm{li}(x^\beta) + O(xe^{-C'\sqrt{\ln x}})).$$

We show that $\mathrm{li}(x^\beta)$ gets absorbed into the O term. Indeed, we have

$$\begin{aligned} x^{-C_\varepsilon N^{-\varepsilon}} &\leq e^{-C'\sqrt{\ln x}} \\ \iff (\ln x) C_\varepsilon N^{-\varepsilon} &\geq C'\sqrt{\ln x} \\ \iff \sqrt{\ln x} &\geq \frac{C'}{C_\varepsilon} N^\varepsilon \\ \iff \left(\frac{C_\varepsilon}{C'}\right)^{\frac{1}{\varepsilon}} (\ln x)^{\frac{1}{2\varepsilon}} &\geq N. \end{aligned}$$

Now given $N \leq (\ln x)^C$, choose $\varepsilon = \frac{1}{2C}$. For large enough C', the equivalences above give $x^{-C_\varepsilon N^{-\varepsilon}} \leq e^{-C'\sqrt{\ln x}}$. Therefore,

$$\mathrm{li}(x^\beta) = O\left(x \frac{x^{\beta-1}}{\beta \ln x}\right) = O(x \cdot x^{-C_\varepsilon N^{-\varepsilon}}) = O(x e^{-C'\sqrt{\ln x}})$$

for some $C' > 0$, as needed. □

Appendix A

Arithmetic over Finite Fields

arith-over-ff This section is from my final paper in 18.784... need to integrate Our main goal in this chapter is to find a way to find the number of solutions for equations over finite fields. One problem we will look at in detail is, for a fixed b, how many solutions are there to

$$b = y_1^d + \cdots + y_n^d$$

over a finite field? We encapsulate the number of representations as a sum of n dth powers in a sum of orthonormal functions on $\mathbb{F}_q$ called the additive characters χ. We consider the product

$$\text{sum2} \left(\sum_{y \in \mathbb{F}_q} \chi(y^d) \right)^n = \sum_{y_1, \ldots, y_n \in \mathbb{F}_q} \chi(y_1^d + \cdots + y_n^d). \tag{A.1}$$

(The additive characters have the nice property that $\chi(a+b) = \chi(a)\chi(b)$.) Note (A.1) is true for all characters. To extract out the coefficient of $\chi(b)$, we multiply by $\overline{\chi(b)}$, average over all distinct characters χ, and take advantage of orthonormality to get

$$\text{thesum} r_{d,n}(b) = \frac{1}{q} \sum_{\chi} \left\{ \left(\sum_{y \in \mathbb{F}_q} \chi(y^d) \right)^n \overline{\chi(b)} \right\}. \tag{A.2}$$

In the next section we will give define and give properties of characters that help us estimate (A.2).

1 Characters

To evaluate (A.2) it would be helpful if $\chi(y^d) = \chi(y)^d$. However, this cannot hold as we defined χ so that it would preserve additive structure, not multiplicative structure. Thus to evaluate (A.2) we would like to rewrite it as a sum of functions ψ such that $\psi(ab) = \psi(a)\psi(b)$, and such that the set of ψ are orthonormal. Thus we will need both the concepts of additive and multiplicative characters. We make this precise below.

Definition 1.1.1: chardef Let G be an abelian group. A **character** of G is a homomorphism from G to $\mathbb{C}^\times$. A character is trivial if it is identically 1. We denote the trivial character by χ_0 or ψ_0.

Definition 1.1.2: Let R be a given finite ring. An additive character $\chi : R^+ \to \mathbb{C}$ is a character χ with R considered as an additive group. A multiplicative character $\psi : R^\times \to \mathbb{C}$ is a character with $R^\times$, the units of R, considered as a multiplicative group.

The two cases we will be working with are $R = \mathbb{Z}/N\mathbb{Z}$ (Section 1.1), and $R = \mathbb{F}_q$ (Section 1.2). We extend multiplicative characters ψ to R by defining $\psi(x) = 0$ for $x \in R \backslash R^\times$, except we follow the convention of setting $\psi_0(0) = 1$ when $R = \mathbb{F}_q$. Note that in any case the extended ψ still preserves multiplication.

We proceed to give an explicit description of characters for abelian groups. First, recall the following theorem.

