MAT 415 Notes

Max Chien

Fall 2025

Contents

Introduction

Chapter 1

Arithmetic Progressions and Finite Fourier Analysis

1.1 Dirichlet's Theorem on Primes in Progression

Theorem 1.1: Dirichlet

Given $a, q \in \mathbb{N}$ with $(a, q) = 1$, there are infinitely many primes p such that $p \equiv a \pmod{q}$.

We begin by considering Euler's proof of the infinitude of primes. Recall that the zeta function,

$$\sum_{n=1}^{\infty} n^{-s}$$

converges absolutely for $s > 1$ (for now we will work with real s), and diverges for $n = 1$. Moreover, consider the product

$$\prod_{p} (1 - p^{-s})^{-1} = \prod_{p} \sum_{k=0}^{\infty} p^{-ks}$$

Since we still have absolute convergence, we may rearrange the generic terms in the product. Each such term is of the form $(p_1^{k_1} \cdots p_m^{k_m})^{-s}$, and by unique factorization thsi means the term n^{-s} shows up exactly once for each n. Hence

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

for $\mathrm{Re}(s) > 1$. Taking $s \to 1$, the right hand side diverges so the left hand side does as well. Hence it is clear that there are infinitely many primes. So our goal will be to use a similar

strategy which demonstrates that

$$\sum_{p \equiv a \pmod{m}} \frac{1}{p} = \infty$$

To do this, consider the ring $\mathbb{Z}/m\mathbb{Z}$, as well as its group of units $(\mathbb{Z}/m\mathbb{Z})^*$. Recall that the totient function is

$$\phi(m) = \left|(\mathbb{Z}/m\mathbb{Z})^*\right| = \#\{\text{numbers } \leq m \text{ rel. prime to } m\}$$

We will work with the space of functions $f : (\mathbb{Z}/m\mathbb{Z})^* \to \mathbb{C}$. The goal is to define a sense of Fourier expansion for this vector space. We define

$$e(z) \coloneqq e^{2\pi i z}$$

For $\mathbb{R}/\mathbb{Z}$, we can perform such an expansion by observing that the set of $e(mx)$ for $m \in \mathbb{Z}$ defines an orthonormal basis for $L^2(\mathbb{R}/\mathbb{Z} \to \mathbb{C})$. More generally, if G is a finite abelian group, then we denote by $\hat{G}$ the group of its characters; that is, homomorphisms $\chi : G \to \mathbb{C}^*$. Since every element in G has finite order, each character maps into the roots of unity.

Remark

We denote additive characters by ψ and multiplicative ones by χ.

Proposition 1.2

$G \cong \hat{G}$.

Proof. First suppose G is cyclic. Then we can assume we are working with $(\mathbb{Z}/r\mathbb{Z}, +)$. Any additive character is determined by $\psi(1)$, and $\psi(1)$ is necessarily an rth root of unity. So the characters are precisely those of the form

$$\psi_\nu(x) = e\left(\frac{\nu x}{r}\right)$$

for $\nu \in \mathbb{Z}/r\mathbb{Z}$. So clearly $\left|\hat{G}\right| = |G|$. Also, $\psi_\nu \psi_\mu = \psi_{\nu+\mu}$, so the map $\nu \mapsto \psi_\nu$ is an onto homomorphism from G to $\hat{G}$, hence an isomorphism.

Homework: in the general case, use the classification of finite groups. □

Note that this isomorphism is not canonical, however the isomorphism $G \cong \hat{\hat{G}}$ is.

Definition 1.1

Suppose $\chi \in \hat{G}$, and $f : G \to \mathbb{C}$ is a function. Then we define the **Fourier transform** $\hat{f} : \hat{G} \to \mathbb{C}$ of f on G by

$$\hat{f}(\chi) = \sum_{g \in G} f(g)\chi(g)$$

Proposition 1.3

If χ_e denotes the trivial character which takes all elements to $1 \in \mathbb{C}$, then

$$\sum_{g \in G} \chi(g) = \begin{cases} |G|, & \chi = \chi_e \\ 0 \end{cases}$$

$$\sum_{\chi \in \hat{G}} \chi(g) = \begin{cases} \left|\hat{G}\right| = |G|, & g = e \\ 0 \end{cases}$$

Proof. For the first, the formula is obvious when $\chi = \chi_e$. Otherwise, there is an element a where $\chi(a) \neq 1$. But then

$$S = \sum_{g \in G} \chi(g) = \sum_{g \in G} \chi(ag) = S = \chi(a)S$$

But $\chi(a) \neq 1$, so we must have $S = 0$. Similarly, for the second, if $g \neq e$, then the canonical double dual $\hat{\hat{g}} \in \hat{\hat{G}}$ is not the identity. So there is $\chi' \in \hat{G}$ with $\chi'(g) \neq 1$. Then

$$S = \sum_{\chi \in \hat{G}} \chi(g) = \sum_{\chi\chi' \in \hat{G}} (\chi\chi')(g) = \chi'(g)S$$

so $S = 0$. □

Theorem 1.4: Fourier Inversion

For any $f : G \to \mathbb{C}$,

$$f(g) = \frac{1}{|G|} \sum_{\chi \in \hat{G}} \hat{f}(\chi)\overline{\chi(g)}$$

Proof. By orthogonality,

$$\begin{aligned} \frac{1}{|G|} \sum_{\chi \in \hat{G}} \hat{f}(\chi)\overline{\chi(g)} &= \frac{1}{|G|} \sum_{\chi \in \hat{G}} \sum_{h \in G} f(h)\chi(h)\overline{\chi(g)} = \frac{1}{|G|} \sum_{h \in G} f(h) \sum_{\chi \in G} \overline{\chi(g)}\chi(h) \\ &= \frac{1}{|G|} \sum_{h \in G} f(h) \sum_{\chi \in G} \chi(g^{-1}h) = \frac{1}{|G|} f(g)|G| = f(g) \end{aligned}$$

□

Now, we want to calculate the transform of the indicator function I_a for some $a \in G$. Then by Fourier inversion,

$$\hat{I}_a(\chi) = \sum_{g \in G} I_a(g)\chi(g) = \chi(a)$$

$$I_a(g) = \frac{1}{|G|} \sum_{\chi \in \hat{G}} \hat{I}_a(\chi)\overline{\chi(g)} = \frac{1}{|G|} \sum_{\chi \in \hat{G}} \chi(a)\overline{\chi(g)} = \frac{1}{|G|} \sum_{\chi \in \hat{G}} \chi(ag^{-1})$$

and by orthogonality we verify that this is correct.

Definition 1.2

Let $m \in \mathbb{N}$ and $\chi \in \widehat{(\mathbb{Z}/m\mathbb{Z})^*}$. Then the **Dirichlet L-function** associated with χ is the function $L(\cdot, \chi) : \mathbb{C} \to \mathbb{C}$, where

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

where we extend χ to the integers by defining

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

Note that because $|\chi| \leq 1$, the series converges absolutely for $\mathrm{Re}(s) > 1$, and the Euler product is given by

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

because $\chi(mn) = \chi(m)\chi(n)$. Actually, we can make do with a slightly weaker assumption so an alternate form of the Euler product holds.

Definition 1.3

Let $f : \mathbb{N} \to \mathbb{C}$. Then f is called **multiplicative** if $f(1) = 1$ and $f(mn) = f(m)f(n)$ whenever $(m, n) = 1$. It is called **totally multiplicative** if $f(mn) = f(m)f(n)$ for all m, n.

For any multiplicative f,

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

Proposition 1.5

1. $L(s, \chi)$ converges absolutely on $\mathrm{Re}(s) > 1$, and is analytic there. If $\chi \neq \chi_0$ then it is analytic on $\mathrm{Re}(s) > 0$ as well, though it converges conditionally ($L(s, \chi_e)$ has a pole at 0).
2. $L(s, \chi)$ is nonzero on $\mathrm{Re}(s) > 1$.
3. $L(s, \chi_e) = \zeta(s) \prod_{q|m} (1 - q^{-s})$.

Proof. 1. Convergence is easy since $|\chi(n)| \leq 1$. If $\chi \neq \chi_e$,

$$\sum_{n=1}^{m} \chi(n) = 0$$

so

$$\left|\sum_{n=1}^{T}\right|\chi(n) \leq m$$

for any T. Then we employ summation by parts:

$$\sum_{n\leq T}\chi(n)n^{-s} = \sum_{n\leq T}\left(\sum_{\nu\leq n}\chi(\nu)\right)[n^{-s}-(n+1)^{-s}] = \sum_{n\leq T}\left(\sum_{\nu\leq n}\chi(\nu)\right)\frac{s}{n^{-(s+1)}}$$

Since the factor $\sum_{\nu\leq n}\chi(\nu)$ is bounded, this sum converges for $\mathrm{Re}(s)>0$.

2. For $\mathrm{Re}(s)>1$, the Euler product

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

converges. But no factor is zero, so the whole product is not either.

3. Using the formula from part 2,

$$L(s,\chi_e) = \prod_{p\not| m}(1-p^{-s})^{-1} = \prod_p \left(1-p^{-s}\right)^{-1}\prod_{p|m}\left(1-p^{-s}\right) = \zeta(s)\prod_{p|m}\left(1-p^{-s}\right) \quad \square$$

For $\mathrm{Re}(s)>1$, we then have

$$\begin{aligned}\log L(s,\chi) &= -\sum_p \log(1-\chi(p)p^{-s}) = \sum_p\sum_{k=1}^{\infty}\frac{\chi(p^k)}{kp^{ks}} = \sum_p\frac{\chi(p)}{p^s} + \sum_p\sum_{k=2}^{\infty}p^{-ks}\chi(p^k)k^{-1}\\ &\leq \sum_p\frac{\chi(p)}{p^s} + \sum_p\sum_{k=2}^{\infty}\left(p^{-s}\chi(p)\right)^k = \sum_p\frac{\chi(p)}{p^s} + \sum_p\frac{p^{-2s}\chi(p)^2}{1-p^{-s}\chi(p)}\end{aligned}$$

The second term is uniformly bounded on $\mathrm{Re}(s)\geq\sigma_0>1/2$. By applying Fourier inversion and evaluating at the indicator function of a mod m,

$$\sum_{\chi\in\hat{G}}\chi(a)\log L(s,\overline{\chi}) = |G|\sum_{p\equiv a(m)}p^{-s} + O_{m,\sigma_0}(1)$$

The notation O_{m,σ_0} means that the last quantity is uniformly bounded, but the constant depends on m,σ_0. Now we extract the trivial character as

$$|G|\sum_{p\equiv a(m)}p^{-s} = \log L(s,\chi_e) + \sum_{\chi\neq\chi_e}\chi(a)\log L(s,\overline{\chi}) + O_{m,\sigma_0}(1)$$

Here we would like to take the limit $s\to 1^+$. However, in order to conclude divergence, we need to know that $L(1,\chi)\neq 0$ if $\chi\neq\chi_0$, so that the second term of the right hand side is finite.

Result (3) suggests that we should look at analytic continuations of $\zeta(s)$. Recall that

$$\zeta(s) = 1+2^{-s}+3^{-s}+\ldots$$

Consider

$$A(s) \coloneqq 1 - 2^{-s} + 3^{-s} - \dots$$

Then

$$\zeta(s) - A(s) = 2^{1-s}(1 + 2^{-s} + 3^{-s} + \dots) = 2^{1-s}\zeta(s)$$

Thus

$$\zeta(s) = \left[1 - 2^{-(s-1)}\right]^{-1} A(s)$$

$A(s)$ is analytic for $\mathrm{Re}(s) > 0$, and the factor in front only has a pole at $s = 1$. Thus $\zeta(s)$ may be continued to $\mathrm{Re}(s) > 0$ so that it has a simple pole at $s = 1$ only.

Returning to the formula in terms of L-functions, we have

$$\log L(s, \chi_e) = \log \zeta(s) + O(1)$$

as $s \to 1$ for $s > 0 \in \mathbb{R}$. Since we know $s = 1$ is a simple pole, we can expand ζ about 1 as

$$\zeta(s) = \alpha(s-1)^{-1} + \beta + \gamma(s-1) + \dots$$

(In fact the residue $\alpha = 1$). Then

$$\log \zeta(s) + O(1) = -\log(s-1) + O(1) \to \infty$$

Combining with our previous work, we get

$$|G| \sum_{p \equiv a(m)} p^{-s} = \sum_{\chi \in \hat{G}} \chi(a) \log L(s, \overline{\chi}) + O(1) = -\log(s-1) + \sum_{\chi \neq \chi_e} \chi(a) \log L(s, \overline{\chi}) + O(1)$$

We still need to show that $L(1, \chi) \neq 0$ for $\chi \neq \chi_e$. Indeed, note that if $L(1, \chi) = 0$ then $L(1, \overline{\chi}) = 0$ as well. So if $\chi \neq \overline{\chi}$ and $L(1, \chi) = 0$, then we write

$$L(s, \chi) = A(s-1)^{\nu}$$

Here ν is the order of the zero at 1. Then $\overline{\chi}$ also has the same order zero.

Our formula is valid for all a. In particular we may choose $a = 1$, so that we are looking for primes which have remainder 1 mod m. Then $\chi(1) = 1$ for all χ, which simplifies to

$$|G| \sum_{p \equiv 1(m)} p^{-s} = -\log(s-1) + \sum_{\chi \neq \chi_0} \log L(s, \overline{\chi}) + O(1)$$

In this case, if $\chi \neq \overline{\chi}$ and $L(s, \chi)$ vanishes with order ν at 1, then the sum at least contains the term

$$2\nu \log(s-1)$$

On the right hand side the term $-\log(s-1)$ tends to $+\infty$, but the terms in the sum may only diverge to $-\infty$ (since they can only have zeros, not poles). Indeed, if $\nu \geq 1$ for at least one nonreal, nontrivial character, then the RHS tends to $-\infty$, but the left hand side is nonnegative. Thus no nonreal character vanishes at $s = 1$. Note that this is the case regardless of a; we simply use $a = 1$ in order to generate a contradiction.

