3 Zeta and L-functions

In this section we will use analytic methods to (i) develop a formula for class numbers, and (ii) use this to prove Dirichlet's theorem in arithmetic progressions: that any arithmetic progression: $a+m, a+2m, a+3m, \ldots$ contains infinitely many primes $\gcd(a,m)=1$.

This chapter follows [Cohn], though our presentation is reversed from his, together with some supplementary material taken from various other sources. More general treatments are found in [Marcus] and [Neukirch], though they do not do everything we will do here.

3.1 Zeta functions

Recall one defines the **Riemann zeta function** by

$$\zeta(s) = \sum \frac{1}{n^s}. \tag{3.1}$$

One knows from calculus that this converges for $s>1$ (compare with

$$\int_1^\infty \frac{1}{x^s}dx = \frac{x^{1-s}}{1-s}\bigg]_{x=1}^{\infty} = \frac{1}{1-s} < \infty.)$$

Euler observed that (for $s>1$) one also has the product expansion

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

Here p runs over all primes of $\mathbb{N}$. The last equality just follows from the formula for a geometric series: $\sum_{n=0}^\infty a^n = \frac{1}{1-a}$ if $|a|<1$. To see the why product expansion (middle equality) is valid, it's perhaps easiest to first notice that it is *formally* true for $s=1$,[*] where it says

$$\sum \frac{1}{n} = \left(1+\frac{1}{2}+\frac{1}{2^2}+\frac{1}{2^3}+\cdots\right)\left(1+\frac{1}{3}+\frac{1}{3^2}+\frac{1}{3^3}+\cdots\right)\left(1+\frac{1}{5}+\frac{1}{5^2}+\frac{1}{5^3}+\cdots\right)\cdots \tag{3.2}$$

What does this (formal) infinite product on the right mean? It just means a (formal) limit of the sequence of finite subproducts:

$$1+\frac{1}{2}+\frac{1}{2^2}+\frac{1}{2^3}+\cdots = \sum_{n\in\mathbb{N}_2} \frac{1}{n}$$

$$\begin{aligned}\left(1+\frac{1}{2}+\frac{1}{2^2}+\frac{1}{2^3}+\cdots\right)\left(1+\frac{1}{3}+\frac{1}{3^2}+\frac{1}{3^3}+\cdots\right) &= 1+\frac{1}{2}+\frac{1}{3}+\frac{1}{2\cdot 3}+\frac{1}{2^2}+\frac{1}{3^2}+\frac{1}{2^2\cdot 3}+\frac{1}{2\cdot 3^2}+\cdots \\ &= \sum_{n\in\mathbb{N}_{2,3}} \frac{1}{n}\end{aligned}$$

[*] When $s=1$, neither side of the equality actually converges, but the explanation for why both sides should be equal is perhaps more transparent. Here "formally" is not to be confused with rigorously—we mean we can formally manipulate one side to get to the other.

$$\left(1+\frac{1}{2}+\frac{1}{2^2}+\frac{1}{2^3}+\cdots\right)\left(1+\frac{1}{3}+\frac{1}{3^2}+\frac{1}{3^3}+\cdots\right)\left(1+\frac{1}{5}+\frac{1}{5^2}+\frac{1}{5^3}+\cdots\right)=\sum_{n\in\mathbb{N}_{2,3,5}}\frac{1}{n}$$

$$\vdots$$

where $\mathbb{N}_{p_1,p_2,\ldots,p_k} = \left\{p_1^{e_1}p_2^{e_2}\cdots p_k^{e_k} : e_i \in \mathbb{N}\cup\{0\}\right\}$, i.e., $\mathbb{N}_{p_1,p_2,\ldots,p_k}$ is the set of natural numbers which only contain the primes $p_1, \ldots p_k$ in their prime decomposition.

We now prove rigorously that the formal product expansion for $\zeta(s)$ given above is valid for $s > 1$.

Definition 3.1.1. *Let $\{p\}$ denote the set of primes of $\mathbb{N}$. Let $a_p \in \mathbb{C}$ for each p. We define*

$$\prod_p a_p = \lim_{n\to\infty}\prod_{n<x} a_p.$$

Hence we will say $\prod a_p$ **converges (diverges)** *if the limit on the right does. We say $\prod a_p$* **converges absolutely** *if*

$$\lim_{n\to\infty}\prod_{i<n} a_{p_i}$$

converges for any ordering $\{p_1, p_2, p_3, \ldots\}$ of the set of primes $\{p\}$.

In other words, a product converges absolutely if it converges regardless of the way we order the terms in the product. One can of course similarly define infinite product over any denumerable index set

Example 3.1.2. *If some $a_p = 0$, then after some point (no matter how the p's are ordered), we will have a finite subproduct of $\prod a_p = 0$. Thus $\prod a_p$ will converge to 0 absolutely.*

Note that if every $a_p > 0$, then $\log(\prod a_p) = \sum \log a_p$. An immediate consequence is that $\prod a_p$ converges (absolutely) if and only if the series $\sum \log a_p$ converges (absolutely).

Proposition 3.1.3. *Let $(a_n)_{n=1}^{\infty}$ be a totally multiplicative sequence of complex numbers, i.e., $a_{mn} = a_m a_n$ for any $m, n \in \mathbb{N}$, and assume $a_1 = 1$. If $\sum a_n$ converges absolutely, then so does $\prod_p \frac{1}{1-a_p}$ and*

$$\sum_{n=1}^{\infty} a_n = \prod_p \frac{1}{1-a_p},$$

where the product is taken over all primes p of $\mathbb{N}$.

Proof. Suppose $\sum a_n$ converges absolutely. Let $\epsilon > 0$. Then for some $N \in \mathbb{N}$ we can say

$$\sum_{n>N} |a_n| < \epsilon.$$

Let $\{p_1, p_2, \ldots\}$ be any ordering of the set of primes of $\mathbb{N}$. Then there is some $K \in \mathbb{N}$ such that $\{p_1, \ldots, p_K\}$ contains all $\leq N$. Observe

$$\prod_{i=1}^{K}\frac{1}{1-a_{p_i}} = \prod_{i=1}^{K}\left(1+a_p+a_p^2+\cdots\right) = \sum_{n\in\mathbb{N}_{p_1,\ldots,p_K}} a_n.$$

Since $\mathbb{N}_{p_1,\ldots,p_K}$ contains $1,\ldots,N$, we have

$$\left|\sum_{n=1}^{\infty} a_n - \prod_{i=1}^{K} \frac{1}{1-a_{p_i}}\right| \leq \left|\sum_{n>N} a_n\right| \leq \sum_{n>N} |a_n| < \epsilon.$$

□

Corollary 3.1.4. *For any $s > 1$ the* **Euler product expansion**

$$\zeta(s) = \prod \frac{1}{1-p^{-s}} \tag{3.3}$$

is valid.

Proof. Apply the proposition with $a_n = n^{-s}$. □

The Euler product expansion demonstrates that the zeta function captures information about primes. In fact, it contains a surprising amount of information about primes. The simplest application of the zeta function to the study of primes is Euler's proof of the infinitude of primes.

Theorem 3.1.5. *There are infinitely many primes.*

Proof. Assume there are finitely many primes, $p_1, \ldots, p_k$. Then

$$\zeta(s) = \frac{1}{1-p_1^{-s}} \cdot \frac{1}{1-p_2^{-s}} \cdots \frac{1}{1-p_k^{-s}} \to \frac{1}{1-1/p_1} \cdot \frac{1}{1-1/p_2} \cdots \frac{1}{1-1/p_k} < \infty$$

as $s \to 1$. On the other hand

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

as $s \to 1$. Contradiction. □

Exercise 3.1. *For any integer $k > 1$, one can show $1/\zeta(k)$ represents the probability that k "randomly chosen" integers are coprime (have gcd 1). Let $f(x) = x$ on $[-\pi, \pi)$, compute the Fourier coefficients and apply Parseval's identity. Use this to compute $\zeta(2)$, and hence determine the probability that 2 randomly chosen integers are coprime. (Alternatively, you can try to derive the product expansion*

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

and look at the x^2 coefficient to find $\zeta(2)$.)