Theorem 1.1.3 (Structure Theorem for Abelian Groups)**.** thm:structure-abelian *Let G be a finite abelian group. Then there exist positive integers $m_1, \dots, m_k$ so that*

$$G \cong \mathbb{Z}/m_1\mathbb{Z} \times \cdots \times \mathbb{Z}/m_k\mathbb{Z}.$$

Theorem 1.1.4 (Characters of abelian groups)**.** char *The group $G = \mathbb{Z}/m_1\mathbb{Z} \times \cdots \times \mathbb{Z}/m_k\mathbb{Z}$ has $|G|$ characters and each is given by an element $(r_1, \dots, r_k) \in \mathbb{Z}/m_1\mathbb{Z} \times \cdots \times \mathbb{Z}/m_k\mathbb{Z}$:*

$$\chi_{r_1,\dots,r_k}(n_1, \dots, n_k) = \prod_{j=1}^{k} e^{\frac{2\pi i r_j n_j}{m_j}}.$$

Moreover the set of characters $\widehat{G}$ form a multiplicative group isomorphic to G.[1]

Proof. Use Theorem 1.1.3. It is easy to check that $\chi = \chi_{r_1,\dots,r_k}$ is a homomorphism. Let e_j be the element in G with 1 in the jth coordinate and 0's elsewhere. Since $\chi(e_j)^{m_j} = 1$, we must have $\chi(e_j) = e^{\frac{2\pi i r_j}{m_j}}$ for some r_j. Each element of G can be expressed as a combination of the e_j, so this shows all characters are in the above form.

This shows that $(r_1, \dots, r_k) \mapsto \chi_{r_1,\dots,r_k}$ is surjective and hence an isomorphism. □

Corollary 1.1.5. numchar *Every finite abelian group G has $|G|$ characters.*

Theorem 1.1.6 (Orthogonality relations)**.** orth *Let G be a finite abelian group and χ_j, $1 \le j \le n$ be all characters of G. Then*

1. *(Row orthogonality)* $\langle \chi_j, \chi_k \rangle := \dfrac{1}{|G|} \displaystyle\sum_{g \in G} \chi_j(g)\overline{\chi_k(g)} = \begin{cases} 0, & j \neq k \\ 1, & j = k \end{cases}.$

2. *(Column orthogonality)* $\displaystyle\sum_{j=1}^{n} \chi_j(g)\overline{\chi_j(h)} = \begin{cases} 0, & g \neq h \\ |G|, & g = h \end{cases}.$

[1] This is a noncanonical isomorphism.

Proof. Write G as $\mathbb{Z}/m_1\mathbb{Z} \times \cdots \times \mathbb{Z}/m_k\mathbb{Z}$. Let $(r_1, \dots, r_k)$ and $(s_1, \dots, s_k)$ be in G. Then

$$\langle \chi_{r_1,\dots,r_k}, \chi_{s_1,\dots,s_k} \rangle = \sum_{(p_1,\dots,p_k) \in G} \prod_{j=1}^{k} e^{\frac{2\pi i (r_j - s_j) p_j}{m_j}} \text{yayy} \tag{A.3}$$

$$= \sum_{(p_1,\dots,p_{k-1}) \in G} \left[\left(\prod_{j=1}^{k-1} e^{\frac{2\pi i (r_j - s_j) p_j}{m_j}} \right) \sum_{p_k=0}^{m_k - 1} e^{2\pi i (r_k - s_k) p_k} \right] . \text{insum} \tag{A.4}$$

If $(r_1, \dots, r_k) = (s_1, \dots, s_k)$ then (A.3) evaluates to $|G|$. Otherwise, we may assume without loss of generality that $r_k \neq s_k$; then the inner sum in (A.4) evaluates to 0 by writing it as a geometric series.

The proof for column orthogonality is similar. □

The most useful case of row orthogonality is when we set $\chi_k = \chi_0$:

Corollary 1.1.7. sum0 *If χ is a character of G and $\chi \neq \chi_0$ then*

$$\sum_{g \in G} \chi(g) = 0.$$

Having established the basic properties of characters of abelian groups, we now turn to the specific cases $\mathbb{Z}/N\mathbb{Z}$ and $\mathbb{F}_q$.