So we have reduced to

$$|G| \sum_{p \equiv a(m)} p^{-s} = -\log(s-1) + \sum_{\chi = \overline{\chi} \neq \chi_e} \chi(a) \log L(s, \overline{\chi}) + O(1)$$

and merely need to show that $L(1, \chi) \neq 0$ for real characters $\chi = \overline{\chi}, \chi \neq \chi_e$. We form the function

$$F(s) = \prod_{\chi \in \hat{G}} L(s, \chi)$$

This function is analytic on $\mathrm{Re}(s) > 1$. On $\mathrm{Re}(s) > 0$, it is analytic except possibly at $s = 1$. Here, if $L(s, \chi) \neq 0$ for $\chi \neq \chi_e$, then there is a simple pole. Otherwise, $F(1)$ is finite. Since F is a product of Dirichlet series, we can write

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

Proposition 1.6

$a_n \geq 0$ for all n. Moreover, $n \mapsto a_n$ is multiplicative (but not necessarily totally multiplicative).

Note that this is because $F(s)$ shows up as the **Dedekind zeta function** $\zeta_K(s)$ of a number field K, but we don't need that for this proof, since we compute the coefficients directly.

Proof. Multiplicativity comes about since F is the Dirichlet convolution of the L-functions. Write

$$\sum_{n=1}^{\infty} \frac{a_n}{n^s} = \prod_{(p,m)=1} (1 + a_p p^{-s} + a_{p^2} p^{-2s} + \ldots) = \prod_{(p,m)=1} \prod_{\chi \in \hat{G}} (1 - \chi(p) p^{-s})^{-1}$$

For $(p, m) = 1$,

$$\prod_{\chi \in \hat{G}} (1 - \chi(p) p^{-s})^{-1} = (1 - (p^{-s})^{f(p)})^{-g(p)}$$

where $f(p)$ is the order of p in $(\mathbb{Z}/m\mathbb{Z})^*$ and

$$g(p) = \frac{\phi(m)}{f(p)}$$

and expanding this gives a series with nonnegative coefficients by the binomial theorem. □

Recall that for a power series

$$\sum_{n=0}^{\infty} c_n z^n$$

with $c_n \geq 0$, there is a pole at $z = \rho_0$, where ρ_0 is the radius of convergence. More generally, for a power series with complex coefficients, there is a pole or other singularity somewhere on the boundary. In the case of Dirichlet series this may not be true, but it does hold specifically in the case of nonnegative coefficients.

Definition 1.4

For a Dirichlet series

$$\sum_{n=1}^{\infty} \frac{a_n}{n^{-s}}$$

the **abscissa of absolute convergence** is

$$\sigma_0 = \inf \left\{ \sigma \in \mathbb{R} : \sum_{n=1}^{\infty} \left| \frac{a_n}{n^{-s}} \right| < \infty (\mathrm{Re}(s) > \sigma) \right\}$$

Lemma 1.7: Landau

Let

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

with $a_n \geq 0$, and let $\sigma = \rho_0$ be the abscissa of absolute convergence. If f is analytic for $\mathrm{Re}(s) > \rho$ then $\rho_0 \leq \rho$.

In other words, the above lemma says that the first pole of a Dirichlet series with nonnegative coefficients is on the real axis.

Now suppose that F has no pole at $s = 1$. Then our Dirichlet series representation of F must converge absolutely for $\mathrm{Re}(s) > 0$. Now note that

$$(1 - p^{f(p)s})^{-g(p)} \geq 1 + p^{-\phi(m)s} + p^{-2\phi(m)s} + \dots$$

So

$$\sum_{n=1}^{\infty} a_n n^{-s} \geq \sum_{n=1}^{\infty} n^{-\phi(m)s}$$

for $s > 0$. But taking $s = 1/\phi(m) > 0$, F must diverge, which is a contradiction. This concludes the proof of Dirichlet's theorem.

The proof of Dirichlet's theorem made use of the fact that $\zeta(1) \neq 0$. In fact a stronger theorem is true, which uses the results on nonnegative coefficient Dirichlet series we derived.

Theorem 1.8

$\zeta(s) \neq 0$ for $\mathrm{Re}(s) = 1$.

Proof. For $\nu \in \mathbb{C}$ we define

$$\sigma_\nu(n) = \sum_{d|n} d^\nu$$

For $t_0 = \mathrm{Im}(s) \neq 0$ (we showed this for $t_0 = 0$ already) we also define

$$F(s) = \sum_{n=1}^{\infty} \frac{|\sigma_{it_0}(n)|^2}{n^s}$$

This is equal to

$$\frac{\zeta^2(s)\zeta(s+it_0)\zeta(s-it_0)}{\zeta(2s)}$$

But since this is a Dirichlet series with positive coefficients, we just need to show that there is no pole at $s=1$. Suppose for contradiction that there is a zero at $\zeta(1+it_0)$. Then there is also a zero at $\zeta(1-it_0)$, which cancels with the order 2 pole for $\zeta^2(s)$. Moreover, since ζ only has poles at $s=0,1$, and it is nonzero for $\mathrm{Re}(s)>1$, this represents the series for all $\mathrm{Re}(s)>1/2$. At $s=1/2$ the denominator forces $F(1/2)=0$. But by inspection it is plainly untrue that $F(1/2)=0$. □

1.2 Class Number Field

Consider $\mathbb{F}$ a finite field. Then $(\mathbb{F},+)$ and $(\mathbb{F}^*,\cdot)$ are finite abelian groups, so we may consider their dual groups. Since both structures are in place, we may think about the additive properties of multiplicative characters, or multiplicative properties of additive characters.

Definition 1.5

Given $\psi \in \widehat{(\mathbb{F},+)}, \chi \in \widehat{(\mathbb{F}^*,\cdot)}$, the **Gauss sum** of ψ,χ is

$$G(\psi,\chi)=\sum_{a\in\mathbb{F}^*}\psi(a)\chi(a)=\hat{\chi}(\psi)=\hat{\psi}(\chi)$$

The Gauss sum connects the additive and multiplicative structure of $\mathbb{F}$, and we observe that it provides the Fourier coefficients of both additive and multiplicative characters, when the transform is taken against the opposing structure.

Example 1.1

Let $\mathbb{F}=\mathbb{R}$. Then the additive characters are those of the form

$$\psi(x)=e(\alpha x)$$

for $\alpha\in\mathbb{C}$, and the multiplicative characters are

$$\chi(a)=a^s|a|$$

for $s\in\mathbb{C}, a\in\mathbb{R}^*$. Since $\mathbb{R}^*$ is infinite, we will need to convert our sum to an integral over an appropriate measure. If we try to assign a "translation invariant" measure to a group, we will need to look at the measure

$$\frac{\mathrm{d}a}{a}$$

which is called the **Haar measure**. So the Gauss sum becomes

$$G(\alpha,s)=\int_0^\infty e^{\alpha x}x^s\frac{\mathrm{d}x}{x}$$

which is the Gamma function.

Over finite fields $\mathbb{F}_p$, $p > 2$, then there is an additive group isomorphism $\mathbb{F}_p \to \hat{\mathbb{F}}_p$ given by

$$a \mapsto \psi_a(x) = e\left(\frac{ax}{p}\right)$$

Note that the choice of pth root of unity implicit in this statement shows why the isomorphism is noncanonical. So then for $b \in \mathbb{Z}/p\mathbb{Z}$ and $\chi \in \widehat{(\mathbb{Z}/p\mathbb{Z})}^*$, the Gauss sum is given by

$$\tau(b, \chi) = \sum_{a=1}^{p-1} \chi(a) e\left(\frac{ab}{p}\right)$$

Proposition 1.9

For $\chi \in \widehat{(\mathbb{Z}/p\mathbb{Z})}^*$,

1. $\tau(a, \chi) = \overline{\chi(a)}\tau(1, \chi)$,
2. For $\tau(\chi) \coloneqq \tau(1, \chi)$, $|\tau(\chi)|^2 = p$ (so long as $\chi \neq \chi_e$)

1. *Proof.* If $a \neq 1$, then writing $ca = w$,

$$\tau(a, \chi) = \sum_{c=1}^{p-1} \chi(c) e\left(\frac{ca}{p}\right) = \sum_{c=1}^{p-1} \chi(wa^{-1}) e\left(\frac{w}{p}\right) = \overline{\chi(a)}\tau(1, \chi) \qquad \square$$

2. *Proof.* We have

$$|\tau(\chi)|^2 = \sum_{a \in \mathbb{F}^*} \chi(a)\psi(a) \overline{\sum_{b \in \mathbb{F}^*} \chi(b)\psi(b)}$$
$$= \sum_{a,b \in \mathbb{F}^*} \chi(a)\overline{\chi(b)}\psi(a)\overline{\psi(b)} = \sum_{a,b} \chi(ab^{-1})\psi(a - b)$$

Set $ab^{-1} = w$. Then this becomes

$$\sum_{b,w} \chi(w)\psi(b(w-1))$$

For $w = 1$, each term is 1 and the sum over b is $p - 1$. If $w \neq 1$, then the sum over b is

$$\sum_{b \in \mathbb{F}^*} \psi(b(w-1)) = \underbrace{\sum_{b \in \mathbb{F}} \psi(b(w-1))}_{=0} - \psi(0) = -1$$

So the sum is now

$$p - 1 + \sum_{\substack{w \in \mathbb{F}^* \\ w \neq 1}} (-\chi(w)) = p - 1 + \chi(1) = p$$

(Here we use the fact that χ is nontrivial to cancel out the sum over $\mathbb{F}^*$). $\square$

Definition 1.6

Let $p > 2$. Then define the Legendre symbol as

$$\chi(n) = \left(\frac{n}{p}\right) = \begin{cases} 1, & n = x^2, x \in \mathbb{F}_p \\ -1 \end{cases}$$

This is a real multiplicative character of $\mathbb{F}_p$.

Proposition 1.10

If χ is the Legendre symbol for $\mathbb{F}_p$, $p > 2$,

$$\tau(\chi)^2 = \begin{cases} p, & p \equiv 1 \pmod 4 \\ -p, & p \equiv 3 \pmod 4 \end{cases}$$

Proof. We write out the Gauss sum as in the previous theorem, noting that $\chi = \chi^{-1}$:

$$\begin{aligned} \tau(\chi)^2 &= \sum_{a,b \in \mathbb{F}_p^*} \chi(a)\chi^{-1}(b) e\left(\frac{a}{p}\right) e\left(\frac{b}{p}\right) = \sum_{a,b \in \mathbb{F}_p^*} \chi(ab^{-1}) e\left(\frac{a+b}{p}\right) \\ &= \sum_{w,b \in \mathbb{F}_p^*} \chi(w) e\left(\frac{b(w+1)}{p}\right) = \sum_{w=1}^{p-2} -\chi(w) + \chi(-1)(p-1) \\ &= \chi(-1) + \chi(0) + \chi(-1)(p-1) = p\chi(-1) \end{aligned}$$

$\chi(-1) = +1$ when $p \equiv 1 \pmod 4$ and -1 otherwise. □

Theorem 1.11: Gauss

1. Let p be an odd prime. Then

$$\sum_{n=0}^{p-1} e\left(\frac{n^2}{p}\right) = \begin{cases} \sqrt{p}, & p \equiv 1 \pmod 4 \\ i\sqrt{p}, & p \equiv 3 \pmod 4 \end{cases}$$

2. If p, q are odd primes,

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{(p-1)(q-1)}{4}}$$

Partial Proof of 1.11, (1). Letting χ denote the Legendre symbol,

$$\tau(\chi) = \sum_{n=1}^{p-1} \left(\frac{n}{p}\right) e\left(\frac{n}{p}\right) = \sum_{n=1}^{p-1} \left[\left(\frac{n}{p}\right) + 1\right] e\left(\frac{n}{p}\right) - \underbrace{\sum_{n=1}^{p-1} e\left(\frac{n}{p}\right)}_{=-1}$$

$$= \sum_{n=0}^{p-1} e\left(\frac{n}{p}\right) \left[\begin{cases} 2, & n = x^2 \\ 0 \end{cases}\right]$$

The Legendre symbol factor turns this sum into twice the sum over the quadratic residues, and since half the numbers are quadratic residues, we can double count by simply summing:

$$\tau(\chi) = \sum_{n=0}^{p-1} e\left(\frac{n^2}{p}\right)$$

So it is clear that $\tau(\chi) = \pm\sqrt{p}$ when $p \equiv 1 \pmod 4$, and $\pm i\sqrt{p}$ otherwise. But the sign of the sum is more subtle. We develop Poisson summation to do this, but the actual proof (which uses results from the Poisson sum) is left as homework after the next section. □

We briefly remark that result (2) is quite powerful. For instance, a consequence of this is that if $p \equiv q \equiv 1 \pmod 4$, then p has a square root mod q if and only if q has a square root mod p.

1.3 Poisson Sums and the Theta Function

Typically, we are looking at $\mathbb{R}$ and the subgroup $\mathbb{Z}$, so that we want to consider the quotient $\mathbb{R}/\mathbb{Z} = [0, 1)$. In general we can work with L a rank n lattice, meaning

$$L = \mathbb{Z}v_1 \oplus \mathbb{Z}v_2 \oplus \ldots \oplus \mathbb{Z}v_n \leq \mathbb{R}^n$$

such that the quotient is topologically a torus. We will consider $S(\mathbb{R}^n)$, the **Schwartz space** of functions which are smooth and for which all derivatives decay at ∞ faster than $|x|^{-A}$.