We will briefly discuss some deeper connections of $\zeta(s)$ to the study of primes, but first let us give a generalization of the Riemann zeta function.

Definition 3.1.6. *Let K be a number field. The* **Dedekind zeta function** *for K is*

$$\zeta_K(s) = \sum_{\mathfrak{a}} \frac{1}{N(\mathfrak{a})^s},$$

for $s > 1$ where $\mathfrak{a}$ runs over all (nonzero) ideals of $\mathcal{O}_K$.

As before, one can show this series indeed converges for all $s > 1$, and we have an Euler product expansion

$$\zeta_K(s) = \prod_{\mathfrak{p}} \frac{1}{1 - N(\mathfrak{p})^{-s}}$$

valid for $s > 1$ as above. Hence the Dedekind zeta function can be used to study the prime ideals of of K.

We remark that another way to write the above definition is

$$\zeta_K(s) = \sum \frac{a_n}{n^s}$$

where a_n denotes the number of ideals of K with norm n (convince yourself of this). Consequently, the Dedekind zeta function can be used to study the number of ideals of norm n. However, we will be interested in it for its applications to the class number h_K of K.

3.2 Interlude: Riemann's crazy ideas

Riemann published a single paper in number theory, *On the Number of Primes Less Than a Given Magnitude* in 1859, which was 8 pages long, contained no formal proofs, and essentially gave birth to all of analytic number theory. We will summarize the main ideas here.

We only defined the Riemann zeta function for real $s > 1$, but in fact Riemann considered it for complex values of s. In general if $a > 0$ and $z \in \mathbb{C}$, then one defines

$$a^z = e^{z \ln a}$$

where

$$e^z = \sum \frac{z^n}{n!}.$$

This allows one formally to make sense of the definition

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

for $s \in \mathbb{C}$, and one can show the sum actually converges provided $\mathrm{Re}(s) > 1$. Riemann showed that $\zeta(s)$ can be extended (uniquely) to a differentiable function on all of $\mathbb{C}$ except at $s = 1$, where $\zeta(s)$ has a *pole* (must be ∞). However the above series expression is only valid for $\mathrm{Re}(s) > 1$.

Riemann showed that $\zeta(s)$ has a certain symmetry around the line $\mathrm{Re}(s) = \frac{1}{2}$, namely one has the *functional equation*

$$\zeta(1-s) = \Gamma^*(s)\zeta(s)$$

where $\Gamma^*(s)$ is a function closely related to the Γ function. The functional equation says one can compute $\zeta(1-s)$ in terms of $\zeta(s)$, so we can indirectly use the series for $\zeta(s)$ to compute $\zeta(s)$ when $\mathrm{Re}(s) < 0$. The region $0 < \mathrm{Re}(s) < 1$ is called the *critical strip*, and the central line of symmetry $\mathrm{Re}(s) = \frac{1}{2}$ is called the *critical line.*

Let $\{\rho\}$ denote the set of zeroes of $\zeta(s)$ inside of the critical strip. There are countably (infinitely) many, and let us order them by their absolute value. Let

$$f(x) = \pi(x) + \frac{1}{2}\pi(x^{1/2}) + \frac{1}{3}\pi(x^{1/3}) + \cdots$$

where $\pi(x)$ is the number of primes less than x. Riemann discovered the following formula for $f(x)$

$$f(x) = \mathrm{Li}(x) - \sum_{\rho} \mathrm{Li}(x^{\rho}) - \log(2) + \int_x^{\infty} \frac{dt}{t(t^2-1)\ln t}$$

where $\mathrm{Li}(x) = \int_0^x \frac{dt}{\ln t}$. Hence this formula relates $\pi(x)$ with the (values of Li at the) zeroes of $\zeta(s)$. In fact, using Möbius inversion, one can rewrite $\pi(x)$ in terms of $f(x)$ (and therefore the zeroes of $\zeta(s)$) as

$$\pi(x) = f(x) - \frac{1}{2}f(x^{1/2}) - \frac{1}{3}(x^{1/3}) - \cdots$$

Essentially this says the following: if we know exactly where all the zeroes ρ of $\zeta(s)$ are we know exactly where the primes are (these are the places on the real line where $\pi(x)$ jumps).

Here is where Riemann made his famous conjecture, the *Riemann hypothesis*, that all the zeroes of $\zeta(s)$ lying in the critical strip actually lie on the critical line. (It is easy to see from series expansion that $\zeta(s) \neq 0$ for $\mathrm{Re}(s) \geq 1$. Then by the functional equation, $\zeta(1-s) = 0$ for $\mathrm{Re}(s) > 1$ if and only if $\Gamma^*(s) = 0$, which happens precisely for s a positive odd integer. Thus the only zeroes of $\zeta(s)$ outside of the critical strip, are the so-called *trivial zeroes* occurring when $s = -2k$, $k \in \mathbb{N}$.)

In 1896, Hadamard and de la Vallée Poussin used Riemann's ideas to prove the *prime number theorem*, that

$$\pi(x) \sim \mathrm{Li}(x) \sim \frac{x}{\ln x}$$

as $x \to \infty$. (This notation means $\pi(x)$ is approximately $\frac{x}{\ln x}$ for x large.) This is important, for example, in cryptography where one wants to know that the primes don't get too thinly spread out, so that large primes provide suitably secure keys for RSA. The Riemann hypothesis is equivalent to the "best possible bound" for the error term in the prime number theorem, precisely that

$$|\pi(x) - \mathrm{Li}(x)| < \frac{1}{8\pi}\sqrt{x}\ln(x)$$

for $x \geq 2657$. The Riemann hypothesis has natural generalizations to Dedekind zeta functions and L-functions (see below). Due to a host of applications, the generalized Riemann hypothesis is considered one of the most important open problems in mathematics.

3.3 Dirichlet L-functions

Let $m \in \mathbb{N}$. In 1837 (before Riemann!)*, Dirichlet introduced L-functions as a generalization of the Riemann zeta function in order to study the primes mod m. In particular, Dirichlet used these L-functions to show that there are infinitely many primes $\equiv a$ mod m, which will be one of the main results of this chapter. This result had been conjectured by Euler for $a = 1$ and by Legendre in general.

Let's start with the example of $p \equiv 1$ mod 4. (Last semester we were able to use a trick together with the first supplemental law of quadratic reciprocity to show there are infinitely many primes $p \equiv 1$ mod 4, and the case of $p \equiv 3$ mod 4 was an exercise using a different trick. Legendre tried to use quadratic reciprocity to treat the general case, but was unsuccessful, and as far as I know, there is no proof of the general case which does not use Dirichlet L-functions.)

*Riemann's zeta function was studied before Riemann as a function of natural numbers by Euler.

Knowing Euler's proof of the infinitude of primes, one might be tempted to try to define a series by

$$\prod_{p\equiv 1 \bmod 4} \frac{1}{1-p^{-s}}.$$

This expands out as a sum

$$\sum_{n\in\mathbb{N}_1} \frac{1}{n^s}$$

where $\mathbb{N}_1$ is the set of all natural numbers which only contain primes $\equiv 1 \bmod 4$ in their prime factorization. If the number of such primes is finite then the product expansion converges as $s \to 1$, and one would like to show the series expansion diverges to obtain a contradiction. However summing over $\mathbb{N}_1$ is not a natural thing to do and there is no simple way to directly analyze it. Hence we will have to be a little more subtle than this.

Dirichlet, being smarter than this, had the following idea using characters. Let's first recall a couple things about characters of finite abelian groups.