1.1 Dirichlet characters

For our applications, it is helpful to think of consider characters on $\mathbb{Z}/N\mathbb{Z}$ as functions on $\mathbb{Z}$. From Theorem 1.1.4, the additive characters are simply given by

$$\chi_a(g) = e^{\frac{2\pi i a g}{N}}.$$

Next we consider multiplicative characters.

Definition 1.1.8: A **Dirichlet character** of level N is a function $\chi : \mathbb{Z} \to \mathbb{C}$ that induces a group homomorphism

$$\chi : (\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{C},$$

and such that $\chi(n) = 0$ for any n sharing a common factor with N. In other words, it induces a multiplicative character $\mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$.

We say χ is **principal** if $\chi(n) = 1$ for all $(\mathbb{Z}/N\mathbb{Z})^\times$, and **primitive** if χ does not induce a group homomorphism $(\mathbb{Z}/M\mathbb{Z})^\times \to \mathbb{C}$ for any $M < N$.

We say χ is **even** or **odd** if $\chi(-1) = 1$ or $\chi(-1) = -1$, respectively; we say χ is **real** when $\mathrm{im}(\chi) \subset \mathbb{R}$ and say it is **nonreal** otherwise.

Any character can be written uniquely as a product of a primitive character χ_1 of level $M \mid N$ and the principal character of level N:

$$\chi = \chi_1 \chi_0.$$

1.2 Characters on finite fields

field-char We give the additive and multiplicative characters on $\mathbb{F}_q$ explicitly. We know that $\mathbb{F}_q^\times$ is cyclic; let ξ be a generator.

Theorem 1.1.9 (Multiplicative characters of $\mathbb{F}_q$)**.** multc *The multiplicative characters of $\mathbb{F}_q$ are given by*

$$\psi_j(\xi^n) = e^{\frac{2\pi i j n}{q-1}}$$

for $0 \le j < q-1$.

Proof. By identifying $\xi \in \mathbb{F}_q^\times$ with $1 \in \mathbb{Z}/(q-1)\mathbb{Z}$, this follows directly from Theorem 1.1.4. □

Describing the additive characters takes slightly more creativity, since it is inconvenient to decompose $\mathbb{F}_q^+$ into cyclic groups.

Theorem 1.1.10 (Additive characters of $\mathbb{F}_q$)**.** *Suppose $q = p^r$ with p prime. The additive characters of $\mathbb{F}_q$ are given by*

$$\text{add}\chi_a(g) = e^{\frac{2\pi i}{p}\operatorname{tr}(ag)} \tag{A.5}$$

for $a \in \mathbb{F}_q$ where[2]

$$\operatorname{tr}(g) = g + g^p + \cdots + g^{p^{r-1}}.$$

Proof. The automorphisms of $\mathbb{F}_q$ fixing $\mathbb{F}_p$ are generated by the Frobenius automorphism σ sending g to g^p. Since $\operatorname{tr}(g)$ is fixed under this operation, it must be in the ground field $\mathbb{F}_p$. This makes (A.5) well-defined since only the value of $\operatorname{tr}(ag)$ modulo p matters in (A.5). The fact that χ_a is a homomorphism comes directly from the fact that σ is a homomorphism.

Since $\chi_1(ag) = \chi_a(g)$, if $\chi_a = \chi_b$ then $\chi_1(ag) = \chi_1(bg)$ and $\chi_1((a-b)g) = 0$. However, χ_1 is not trivial (identically equal to 1) since there are at most p^{r-1} values of g such that $g + \cdots + g^{p^{r-1}} = 0$. Thus $a = b$. This shows all characters in our list are distinct. Since we have found $|G|$ characters we have found all of them. □

Remark 1.1.11: In general, a n-dimensional complex representation of a group G is a homomorphism ρ from G into $GL_n(\mathbb{C})$, and the character χ of a representation is defined by $\chi(g) = \operatorname{tr}(\rho(g))$. This coincides with Definition 1.1.1 for abelian G, if we just consider 1-dimensional representations, since ρ is multiplication by a constant and χ is just that constant.

The general case of Corollary 1.1.5 is replaced by the following: every finite group has a number of irreducible characters equal to the number of conjugacy classes. The orthogonality relations hold when we consider just irreducible characters, and with $|G|$ replaced by the size of the centralizer of g in the equation for column orthogonality.

[2]For the general definition of trace see Definition **??.??**.