Definition 1.7

Let $f \in S(\mathbb{R})$. Then define the **Fourier transform** of f to be $\hat{f} : \mathbb{R} \to \mathbb{C}$ given by

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x) e(-x\xi)\, \mathrm{d}x$$

If $f \in S(\mathbb{R}^n)$, then $\hat{f} : \mathbb{R}^n \to \mathbb{C}$ is

$$\hat{f}(\xi) = \int_{\mathbb{R}^n} f(x) e(-\langle x, \xi \rangle)\, \mathrm{d}x$$

Proposition 1.12

$\hat{f} \in S(\mathbb{R})$.

Proof. We integrate by parts

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x)e(-x\xi)\,\mathrm{d}x = \underbrace{-\frac{1}{\xi}f(x)e(-x\xi)\Big|_{x=-\infty}^{\infty}}_{=0} + \frac{1}{\xi}\int_{-\infty}^{\infty} f'(x)e(-x\xi)\,\mathrm{d}x$$

The integral is bounded by the decay of f', so $\hat{f}$ decays at least as fast as ξ^{-1}. Integrating by parts again and using the decay for f'' shows that $\hat{f}$ decays as ξ^{-2} as well, and by induction for all ξ^{-n}.

For the derivatives,

$$\hat{f}^{(m)}(\xi) = \int_{-\infty}^{\infty} (-x)^m f(x)e(-x\xi)\,\mathrm{d}x = (-1)^m \widehat{x^m f}(\xi)$$

Since $x^m f$ is also Schwartz, the first result shows that $\hat{f}^{(m)}$ also decays. □

Theorem 1.13: Poisson Summation

For $f \in S(\mathbb{R})$,

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

More generally, if $f \in S(\mathbb{R}^n)$ and L is a lattice, then

$$\sum_{l\in L} f(l) = \frac{1}{\mathrm{Vol}(\mathbb{R}^n/\mathrm{L})} \sum_{v\in\hat{L}} \hat{f}(v)$$

where

$$\hat{L} = \{\xi \in \mathbb{R}^n : l \in L \implies \langle \xi, l\rangle \in \mathbb{Z}\}$$

The following argument essentially uses the fact that the Fourier transform of the Dirac comb

$$\sum_{n\in\mathbb{Z}} \delta(x-n)$$

is itself. Convolving with arbitrary functions, we get the below proof.

Proof. Define

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

The sum converges absolutely because of the decay of f, so F is smooth and has period 1. Thus we can expand it as a Fourier series:

$$F(z) = \sum_{m\in\mathbb{Z}} \hat{F}(m)e(mx)$$

where

$$\hat{F}(m) = \int_0^1 F(x)e(-mx)\,\mathrm{d}x = \int_0^1 e(-mx)\sum_{k\in\mathbb{Z}} f(x+k)\,\mathrm{d}x = \int_{-\infty}^{\infty} f(x)e(-mx)\,\mathrm{d}x = \hat{f}(m)$$

So

$$F(x) = \sum_{m\in\mathbb{Z}} \hat{f}(m)e(-mx)\,\mathrm{d}x$$

Substitute $x = 0$:

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

The following proof instead uses the functional analysis fact that an operator's trace is invariant under change of basis.

Alternate Proof of Poisson Summation. Let $K(x,y) : S \times S \to \mathbb{C}$ be such that $K(x,y) = g(x-y)$ for some periodic g, and define the operator T_K on $L^2(S)$ by

$$T_K f(x) = \int_S K(x,y)f(y)\,\mathrm{d}y$$

If K is continuous, then T_K is a compact operator. In this case we define

$$\mathrm{tr}(T_K) = \int_S K(x,x)\,\mathrm{d}x$$

Alternatively, if $\phi_1, \phi_2, \ldots$ are an orthonormal eigenbasis for $L^2(S)$ and T_K, we can also compoute the trace by diagonalization as

$$\mathrm{tr}(T_K) = \sum_j \lambda_j$$

We diagonalize this by observing that

$$\phi_\nu(x) = e(\nu x), \quad \nu \in \mathbb{Z}$$

defines an orthonormal eigenbasis for T_K. Indeed,

$$\begin{aligned} T_K\phi_\nu(x) &= \int_S K(x,y)\phi_\nu(y)\,\mathrm{d}y = \int_0^1 g(x-y)e(\nu y)\,\mathrm{d}y \\ &= \int_0^1 g(t)e(\nu(x-t))\,\mathrm{d}t = e(\nu x)\hat{g}(\nu) = \hat{g}(\nu)\phi_\nu(x) \end{aligned}$$

Then

$$\mathrm{tr}(T_K) = \sum_{\nu\in\mathbb{Z}} \hat{g}(\nu)$$

which makes it plain that we should choose K in such a way that we recover the right hand side of Poisson summation. Specifically, for $f \in S(\mathbb{R})$, define

$$K_f(x,y) = \sum_{m\in\mathbb{Z}} f(x-y+m)$$

Then by integrating over the diagonal, we have

$$\operatorname{tr}(T_K) = \int_0^1 \sum_{m\in\mathbb{Z}} f(m)\,\mathrm{d}x = \sum_{m\in\mathbb{Z}} f(m)$$

On the other hand, $K(x,y) = g(x-y)$, where

$$g(z) = \sum_{m\in\mathbb{Z}} f(z+m)$$

Then

$$\hat{g}(\nu) = \int_0^1 \sum_{m\in\mathbb{Z}} f(u+m)e(\nu u)\,\mathrm{d}u = \int_{\mathbb{R}} f(u)e(\nu u) = \hat{f}(\nu)$$

which means

$$\operatorname{tr}(T_K) = \sum_{\nu\in\mathbb{Z}} \hat{g}(\nu) = \sum_{\nu\in\mathbb{Z}} \hat{f}(\nu)$$ □

Definition 1.8

For $\operatorname{Re}(t) > 0$, define the **theta function** by

$$\Theta(t) \coloneqq \sum_{n\in\mathbb{Z}} e^{-\pi n^2 t}$$

Proposition 1.14

For $t > 0$, $\Theta(1/t) = \sqrt{t}\Theta(t)$.

Proof. Define

$$f(x) = e^{-\pi x^2}$$

and

$$f_{\sqrt{t}}(x) = f(\sqrt{t}x)$$

We apply Poisson summation to $f_{\sqrt{t}}$:

$$\sum_n f_{\sqrt{t}}(n) = \sum_n e^{-t\pi n^2} = \sum_{m\in\mathbb{Z}} \widehat{f_{\sqrt{t}}}(m)$$

Recall that by change of variables, if $f_\lambda(x) = f(\lambda x)$,

$$\hat{f}_\lambda(\xi) = \frac{1}{\lambda}\hat{f}(\xi/\lambda)$$

so

$$\hat{f}(\xi) = \int_{\mathbb{R}} e^{-\pi x^2 - 2\pi i x\xi}\,\mathrm{d}x = e^{-\pi\xi^2}\int_{\mathbb{R}} e^{-\pi(x+i\xi)^2}\,\mathrm{d}x$$

By shifting the contour, this is

$$e^{-\pi\xi^2}\int_{\mathbb{R}} e^{-\pi\nu^2}\,\mathrm{d}\nu = e^{-\pi\xi^2} = f(\xi)$$

Then applying change of variables,

$$\Theta(t) = \sum_n e^{-t\pi n^2} = \sum_{m\in\mathbb{Z}} \frac{1}{\sqrt{t}} e^{-\pi m^2/t} = \frac{1}{\sqrt{t}}\Theta\left(\frac{1}{t}\right) \qquad \square$$

Proposition 1.15

$\Theta(z+2li) = \Theta(z)$ for $l \in \mathbb{Z}$.

By analytic continuation, the above holds for $\mathrm{Re}(t) > 0$ (with the root continued appropriately). If we consider $\Theta(it)$, then we are working with the upper half plane, where the transformations are given by

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} z = \frac{az+b}{cz+d}, \qquad \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{PSL}_2(\mathbb{R})$$

We can recast the above two properties as:

Proposition 1.16

1. $\Theta\left(i\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} z\right) = \Theta(it)$
2. $\Theta\left(i\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} z\right) = \sqrt{iz}\Theta(iz)$

Definition 1.9

Let $a, b \in \mathbb{N}$ with $(a,b) = 1$. Then define the **quadratic Gauss sum** to be

$$S(a,b) = \sum_{x=0}^{b-1} e\left(\frac{ax^2}{b}\right)$$

The following relation will allow us to prove quadratic reciprocity.

Theorem 1.17

$$\frac{1}{\sqrt{b}} S(-a,b) = \frac{e^{i\pi/4}}{2\sqrt{2a}} S(b,4a)$$

Proof. Pick $t = \varepsilon + 2ai/b$. Then

$$\Theta(t) = 1 + 2\sum_{n=1}^{\infty} e^{-\pi n^2(\varepsilon + \frac{2ai}{b})} = 1 + 2\sum_{n=1}^{\infty} e^{-\pi n^2\varepsilon} e^{-2\pi i n^2 \frac{a}{b}}$$

The sum has b dependence based on $n \mod b$. Writing $n = r + sb$, we have

$$\Theta(t) = 1 + 2 \underbrace{\sum_{r=1}^{b} e^{-2\pi i r^2 \frac{a}{b}}}_{=S(-a,b)} \sum_{s=0}^{\infty} e^{-\pi\varepsilon(r+sb)^2}$$

For fixed r, as $\varepsilon \to 0$ we have

$$\sum_{s=0}^{\infty} e^{-\pi\varepsilon(r+sb)^2} = \sum_{s=0}^{\infty} e^{-\left(\sqrt{\pi\varepsilon}r + \sqrt{\pi\varepsilon}bs\right)^2}$$

This is asymptotically equal to a Riemann integral, that is it is

$$\sum_{s=0}^{\infty} e^{-\left(\sqrt{\pi\varepsilon}r + \sqrt{\pi\varepsilon}bs\right)^2} \sim \frac{1}{\sqrt{\pi\varepsilon}b} \frac{\sqrt{\pi}}{2}$$

So as $\varepsilon \to 0$, Θ blows up as

$$\Theta\left(\varepsilon + \frac{2ai}{b}\right) \sim \frac{1}{b\sqrt{\varepsilon}} S(-a, b)$$

We will recompute this using the reflection formula:

$$\Theta\left(\varepsilon + \frac{2ai}{b}\right) = \frac{1}{\sqrt{\varepsilon + \frac{2ai}{b}}} \Theta\left(\frac{1}{\varepsilon + \frac{2ai}{b}}\right)$$

The square root is not at a branch point, so simple substitution gives

$$\frac{1}{\sqrt{\varepsilon + \frac{2ai}{b}}} \to \frac{e^{i\pi/4}\sqrt{b}}{\sqrt{2a}}$$

On the other hand, $\Theta(1/t)$ will blow up as $\varepsilon \to 0$. To study how this happens, we expand its argument as

$$\frac{1}{\varepsilon + \frac{2ai}{b}} = -\frac{bi}{2a} + \frac{\varepsilon b^2}{4a^2} + O(\varepsilon^2)$$

so by the same work as above,

$$\Theta\left(\varepsilon \frac{b^2}{4a^2} - \frac{bi}{2a}\right) \sim \frac{1}{\sqrt{\varepsilon}\sqrt{\frac{b^2}{4a^2}}} \frac{1}{2a} S(-b/2, 2a) = \frac{1}{\sqrt{\varepsilon}} \frac{1}{b} \frac{1}{2} S(-b, 4a)$$

So equating coefficients of the $1/\sqrt{\varepsilon}$ blowup, we get

$$\frac{1}{b} S(-a, b) = \frac{e^{-\pi/4}\sqrt{b}}{\sqrt{2a}} \frac{1}{2b} S(-b, 4a) = \frac{1}{2\sqrt{2ab}} S(-b, 4a) \qquad \square$$

Theorem 1.18

If p, q are primes and $(p, q) = 1$, then

$$S(a, pq) = S(ap, q)S(aq, p)$$

Proof. Homework. □

Now that we have computed a relation for quadratic sums, which we previously saw were related to the Gauss sum of the Legendre symbol, we can prove 1.11.

Proof of 1.11. Homework. □

1.4 Applications of Gauss Sums

Suppose we consider

$$\sum_{x \leq T} \chi(x)$$

where χ is a nontrivial Dirichlet character mod q (q not necessarily prime). Trivially,

$$\left| \sum_{x \leq T} \chi(x) \right| \leq q$$

However, it is possible to improve on the naive bound.

Definition 1.10

Let f be a complex function and g nonnegative. Then

$$f \ll g$$

means $|f| \leq Cg$. This may also be parameterized, as in

$$f \ll_\varepsilon g$$

which means that $|f| \leq C_\varepsilon g$, so that the constant depends on the parameters.

Definition 1.11

Let $\xi \in \mathbb{R}$. Define $\|\xi\|$ to be the distance from ξ to $\mathbb{Z}$.