Definition 3.3.1. *Let G be a finite abelian group. A* **character** *(or* **1-dimensional representation***) of G is a group homomorphism into $\mathbb{C}^\times$. The set of characters of G is denoted by $\hat{G}$, and is called the* **dual** *of G.*

Exercise 3.2. *Let G be a finite abelian group. Let $\chi, \lambda \in \hat{G}$.*

(i) Show $\chi\lambda$, defined by $(\chi\lambda)(g) = \chi(g)\lambda(g)$, is also in $\hat{G}$.

(ii) Show $\overline{\chi}$, defined by $\overline{\chi}(g) = \overline{\chi(g)}$, is also in $\hat{G}$. (Here the bar denotes complex conjugation.)

(iii) For any $g \in G$, show $\chi(g)$ is a (not necessarily primitive) n-th root of unity[†] where n is the order of g in G.

(iv) Deduce that $\overline{\chi}\chi = \chi_0$, where χ_0 denotes the **trivial character***, i.e., $\chi_0(g) = 1$ for all $g \in G$.*

(v) Conclude that $\hat{G}$ is an abelian group.

Proposition 3.3.2. *Let G be finite abelian group. Then $\hat{G} \simeq G$.*

Proof. Let us first prove the proposition in the case $G = C_n$ (the cyclic group of order n). Let α be a generator of G. By (iii) of the exercise above, if $\chi \in \hat{G}$, then $\chi(\alpha)$ must be an n-th root of unity ζ. Furthermore any n-th root of unity ζ defines (uniquely) a character on G by setting $\chi(\alpha^k) = \zeta^k$. (Observe this is character, and that nothing else can be.) In other words, a character is determined by what it does to a generator of G.

Let ζ_n denote a primitive n-th root of unity (take $\zeta_n = e^{2\pi i/n}$ if you wish). There are n n-th roots of unity, given by $\zeta_n^0 = 1, \zeta_n, \zeta_n^2, \dots, \zeta_n^{n-1}$. Hence there are precisely n distinct character of G, $\chi_0, \chi_1, \chi_2, \dots, \chi_{n-1}$ given by $\chi_i(\alpha) = \zeta_n^i$. It is obvious then that $(\chi_i\chi_j)\alpha = \zeta_n^{i+j}$, i.e., $\chi_i\chi_j = \chi_{i+j \bmod n}$. Hence $\hat{G} \simeq \mathbb{Z}/n\mathbb{Z} \simeq C_n$.

Now that we have prove the proposition in the case where G is cyclic, let us assume G is an arbitrary finite abelian group. Then we know by the classification of such groups, $G \simeq \prod C_{n_i}^{r_i}$. Set $\zeta_{n_i} = e^{2\pi i/n_i}$ and let $\alpha_i \in G$ be a generator for C_{n_i}. As above, we can define characters χ_{ij} on C_{n_i} by $\chi_{ij}(\alpha_i) = \zeta_{n_i}^j$. We can extend each χ_{ij} to a character on G by setting $\chi_{ij}(\alpha_k) = 1$ whenever $i \neq k$ and using multiplicativity. In other words, view and $g \in G$ as

$$g = (\alpha_1^{e_1}, \alpha_2^{e_2}, \dots, \alpha_t^{e_t})$$

[†] $\zeta \in \mathbb{C}$ is a *primitive* n-th root of unity if $\zeta^n = 1$ but $\zeta^d \neq 1$ for any $d|n$. For example there are four 4-th roots of unity: $\pm 1, \pm i$, but only $\pm i$ are primitive.

and set

$$\chi_{ij}(g) = \zeta_{n_i}^{je_i}.$$

Hence for each i, we get disjoint subgroups of $\hat{G}$ each isomorphic to $\hat{C}_{n_i} \simeq C_{n_i}$. It is straightforward to check that any character of G is a product of some χ_{ij}'s, i.e., we have

$$\hat{G} \simeq \prod \hat{C}_{n_i} \simeq \prod C_{n_i} \simeq G.$$

□

We will be interested in the case where $G = (\mathbb{Z}/m\mathbb{Z})^\times$.

Example 3.3.3. *Suppose $G = (\mathbb{Z}/2\mathbb{Z})^\times$. Then $G = \{1\} = C_1$ so $\hat{G} = \{\chi_0\}$. In other words, the only character of G is the trivial one χ_0 which sends 1 to 1.*

Example 3.3.4. *Suppose $G = (\mathbb{Z}/3\mathbb{Z})^\times$. Then $G = \{1, 2\} \simeq C_2$ (here 1 and 2 represent the corresponding congruence classes in $\mathbb{Z}/3\mathbb{Z}$). Since 2 generates G and has order 2, there are two possibilities for characters:*

$$\chi_0 : 1 \mapsto 1, \;\; 2 \mapsto 1$$

and

$$\chi_1 : 1 \mapsto 1, \;\; 2 \mapsto -1$$

So $\hat{G} = \{\chi_0, \chi_1\} \simeq C_2$.

Note that $G = (\mathbb{Z}/4\mathbb{Z})^\times \simeq C_2$ also, so this case is essentially the same as $(\mathbb{Z}/3\mathbb{Z})^\times$.

Example 3.3.5. *Suppose $G = (\mathbb{Z}/5\mathbb{Z})^\times$. Then $G = \{1, 2, 3, 4\} \simeq C_4$. Here 2 generates G, so a character of G is determined by what 4-th root of unity 2 maps to. Explicitly, we have 4 characters, whose values are read off of the following* **character table***:*

	1	2	3	4
χ_0	1	1	1	1
χ_1	1	-1	-1	1
χ_2	1	i	$-i$	-1
χ_3	1	$-i$	i	-1

Looking at this table, it is easy to see $\hat{G}$ is cyclic of order four, generated either by χ_2 or χ_3.

Exercise 3.3. *Determine all characters of $(\mathbb{Z}/15\mathbb{Z})^\times$ and $(\mathbb{Z}/16\mathbb{Z})^\times$. (Write them down in character tables like our $(\mathbb{Z}/5\mathbb{Z})^\times$ case. Feel free to order the columns and rows however you find easiest.)*

Theorem 3.3.6. *Let $\mathbb{C}G = \{f : G \to \mathbb{C}\}$ denote the space of complex valued functions from a finite abelian group G into $\mathbb{C}$ (the* **group algebra** *of G). This is a $|G|$-dimensional vector space over $\mathbb{C}$. The characters $\chi \in \hat{G}$ form a $\mathbb{C}$-basis for $\mathbb{C}G$.*

We omit the proof, and in fact we do not need this precise result, but it is helpful for motivation. What this means for us is the following. We want to study the primes p in a congruence class mod m. As long as $p \nmid m$, we have $p \equiv a \bmod m$ for some $a \in (\mathbb{Z}/m\mathbb{Z})^\times$. Let $\{\chi\}$ denote the the set of characters of $G = (\mathbb{Z}/m\mathbb{Z})^\times$ and fix $a \in (\mathbb{Z}/m\mathbb{Z})^\times$. Consider the function $f \in \mathbb{C}G$ given by

$$f : (\mathbb{Z}/m\mathbb{Z})^\times \to \mathbb{C}$$

$$f(x) = \begin{cases} 1 & x = a \\ 0 & \text{else.} \end{cases}$$

In other words, f tells me if p mod m is a or not. What the above says is that

$$f(x) = \sum c_\chi \chi(x)$$

for some $c_\chi \in \mathbb{C}$. Hence knowing $\chi(p \bmod m)$ for all χ tells us whether $p \equiv a \bmod m$ or not. Slightly more generally, the above theorem says knowing $\chi(p \bmod m)$ for all χ and $p \nmid m$ is the same as knowing p mod m for all $p \nmid m$. Put another way, the characters of $(\mathbb{Z}/m\mathbb{Z})^\times$ distinguish all the (invertible) congruence classes mod m. This suggest it is possible to study the primes $\equiv a$ mod m using the characters of $(\mathbb{Z}/m\mathbb{Z})^\times$. To somehow connect this with something like the zeta function, one actually wants to think of the characters of $(\mathbb{Z}/m\mathbb{Z})^\times$ as "characters" of $\mathbb{Z}$.