2 Gauss Sums

gauss-sums To relate additive characters to multiplicative characters, we need to evaluate sums in the form

$$\text{gauss}G(\psi,\chi)=\sum_{y\in R^\times}\psi(y)\chi(y). \tag{A.6}$$

where ψ is a multiplicative character and χ is an additive character.

Suppose we wanted to write an additive character on $\mathbb{F}_q$ in terms of multiplicative characters. By row orthogonality, $\frac{1}{q-1}\sum_{\psi\in\widehat{\mathbb{F}_q^\times}}\psi(y)\overline{\psi(g)}$ equals 1 if $y=g$ and is 0 otherwise. This allows us to introduce multiplicative characters as follows: for $y\in\mathbb{F}_q^\times$,

$$\begin{aligned}\chi(y)&=\frac{1}{q-1}\sum_{g\in\mathbb{F}_q^\times}\chi(g)\sum_{\psi\in\widehat{\mathbb{F}_q^\times}}\psi(y)\overline{\psi(g)}\\&=\frac{1}{q-1}\sum_{\psi\in\widehat{\mathbb{F}_q^\times}}\psi(y)\sum_{g\in\mathbb{F}_q^\times}\overline{\psi}(g)\chi(g)\\&\text{am}=\frac{1}{q-1}\sum_{\psi\in\widehat{\mathbb{F}_q^\times}}G(\overline{\psi},\chi)\psi(y).\end{aligned} \tag{A.7}$$

The Gauss sums are the coefficients of the expansion of χ in terms of multiplicative characters. The next theorem tells us how to calculate Gauss sums.

Theorem 1.2.1 (Magnitude of Gauss sums)**.** egau *Let ψ_0 and χ_0 denote the trivial multiplicative and additive characters on $\mathbb{F}_q$, respectively. Then for multiplicative and additive characters ψ and χ on $\mathbb{F}_q$, we have*

$$G(\psi,\chi)=\begin{cases}q-1, & \psi=\psi_0,\chi=\chi_0\\ -1, & \psi=\psi_0,\chi\neq\chi_0\\ 0, & \psi\neq\psi_0,\chi=\chi_0\end{cases}$$

and

$$|G(\psi,\chi)|=\sqrt{q},\ \psi\neq\psi_0,\chi\neq\chi_0.$$

If ψ is a nontrivial multiplicative character and χ is a primitive additive character on $\mathbb{Z}/N\mathbb{Z}$, then

$$|G(\psi,\chi)|=\sqrt{N}.$$

Proof. The first case is trivial. For the second case,

$$G(\psi_0,\chi)=\sum_{y\in\mathbb{F}_q^\times}\chi(y)=\left(\sum_{y\in\mathbb{F}_q}\chi(y)\right)-1=-1$$

by Corollary 1.1.7. The third case directly from Corollary 1.1.7 with ψ.

Now we consider the case case when ψ is nontrivial, and either $\chi \neq \chi_0$ (in the case $R = \mathbb{F}_q$) or χ is primitive (in the case $R = \mathbb{Z}/N\mathbb{Z}$), respectively. We have

$$\begin{aligned}
|G(\psi,\chi)|^2 &= \sum_{g_1,g_2 \in R^\times} \overline{\psi(g_1)}\psi(g_2)\overline{\chi(g_1)}\chi(g_2) \\
&= \sum_{g_1,g_2 \in R^\times} \psi(g_1^{-1}g_2)\chi(g_2 - g_1) \\
&= \sum_{h \in R^\times} \sum_{g_1 \in R^\times} \psi(h)\chi(g_1(h-1)) && \text{setting } h = g_1^{-1}g_2 \\
&= \sum_{h \in R^\times} \psi(h) \left[\left(\sum_{g_1 \in R} \chi(g_1(h-1)) \right) - \sum_{y \in R \setminus R^\times} \chi(y) \right] \\
&= \sum_{h \in R^\times} \psi(h) \left(\sum_{g_1 \in R} \chi(g_1(h-1)) \right) && \text{by Corollary 1.1.7 with } \psi
\end{aligned}$$

Now we note the following: when $h = 1$ all terms in the inner sum are 1, so it equals q or N, respectively. When $h \neq 1$, consider two cases.