For now, we will prove the following:

Proposition 1.19: Polya-Vinogradov

If $\chi \neq \chi_e$ is a Dirichlet character mod q, then

$$\sum_{m \leq T} \chi(m) \ll \sqrt{q} \log q$$

Proof. The left hand side is periodic, so we can assume $T < q$. Define

$$I_T(x) = \begin{cases} 1, & 0 \leq x \leq T \\ 0 \end{cases}$$

We will Fourier expand I_T:

$$\hat{I}_T(a) = \sum_{m \leq T} e\left(-\frac{ma}{q}\right) = \begin{cases} T+1, & a \equiv 0 \pmod{q} \\ \frac{1-e\left(-\frac{Ta}{q}\right)}{1-e\left(-\frac{a}{q}\right)}, & a \not\equiv 0 \pmod{q} \end{cases}$$

For $a \neq 0$, we bound this as

$$\left|\hat{I}_T(a)\right| \leq \frac{2}{\sin(2\pi a/q)} = \frac{2}{\frac{2\pi a}{q} + O(1)} = \frac{q}{\pi a} + O(1)$$

By Fourier inversion,

$$I_t(x) = \frac{1}{q} \sum_{a \mod q} \hat{I}_T(q) e\left(\frac{ax}{q}\right)$$

Then

$$\begin{aligned} \left|\sum_{x \leq T} \chi(x)\right| &= \left|\sum_{x=0}^{\infty} I_T(x)\chi(x)\right| = \frac{1}{q}\left|\sum_{a \mod q} \hat{I}_T(a) \sum_{x=0}^{\infty} \chi(x) e\left(\frac{ax}{q}\right)\right| \\ &= \frac{1}{q}|G(x,q)|\left|\sum_{a \mod q} \hat{I}_T(a)\right| \leq \frac{1}{\sqrt{q}}\left(\sum_{\substack{a \mod q \\ a \neq 0}} \left|\hat{I}_T(a)\right| + q\right) \\ &\leq \frac{1}{\sqrt{q}}\left(\sum_{a=1}^{q-1} \frac{q}{\pi a} + q\right) = \frac{1}{\sqrt{q}}\left(q \log q + O(1)\right) = q^{1/2} \log q + O(1) \end{aligned}$$

since the sum is $\log q + O(1)$. □

Chapter 2

L-Functions

2.1 The Riemann Hypothesis over Finite Fields

Cnosider a finite field $\mathbb{F}_p$, (so $\mathbb{F}_p^*$ is cyclic). Consider a polynomial $f(x,y)$ over $\mathbb{F}_p$. We consider the problem of counting the number of solutions N_f of f. We will find that as $p \to \infty$ (for fixed polynomials),

$$N_f = p + O_f\left(p^{\frac{1}{2}}\right)$$

As before, we will let χ be a multiplicative character of $\mathbb{F}_p$. We will denote by ε the trivial multiplicative character, and we will define $\varepsilon(0) = 1$, with $\chi(0) = 0$ for all nontrivial characters.

Example 2.1

Consider the Fermat curve $x^n + y^n = 1$ over $\mathbb{F}_p$. We can analyze this by keeping n fixed and sending $p \to \infty$. This may be diagonalized as

$$\#\left\{x^n + y^n = 1\right\} = \sum_{a+b=1} \#\left\{x^n = a\right\} \#\left\{y^n = b\right\}$$

We have

$$\#\left\{x : x^n = a\right\} = \sum_{\chi^n = \varepsilon} \chi(a)$$

The proof of this is somewhat long, but one observes that when $a = x^n$, then $\chi(a) = 1$ for all $\chi, \chi^n = \varepsilon$, so it is just a matter of verifying that 1 and a have the same number of nth roots in $\mathbb{F}_p^*$ using its cyclic structure. If a is not an nth power, then the right side is the sum over cyclic characters (perhaps multiple times around the circle) and cancels. So

$$\#\left\{x^n + y^n = 1\right\} = \sum_{a+b=1} \sum_{\chi^n=\varepsilon} \sum_{\lambda^n=\varepsilon} \chi(a)\lambda(b) = \sum_{\substack{\chi^n=\varepsilon \\ \lambda^n=\varepsilon}} \sum_{a+b=1} \chi(a)\lambda(b)$$

For fixed n, the first sum just involves n^2 terms, while the second sum varies with p. We denote it by the Jacobi sum $J(\chi, \lambda)$. Note the characters involved will change as p varies. But since $\widehat{\mathbb{F}_p^*} \cong \mathbb{F}_p^*$, and $x^n - 1$ has at most n roots in $\mathbb{F}_p^*$, there are only at most n^2 terms.

Definition 2.1

The **Jacobi sum** of two multiplicative characters χ, λ over a finite field $\mathbb{F}_p$ is

$$J(\chi, \lambda) = \sum_{a+b=1} \chi(a)\lambda(b)$$

Just as the Gauss sum (which considers an additive character against a multiplicative one) was an analogue of the gamma function, the Jacobi sum (which takes two multiplicative characters against one another) is the analogue of the beta function.

$$B(x, y) = \int_0^1 t^{x-1}(1-t)^{y-1}\,\mathrm{d}t = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$$

Proposition 2.1

(a) $J(\varepsilon, \varepsilon) = p$.

(b) $J(\varepsilon, \chi) = 0$ when $\chi \neq \varepsilon$.

(c) $J(\chi, \overline{\chi}) = -\chi(-1)$ when $\chi \neq \varepsilon$.

(d) If $\chi, \lambda, \chi\lambda \neq \varepsilon$, then

$$J(\chi, \lambda) = \frac{\tau(\chi)\tau(\lambda)}{\tau(\chi\lambda)}$$

(e) If $\chi, \lambda \neq \varepsilon$ and $\chi\lambda \neq \varepsilon$, then

$$|J(\chi, \lambda)| = \sqrt{p}$$

We only prove (d).

Proof. For $\chi\lambda \neq \varepsilon$, we write

$$\tau(\chi)\tau(\lambda) = \sum_{x,y} \chi(x) e\left(\frac{x}{p}\right) \lambda(y) e\left(\frac{y}{p}\right) = \sum_{x,y} \chi(x)\lambda(y) e\left(\frac{x+y}{p}\right)$$

We set $t = x + y$ and get

$$\sum_{t \in \mathbb{F}_p} \left(\sum_{x+y=t} \chi(x)\lambda(y) \right) e\left(\frac{t}{p}\right)$$

The case $t = 1$ is the Jacobi sum. For the case $t = 0$ we get zero by symmetry. For $t \neq 0$, we set

$$x = tx'$$
$$y = ty'$$
$$x' + y' = 1$$

and we once again get the Jacobi sum, scaled by $\chi(t)\lambda(t)$. So

$$\tau(\chi)\tau(\lambda) = \left(\sum_{t \in \mathbb{F}_p^*} \chi(t)\lambda(t) e\left(\frac{t}{p}\right) \right) J(\chi, \lambda) = J(\chi, \lambda)\tau(\chi\lambda) \qquad \square$$

Example 2.2

Continuing our example of $\#\{x^n + y^n = 1\}$, we have

$$N_f = \sum_{\substack{\chi^n = \varepsilon \\ \lambda^n = \varepsilon}} J(\chi, \lambda) = J(\varepsilon, \varepsilon) - \sum_{\substack{\chi^n = \varepsilon \\ \chi \neq \varepsilon}} \chi(-1) + \sum_{\substack{\chi^n = \lambda^n = \varepsilon \\ \chi\lambda \neq \varepsilon \\ \chi, \lambda \neq \varepsilon}} J(\chi, \lambda)$$

The second sum has at most n terms, so it is $O_n(1)$. The third sum has at most n^2 terms, and each term has absolute value $\sqrt{p}$, so the whole sum is

$$N_f = p + O_n(\sqrt{p})$$

The **Riemann hypothesis over finite fields** is the statement that $N_f = p + O_n(\sqrt{p})$ for any irreducible affine plane curve f over $\mathbb{F}_p$. It is known to be true.

2.2 Riemann's Paper

Now we consider Riemann's paper "On the Number of Prime Numbers Less than a Given Quantity." Consider $f \in C_c^\infty(\mathbb{R}_{>0}) \subseteq S(\mathbb{R})$ with $\hat{f}(0) = 0$. Then

$$\sum_{n=1}^{\infty} f(nx)$$

is rapidly decreasing as $x \to \infty$ since f is Schwartz. As $x \to 0$, the function is approximately $1/x$ times the integral of f from 0 to ∞, hence $O(1/x)$. So

$$\int_0^\infty \sum_{n=1}^{\infty} f(nx) x^s \frac{\mathrm{d}x}{x}$$

is analytic on $\mathrm{Re}(s) > 1$. This integral is

$$\sum_{n=1}^{\infty} \int_0^\infty f(nx) x^s \frac{\mathrm{d}x}{x} = \sum_{n=1}^{\infty} \frac{1}{n^s} \int_0^\infty f(nx) nx^s \frac{\mathrm{d}x}{x} = \zeta(s)\mathcal{M}\{f\}(s) = \zeta(s)\tilde{f}(s)$$

Definition 2.2

The **Mellin transform** of a function f is

$$\tilde{f}(s) = \int_0^\infty f(x) x^s \frac{\mathrm{d}x}{x}$$

Proposition 2.2

If f is $O(x^a)$ as $x \to 0$ and $O(x^b)$ as $x \to \infty$, then $\tilde{f}$ is analytic for $-a < \sigma < -b$.

By Poisson summation,

$$F(x) := \sum_{n=1}^\infty f(nx) = \sum_{n \in \mathbb{Z}} f(nx) = \frac{1}{x} \sum_{m \in \mathbb{Z}} \hat{f}\left(\frac{m}{x}\right)$$

F is compactly supported by assumption, and as $x \to 0$, $\hat{f}$ is Schwartz so F is $O_A(x^A)$ for all large A. So $\zeta \tilde{f} = \tilde{F}$ is entire. This holds for all f satisfying our assumptions, so we can probe ζ with different functions to obtain analytic continuation and the functional equation. We know $\tilde{f}(1) = \hat{f}(0) = 0$, so ζ may have a pole at 1, but it admits a meromorphic continuation to the rest of $\mathbb{C}$.

Definition 2.3

We define the **completed zeta function** to be the function

$$\Lambda(s) = \pi^{-s/2} \Gamma\left(\frac{s}{2}\right) \zeta(s)$$

Proposition 2.3

Λ admits an analytic continuation to $\mathbb{C}$ with simple poles at $s = 0, 1$, and

$$\Lambda(s) = \Lambda(1 - s)$$

Proof. For $\mathrm{Re}(s) > 1$ we have

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

We can also solve this integral in a different way. Define

$$w(x) = \sum_{n=1}^\infty e^{-\pi n^2 x} = \frac{\Theta(x) - 1}{2}$$

Then

$$\int_0^\infty x^{\frac{s}{2}} w(x) \frac{\mathrm{d}x}{x} = \int_0^1 x^{\frac{s}{2}} w(x) \frac{\mathrm{d}x}{x} + \int_1^\infty x^{\frac{s}{2}} w(x) \frac{\mathrm{d}x}{x}$$

The second term is entire since w is rapidly decreasing. The first term is

$$\int_1^\infty x^{-\frac{s}{2}} w\left(\frac{1}{x}\right)\frac{\mathrm{d}x}{x}$$

Using the functional equation for Θ, we have

$$w\left(\frac{1}{x}\right) = -\frac{1}{2} + \frac{1}{2}x^{\frac{1}{2}} + x^{\frac{1}{2}}w(x)$$

So the first integral is

$$\int_1^\infty x^{-\frac{s}{2}}\left[-\frac{1}{2} + \frac{1}{2}x^{\frac{1}{2}} + x^{\frac{1}{2}}w(x)\right]\frac{\mathrm{d}x}{x} = -\frac{1}{s} + \frac{1}{s-1} + \int_1^\infty x^{-\frac{s-1}{2}}w(x)\frac{\mathrm{d}x}{x}$$
$$\implies \Lambda(s) = -\frac{1}{s} + \frac{1}{s-1} + \int_1^\infty w(x)\left[x^{\frac{s}{2}} + x^{\frac{1-s}{2}}\right]\frac{\mathrm{d}x}{x}$$

From this formula it is plain that Λ is invariant under $s \to 1-s$ and has two poles. □

Definition 2.4

The **prime counting function** is $\pi(x) = \#\{p \le x\}$.

Definition 2.5

The **logarithmic integral** is

$$\mathrm{Li}(x) = \int_1^x \frac{\mathrm{d}t}{\log t}$$

The prime number theorem says that $\pi(x) \sim x/\log x$. The local density of primes is $1/\log t$, so it is also true that $\mathrm{Li}(x) \sim \pi(x)$, but Li is a better approximation.

Definition 2.6

The **first Chebyshev function** is

$$\Theta(x) = \sum_{p \le x} \log p$$

Θ is asymptotically x if and only if the prime number theorem holds.

Definition 2.7

The **von Mangoldt function** is

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

The **second Chebyshev function** is

$$\psi(x) = \sum_{n \le x} \Lambda(n)$$

Θ is asymptotically x if and only if ψ is as well. Essentially the only difference is that Ψ adds in the square and higher terms, but these are asymptotically less than the linear term.

Theorem 2.4: Riemann's Explicit Formula

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

where ρ ranges over the zeros of ζ.

Note that since Ψ is discontinuous, the sum must have infinitely many terms. Moreover, by examining the completed zeta function, we see that the only zeros of ζ outside the strip $0 \le \operatorname{Re}(s) < 1$ are the trivial zeros $\rho = -2, -4, \ldots$. Summing over these zeros gives $\frac{1}{2}\log(1 - x^{-2})$, so from the formula we then conclude that there are also infinitely many zeros in the critical strip (but only finitely many for a given height).