Definition 3.3.7. *Let χ be a character of $(\mathbb{Z}/m\mathbb{Z})^\times$. We extend χ to a function*

$$\chi : \mathbb{Z} \to \mathbb{C}$$

by

$$\chi(a) = \begin{cases} \chi(a \bmod m) & \textit{if } a \textit{ is invertible} \bmod m \\ 0 & \textit{else.} \end{cases}$$

The resulting function $\chi : \mathbb{Z} \to \mathbb{C}$ is called a **Dirichlet character** mod *m.*

Note that the values of χ only depend upon congruence classes mod m. Then our remarks before the definition can be rephrased as follows: knowing the value of $\chi(p)$ for every Dirichlet character χ mod m and every $p \nmid m$ is equivalent to knowing the value of p mod m for every $p \nmid m$.

Example 3.3.8. *Consider χ_2 from our earlier $(\mathbb{Z}/5\mathbb{Z})^\times$ example. Then the corresponding Dirichlet character is*

$$\chi_2(a) = \begin{cases} 0 & a \equiv 0 \bmod 5 \\ 1 & a \equiv 1 \bmod 5 \\ i & a \equiv 2 \bmod 5 \\ -i & a \equiv 3 \bmod 5 \\ -1 & a \equiv 4 \bmod 5. \end{cases}$$

Exercise 3.4. *Let χ be a Dirichlet character mod m. Then $\chi(ab) = \chi(a)\chi(b)$ for any $a, b \in \mathbb{Z}$.*

Definition 3.3.9. *Let χ be a Dirichlet character mod m. Then* **Dirichlet L-function** *(or* **L-series***) for χ is given by*

$$L(s, \chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}. \tag{3.4}$$

for $s > 1$.

Note that since $|\chi(n)| \leq 1$ for all n (see Exercise 3.2(iii)), we can say

$$|L(s, \chi)| \leq \sum \frac{1}{n^s} = \zeta(s) \ \ (s > 1),$$

hence the series defining $L(s, \chi)$ converges absolutely for $s > 1$. Because χ is multiplicative by the exercise above, the same expansion trick we used for $\zeta(s)$ above works to give a product expansion

$$L(s, \chi) = \prod_p \frac{1}{1 - \chi(p)p^{-s}}, \tag{3.5}$$

which again is valid for $s > 1$. (Just apply Proposition 3.1.3 with $a_n = \frac{\chi(n)}{n^s}$.)

Just like $\zeta(1) = \infty$ told us there were infinitely many primes, the most important value of $L(s, \chi)$ for primes in arithmetic progressions is $L(1, \chi)$ (which is not typically ∞—it will be if and only if $\chi = \chi_0$ is the trivial character mod m).

Example 3.3.10. *To return to the case of studying primes* $\equiv 1 \bmod 4$*, look at the Dirichlet characters mod* 4*. Since* $(\mathbb{Z}/4\mathbb{Z})^\times \simeq C_2$*, there are only two Dirichlet characters mod* 4*. We can write them as*

$$\chi_0(a) = \begin{cases} 0 & a \equiv 0, 2 \bmod 4 \\ 1 & a \equiv 1, 3 \bmod 4 \end{cases}$$

(the trivial character) and

$$\chi_1(a) = \begin{cases} 0 & a \equiv 0, 2 \bmod 4 \\ 1 & a \equiv 1 \bmod 4 \\ -1 & a \equiv 3 \bmod 4 \end{cases}$$

(the nontrivial character). Then, for $s > 1$*, we have*

$$L(s, \chi_0) = \sum_{n \text{ odd}} \frac{1}{n^s} = \prod_{p \neq 2} \frac{1}{1 - p^{-s}}$$

and

$$L(s, \chi_1) = \sum_{n \equiv 1 \bmod 4} \frac{1}{n^s} - \sum_{n \equiv 3 \bmod 4} \frac{1}{n_s} = \prod_{p \equiv 1 \bmod 4} \frac{1}{1 - p^{-s}} \cdot \prod_{p \equiv 3 \bmod 4} \frac{1}{1 + p^{-s}}.$$

Hence these L-functions are not too far from our original naive suggestion, and they are not too difficulty (though not trivial) to analyze. We also note that $L(s, \chi_0)$ *is essentially* $\zeta(s)$*—it is off by a single factor*

$$\zeta(s) = \frac{1}{1 - 2^{-s}} L(s, \chi_0),$$

so $L(1, \chi_0) = (1 - 2^{-1})\zeta(1) = \infty$*.*

From the series expansion of $L(s, \chi_1)$ *it's not to hard to see that*

$$L(1, \chi_1) = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \cdots = \frac{\pi}{4}$$

(The latter equality, known as Leibnitz's formula, follows from let $x \to 1$ *in the Taylor series for* $\arctan(x)$*.) Now we can use this to show there are infinitely many primes* $\equiv 1 \bmod 4$ *and* $\equiv 3 \bmod 4$*.*

Suppose there were finitely many primes $p \equiv 3 \bmod 4$*. Then the above formula for* $L(s, \chi_1)$ *shows*

$$\zeta(s)/L(s, \chi_1) = \frac{1}{1 - 2^{-s}} \prod_{p \equiv 3 \bmod 4} \frac{1 + p^{-s}}{1 - p^{-s}}$$

for $s > 1$. But then letting $s \to 1$ yields

$$\frac{1}{1-\frac{1}{2}} \cdot \prod_{p \equiv 3 \bmod 4} \frac{1+\frac{1}{p}}{1-\frac{1}{p}} = \zeta(1) \cdot \frac{4}{\pi} = \infty,$$

which is a contradiction since the left hand side must be finite if there are only finitely many $p \equiv 3 \bmod 4$.

Similarly suppose there were finitely many primes $p \equiv 1 \bmod 4$. Then

$$\begin{aligned}
\zeta(s)L(s,\chi_1) &= \frac{1}{1-2^{-s}} \left(\prod_{p \equiv 1 \bmod 4} \frac{1}{1-p^{-s}} \right)^2 \prod_{p \equiv 3 \bmod 4} \frac{1}{1-p^{-2s}} \\
&= \frac{1}{1-2^{-s}} \left(\prod_{p \equiv 1 \bmod 4} \frac{1}{1-p^{-s}} \right)^2 \left(\zeta(2s)(1-2^{-2s}) \prod_{p \equiv 1 \bmod 4} (1-p^{-2s}) \right) \\
&= (1+2^{-s}) \prod_{p \equiv 1 \bmod 4} \frac{1-p^{-2s}}{(1-p^{-s})^2} \cdot \zeta(2s).
\end{aligned}$$

for $s > 1$. Letting $s \to 1$, we see the right hand side is finite, since $\zeta(2) = \frac{\pi^2}{6} < \infty$ and the product has only finitely many terms, but the left hand side approaches $\zeta(1)/L(s,\chi_1) = \frac{4}{\pi}\infty = \infty$, a contradiction.

Hence just knowing that $0 < |L(1,\chi_1)| < \infty$ allows us to conclude there are infinitely many primes $p \equiv 1 \bmod 4$ and infinitely many $p \equiv 3 \bmod 4$.