1. $R = \mathbb{F}_q$: As g_1 ranges over $\mathbb{F}_q$, $g_1(h-1)$ ranges over $\mathbb{F}_q$.
2. $R = \mathbb{Z}/N\mathbb{Z}$: As g_1 ranges over $\mathbb{Z}/N\mathbb{Z}$, $g_1(h-1)$ ranges over a subgroup $H \subseteq \mathbb{Z}/N\mathbb{Z}$, hitting each element $\frac{N}{|H|}$ times. Since χ is primitive, $\chi|_H$ is nontrivial.

In either case, Corollary 1.1.7 gives the inner sum to be 0. Hence $|G(\psi,\chi)|^2$ evaluates to $\psi(1)q = q$ or $\psi(1)N = N$, respectively. □

We will need the following fact later on.

Proposition 1.2.2 (Gauss sum dependence on additive character)**:** gaussprop Let $R = \mathbb{F}_q$ or $\mathbb{Z}/N\mathbb{Z}$. For $a \in R^\times$ and $b \in R$,

$$G(\psi,\chi_{ab}) = \overline{\psi(a)}G(\psi,\chi_b).$$

Proof. Using the fact that $\chi_c(g) = \chi_1(cg)$,

$$\begin{aligned}
G(\psi,\chi_{ab}) &= \sum_{y \in R^\times} \psi(y)\chi_{ab}(y) \\
&= \sum_{y \in R^\times} \psi(y)\chi_b(ay) \\
&= \sum_{y \in R^\times} \psi(a^{-1}y)\chi_b(y) && \text{replacing } y \to a^{-1}y \\
&= \psi(a)^{-1} \sum_{y \in R^\times} \psi(y)\chi_b(y) \\
&= \overline{\psi(a)}G(\psi,\chi_b) && \square
\end{aligned}$$

3 Enumerating Solutions

We return to our original problem. Rather than just work with sums of dth powers, we work with diagonal equations

$$_{\text{diag}}a_1y_1^{d_1} + \cdots + a_ny_n^{d_n} = b \tag{A.8}$$

where $a_i \in \mathbb{F}_q^\times$ and $d_i \in \mathbb{N}$. First, note that because of the following lemma, we can restrict to case where $d_i | q-1$.

Lemma 1.3.1. *gcd The multisets $\{y^d | y \in \mathbb{F}_q\}$ and $\{y^{\gcd(d,q-1)} | y \in \mathbb{F}_q\}$ are equal.*

Proof. Let ξ be a generator for $\mathbb{F}_q^\times$, and write $d = k\gcd(d, q-1)$ where $\gcd(k, q-1) = 1$. Then removing the one occurrence of 0 in the two sets, we get $\{\xi^{jd} | 0 \le j < q-1\}$ and $\{\xi^{j\gcd(d,q-1)} | 0 \le j < q-1\}$. The lemma follows from the fact that as multisets,

$$\{jd \pmod{q-1} | 0 \le j < q-1\} = \{j\gcd(d, q-1) \pmod{q-1} | 0 \le j < q-1\}.$$

Indeed, each multiple of $\gcd(d, q-1)$ appears $\frac{q-1}{\gcd(d,q-1)}$ times on both sides. □

As (A.8) always has the trivial solution when $b = 0$, we just need to estimate the number of solutions to (A.8) when $b \neq 0$.

Theorem 1.3.2. *[?, 6.37]mainthm Fix $b \neq 0, d_i | q-1$ and let N be the number of solutions to (A.8) when $b \neq 0$ is fixed. Then*

$$|N - q^{n-1}| \le [(d_1 - 1)\cdots(d_n - 1) - (1 - q^{-\frac{1}{2}})M(d_1, \ldots, d_n)]q^{\frac{n-1}{2}}$$

where $M(d_1, \ldots, d_n)$ is the number of n-tuples in the set

$$S := \left\{(j_1, \ldots, j_n) \in \mathbb{Z}^n | 1 \le j_i \le d_i - 1 \text{ and } \sum_{i=1}^n \frac{j_i}{d_i} \in \mathbb{Z}\right\}.$$

Note that we would expect N to be close to q^{n-1}, because there are q^n possible choices for $(y_1, \ldots, y_n)$ and q possible values for their sum.