Theorem 2.5: Perron's Formula

$$\psi(x) = \sum_{n=x} \Lambda(n) = \frac{1}{2\pi i} \int_{\operatorname{Re}(s)=2} \left[-\frac{\zeta'}{\zeta}(s)\right] \frac{x^s \, \mathrm{d}s}{s}$$

Proof. By taking the logarithmic derivative of ζ, we get

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

To see that the integral converges, we note that on $\operatorname{Re}(s) = 2$, ζ'/ζ is bounded. Each term is periodic in $2\pi/\log p$, so it is almost periodic as a function of t. The integral is only conditionally convergent, but we take the integral from $2 - iT$ to $2 + iT$ and take $T \to \infty$.

$$\begin{aligned}\frac{1}{2\pi i} \int_{\operatorname{Re}(s)=2} \left[-\frac{\zeta'}{\zeta}(s)\right] \frac{x^s \, \mathrm{d}s}{s} &= \frac{1}{2\pi i} \int_{\operatorname{Re}(s)=2} \sum_{n=1}^{\infty} \Lambda(n) n^{-s} x^s \frac{\mathrm{d}s}{s} \\ &= \frac{1}{2\pi i} \sum_{n} \Lambda(n) \int_{\operatorname{Re}(s)=2} n^{-s} x^s \frac{\mathrm{d}s}{s}\end{aligned}$$

We delay the justification of the change of order. For fixed n the inner integral may be evaluated using the residue theorem

$$\frac{1}{2\pi i} \int_{\operatorname{Re}(s)=2} y^s \frac{\mathrm{d}s}{s} = \begin{cases} 1, & y > 1 \\ 0, & y < 1 \end{cases}$$

So the integral is 1 whenever $n/x < 1$, and we get (ignoring what happens when x is a discontinuity, which nevertheless be fixed)

$$\sum_{n \leq x} \Lambda(n) \qquad \square$$

Proof of 2.4. Shifting the contour further to $-\infty$ picks up residues at the zeros of ζ (note ζ' has no poles other than the pole of ζ), which produces Riemann's formula. The x term is the residue of ζ'/ζ at $s = 1$. $\square$

Observe from the formula that if $\mathrm{Re}(\rho) < 1$ for all zeros, then each term individually is $o(x)$ (which is certainly necessary for the prime number theory). Of course the Riemann hypothesis is that the real part of the nontrivial zeros is $1/2$. This is equivalent to the statement that for all $\varepsilon > 0$,

$$\psi(x) = x + O_\varepsilon(x^{\frac{1}{2}+\varepsilon})$$

2.3 Computing Zeros

Consider integrating ζ'/ζ over some box in the critical strip, symmetric around $1/2$, away from the real line and the lines $\mathrm{Re}(s) = 0, 1$. If all zeros of ζ are simple, then the residues are all 1 and this integral computes the number of zeros in the box. Thus we can approximate ζ and still find the number of zeros. Moreover if there is exactly one zero, then by the functional equation it has to be on the critical line, otherwise it would come in a pair on the other side. This method is the basis of most attempts to prove the Riemann hypothesis. The first 10^{12} or so zeros have been verified.

2.4 Functional Equations for L-Functions

We return to studying the functions $L(s, \chi)$ for $\chi \neq \chi_e$, which are analytic in $\mathrm{Re}(s) > 0$. We will provide a functional equation for them which is similar to the one for ζ. We work mod q (q not necessarily prime).

Definition 2.8

Suppose there is $q_1 | q$ such that $\chi(n) = \chi_1(n)$ for $(n, q) = 1$ with $\chi_1(n)$ a character mod q_1. Then χ is said to be induced by χ_1 and not primitive. Otherwise it is a **primitive character**.

Recall that when χ_e is the principal character mod q, we have

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

so the analytic proprties of $L(s, \chi_e)$ are determined by the properties of the L-function from the function which is identitically 1. Similarly, if $q_1 | q$ and $q_1 \neq q$, and χ is induced from χ_1, then

$$L(s, \chi) = L(s, \chi_1) \prod_{p|q} (1 - \chi_1(p) p^{-s})$$

So it suffices to study primitive characters.

Definition 2.9

If χ is a primitive character mod q, then q is called the **conductor** of χ.

The conductor can be thought of as the smallest number for which χ is periodic. To apply Poisson summation to a primitive character we need to convert its multiplicative structure into additive structure. Recall that the Gauss sum is defined as

$$\tau(\chi) \coloneqq \sum_{m=1}^{q} \chi(m) e\left(\frac{m}{q}\right)$$

and that if $(n, q) = 1$, then by change of variables,

$$\chi(n)\tau(\overline{\chi}) = \sum_{h=1}^{q} \overline{\chi}(h) e\left(\frac{nh}{q}\right)$$

This gives

$$\chi(n) = \frac{1}{\tau(\overline{\chi})} \sum_{h=1}^{q} \overline{\chi}(h) e\left(\frac{nh}{q}\right)$$

Lemma 2.6

If χ is primitive mod q then

$$\chi(n)\tau(\overline{\chi}) = \sum_{h=1}^{q} \overline{\chi}(h) e\left(\frac{nh}{q}\right)$$

for all n (not only those for which $(n, q) = 1$).

This will allow us to integrate χ against a Schwartz function.

Proposition 2.7

If χ is primitive mod $q \neq 1$ then $|\tau(\chi)| = \sqrt{q}$ and $\tau(\chi)\tau(\overline{\chi}) = \chi(-1)q$.

Proof. For all n we have

$$|\chi(n)|^2 |\tau(\chi)|^2 = \sum_{h_1, h_2} \overline{\chi}(h_1)\chi(h_2) e\left(\frac{n(h_1 - h_2)}{q}\right)$$

Summing over n the left hand side is

$$\sum_{n=1}^{q} |\chi(n)|^2 |\tau(\chi)|^2 = \phi(q) |\tau(\chi)|^2$$

and on the right the terms cancel except when $h_1 = h_2$, so we get

$$q \sum_h \chi(h)\overline{\chi}(h) = \phi(q)q$$

so $|\tau(\chi)| = \sqrt{q}$. □

If $\chi(-1) = 1$ then χ is even, otherwise it is odd. When it is even we can write the Poisson sum as the sum over $\mathbb{N}$ rather than $\mathbb{Z}$ as in the case of the Riemann zeta function.

For $x > 0$ and q the conductor of χ, we define the theta function

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

Even Case: $\chi(-1) = 1$.

We can write

$$\begin{aligned}\psi(x, \chi) &= \sum_{n=-\infty}^{\infty} \frac{1}{\tau(\overline{\chi})} \sum_{m=1}^{q} \overline{\chi}(m) e\left(\frac{nm}{q}\right) e^{-\pi n^2 x/q} \\ &= \frac{1}{\tau(\overline{\chi})} \sum_{m \bmod q} \overline{\chi}(m) \sum_{n=-\infty}^{\infty} e^{\frac{-\pi n^2 x + 2\pi i m n}{q}}\end{aligned}$$

For $x > 0$ and $\alpha \in \mathbb{R}$, we pick

$$f_{\alpha,x}(y) = e^{-\frac{(y+\alpha)^2 \pi}{x}} \in S(\mathbb{R})$$

Applying Poisson summation,

$$\sum_{n=-\infty}^{\infty} e^{-\frac{(n+\alpha)^2 \pi}{x}} = \sqrt{x} \sum_{m=-\infty}^{\infty} e^{-\pi m^2 x + 2\pi m \alpha}$$

So our running expression is

$$\begin{aligned}\psi(x, \chi) &= \frac{1}{\tau(\overline{\chi})} \sum_{m \bmod q} \overline{\chi}(m) \sqrt{\frac{q}{x}} \sum_{n=-\infty}^{\infty} e^{-\frac{\left(n + \frac{m}{q}\right)^2 \pi q}{x}} \\ &= \frac{1}{\tau(\overline{\chi})} \sqrt{\frac{q}{x}} \sum_{m \bmod q} \overline{\chi}(m) \sum_n e^{-\pi \frac{(qn+m)^2}{qx}}\end{aligned}$$

Since χ is periodic mod q, we get

$$\frac{1}{\tau(\overline{\chi})} \sqrt{\frac{q}{x}} \sum_t \overline{\chi}(t) e^{-\pi \frac{t^2}{qx}}$$

Thus we can write

$$\psi(x, \chi) = \varepsilon_\chi x^{-\frac{1}{2}} \psi(x^{-1}, \overline{\chi})$$

where

$$\varepsilon_\chi = \frac{\sqrt{q}}{\tau(\overline{\chi})}, \quad |\varepsilon_\chi| = 1$$

is called the **root number** of χ. When χ is even we have $\varepsilon_\chi \varepsilon_{\overline{\chi}} = 1$.

We write

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

We switch the order and change variables as $n\sqrt{x} = y$ to get

$$I(s) = \left(\frac{\pi}{q}\right)^{-\frac{s}{2}} \Gamma\left(\frac{s}{2}\right) L(s, \chi)$$

Definition 2.10

For an even primitive character χ with conductor q, we define the **completed L-function**

$$\Lambda(s, \chi) = \left(\frac{\pi}{q}\right)^{-\frac{s}{2}} \Gamma\left(\frac{s}{2}\right) L(s, \chi)$$

Since χ is even this becomes

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

We split this into the integral term to 0 (I) and the term to ∞ (II). Since ψ is Schwartz, II is entire. For I, we use the functional equation for ψ to get

$$\begin{aligned} I(s) &= \frac{1}{2} \int_0^1 x^{\frac{s}{2}} \varepsilon_\chi x^{-\frac{1}{2}} \psi(x^{-1}, \overline{\chi}) \frac{\mathrm{d}x}{x} + \frac{1}{2} \int_1^\infty x^{\frac{s}{2}} \psi(x, \chi) \frac{\mathrm{d}x}{x} \\ &= \frac{1}{2} \int_1^\infty x^{\frac{s}{2}} \psi(x, \chi) \frac{\mathrm{d}x}{x} + \frac{\varepsilon_\chi}{2} \int_1^\infty x^{\frac{1-s}{2}} \psi(x, \overline{\chi}) \frac{\mathrm{d}x}{x} \end{aligned}$$

Now both terms are entire. We use the fact that $\varepsilon_\chi \varepsilon_{\overline{\chi}} = 1$ to see that

Theorem 2.8

When χ is an even primitive character with conductor $q \neq 1$,

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

Odd Case: $\chi(-1) = 1$.

When χ is odd most of the above work still holds, but the final integral cannot be related

to ψ. Instead, we modify the Poisson summation formula by differentiating in α:

$$\sum_{n=-\infty}^{\infty} e^{-\frac{(n+\alpha)^2\pi}{x}} = \sqrt{x} \sum_{m=-\infty}^{\infty} e^{-\pi m^2 x + 2\pi m\alpha}$$

$$2\pi i\sqrt{x} \sum_{n=-\infty}^{\infty} n e^{-\pi n^2 + 2\pi n\alpha} = -2\pi \sum_{n} (n+\alpha) e^{-\pi \frac{(n+\alpha)^2}{x}}$$

We now have an odd function which we can combine with our odd character, proceeding with all work as before.

Definition 2.11

When χ is an odd primitive character with conductor q, the completed L-function is defined as

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

Theorem 2.9

When χ is an odd primitive character with conductor q,

$$\Lambda(s,\chi) = \frac{\tau(\chi)}{i\sqrt{q}} \Lambda(1-s,\overline{\chi}) = \varepsilon_\chi \Lambda(1-s,\overline{\chi})$$

We can consolidate these into the following equation:

Theorem 2.10

If χ is a primitive character with conductor q, let $a = 1$ if χ is odd and 0 otherwise. Then

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

where

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

$$\varepsilon_\chi = \frac{i^a\sqrt{q}}{\tau(\chi)}$$

The functional equations tell us that when χ is even, $L(s,\chi)$ has simple zeros at 0 and the negative even integers. When χ is odd $L(s,\chi)$ has simple zeros at the negative odd integers.

Theorem 2.11: Stirling's Formula

For fixed σ, as $|t| \to \infty$,

$$|\Gamma(\sigma + it)| \sim \sqrt{2\pi}|t|^{\sigma - \frac{1}{2}} e^{-\frac{\pi}{2}|t|}$$

2.5 Consequences of the Riemann Hypothesis

The Riemann hypothesis has a natural extension to more general L-functions. For Dirichlet L-functions, which arise from multiplicative characters, $L(s, \chi)$ is not necessarily symmetric. (When $\chi = \overline{\chi}$, the functional equation means that it is symmetric, and that χ is even).

The generalized Riemann hypothesis for Dirichlet L-functions says that the zeros of all completed L-functions $\Lambda(s, \chi)$ with χ primitive are on $\mathrm{Re}(s) = 1/2$.

It is also the case that there are approximately $\log q$ zeros of $L(s, \chi)$ up to height 1. We can use this to provide a polynomial-time algorithm (in the number of digits $\log n$) for testing whether a number is prime.

Theorem 2.12: Miller

Assuming the generalized Riemann hypothesis for Dirichlet L-functions, there is a polynomial time algorithm for primality testing in $\log n$.

Theorem 2.13: Agrawal-Kayal-Saxena

The above theorem is true without assuming GRH.

We say that an algorithm is efficient if it can be computed in polynomial time of $\log n$. Take n to be a large odd number. If n is prime, then Fermat's little theorem says that

$$x^{n-1} \equiv 1 \pmod{n} \qquad (*)$$

for all $(x, n) = 1$ (the gcd may be obtained efficiently). Given x, we can compute x^{n-1} efficiently by repeatedly squaring. The converse is not a prime test; there are numbers which are not prime but for which $(*)$ holds (these are called Carmichael numbers; there are infinitely many nonprime).

Proposition 2.14

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

$$\left(\frac{2}{p}\right) = \begin{cases} 1, & p \equiv \pm 1 \pmod 8 \\ -1, & p \equiv \pm 5 \pmod 8 \end{cases}$$

Proof. We only prove the second. Take α a primitive 8th root of 1 in $\overline{\mathbb{F}_p}$. Then $\alpha^4 = -1$, so $\alpha^2 + \alpha^{-2} = 0$. If $y = \alpha + \alpha^{-1}$ then $y^2 = 2$. We have

$$y^p = \alpha^p + \alpha^{-p}$$

If $p \equiv \pm 1(8)$ then $y^p = y$ so $y \in \mathbb{F}_p$ and $\left(\frac{2}{p}\right) = 1$. □

Definition 2.12

For n odd and $(a, n) = 1$, we define the **Jacobi symbol** by

$$\left(\frac{a}{n}\right) \equiv a^{\frac{n-1}{2}} \pmod{n}$$

When n is prime this is equivalent to the Legendre symbol. For other odd n it is multiplicatively defined as

$$\left(\frac{a}{n_1 \cdots n_t}\right) = \left(\frac{a}{n_1}\right) \cdots \left(\frac{a}{n_t}\right)$$

Theorem 2.15: Jacobi Reciprocity

If m, n are odd and $(m, n) = 1$, then

$$\left(\frac{m}{n}\right)\left(\frac{n}{m}\right) = (-1)^{\frac{m-1}{2}\frac{n-1}{2}}$$

and

$$\left(\frac{-1}{m}\right) = (-1)^{\frac{m-1}{2}}$$
$$\left(\frac{2}{m}\right) = (-1)^{\frac{m^2-1}{8}}$$

Theorem 2.16: Lehmer

For n odd,

$$\left(\frac{a}{n}\right) \equiv a^{\frac{n-1}{2}} \pmod{n}$$

for all $(a, n) = 1$ if and only if n is prime.