Theorem 3.3.11. (Dirichlet's theorem on arithmetic progressions) *Suppose* $\gcd(a,m) = 1$. *Then there are infinitely many primes $p \equiv a \bmod m$.*

Proof. Let $\{\chi\}$ be the set of Dirichlet characters mod m. We consider the complex logarithm defined by

$$\log(1+z) = z - \frac{z^2}{2} + \frac{z^3}{3} - \cdots .$$

for $|z| < 1$. Then for any χ, we have

$$\left| \log\left(\frac{1}{1-\chi(p)p^{-s}} \right) \right| = |\log(1-\chi(p)p^{-s})| = \left| \frac{\chi(p)}{p^s} + \frac{\chi(p^2)}{2p^{2s}} + \frac{\chi(p^3)}{3p^{3s}} + \cdots \right| .$$

It follows from the Taylor expansion of $\log(1+x)$ that

$$\log\left(\frac{1}{1-\chi(p)p^{-s}} \right) = -\log(1-\chi(p)p^{-s}) = \frac{\chi(p)}{p^s} + \epsilon_p(s)$$

where $\epsilon_p(s)$ is an error term satisfying $|\epsilon_p(s)| < \frac{1}{p^{2s}}$. Hence

$$\log L(s,\chi) = \sum_p \frac{\chi(p)}{p^s} + \epsilon_\chi(s)$$

where $\epsilon_\chi(s) = \sum_p \epsilon_p(s)$ so $|\epsilon_\chi(s)| < \sum \frac{1}{p^{2s}}$.

Now consider the sum

$$\sum_{\chi} \chi^{-1}(a) \log L(s,\chi) = \sum_{\chi,p} \frac{\chi^{-1}(a)\chi(p)}{p^s} + \epsilon(s) \tag{3.6}$$

where $\epsilon(s) = \sum_\chi \chi^{-1}(a)\epsilon_\chi(s)$ is an error term. Now we appeal to a fundamental result from representation theory (see exercise below), the orthogonality relation for characters. This is essentially a refinement of Theorem 3.3.6, saying that the characters $\hat{G}$ of a group G form an *orthogonal* basis for $\mathbb{C}G$. In our case, the orthogonality relation says

$$\sum_{\chi} \chi^{-1}(a)\chi(p) = \begin{cases} \phi(m) & a \equiv p \bmod m \\ 0 & \text{else.} \end{cases}$$

(Here $\phi(m) = |(\mathbb{Z}/m\mathbb{Z})^\times|$.) Hence by first summing over χ, we see the only p that contribute are those $\equiv a$ mod m. In other words

$$\sum_{\chi} \chi^{-1}(a) \log L(s,\chi) = \phi(m) \sum_{p \equiv a \bmod m} \frac{\chi(p)}{p^s} + \epsilon(s). \tag{3.7}$$

It is straightforward to check that the error term $\epsilon(s)$ remains bounded as $s \to 1$ (exercise below). Thus if there are only finitely many primes $p \equiv a$ mod m, then the right hand side converges as $s \to 1$. In other words, it suffices to show the left hand diverges when $s \to 1$.

Consider the trivial character χ_0. Then

$$L(s,\chi_0) = \prod_{p \nmid m} \frac{1}{1-p^{-s}} = \prod_{p \mid m} (1-p^{-s}) \cdot \zeta(s),$$

hence $L(s,\chi_0) \to \prod_{p|m}(1-p^{-1}) \cdot \zeta(1) = \infty$ as $s \to 1$. Thus $\log L(s,\chi_0) \to \infty$ as $s \to 1$. This means the sum (3.6) must tend to ∞ as $s \to 1$ *provided* no single term $\chi^{-1}(a) \log L(s,\chi) \to -\infty$ as $s \to 1$. This follows from the fact that $L(s,\chi) \neq 0$, which is the content of Proposition 3.4.7 in the next section. □

The fact that $L(s,\chi) \neq 0$ follows from *Dirichlet's class number formula*, which is itself of great interest. Historically, Dirichlet proved his class number formula, and used this to prove his theorem on arithmetic progressions, though now there are other proofs that $L(1,\chi) \neq 0$. We will follow Dirichlet's approach (at least the presentation in [Cohn]) and prove the class number formula, and use this to conclude the proof of Theorem 3.3.11 in the next section.

Exercise 3.5. *Let $G = \mathbb{Z}/n\mathbb{Z}$. Let $a, b \in G$. Show*

$$\sum_{\chi \in \hat{G}} \chi^{-1}(a)\chi(b) = \begin{cases} |G| & a = b \\ 0 & \textit{else.} \end{cases}$$

(Hint: use the fact that we know explicitly what the characters are as in the proof of Proposition 3.3.2.) This proves the orthogonality relation we used above in the case of cyclic groups.

Exercise 3.6. *Check that the error term $\epsilon(s)$ appearing in the proof is bounded as $s \to 1$.*

3.4 The class number formula

The class number formula will fall out of analysis of the Dedekind zeta function. Let $K = \mathbb{Q}(\sqrt{d})$ where $d \neq 1$ is squarefree. Let $\Delta = \Delta_K$. Recall that for $s > 1$,

$$\zeta_K(s) = \prod_{\mathfrak{p}} \frac{1}{1 - N(\mathfrak{p})^{-s}}.$$

If $\mathfrak{p}|p$ where p is a prime of $\mathbb{N}$, then $N(\mathfrak{p}) = p$ if p is split or ramified and $N(\mathfrak{p}) = p^2$ if p is inert. If p splits in K there are 2 prime ideals $\mathfrak{p}$ lying above p, and otherwise there is just 1. So we can rewrite

$$\begin{aligned}
\zeta_K(s) &= \prod_{p \text{ ramified}} \frac{1}{1-p^{-s}} \prod_{p \text{ split}} \left(\frac{1}{1-p^{-s}}\right)^2 \prod_{p \text{ inert}} \frac{1}{1-p^{-2s}} \\
&= \prod_{p \text{ ramified}} \frac{1}{1-p^{-s}} \prod_{p \text{ split}} \left(\frac{1}{1-p^{-s}}\right)^2 \prod_{p \text{ inert}} \left(\frac{1}{1-p^{-s}}\right)\left(\frac{1}{1+p^{-s}}\right) \\
&= \prod_{p} \frac{1}{1-p^-s} \prod_{p \text{ split}} \frac{1}{1-p^{-s}} \prod_{p \text{ inert}} \frac{1}{1+p^{-s}} \\
&= \zeta(s) \prod_{p \text{ split}} \frac{1}{1-p^{-s}} \prod_{p \text{ inert}} \frac{1}{1+p^{-s}}.
\end{aligned}$$

Note that the two products on the right, look like one of Dirichlet's L-functions. In fact, we know p splits in K if and only if $\left(\frac{\Delta}{p}\right) = 1$, and p is inert in K if and only if $\left(\frac{\Delta}{p}\right) = -1$. Hence if we extend the Kronecker symbol $\left(\frac{\Delta}{p}\right)$ to a function from $\mathbb{Z} \to \mathbb{C}$ by

$$\chi_\Delta(n) = \begin{cases} 0 & \gcd(n, \Delta) > 1 \\ \left(\frac{\Delta}{p_1}\right)^{e_1}\left(\frac{\Delta}{p_2}\right)^{e_2} \cdots \left(\frac{\Delta}{p_k}\right)^{e_k} & n > 0, n = p_1^{e_1} p_2^{e_2} \ldots p_k^{e_k} \text{ and } \gcd(n, \Delta) = 1 \\ \chi_\Delta(-n)\chi_\Delta(|\Delta| - 1) & n < 0. \end{cases}$$

One easily checks that χ_Δ is a totally multiplicative function on $\mathbb{Z}$ which depends only upon congruences classes mod Δ. Restricting to $(\mathbb{Z}/\Delta\mathbb{Z})^\times$ gives only nonzero values, so we get a (group) character of $(\mathbb{Z}/\Delta\mathbb{Z})^\times$. In other words, χ_Δ is a Dirichlet character mod Δ, and one has (for $s > 1$)

$$L(s, \chi_\Delta) = \prod_{\chi_\Delta(p) = \left(\frac{\Delta}{p}\right) = 0} 1 \cdot \prod_{\chi_\Delta(p) = \left(\frac{\Delta}{p}\right) = 1} \frac{1}{1-p^{-s}} \cdot \prod_{\chi_\Delta(p) = \left(\frac{\Delta}{p}\right) = -1} \frac{1}{1+p^{-s}}.$$

In other words, we have shown

Lemma 3.4.1. *For $s > 1$,*

$$\zeta_K(s) = \zeta(s) L(s, \chi_\Delta).$$

Theorem 3.4.2. (Dirichlet's class number formula for imaginary quadratic fields) *Let $K = \mathbb{Q}(\sqrt{d})$ where $d < 0$ is squarefree. Then*

$$L(1, \chi_\Delta) = \lim_{s \to 1} \frac{\zeta_K(s)}{\zeta(s)} = \frac{2\pi h_K}{w\sqrt{-\Delta}},$$

where w is the number of roots of unity in $\mathcal{O}_K$, i.e., $w = 6$ if $d = -3$, $w = 4$ if $d = -1$ and $w = 2$ otherwise.