Proof. We use the idea mentioned in the introduction. We have

$$N = \frac{1}{q} \sum_{y_1,\ldots,y_n \in \mathbb{F}_q,\, \chi \in \widehat{\mathbb{F}_q^+}} \chi(a_1y_1^{d_1} + \cdots + a_ny_n^{d_n})\overline{\chi}(b) = \frac{1}{q} \sum_{y_1,\ldots,y_n \in \mathbb{F}_q,\, \chi \in \widehat{\mathbb{F}_q^+}} \chi(a_1y_1^{d_1}) \cdots \chi(a_ny_n^{d_n})\overline{\chi}(b)$$

since by row orthogonality the inner sum is 1 if $a_1y_1^{d_1} + \cdots + a_ny_n^{d_n} = b$ and 0 otherwise. Note that χ_0 contributes q^n to the sum. Taking it out and factoring the remaining terms gives

$$_{\text{N1}}N = q^{n-1} + \frac{1}{q} \sum_{\chi \in \widehat{\mathbb{F}_q^+}, \chi \neq \chi_0} \left(\overline{\chi}(b) \prod_{j=1}^n \sum_{y_j \in \mathbb{F}_q} \chi(a_jy_j^{d_j})\right) \tag{A.9}$$

We write the sums of additive characters as sums of multiplicative characters using the following lemma.

Lemma 1.3.3. ayn *Let χ be a nontrivial additive character and λ a multiplicative character of order d dividing $q-1$. Then*

$$\sum_{y\in\mathbb{F}_q}\chi(ay^d)=\sum_{j=1}^{d-1}\overline{\lambda}(a)^jG(\lambda^j,\chi).$$

Proof. Note that λ exists since the group of multiplicative characters is isomorphic to $\mathbb{Z}/(q-1)\mathbb{Z}$ by Theorem 1.1.4. Suppose $\chi=\chi_c$. We write χ as a sum of multiplicative characters using (A.7), get the Gauss sum to be independent of a by using Proposition 1.2.2, and take out the exponent as we were hoping to do:

$$\begin{aligned}\sum_{y\in\mathbb{F}_q}\chi(ay^d)&=\sum_{y\in\mathbb{F}_q}\chi_{ac}(y^d)\\&=1+\sum_{y\in\mathbb{F}_q^\times}\chi_{ac}(y^d)\\&=1+\frac{1}{q-1}\sum_{\psi\in\widehat{\mathbb{F}_q^\times}}\sum_{y\in\mathbb{F}_q^\times}G(\overline{\psi},\chi_{ac})\psi(y^d)\\&\overset{\text{insummy}}{=}1+\frac{1}{q-1}\sum_{\psi\in\widehat{\mathbb{F}_q^\times}}\overline{\psi}(a)G(\overline{\psi},\chi_c)\sum_{y\in\mathbb{F}_q}\psi(y)^d \qquad (A.10)\\&\overset{\text{eq1}}{=}1+\sum_{j=0}^{d-1}\overline{\lambda}(a)^jG(\lambda^j,\chi) \qquad (A.11)\\&\overset{\text{eq2}}{=}\sum_{j=1}^{d-1}\overline{\lambda}(a)^jG(\lambda^j,\chi) \qquad (A.12)\end{aligned}$$

Note (A.11) follows since by Corollary 1.1.7, $\sum_{y\in\mathbb{F}_q^\times}\psi(y)^d=0$ unless ψ^d is the trivial character, which is true iff ψ is a power of λ. In that case, the inner sum in (A.10) is $q-1$. In (A.12) we used $G(\psi_0,\chi)=-1$ (Theorem 1.2.1). □

Using Lemma 1.3.3 and letting λ_j be the multiplicative character with $\lambda_j(\xi^t)=e^{\frac{2\pi it}{d_j}}$ we rewrite (A.9) as