Proof. Exercise: show that the test fails when n is divisible by a square.

When $n = p_1 \cdots p_r$ is squarefree, we want to show that $r = 1$. By assumption,

$$a^{n-1} \equiv 1 \pmod{n}$$

so by the Chinese Remainder theorem,

$$a^{n-1} \equiv 1 \pmod{a}_j$$

for each j. We will pick a which is a primitive root for each p_j. $\square$

For a given a we can efficiently test Lehmer's criterion. The left hand side, as before, is done using repeated squares, and the right hand side is evaluated efficiently using reciprocity to reduce the numbers exponentially quickly. When assuming GRH, we need only check this condition for polylogarithmically many a.

Theorem 2.17

Assume GRH. Then

$$\left(\frac{a}{n}\right) \equiv a^{\frac{n-1}{2}} \pmod{n}$$

for all $1 < a \ll (\log n)^2$ if and only if n is prime.

This is an immediate consequence of the following theorem:

Theorem 2.18: Ankeny

Under GRH, if G is a proper subgroup of $(\mathbb{Z}/n\mathbb{Z})^*$, then there is $a \notin G$ with $1 < a \ll (\log n)^2$.

Then the previous theorem follows by considering

$$G = \left\{a \in (\mathbb{Z}/n\mathbb{Z})^* : \left(\frac{a}{n}\right) \equiv a^{\frac{n-1}{2}} \pmod{n}\right\}$$

Proof. Since G is proper, $(\mathbb{Z}/n\mathbb{Z})^*/G$ and $\widehat{(\mathbb{Z}/n\mathbb{Z})^*/G}$ are both nontrivial. Let χ be a nontrivial character of $(\mathbb{Z}/n\mathbb{Z})^*/G$. Then χ lifts to a nontrivial Dirichlet character of $(\mathbb{Z}/n\mathbb{Z})^*$ which is equal to 1 on G. So we need to show that if $\chi(m) = 1$ for $1 \le m \ll (\log q)^2$ with (m, q), then $\chi = \chi_0$.

We consider the smoothed function

$$\sum_{m \le x} \chi(m)\Lambda(m)\left(1 - \frac{m}{x}\right) = -\frac{1}{2\pi i}\int_{\mathrm{Re}(s)=2} \frac{L'}{L}(s,\chi)x^s \frac{\mathrm{d}s}{s(s+1)}$$

The integrand has poles at the zeros of L, which are those in the critical strip and the trivial zeros. The sum of the contributions for the trivial zeros is $\ll \log p$. We know there are infinitely many nontrivial zeros, and that the sum is conditionally convergent:

$$\sum_\rho \frac{1}{|\rho|} = \infty, \qquad \sum_\rho \frac{x^\rho}{\rho(\rho+1)} < \infty$$

One can prove (we delay this proof) using only complex analysis that the number of zeros with imaginary part $|\gamma - t| \le 1$ is $\ll \log(q(|t|+1))$. Under the generalized Riemann hypothesis, we then get

$$\sum_\rho \frac{x^\rho}{\rho(\rho+1)} = \sum_\rho \frac{x^{\frac{1}{2}+i\gamma}}{\rho(\rho+1)} \ll x^{\frac{1}{2}} \sum_\rho \frac{1}{|\rho(\rho+1)|} \ll x^{\frac{1}{2}} \int \frac{\log q + \log(|t|+1)}{|t|^2}\,\mathrm{d}t \ll x^{\frac{1}{2}} \log q$$

Now assume that $\chi(m) = 1$ for all $(m, q) = 1$, $m \leq X$. Then

$$\sum_{\substack{(p,q)=1 \\ p \leq X}} \log p \left(1 - \frac{p}{X}\right) = \sum_{p \leq X} \log p \left(1 - \frac{p}{X}\right) - \sum_{p \mid q} \log p$$

so

$$\sum_{p \leq X} \left(1 - \frac{p}{X}\right) \log p \ll X^{\frac{1}{2}} \log q \implies \frac{X}{2} \ll X^{\frac{1}{2}} \log q \implies X \ll (\log q)^2 \qquad \square$$

2.6 Partial Fraction Expansion of ζ

Theorem 2.19: Jensen's Formula

Suppose $f(0) \neq 0$ and $z_1, \ldots, z_n$ are the zeros of $f(z)$ in $\mathbb{D}_R$ with multiplicity, and f does not vanish on $\partial \mathbb{D}_R$. Then

$$\frac{1}{2\pi} \int_0^{2\pi} \log \left| f(Re^{i\theta}) \right| \mathrm{d}\theta = \log |f(0)| + \log \frac{R^n}{|z_1| \cdots |z_n|}$$

Proof. The function

$$F(z) = \frac{f(z)}{(z - z_1) \cdots (z - z_n)}$$

is nonvanishing on $\mathbb{D}_R$, so $\log |F(z)|$ is harmonic. By the mean value property,

$$\begin{aligned} \frac{1}{2\pi} \int_0^{2\pi} \log \left| f(Re^{i\theta}) \right| \mathrm{d}\theta &= \log |f(0)| - \sum_{j=1}^n \log |z_j| + \sum_{j=1}^n \frac{1}{2\pi} \int_0^{2\pi} \log \left| Re^{i\theta} - z_j \right| \mathrm{d}\theta \\ &= \log |f(0)| + \sum_{j=1}^n \log \frac{R}{|z_j|} \end{aligned} \qquad \square$$

The remainder term of Jensen's formula can be expressed in terms of counting the number of zeros. Define

$$n(r) = \#\{z : f(z) = 0, 0 < |z| < r\}$$

Then we write

$$\log \frac{R^n}{|z_1| \cdots |z_n|} = \int_0^R \frac{1}{t} \, \mathrm{d}n(t) = \int_0^R n(r) \frac{\mathrm{d}r}{r}$$

If

$$\sup_{|z| \leq R} |f(z)| \ll e^{R^\alpha}$$

for all $R \geq 1$ then we say that f is of order at most α. If this is the case, then

$$\int_0^R n(r) \frac{\mathrm{d}r}{r} \ll R^\alpha$$

Similarly

$$\int_R^{2R} n(r)\frac{\mathrm{d}r}{r} \ll R^\alpha$$

so

$$n(R) \ll R^\alpha$$

Theorem 2.20

Suppose that $f(z)$ has order less than $\alpha < 2$. Then

$$\sum_\rho \frac{1}{|\rho|^2} < \infty$$

Proof. There are finitely many zeros in $\mathbb{D}$, so we just consider the rest. This is

$$\int_1^\infty \frac{1}{t^2}\,\mathrm{d}n(t) = \left.\frac{n(t)}{t^2}\right|_1^\infty + 2\int_0^\infty t^{-3}n(t)\,\mathrm{d}t < \infty$$

□

Theorem 2.21: Hadamard Factorization

Let ρ be a sequence of zeros of f with $|\rho \to \infty|$. Define

$$g(z) = \prod_\rho \left(! - \frac{z}{\rho}\right) e^{\frac{z}{\rho}}$$

Let $z \in K \subseteq \mathbb{C}$ compact. Then

$$\log\left(1 - \frac{z}{\rho}\right) + \frac{z}{\rho} = -\frac{z}{\rho} + \frac{z}{\rho} + O\left(\left|\frac{z^2}{\rho^2}\right|\right)$$

so g is entire. Then f/g is entire and nonvanishing, so $f = ge^p$, where

$$\sup_{|z|=R} |P(z)| \ll R^\alpha$$

and therefore p is a polynomial of degree at most α.

When f has order $\alpha = 1$ we have

$$f(z) = e^{A+Bz}z^a \prod_\rho \left(1 - \frac{z}{\rho}\right) e^{\frac{z}{\rho}}$$

We can also take the logarithmic derivative to get

$$\frac{f'}{f}(z) = B + \frac{a}{z} + \sum_\rho \left(\frac{1}{z-\rho} + \frac{1}{\rho}\right)$$

We apply this to the entire function

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

or more generally for primitive characters χ with conductor $q > 1$,

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

We need to show this is order 1. By Stirling's formula, for s away from the negative real axis,

$$\log\Gamma(s) = \left(s-\frac{1}{2}\right)\log s + \text{lower order terms}$$

$\log(\pi/q)^{-(s+a)/2}$ is linear in s. So the last term we need to evaluate is $\log|L(s,\chi)|$. By summing by parts, we get $\log|L(s,\chi)| \ll_q \log|s|$ for $\mathrm{Re}(s) \geq 1/2$. So for $\mathrm{Re}(s) \geq 1/2$.

$$\log|\Lambda(s,\chi)| \ll_q |s|\log(|s|+2)$$

So $\lambda(s,\chi)$ is entire of order 1, and similarly so is $\xi(s)$. Thus we write

$$\Lambda(s,\chi) = e^{A+Bs}\prod_{\rho}\left(1-\frac{s}{\rho}\right)e^{\frac{s}{\rho}}$$

where ρ ranges over the zeros of $L(s,\chi)$ in the critical strip. Then

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

When working with ξ rather than Λ, the sum can be shown to be converge by grouping ρ with $\overline{\rho}$. Therefore by evaluating at certain values,

$$B(\zeta) = -\sum_{\rho}\frac{1}{\rho} = -2\sum_{\rho=\beta+i\gamma}\frac{\beta}{\beta^2+\gamma^2} = -\frac{\gamma}{2} - 1 + \frac{1}{2}\log 4\pi \approx -0.023$$

where the γ in the fourth expression is the Euler-Mascheroni constant. Finally we write

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

It is still true that

$$\mathrm{Re}(B(\chi)) = -\sum_{\rho}\mathrm{Re}\left(\frac{1}{\rho}\right)$$

converges. When evaluating for $s = 2+it$,

$$-\mathrm{Re}\,\frac{L'}{L}(2+it,\chi)$$

is uniformly bounded in t. Specifically,

$$-\operatorname{Re}\frac{L'}{L}(2+it,\chi) \ll \log q + \log(|t|+2) - \sum_\rho \frac{2-\beta}{(2-\beta)^2(t-\gamma)^2}$$

So we apply boundedness and can count zeros as

$$\sum_\rho \frac{2-\beta}{(2-\beta)^2+(t-\gamma)^2} \ll \log[q(2+|t|)]$$

Therefore

$$\#\{\rho : t-1 \le \gamma \le t+1\} \ll \log q(|t|+2)$$

To actually count the zeros, let C be the contour from $2 \to 2+iT \to -1+iT \to -1 \to 2$, bounding the box B. Then by winding number arguments,

$$\frac{1}{2\pi}\Delta_C \arg \xi(s) = \#\{\rho \in B : \xi(\rho) = 0\} = N(T)$$

where $\Delta_C \arg$ denotes the change in argument along C. But $\xi(s)$ is real when s is real, so there is no change in argument along the bottom segment, and by the functional equation we have

$$N(T) = \frac{1}{\pi}\Delta_L \arg \xi(s)$$

where L is the contour from $2 \to 2+iT \to 1/2+iT$. Then

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

The first two terms are small. The Γ term is given by applying Stirling's formula, which gives terms looking like $T \log T$. The ζ term is shown to be $O(\log T)$ since zeros with height close to T have contribution bounded by π, and terms farther away are negligible. So by our bound on the number of these zeros, this is negligible.

Theorem 2.22

The number of zeros up to height T for $L(s,\chi)$, χ primitive mod q is

$$N(T,\chi) = \frac{T}{2\pi}\log\left(\frac{qT}{2\pi}\right) - \frac{T}{2\pi} + O(\log(|T+2|q))$$

Theorem 2.23: Selberg

The number of zeros of ζ up to height T with real part $1/2$ is

$$\gg T \log T$$

Chapter 3

Dirichlet's Class Number Formula

Consider a quadratic form (which is a homogeneous polynomial in degree 2), and without loss of generality we may write it in symmetric form:

$$\sum_{i,j} a_{ij}x_ix_j, \qquad a_{ij} = a_{ji}$$

We can consider what numbers may be represented by quadratic forms.

Theorem 3.1: Gauss

A number n may be written as a sum of 3 squares if and only if n is not of the form

$$n = 4^a(8b + 7)$$

Quadratic forms may be nicely analyzed because the level sets of a quadratic form over $\mathbb{R}$ form a sphere, or over more general fields an orthogonal group.

Definition 3.1

Let f be a quadratic form. Then the **orthogonal group** of F over $\mathbb{R}$ is

$$O_f(\mathbb{R}) = \{B \in \mathrm{GL}_n(\mathbb{R}) : f(Bx) = f(x) \forall x \in \mathbb{R}^n\}$$

Equivalently, if A is the matrix representing f, then

$$O_f(\mathbb{R}) = \left\{B \in \mathrm{GL}_n(\mathbb{R}) : B^TAB = A\right\}$$

In particular elements of $O_f(\mathbb{R})$ have determinant ± 1. We similarly define $O_f(\mathbb{Z}) \subseteq \mathrm{GL}_n(\mathbb{Z})$, $\mathrm{SO}_f(\mathbb{R}) \subseteq \mathrm{SL}_n(\mathbb{R}), \mathrm{SO}_f(\mathbb{Z}) \subseteq \mathrm{SL}_n(\mathbb{Z})$. We also use the alternative notation $\mathrm{Aut}_f = \mathrm{SO}_f$.

The function $f \circ A$ is a quadratic form as well, so we consider equivalence classes of integer quadratic forms up to integer invertible linear transformations.