Theorem 3.4.3. (Dirichlet's class number formula for real quadratic fields) *Let $K = \mathbb{Q}(\sqrt{d})$ where $d > 1$ is squarefree. Then*

$$L(1, \chi_\Delta) = \lim_{s \to 1} \frac{\zeta_K(s)}{\zeta(s)} = \frac{2 \log \eta \, h_K}{\sqrt{\Delta}},$$

where η is the fundamental unit in $\mathcal{O}_K$.

(For us, the fundamental unit η of $\mathcal{O}_K$ where K is real quadratic is the unique unit $\eta > 1$ such that any unit of $\mathcal{O}_K$ is of the form $\pm\eta^m$ for some $m \in \mathbb{Z}$.)

The quantity $\lim_{s \to 1} \frac{\zeta_K(s)}{\zeta(s)}$ is called the *residue of* $\zeta_K(s)$ at $s = 1$. The idea is that while $\zeta_K(s)$ and $\zeta(s)$ both have a simple pole at $s = 1$*, these poles should cancel in the quotient to give us a finite number that tells us about the arithmetic of K. Indeed, the quotient $\zeta_K(s)/\zeta(s) = L(1, \chi_\Delta)$ is (or can be continued to) a well-defined continuous function at $s = 1$.

Example 3.4.4. *Let $\Delta = -4$. The quadratic field of discriminant $\Delta = -4$ is $K = \mathbb{Q}(i)$, and Dirichlet's class number formula says*

$$L(1, \chi_{-4}) = \frac{2\pi h_K}{4\sqrt{4}} = \frac{\pi h_K}{4}.$$

But we saw last section, that

$$L(1, \chi_{-4}) = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots = \frac{\pi}{4},$$

hence the class number formula provides another proof that $h_K = 1$. Or if one wishes to approach things from the opposite direction, we see that the fact that $h_K = 1$ (which we proved in two other ways before: by showing $\mathbb{Z}[i]$ is a Euclidean domain and by using Minkowski's bound) implies Leibnitz's formula for $\pi = 4(1 - 1/3 + 1/5 - \cdots)$.

Exercise 3.7. *Using the fact that $\mathbb{Q}(\sqrt{-3})$ has class number 1, determine value of the series*

$$1 - \frac{1}{2} + \frac{1}{4} - \frac{1}{5} + \frac{1}{7} - \frac{1}{8} + \cdots$$

from the class number formula.

Exercise 3.8. *Analogous to the previous exercise, explicate the class number formula in the cases $\mathbb{Q}(\sqrt{-5})$ and $\mathbb{Q}(\sqrt{2})$. (I.e., write down explicitly what the series $L(1, \chi_\Delta)$ is and determine its value from the class number formula.)*

Now you may think above exercises, while interesting, are sort of the opposite of what we want. Instead of using h_K to determine $L(1, \chi_\Delta)$, we would prefer to use $L(1, \chi_\Delta)$ to determine the class number, as we gave an alternate proof for $h_{\mathbb{Q}(i)} = 1$ in the example above. Thus we will need some way of evaluating the series $L(1, \chi_\Delta) = \sum_{n \geq 1} \frac{\chi_\Delta(n)}{n}$ (which converges conditionally).

In fact, one can write down a finite expression for $L(1, \chi_\Delta)$ to actually compute class numbers and we will discuss this later. First we want to prove the class number formula. This immediately implies $L(1, \chi_\Delta) \neq 0$ for any Δ. From this one can deduce that $L(1, \chi) \neq 0$ for any Dirichlet character χ, which will complete the proof of Dirichlet's theorem on primes in arithmetic progressions.

The proof of the class number formula relies on simple geometric lattice point counting problems.

* This means that they go to infinity at the "same rate" as $\frac{1}{1-s}$.

Lemma 3.4.5. (Gauss) *Suppose A, B, C are integers with $B^2 - 4AC < 0$. The number $\lambda(T)$ of lattice points (points in $\mathbb{Z}^2$) contained in the solid ellipse*

$$Ax^2 + Bxy + Cy^2 \leq T$$

satisfies

$$\lambda(T) = \frac{2\pi T}{\sqrt{4AC - B^2}} + O(\sqrt{T})$$

Recall the big-O notation means the following: for $f, g : \mathbb{R} \to \mathbb{R}$, we say

$$f(x) = O(g(x))$$

if $|f(x)| \leq cg(x)$ for all $x > X$ where X and c are some constants. Another way to say this is that $f(x) = O(g(x))$ means $\limsup_{x\to\infty} \frac{|f(x)|}{g(x)} < \infty$ (assuming g is positive).

Proof. First we observe the area of the ellipse is $\frac{2\pi T}{\sqrt{4AC-B^2}}$. To see this, note that there is a change of variables

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

for some θ such that $Ax^2+Bxy+Cy^2 = A'x'^2+C'y'^2$ for some A', C'. (In other words, we just rotate the ellipse so its major and minor axes are on the x' and y' axes.) However since the determinant of this transformation (a rotation matrix) is 1, the determinant B^2-4AC of the form $Ax^2+Bxy+Cy^2$ equals the determinant $-4A'C'$ of the form $A'x'^2 + C'y'^2$. (Just check this explicitly.) Since the major and minor axes of the ellipse must have length $2\sqrt{T/A'}$ and $2\sqrt{T/B'}$, we know the area of the ellipse is $A(T) = \frac{\pi T}{\sqrt{A'C'}} = \frac{2\pi T}{\sqrt{4AC-B^2}}$.

Now we can tile $\mathbb{R}^2$ with squares, with each square having area one and centered about some lattice point $(x, y) \in \mathbb{Z}^2$. Let $m(T)$ be the number of squares completely contained in our ellipse $Ax^2 + Bxy + Cy^2 \leq T$ and $M(T)$ be the smallest number of squares which completely contain the ellipse. Since $\lambda(T)$ is the number of squares whose center is contained in the ellipse, we clearly have

$$m(T) \leq \lambda(T) \leq M(T).$$

The point is that $m(T)$ and $M(T)$ are both roughly the area $A(T)$ of the ellipse. Precisely, let $n(T)$ denote the number of squares which intersect the boundary $Ax^2 + Bxy + Cy^2 = T$ of the ellipse. Note that $m(T) \geq A(T) - n(T)$ and $M(T) \leq A(T) + n(T)$. (Draw a picture.) Thus

$$A(T) - n(T) \leq \lambda(T) \leq A(T) + n(T)$$

for all n. So it suffices to show that

$$n(T) = O(\sqrt{T}).$$

Interchanging x and y if necessary, we may assume that $A \leq C$, so the ellipse is wider in the x-direction than it is tall in the y-direction. Then $n(T) \leq 8 + 8\sqrt{T/A} = O(\sqrt{T})$ by the exercise below. □

Exercise 3.9. *Show there are 4 points of slope ±1 on the ellipse $Ax^2+Bxy+Cy^2 = T$ with $A \leq C$. This breaks the ellipse up into 4 arcs. Show each arc intersects at most $2 + 2\sqrt{T/A}$ squares in the tiling above by considering projections onto the x- and y- axes.*

To put the proof in simpler terms, the number of lattice points in an ellipse is essentially the area of the ellipse, with some error term which is essentially determined by the arclength. If the area increases like T, then the arclength will increase like $\sqrt{T}$, hence the $O(\sqrt{T})$ bound on the error.