$$\begin{aligned}N-q^{n-1}&=\frac{1}{q}\sum_{\chi\in\widehat{\mathbb{F}_q^+},\chi\neq\chi_0}\left(\overline{\chi}(b)\prod_{j=1}^{n}\sum_{k=1}^{d-1}\overline{\lambda_j}(a_j)^kG(\lambda_j^k,\chi)\right)\\&=\frac{1}{q}\sum_{\chi\in\widehat{\mathbb{F}_q^+},\chi\neq\chi_0}\sum_{(k_1,\ldots,k_n),1\le k_i\le d_i-1}\overline{\chi}(b)\overline{\lambda_1}^{k_1}(a_1)\cdots\overline{\lambda_n}^{k_n}(a_n)G({\lambda_1}^{k_1},\chi)\cdots G({\lambda_n}^{k_n},\chi)\\&=\frac{1}{q}\sum_{c\in\mathbb{F}_q^\times}\sum_{(k_1,\ldots,k_n),1\le k_i\le d_i-1}\overline{\chi_c}(b)\overline{\lambda_1}^{k_1}(a_1)\cdots\overline{\lambda_n}^{k_n}(a_n)G({\lambda_1}^{k_1},\chi_c)\cdots G({\lambda_n}^{k_n},\chi_c)\\&\overset{\text{lotsofgauss}}{=}\frac{1}{q}\sum_{(k_1,\ldots,k_n),1\le k_i\le d_i-1}G({\lambda_1}^{k_1},\chi_{a_1})\cdots G({\lambda_n}^{k_n},\chi_{a_n})\sum_{c\in\mathbb{F}_q^\times}\overline{\chi_b}(c)\overline{\lambda_1}^{k_1}(c)\cdots\overline{\lambda_n}^{k_n}(c)\end{aligned} \qquad (A.13)$$

$$\text{log2} = \frac{1}{q} \sum_{(k_1,\dots,k_n),1\le k_i\le d_i-1} G(\lambda_1{}^{k_1},\chi_{a_1})\cdots G(\lambda_n{}^{k_n},\chi_{a_n})G(\overline{\lambda_1}^{k_1}\cdots\overline{\lambda_n}^{k_n},\overline{\chi_b}) \tag{A.14}$$

where in (A.13) we used Proposition 1.2.2 twice, to get

$$\overline{\lambda_j}^{k_j}(a_j)G(\lambda_j{}^{k_j},\chi_c) = \overline{\lambda_j}^{k_j}(c)\overline{\lambda_j}^{k_j}(a_j)G(\lambda_j{}^{k_j},\chi_1) = \overline{\lambda_j}^{k_j}(c)G(\lambda_j{}^{k_j},\chi_{a_j}).$$

Now we apply Theorem 1.2.1 to get that $|G(\lambda_i^{k_i},\chi_{a_i})| = \sqrt{q}$. Note

$$(\overline{\lambda_1}^{k_1}\cdots\overline{\lambda_n}^{k_n})(\xi^t) = e^{(2\pi i)\left(\frac{k_1}{d_1}+\cdots+\frac{k_n}{d_n}\right)t}$$

is the trivial character iff $(k_1,\dots,k_n)\in S$. Hence $|G(\overline{\lambda_1}^{k_1}\cdots\overline{\lambda_n}^{k_n},\overline{\chi_b})| = 1$ if $(k_1,\dots,k_n)\in S$ and $\sqrt{q}$ otherwise. Using this and the triangle inequality, (A.14) becomes

$$|N - q^{n-1}| \le \frac{1}{q}[q^{\frac{n}{2}}|S| + q^{\frac{n+1}{2}}((d_1-1)\cdots(d_n-1) - |S|)],$$

proving the theorem. □

4 Applications to Waring's Problem

Now we derive Small's bound for Waring's constant $g(d,q)$, the minimum n such that (A.8) has a solution with $d_1 = \cdots = d_n = d$ for all b. By Lemma 1.3.1, $g(d,q) = g(\gcd(d,q-1),q)$, so it suffices to consider the case $d|q-1$.

First, note that sufficient condition for Waring's constant to exist is that the set $\{y^d | y\in\mathbb{F}_q\}$ is not contained in a proper subfield of $\mathbb{F}_q$. Since this set is generated multiplicatively by ξ^d, and any subfield is multiplicatively generated by $\xi^{\frac{p^r-1}{p^k-1}}$ for some $k|d$, writing $q = p^r$ with p prime we need

$$\text{badcond}\ \frac{p^r-1}{p^k-1} \nmid d \quad \text{for every proper divisor } k \text{ of } r. \tag{A.15}$$

Apply Theorem 1.3.2 (dropping the term with $M(d_1,\dots,d_n)$) to get

$$\text{aboutsame}\ N \ge q^{n-1} - (d-1)^n q^{\frac{n-1}{2}} \tag{A.16}$$

This is positive when

$$\text{yayit}\ q^{\frac{n-1}{2}} > (d-1)^n \iff \frac{n}{2}(\ln q - 2\ln(d-1)) > \frac{\ln q}{2} \tag{A.17}$$