Definition 3.2

Let f, g be two integral quadratic forms over two variables. Then f is **integrally equivalent** to g if $f(x) = g(Ax)$ for some $A \in \mathrm{GL}_n(\mathbb{Z})$. They are **narrow equivalent** if $A \in \mathrm{SL}_n(\mathbb{Z})$.

Example 3.1

$x_1^2 - x_2^2$ is equivalent to $x_1 x_2$ over $\mathbb{R}$ by the transformation $(x_1, x_2) \mapsto (x_1 + x_2, x_1 - x_2)$.

To work to analyze these quadratic Diophantine equations, we can use the Dirichlet class number formula, which is a special case of the Siegel mass formula.

3.1 Quadratic Forms

First we consider the case $n = 2$, so

$$f = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$G = \mathrm{SL}_2(\mathbb{R})$ has a topology inherited from $\mathbb{R}^4$, and $\Gamma = \mathrm{PSL}_2(\mathbb{Z})$ is a discrete subgroup in $\mathrm{SL}_2(\mathbb{R})$. $\mathrm{SL}_2(\mathbb{R})$ acts transitively on the upper half plane via Mobius transformations

$$z \mapsto \frac{az + b}{cz + d}$$

Then the stabilizer is

$$K = \mathrm{stab}_i(G) = \{d = a, c = -b, ad - bc = 1\} = \left\{ \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}, \theta \in [0, 2\pi) \right\}$$

So

$$\mathrm{PSL}_2(\mathbb{R}) \big/ K \cong \mathbb{H}$$

both as sets and as topological spaces. For $z \in \mathbb{H}$, if we consider the orbit Γz under $\mathrm{PSL}_2(\mathbb{Z})$, we claim that Γz has limit points everywhere on the real line, but nowhere else. This is called a **properly discontinuous action**. We can try to understand this action by searching for representatives of the orbits to create a fundamental domain. A fundamental domain is a set such that any point is equivalent to exactly one point in the fundamental domain under the action of Γ.

Theorem 3.2: Gauss

Let F be the deleted strip

$$F = \{|\mathrm{Re}(z)| < 1/2, |z| > 1\}$$

and let $\mathcal{F}$ include half the boundary:

$$\mathcal{F} = F \cup (\partial F \cap \mathrm{Re}(z) \geq 0)$$

$\mathcal{F}$ is a fundamental domain of Γ acting on $\mathbb{H}$.

Proof. Γ is generated by

$$T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \qquad S = \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix}$$

$T^n z = z + n$, so it certainly suffices to look at this strip. To show that we can avoid the lower parabola, note that

$$\mathrm{Im}(Sz) = \frac{\mathrm{Im}(z)}{|z|^2}$$

So if we begin inside the circle, we will approach the boundary vertically. Since there are no accumulation points, this terminates after a finite number of steps.

Exercise: show that the only points which are Γ equivalent lie on the boundary. □

In higher dimensions we consider either $\Gamma = \mathrm{SL}_n(\mathbb{Z})$ or $\Gamma = \mathrm{GL}_n(\mathbb{Z})$, and $G = \mathrm{SL}_n(\mathbb{R})$ or $\mathrm{GL}_n(\mathbb{R})$ respectively. We can look at the coset space

$$\mathrm{SL}_n(\mathbb{R}) \big/ \mathrm{SL}_n(\mathbb{Z})$$

which is endowed with the quotient topology. For any $g \in \mathrm{GL}_n(\mathbb{R})$, $L_g = g\mathbb{Z}^n$ is a rank n lattice in $\mathbb{R}^n$, and if $g \in \mathrm{SL}_n(\mathbb{R})$ then L_g has volume 1. The map $g \mapsto L_g$ from $\mathrm{SL}_n(\mathbb{R})$ to the space of lattices satisfies

$$\gamma\mathbb{Z}^n = \mathbb{Z}^n$$

for $\gamma \in \mathrm{SL}_n(\mathbb{Z})$ since γ is integrally invertible, so the space of lattices with volume 1 in $\mathbb{R}^n$ is $\mathrm{SL}_n(\mathbb{R})/\mathrm{SL}_n(\mathbb{Z})$ (with no volume restriction for GL_n).

Definition 3.3

Let $L \subseteq \mathbb{R}^n$ be a lattice. Define

$$\mu(L) = \inf\{|v| : v \in L \setminus \{0\}\}$$

μ is strictly positive and continuous with respect to the topology on the space of lattices.

Definition 3.4

A set in a topological space is **precompact** if it has compact closure.

Theorem 3.3: Mahler Compactness

Let $B \subseteq \mathrm{GL}_n(\mathbb{R})/\mathrm{GL}_n(\mathbb{Z})$ be a subset of the space of lattices on $\mathbb{R}^n$. B is precompact if and only if there exists $C_B < \infty, \delta_B > 0$ such that for any $L \in B$,

$$\mu(L) \geq \delta_B, \quad \mathrm{covol}(L) \leq C_B$$

Lemma 3.4

Let L be a lattice with covolume $d(L)$ and $S \subseteq \mathbb{R}^n$ be convex, with $S = -S$ (symmetric about the origin). If $\mathrm{vol}(S) > 2^n d(L)$, then there is $\xi \neq 0$ contained in $S \cap L$.

Proof. We prove it for $d(L) = 1$.

Write $L = g\mathbb{Z}^n$ with $\det g = 1$. Then $g^{-1}S$ is also convex and symmetric, so we can assume $L = \mathbb{Z}$. Reduce v modulo $\mathbb{Z}^n$. Since the volume of $S/2$ is greater than the volume of the fundamental domain $[-1/2, 1/2]^n$, there are $v \neq w \in S/2$ with $\pi(v) = \pi(w)$. Denote

$$\xi = v - w = \frac{1}{2}(2v) + \frac{1}{2}(-2w) \in S$$

(by symmetry and convexity). Then $\xi \neq 0$ but $\xi \in \mathbb{Z}^n$. □

Lemma 3.5

Let $L_\tau = \{m + n\tau : m, n \in \mathbb{Z}\}$ be a lattice for $\tau \in \mathcal{F}$. Then

$$\mu(L_\tau) \leq \frac{1}{\sqrt{\mathrm{Im}(\tau)}}$$

Proof. The covolume of L_τ is $\mathrm{Im}(\tau)$. The disc $\mathbb{D}_r$ contains

$$\left[-\frac{r}{\sqrt{2}}, \frac{r}{\sqrt{2}}\right]^2$$

which has volume $2^2 r^2/\sqrt{2}^2$ (generally $(2r)^n/(\sqrt{n})^n$). For $r = 1/\sqrt{\mathrm{Im}(\tau)}$, we are guaranteed to contain a lattice point. So

$$\mu(L_\tau) \leq \frac{1}{\sqrt{\mathrm{Im}(\tau)}}$$ □

We give the proof of Mahler compactness in the converse direction for SL_n (where there is no covolume restriction since all covolumes are 1). This suffices for our purposes.

Proof of Mahler Compactness. ($\Longleftarrow$) Let $B \subseteq \mathcal{L} = \mathrm{SL}_n(\mathbb{R})/\mathrm{SL}_n(\mathbb{Z})$. By assumption $\mu(L_\tau) \geq \delta_B > 0$, so there is some C_B such that $\mathrm{Im}(\tau) \leq C_B < \infty$ for any $\tau \in B$. So B is bounded, hence precompact. □

Let $f(x) = x^T Ax$ be a real quadratic form ($\det A \neq 0$). If f is positive definite then $O_f(\mathbb{R})$ is compact (since it is the standard orthogonal group, up to transformations). If $f(x_1, \ldots, x_n)$ is integral (so the off diagonal entries of A may be half integers), we define $\Gamma = O_f(\mathbb{Z})$ to be the group of matrices with integer entries that stabilize f. Then $O_f(\mathbb{Z})$ is a discrete subgroup of $O_f(\mathbb{R})$. So if f is integral and definite then $O_f(\mathbb{Z})$ is finite, and the converse is true as well.

Proposition 3.6

If f is a definite integral quadratic form then $O_f(\mathbb{Z})$ is finite.

Definition 3.5

If f is integral or rational, then it is said to be **isotropic** if there is nonzero $x \in \mathbb{Z}^n$ (or $\mathbb{Q}^n$) such that $f(x) = 0$. Otherwise, f is anisotropic.

Theorem 3.7

Let f be anisotropic and integral. Then

$$O_f(\mathbb{R})/O_f(\mathbb{Z})$$

is compact.

Proof. We apply Mahler compactness. The covolume of the lattices is 1, so we just need to show that their shortest vectors are bounded below uniformly. Let $f(0) = 0$, so by continuity there is a set U containing 0 on which $|f| \leq 1/2$.

Now let $g\mathbb{Z}^2 \in \mathcal{L}_0$ and let $v = ge \in g\mathbb{Z}^2$ for some $e \in \mathbb{Z}^2$ nonzero. Then

$$f(v) = f(ge) = f(e) \neq 0$$

so $|f(v)| > 1/2$. Thus $v \notin U$, so there is a uniform bound from below on v. Then we apply Mahler compactness to see that $\mathcal{L}_0$ is precompact. To see that it is compact, we need to show that it is equal to its closure. Take $\mathcal{L}_i \to \mathcal{L}^*$ a sequence of lattices in $\mathcal{L}_0$. We can pick bases (u_i, v_i) tending to (u^*, v^*). Then $f(u_i) \to f(u^*)$ by continuity, but since $f(u_i)$ are integer values, $f(u^*)$ is also integral. Similarly for v^*, so $\mathcal{L}^* \in \mathcal{L}_0$. Thus $\mathcal{L}_0$ is compact. □

In particular, if f is not definite then $O_f(\mathbb{R})$ must be infinite, and since it is a discrete subgroup it is cyclic as well. This holds for binary quadratic forms with $d > 0$.

Example 3.2

Let $f(x_1, x_2) = x_1^2 - dx_2^2$ for some $d > 0$ with d not a perfect square. Then $f = 0$ only if $(\sqrt{d}x_2) = \pm x_1$, but this cannot be satisfied over $\mathbb{Z}, \mathbb{Q}$, so f is anisotropic.

Example 3.3

Let $f(x_1, x_2, x_3) = x_1^2 + x_2^2 - 7x_3^2$. Then by the classification of numbers which are sums of two squares, there are no solutions over $\mathbb{Z}$, so f is anisotropic.

Example 3.4

Let $f(x_1, x_2, x_3, x_4) = x_1^2 + x_2^2 + x_3^2 - 7x_4^2$ is anisotropic for the same reason.

3.2 Binary Quadratic Forms

We specialize to the case of binary quadratic forms

$$f(x, y) = ax^2 + bxy + cy^2, \quad (a, b, c) = 1$$

Definition 3.6

If f is a binary quadratic form, the **discriminant** is defined as

$$d = b^2 - 4ac$$

The discriminant is stabilized under the action $\mathrm{SL}_n(\mathbb{Z})$.

Definition 3.7

Let $C(d)$ denote the set of equivalence classes of binary quadratic forms with discriminant d, where $f \sim f'$ are equivalent if they are equivalent by a linear integral change of variables $\gamma \in \mathrm{SL}_2(\mathbb{Z})$. We denote $h(d) = |C(d)|$.

Proposition 3.8

If d is not a perfect square then $h(d)$ is finite.

Proof. We give a sketch for $d < 0$. The general idea is that any binary quadratic form with $d < 0$ not a square can be transformed so that $|b| \leq a \leq c$ for $a, c > 0$ (with $b = -a$ when $|b| = a$). This is analogous to the fundamental domain for $\mathrm{SL}_2(\mathbb{Z})$ (and in fact is proved by associating a point with $[a, b, c]$ and moving it into the fundamental domain). Then by combinatorics, there are finitely many such $[a, b, c]$. □

We have showed that $\mathrm{Aut}_f(\mathbb{R})/\mathrm{Aut}_f(\mathbb{Z})$ is compact. We have

$$\mathrm{Aut}_f(\mathbb{R}) = \left\{ \begin{bmatrix} \frac{t-bu}{2} & -cu \\ au & \frac{t+bu}{2} \end{bmatrix} : t^2 - du^2 = 4 \right\}$$

The equation $t^2 - du^2 = 4$ is called **Pell's equation**. When $d > 0$ it describes a hyperbola, and when $d < 0$ an ellipse. Either way it is a one-parameter group. Moreover, we can endow this curve with a group structure by lifting to the automorphism group.

Definition 3.8

The **fundamental unit** for $d > 0$ not a square is

$$\varepsilon_d = \frac{t_0 + \sqrt{d}u_0}{2}$$

where (t_0, u_0) is the integer solution with $u_0 \geq 1$ minimal.

As we showed in the previous section, when $d > 0$ the group $O_f(\mathbb{Z})$ is infinite and cyclic, so there are integer solutions to Pell's equation, and $O_f(\mathbb{Z})$ is generated by the minimal solution.

When $d < 0$, the number of integer solutions to $t^2 - du^2 = 4$ is

$$w_d = \begin{cases} 4, & d = -4 \\ 6, & d = -3 \\ 2, & d < -4 \end{cases}$$

Note that d cannot be $-1, -2$ as $d \equiv 0, 1(4)$ always. The value of w_d can be seen by sketching an ellipse with major axis 2 and minor axis $2/\sqrt{d}$; when $d = -4$ there are integer points at the corners of the circle; when $d = -3$ it is solved by $(\pm 1, \pm 1), (\pm 2, 0)$, and when $d < -4$ the minor axis is too short to intersect any lattice points so the only solutions are $(\pm 2, 0)$.