Let $F(T)$ denote the number of (nonzero) ideals $\mathcal{I}$ of $\mathcal{O}_K$ with norm $N(\mathcal{I}) \leq T$. We know $F(T)$ is finite for any T.

Lemma 3.4.6. *We have*

$$F(T) = \kappa h_K T + O(\sqrt{T})$$

where

$$\kappa = \begin{cases} \frac{2\pi}{w\sqrt{-\Delta}} & \Delta < 0 \\ \frac{2\log \eta}{\sqrt{\Delta}} & \Delta > 0, \end{cases}$$

using the notation above.

The number κ is called Dirichlet's structure constant. So in this notation Dirichlet's class number formula reads (in both cases) $L(1, \chi_\Delta) = \kappa h_K$.

Proof. For a nonzero ideal $\mathcal{I}$ of $\mathcal{O}_K$, set $G(\mathcal{I}, T) = \{(\alpha) \subseteq \mathcal{I} : 0 < |N(\alpha)| \leq T\}$, i.e., $G(\mathcal{I}, T)$ is the number of principal ideals contained in $\mathcal{I}$ of (absolute) norm at most T. Suppose $\mathcal{J}$ is an ideal of $\mathcal{O}_K$ equivalent to $\mathcal{I}^{-1}$. Then $\mathcal{I}\mathcal{J} = (\alpha)$ is a principal ideal contained in $\mathcal{I}$. Conversely, any principal ideal $(\alpha) \subseteq \mathcal{I}$ is of this form $\alpha = \mathcal{I}\mathcal{J}$ for $\mathcal{J} \sim \mathcal{I}^{-1}$ and we have

$$N(\mathcal{J}) \leq T \iff |N(\alpha)| = N(\mathcal{J}\mathcal{I}) \leq TN(\mathcal{I}).$$

Hence $G(\mathcal{I}, TN(\mathcal{I}))$ is the number of ideals of norm $\leq T$ which are equivalent to $\mathcal{I}^{-1}$. Thus we may write

$$F(T) = G(\mathcal{I}_1, TN(\mathcal{I}_1)) + G(\mathcal{I}_2, TN(\mathcal{I}_2)) + \cdots + G(\mathcal{I}_h, TN(\mathcal{I}_h))$$

where $\mathcal{I}_1, \ldots, \mathcal{I}_h$ are a set of ideal representatives for the class group $\mathcal{C}l_K$. Consequently, the lemma follows if we can show

$$G(I, TN(\mathcal{I})) = \kappa T + O(\sqrt{T})$$

for any ideal $\mathcal{I}$ of $\mathcal{O}_K$.

Suppose $\Delta < 0$. Let β_1, β_2 be a $\mathbb{Z}$-basis for $\mathcal{I}$. Write $\alpha = \beta_1 x + \beta_2 y$. Then

$$N(\alpha) = \alpha\overline{\alpha} = Ax^2 + Bxy + Cy^2,$$

where $A = N(\beta_1)$, $B = Tr(\beta_1\overline{\beta}_2)$ and $C = N(\beta_2)$. Hence the number of α with norm $\leq TN(\mathcal{I})$ is the number of lattice points (points of $\mathbb{Z}^2$) contained inside the solid ellipse $Ax^2 + Bxy + Cy^2 \leq TN(\mathcal{I})$. Note α and α' generate the same principal ideal if and only if they differ by units. Hence $G(\mathcal{I}, TN(\mathcal{I}))$ is $\frac{1}{w}$ times the number of nonzero lattice points inside the ellipse $Ax^2 + Bxy + Cy^2 \leq TN(\mathcal{I})$. This ellipse has discriminant $B^2 - 4AC = \Delta[\beta_1, \beta_2] = \Delta_K N(\mathcal{I})^2$, so the desired estimate of $G(I, TN(\mathcal{I}))$ follows from Gauss's lemma above.

Suppose $\Delta > 0$. The idea is basically the same. If β_1, β_2 is a $\mathbb{Z}$-basis for $\mathcal{I}$ and $\alpha = \beta_1 x + \beta_2 y$ then

$$|N(\alpha)| = |Ax^2 + Bxy + Cy^2|$$

where $A = N(\beta_1)$, $B = Tr(\beta_1\overline{\beta}_2)$ and $C = N(\beta_2)$. However there will be infinitely many solutions to $|N(\alpha)| \leq TN(\mathcal{I})$ owing to the infinitude of units. But there is a one-to-one correspondence of elements $\alpha \in \mathcal{I}$ satisfying $1 \le |\alpha/\overline{\alpha}| < \eta^2$, $\alpha > 0$ and principal subideals (α) of $\mathcal{I}$. Hence we may write $G(\mathcal{I}, TN(\mathcal{I}))$ as the number of nonzero solutions to the following system of equations:

$$-T \le Ax^2 + Bxy + Cy^2 \le T,$$

$$1 \le |\frac{\beta_1 x + \beta_2 y}{\overline{\beta}_1 x + \overline{\beta}_2 y}| < \eta^2,$$

$$\beta_1 x + \beta_2 y > 0.$$

Since the discriminant $B^2 - 4AC$ of the quadratic form $Ax^2 + Bxy + Cy^2$ is positive, one gets not an ellipse but a hyperbolic region from the first equation. The latter two equations leave us with two finite hyperbolic sectors to count lattice points in. By setting up an appropriate integral one can show the area is $\frac{2\log\eta\, T}{\sqrt{\Delta}}$. For more details, see [Cohn]. □

Now we can prove the class number formula.

Proof. We want to show $L(s, \chi_\Delta) = \kappa h_K$. Recall we can write $\zeta_K(s) = \sum \frac{a_n}{n^s}$ where a_n is the number of ideals of norm n. Hence

$$\begin{aligned}\zeta_K(s) &= \frac{F(1)}{1^s} + \frac{F(2) - F(1)}{2^s} + \frac{F(3) - F(2)}{3^s} + \cdots \\ &= F(1)\left(\frac{1}{1^s} - \frac{1}{2^s}\right) + F(2)\left(\frac{1}{2^s} - \frac{1}{3^s}\right) + \cdots \\ &= \sum_{T=1}^{\infty} F(T)\left\{\frac{1}{T^s} - \frac{1}{(T+1)^s}\right\}.\end{aligned}$$

For a fixed T, we have

$$\begin{aligned}\frac{1}{T^s} - \frac{1}{(T+1)^s} &= \frac{1}{T^s}\left\{1 - \left(\frac{T}{T+1}\right)^s\right\} \\ &= \frac{1}{T^s}\left\{1 - \left(1 + \frac{1}{T}\right)^{-s}\right\} \\ &= \frac{1}{T^s}\left\{1 - \left(1 - \frac{s}{T} + \frac{s(s+1)}{2!T^2} - \frac{s(s+1)(s+2)}{3!T^3} + \cdots\right)\right\} \\ &= \frac{s}{T^{s+1}} + \epsilon(T, s).\end{aligned}$$

The third line follows from the Taylor expansion of $(1+x)^{-s}$ about $x = 0$, and the $\epsilon(T, s)$ is an error term satisfying $|\epsilon(T, s)| < C\frac{s^2}{T^{s+2}}$ for some constant C from Taylor's theorem with remainder. Hence

$$\zeta_K(s) = \left(s\sum_{T=1}^{\infty} \cdot \frac{F(T)}{T^{s+1}} + \sum_{T=1}^{\infty} F(T)\epsilon(T, s)\right).$$