Thus we obtain the following bound for $g(d,q)$:

Theorem 1.4.1. *thm51* *Suppose $d|q-1$ and $q > (d-1)^2$. Then*

$$g(d,q) \le \left\lfloor \frac{\ln q}{\ln q - 2\ln(d-1)} + 1 \right\rfloor.$$

Note that in particular, (A.17) for $n=2$ allows us to make the "inverse" statement that if $q > (d-1)^4$, then the equation $y_1^d + y_2^d = b$ has a solution for any $b\in\mathbb{F}_q$. That is, for any d, in any sufficiently large finite field every element can be written as a sum of 2 dth powers.

5 Finite calculus

Theorem 1.5.1 (Summation by parts, Abel summation). sum-parts *Suppose that u is an arithmetic function, and let*

$$U(x) = \sum_{n \le x} u(x).$$

Then for $m, n \in \mathbb{N}$

$$\sum_{x=m}^{n} u(x)v(x) = U(n)v(n) - U(m-1)v(m-1) - \sum_{x=m}^{n} U(x-1)(v(x) - v(x-1)).$$

If $0 \le a < b$ and v has continuous derivative on $a < x < b$, then

$$\sum_{a \le x \le b} u(x)v(x) = U(b)v(b) - U(a)v(a) - \int_a^b U(x)v'(x).$$

Proof. We imitate the proof of integration by parts. For a function f define the function

$$\Delta_-(f) = f(x) - f(x-1).$$

This is the discrete analogue of differentiation. It is the inverse of summation in the sense that by telescoping,

$$\sum_{x=m}^{n} \Delta_-(f) = f(n) - f(m-1). \tag{A.18}$$

Note that $\Delta_-(U) = u$. We have the "product rule"

$$\begin{aligned} \Delta_-(uv) &= u(x)v(x) - u(x-1)v(x-1) \\ &= (u(x) - u(x-1))v(x) + u(x-1)(v(x) - v(x-1)) \\ &= \Delta_-(u)v + E_- u \Delta_-(v) \end{aligned}$$

where E_- is the left shift operator $(E_- f)(x) = f(x-1)$. Replacing u by U and rearranging gives

$$uv = \Delta_-(Uv) - E_- U \Delta_-(v).$$

Summing over $m \le x \le n$ and telescoping using (A.18) gives

$$\sum_{x=m}^{n} u(x)v(x) = U(n)v(n) - U(m-1)v(m-1) - \sum_{x=m}^{n} U(x-1)(v(x) - v(x-1)).$$

When v has continuous derivative, noting $U(t) = U(\lfloor t \rfloor)$, we have

$$\begin{aligned} \sum_{x=m}^{n} U(x-1)(v(x) - v(x-1)) &= \sum_{x=m}^{n} \int_{x-1}^{x} U(t)v'(t)\,dt \\ &= \int_{m-1}^{n} U(t)v'(t)\,dt. \end{aligned}$$

For general a, b, since U is constant on $(\lfloor b \rfloor, b)$ and $(a, \lfloor a \rfloor + 1)$,

$$\begin{aligned}\sum_{a<x\leq b} u(x)v(x) &= \sum_{x=\lfloor a \rfloor+1}^{\lfloor b \rfloor} u(x)v(x) \\ &= U(\lfloor b \rfloor)v(\lfloor b \rfloor) - U(\lfloor a \rfloor)v(\lfloor a \rfloor) + \int_{\lfloor a \rfloor}^{\lfloor b \rfloor} U(t)v'(t)\,dt \\ &= U(b)v(b) - U(a)v(a) + \int_a^b U(t)v'(t)\,dt \qquad \square\end{aligned}$$

Bibliography

[Ahl79] L. Ahlfors. *Complex Analysis*. McGraw-Hill, 1979.

[Apo94] T. Apostol. *Modular forms and Dirichlet series*. Number 110 in GTM. Springer, 2nd edition, 1994.

[GBGL10] T. Gowers, J. Barrow-Green, and I. Leader. *The Princeton Companion to Mathematics*. Princeton University Press, 2010.

Index

Mittag-Leffler, 6

prime number theorem, 17

Riemann hypothesis, 33

Stirling's approximation, 8

von Mangoldt's formula, 29
von Mangoldt's Theorem, 26

zeta function, 20