Definition 3.9

Let d be nonsquare, and $d \equiv 0, 1 \pmod 4$. Define the **generalized Legendre Kronecker symbol**

$$\chi_d(n) = \left(\frac{d}{n}\right)$$

to be the extension of the Legendre symbol to all integers d, n, with the additional definitions

$$\left(\frac{a}{2}\right) = \begin{cases} 0, & a \text{ even} \\ 1, & a \equiv \pm 1 \pmod 8 \\ -1, & a \equiv \pm 3 \pmod 8 \end{cases}$$

$$\left(\frac{a}{-1}\right) = \begin{cases} -1, & a < 0 \\ 1, & a \geq 0 \end{cases}$$

$$\left(\frac{a}{0}\right) = \begin{cases} 1, & a = \pm 1 \\ 0 \end{cases}$$

When d is squarefree and not 2 then χ_d is a primitive character. We then know that

$$L(1, \chi_d) = \sum_{n=1}^{\infty} \left(\frac{d}{n}\right) n^{-1} > 0$$

But the following theorem of Dirichlet provides an alternative proof, which was his original method.

Theorem 3.9: Dirichlet's Class Number Formula

If $d < 0$ then

$$h(d) = \frac{w_d |d|^{1/2}}{2\pi} L(1, \chi_d)$$

In particular the right hand side is an integer, and it is at least 1. If $d > 0$ then

$$h(d) \log \varepsilon_d = \frac{1}{2} d^{1/2} L(1, \chi_d)$$

As a result,

$$L(1, \chi_d) \gg |d|^{-1/2}$$

We prove that

$$L(1, \chi_d) \geq C_\varepsilon |d|^{-\varepsilon}$$

for all $\varepsilon > 0$. However, the proof of this statement does not produce C_ε. This is because it produces b_ε so that this holds except for at most 1 exception. Taking $d \to -\infty$, this proves that there are only finitely many d with class number 1, but this theorem cannot be used to verify a complete list. A newer theorem says that

$$h(d) \geq 10^{-300} \log |d|, \quad d < 0$$

For $d > 0$, it is difficult to separate $h(d) \log \varepsilon_d$, since ε_d is not well understood. For instance,

$$\sum_{d \leq X} h(d) \sim cX (\log X)^2$$

but

$$\sum_{d \leq X} h(d) \log \varepsilon_d \sim CX^{3/2}$$

For fixed d, the question of finding binary forms with discriminant d is equivalent to finding the number of solutions to the ternary quadratic form

$$d = f(a, b, c) = b^2 - 4ac$$

This is isotropic, so $O_f(\mathbb{R})/O_f(\mathbb{Z})$ is not necessarily compact.

Definition 3.10

Let $K \subseteq \mathbb{R}^3$ be compact. Then we define the **Archimedean measure** to be

$$\mu_{f,\mathbb{R}}(K) = \lim_{\varepsilon \to 0} \frac{\lambda \{x \in K : |f(x) - d| < \varepsilon\}}{2\varepsilon}$$

and for p prime, the **Archimedean density** is

$$\mu_f(p) = \lim_{\nu \to \infty} \frac{\#\{x : f(x) \equiv d \pmod{p^\nu}\}}{p^{2\nu}}$$

Theorem 3.10: Mass Formula

The number of solutions to $f(a,b,c) = d$ for $(a,b,c) \in K$ is

$$\mu_{F,\mathbb{R}}(K) \prod_p \mu_F(p)$$

Theorem 3.11

Consider the solutions to $f(x,y) = n$. Let d be the discriminant and assume $(n,d) = 1$. Let $R_f(n)$ denote the number of solutions over $\mathbb{Z}$.

If $d < 0$ then

$$\sum_{f \in C_d} R_f(n) = w_d \sum_{m|n} \left(\frac{d}{m}\right)$$

Proof of Dirichlet's Class Number Formula. For fixed $d < 0$, we write

$$\sum_{\substack{n \le N \\ (n,d)=1}} \sum_{f \in C_d} R_f(n) = \sum_{\substack{n \le N \\ (n,d)=1}} w_d \sum_{m|n} \left(\frac{d}{n}\right)$$

On the right, we get

$$w_d \sum_{\substack{n \le N \\ (n,d)=1}} \sum_{m|n} \left(\frac{d}{m}\right) = w_d \sum_{\substack{(m_1,m_2) \\ (m_1 m_2, d)=1 \\ m_1 m_2 \le N}} \left(\frac{d}{m_1}\right)$$

$$= w_d \sum_{m_1 \le \sqrt{N}} \sum_{\substack{m_2 \le \frac{N}{m_1} \\ (m_1 m_2, d)=1}} \left(\frac{d}{m_1}\right) + w_d \sum_{m_2 \le \sqrt{N}} \sum_{m_1} \left(\frac{d}{m_1}\right)$$

Since $\left(\frac{d}{m_1}\right)$ is periodic in d, the sum over m_1 in the second term is $O_1(d)$, so the entire second term is $\ll_d (\sqrt{N})$. The sum over m_2 in the first term ignoring the coprime condition is $N/m_1 + O(1)$, and with the condition we have $\phi(|d|)/|d|$ of the terms (ϕ the Euler totient function), so we get

$$w_d \frac{\phi(|d|)}{|d|} N \sum_{m_1 \le \sqrt{N}} \left(\frac{d}{m_1}\right) \frac{1}{m_1} + o(N) \sim w_d \frac{\phi(|d|)}{|d|} L(1, \chi_d) N$$

The left hand side is

$$\sum_{f\in C_d}\sum_{\substack{(x,y)\\(f(x,y),d)=1\\f(x,y)\le N}} 1$$

This counts the lattice points in an ellipse up to a congruence condition on d. The area of the ellipse is

$$A\sim\frac{2\pi N}{|d|^{\frac{1}{2}}}$$

Since f is a quadratic form, it suffices to just consider the congruence classes $x\equiv x_0(d), y\equiv y_0(d)$. The area of each of these is the same up to a constant factor (since it is an $O(1)$ shift). The number of classes for which this admits coprime solutions is $\phi(|d|)|d|$, so we get

$$\sum_{\substack{n\le N\\(n,d)=1}} R_f(n)\sim\frac{2\pi N}{|d|^{\frac{1}{2}}}\cdot\frac{\phi(|d|)|d|}{|d|^2}$$

and

$$w_d\frac{\phi(|d|)}{|d|}L(1,\chi_d)N\sim\sum_{\substack{n\le N\\(n,d)=1}}\sum_{f\in C(d)}R_f(n)\sim h(d)\frac{\phi(|d|)}{|d|}\frac{2\pi N}{|d|^{\frac{1}{2}}}$$

When $d>0$, instead of looking at ellipses we look at hyperoblas. These have infinite area so we have to look at the area mod $\mathrm{Aut}_f(\mathbb{Z})$. The automorphism group brings every point into a sector between the lines of slope $\varepsilon_d, \varepsilon_d^{-1}$. We can intersect with the hyperbola to get a finite area region of area $\log\varepsilon_d N$ and we proceed similarly. □

3.3 Zero Free Regions

We now derive regions in the critical strip on which L functions do not vanish. The strategy is to take a product of L functions such that the resulting Dirichlet series has positive coefficients:

$$\prod_\nu L(s,\chi_\nu)=\phi(s)=\sum_{n=1}^\infty b_n n^{-s},\qquad b_n\ge 0$$

By considering the completed L functions, we can also complete Φ as

$$\Phi(1-s)=\varepsilon\Phi(s)$$

(up to some real factors). Recall that $|\varepsilon|=1$, but since Φ has real and positive coefficients, $\varepsilon=\pm 1$. For $t_1\in\mathbb{R}$, we also allow a vertical shift:

$$L(s,(\chi\otimes t_1))=\sum_{n=1}^\infty\frac{\chi(n)}{n^{it}}n^{-s}$$

Taking the completion, we get

$$\Phi(s,\chi\otimes t_1)$$

which has **analytic conductor**

$$N_\Phi:=q_1(1+|t_1|)\cdots q_\nu(1+|t_\nu|)$$

Lemma 3.12

There is a family of absolute constants C_m such that if m is the order of the pole of Φ at $s = 1$, then Φ has at most m zeros in the real interval

$$\left(1 - \frac{C_m}{\log N_\Phi}, 1\right)$$

Proof. For $1 < s < 2$, we use the partial fraction expansion

$$\frac{\Phi'}{\Phi}(s) = \sum_\rho \frac{1}{s-\rho} - \left[\frac{m}{s-1} + \frac{m}{s}\right]$$

The ρ occur in conjugate pairs, so by gropuing the terms we see that $\Phi'/\Phi \leq 0$ for $1 < s < 2$. Since Φ differs from ϕ by a factor, the logarithmic derivative ϕ'/ϕ is Φ'/Φ plus terms of the form $\Gamma'/\Gamma((s+it_j)/2), \log q_j$. The Γ terms act like $\log t_j$, so the whole sum acts as $\log N_\Phi$. We can also discard ϕ'/ϕ. So we get

$$\sum_\rho \frac{1}{s-\rho} \leq \frac{m}{s-1} + C \log N_\phi$$

Plug in

$$s = \frac{1}{2C \log N_\Phi} + 1$$

and pick

$$C_m = \frac{1}{2c(m+1)}$$

Then it follows that if there are more than m zeros in the interval, the inequality is violated. □

The point of this is that we will construct our Φ such that if any L function has a zero, then the entire product has more zeros than is allowed in the interval.

Example 3.5

Let $\chi \neq \overline{\chi}$ be a nonreal character. Then

$$\phi(s) = \zeta^3(s) L^2(s,\chi) L^2(s,\overline{\chi}) L(s,\chi^2) L(s,\overline{\chi}^2)$$

We can check that this has positive coefficients. This has a pole of order $m = 3$ at $s = 1$ (since χ is not real, $\chi^2, \overline{\chi}^2$ are not trivial so the L-function does not have a pole). Let q be the conductor of χ. Then in

$$\left(1 - \frac{C_3}{6 \log q}, 1\right)$$

has at most 3 zeros. If $L(s,\chi)$ vanishes at any point in this interval then Φ has a zero of order 4, so $L(s,\chi)$ does not vanish on this interval.

Example 3.6

Consider

$$\phi(s) = \zeta^3(s)\zeta^2(s+it_0)\zeta^2(s-it_0)\zeta(s+2it_0)\zeta(s-2it_0)$$

This has positive coefficients. The conductor is $6\log(1+|t_0|)$. So (assuming $t_0 \geq 1$ for simplicity of notation),

$$\left(1 - \frac{c}{6\log|t_0|}, 1\right)$$

has at most 3 zeros. Thus there are no zeros of $\zeta(\sigma + it_0)$ for σ in the interval.

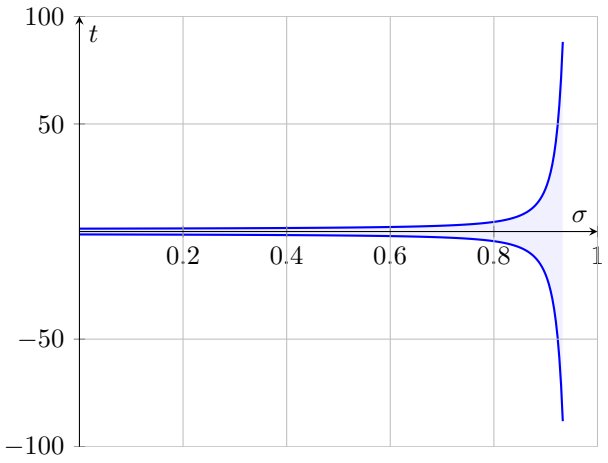

Figure 3.1: Zero-Free Region for ζ

Example 3.7

Let $t_0 \geq 1$, and $\chi = \overline{\chi}$ real. Then

$$\phi(s) = \zeta^3(s)L^2(s+it_0,\chi)L^2(s-it_0,\chi)L(s+2it_0,\chi)L(s-2it_0,\chi)$$

has positive coefficients. It has a pole of order 3 and the standard zero free region is

$$\left(1 - \frac{c}{\log q(t_0+1)}, 1\right), \quad t_0 \geq 1$$

The **isobaric sum** $\boxplus$ is a tool from representation which is defined on characters of the form $\chi, \chi \otimes t_1$ such that

$$L(s, \pi_1 \boxplus \ldots \boxplus \pi_\nu) = L(s,\pi_1)\cdots L(s,\pi_\nu)$$

To get positivity, we use

$$L(s, (\pi_1 \boxplus \ldots \boxplus \pi_\nu) \times \overline{(\pi_1 \boxplus \ldots \boxplus \pi_\nu)}) = L(s, \pi_1 \times \overline{\pi_1} \boxplus \pi_1 \times \overline{\pi_2} \boxplus \ldots)$$

The L-function $L(s, \pi \times \pi')$ is simply $L(s, \chi\overline{\chi})$ when $\pi = \chi, \pi' = \overline{\chi}$, otherwise when we twist by t_0 it is defined as something called a **Rankin-Selberg L-function**. Since we multiply by conjugates, it is easy to check that the coefficients are positive.

Example 3.8

To see how our first example can be produced using this method, we again pick $\chi \neq \overline{\chi}$ and consider

$$L(s, (1 \boxplus \chi \boxplus \overline{\chi}) \times \overline{1 \boxplus \chi \boxplus \overline{\chi}})$$

On multiplying out, we get three $L(s, 1) = \zeta$ factors from the $(1, 1)$ term and the cross diagonal, as well as two factors $L(s, \chi), L(s, \overline{\chi})$, and lastly a factor each of $L(s, \chi^2), L(s, \overline{\chi}^2)$ from the diagonal.

Example 3.9

Let χ_1, χ_2 be two real characters with $\chi_1 \neq \chi_2$. Then let

$$\phi(s) = \zeta(s)L(s, \chi_1)L(s, \chi_2)L(s, \chi_1\chi_2)$$

We can observe that this has positive coefficients since it is the Dedekind zeta function of $K = Q(\sqrt{d_1}, \sqrt{d_2})$. Then $N_\phi \leq q_1 q_2$ so

$$\left(1 - \frac{C_1}{\log q_1 q_2}, 1\right)$$

has at most one zero. Since this is true for every χ_1, χ_2, we can use this later to run arguments which show that there is at least one exception.

Definitions