Bear in mind that we will want to take the limit as $s \to 1^+$. By the above lemma, we know $F(T) \le C_1 T$ for some C_1. So for s in the range $1 < s < 2$, we have

$$\sum_{T=1}^{\infty} F(T)|\epsilon(T, s)| \le CC_1 s^2 T \sum_{T=1}^{\infty} \frac{1}{T^{s+2}} \le 4C\sum_{n=1}^{\infty} \frac{1}{n^2} = 4C\zeta(2) < \infty.$$

Thus we may write

$$\zeta_K(s) = s\sum_{T=1}^{\infty}\frac{F(T)}{T^{s+1}} + \text{f.e.}$$

where f.e. (finite error) represents an error term which remains finite as $s \to 1^+$. Again using the lemma above, we have

$$\zeta_K(s) = s\kappa h_k \underbrace{\sum_{T=1}^{\infty}\frac{1}{T^s}}\zeta(s) + \epsilon_2(T)s\sum_{T=1}^{\infty}\frac{1}{T^{s+1}} + \text{f.e.},$$

where $\epsilon_2(T)$ is an error term satisfying $|\epsilon_2(T)| \leq C_2\sqrt{T}$ for some constant C_2. Then the middle term satisfies

$$|\epsilon_2(T)|s\sum_{T=1}^{\infty}\frac{1}{T^{s+1}} \leq 2C_2\sum_{T=1}^{\infty}\frac{1}{T^{s+\frac{1}{2}}} \leq 2C_2\zeta(s+\frac{1}{2}) \leq 2C_2\zeta(\frac{3}{2}) < \infty$$

for $1 < s < 2$. So we may write

$$\zeta_K(s) = s\kappa h_k\zeta(s) + \text{f.e.}$$

Dividing both sides by $\zeta(s)$ and sending $s \to 1^+$ gives the class number formula. □

Proposition 3.4.7. *Let χ be a nontrivial Dirichlet character mod m. Then $0 < |L(1,\chi)| < \infty$.*

This result completes the proof of Dirichlet's theorem on primes in arithmetic progressions.

Proof. If χ is nontrivial, one can show the series $\sum \frac{\chi(n)}{n}$ for $L(1,\chi)$ converges conditionally as one shows the alternating harmonic series converges in calculus. One shows that the series $L(s,\chi) = \sum \frac{\chi(n)}{n^s}$ converges uniformly for $s \geq 1$ so that $\sum \frac{\chi(n)}{n}$ actually equals $\lim_{s\to 1} L(s,\chi)$. The details are standard analysis and we will omit them. It remains to show $L(1,\chi) \neq 0$.

First suppose χ is a *real character*, i.e., the image of χ is contained in $\mathbb{R}$. In particular, χ can only take on the values ±1 and 0. We claim that $\chi(n) = \chi_\Delta(n)$ for some discriminant Δ of a quadratic field K.

Recall that $(\mathbb{Z}/m\mathbb{Z})^\times = \prod(\mathbb{Z}/p_i^{e_i}\mathbb{Z})^\times$ where $m = \prod p_i^{e_i}$ by the Chinese Remainder Theorem. This means any Dirichlet character mod m is just a product of Dirichlet characters mod $p_i^{e_i}$. Hence it suffices to prove the claim when $m = p^e$ for some prime p. Assume p is odd. (The case $p = 2$ is an exercise below.) In this case $(\mathbb{Z}/m\mathbb{Z})^\times$ is cyclic.

Let ξ be the restriction of χ to $G = (\mathbb{Z}/m\mathbb{Z})^\times$, i.e., ξ is the group character of $G = (\mathbb{Z}/m\mathbb{Z})^\times$ which gives rise to the Dirichlet character $\chi : \mathbb{Z} \to \mathbb{R}$. Note that ξ only takes on values ±1 so $\xi^2 = 1$. If ξ is trivial, then so is χ, contrary to our assumption. Hence ξ must be an element of order 2 in $\hat{G}$. (ξ is called a quadratic character.) But $\hat{G} \simeq G$ is cyclic, so there is only 1 element of order 2 in $\hat{G}$. Let $K = \mathbb{Q}(\sqrt{p})$ if $p \equiv 1 \bmod 4$ and $K = \mathbb{Q}(\sqrt{-p})$ if $p \equiv 3 \bmod 4$, so $|\Delta| = p$ where $\Delta = \Delta_K$. Thus $\left(\frac{\Delta}{\cdot}\right)$ defines a quadratic character of $(\mathbb{Z}/p\mathbb{Z})^\times$. Composing this with the natural map from $(\mathbb{Z}/p^e\mathbb{Z})^\times \to (\mathbb{Z}/p\mathbb{Z})^\times$ (just taking elements mod p) gives a nontrivial quadratic character of $(\mathbb{Z}/p^e\mathbb{Z})^\times$, which must equal ξ since it has order 2 in $\hat{G}$. Hence $\chi_\Delta = \chi$. (χ_Δ is naturally a Dirichlet character mod p, but it may also be regarded as a Dirichlet character mod $m = p^e$.)

This proves the claim that any real Dirichlet character mod m is of the form χ_Δ for some quadratic field discriminant Δ (which will sometimes be positive and sometimes be negative). But the class number formula immediately implies that $L(1,\chi_\Delta) \neq 0$.

Now suppose χ is a *complex character*, i.e., the image of χ is not contained in $\mathbb{R}$. Note the derivative

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

exists and is continuous for $s \geq 1$ (this series converges uniformly). If χ' is a complex Dirichlet character mod m such that $L(s,1) = 0$, then by the mean value theorem for any $s > 1$ there is a $1 < s_0 < s$ such that,

$$L(s,\chi') = L(s,\chi') - L(1,\chi') = L'(s_0,\chi')(s-1).$$

Similarly we have $L(s,\overline{\chi}') = L'(s_0,\overline{\chi}')(s-1)$ where $\overline{\chi}'$ is the complex conjugate of χ' (which is easily seen to also be a Dirichlet character mod m). This means that $L(s,\chi')$ and $L(s,\overline{\chi}')$ go to 0 at least as fast as $s-1$ as $s \to 1$ (faster if $L'(s_0,\chi')$ also goes to 0).

From (3.7) with $a = 1$, we have

$$\sum_{\chi} \log L(s,\chi) = \phi(m) \sum_{p \equiv 1 \bmod m} \frac{1}{p^s} + \text{f.e.}$$

where χ runs over all Dirichlet characters mod m and f.e. represents an error term which remains finite as $s \to 1$. As $s \to 1$, both $\log L(s,\chi')$ and $\log L(s,\overline{\chi}')$ approach $-\infty$ at least as fast as $\log(s-1)$. But there is only one term on the left, $\log L(s,\chi_0)$ where χ_0 is the trivial Dirichlet character mod m, which goes to ∞. It goes to ∞ at the same rate as $\log\zeta(s)$, which is $-\log(s-1)$. Hence the left hand side of the above equation goes to $-\infty$, but the right hand side stays positive (in fact goes to $+\infty$), a contradiction. □

Exercise 3.10. *Suppose χ is a real Dirichlet character mod $m = 2^e$. Using the fact that $(\mathbb{Z}/m\mathbb{Z})^\times \simeq C_2 \times C_{2^{e-2}}$ for $e > 1$, show $\chi = \chi_\Delta$ for Δ the discriminant of some quadratic field.*

Exercise 3.11. *Determine all* real *nontrivial Dirichlet characters mod m for $m = 3, 5, 6, 9, 15$. For each of these characters χ determine a quadratic field discriminant Δ such that $\chi = \chi_\Delta$. For which χ we can choose a $\Delta > 0$ so that $\chi = \chi_\Delta$ and when we can choose a $\Delta < 0$?*

In the proof for real characters, we used a Dirichlet character mod p to get a Dirichlet character mod p^e.

Exercise 3.12. *If χ is a Dirichlet character mod m, is it a Dirichlet character mod mn for any n?*

We have now seen how to prove the class number formula (omitting some details in the real quadratic case), and how one can use this to prove Dirichlet's theorem on arithmetic progressions. Of course, the natural thing to try to use this formula for is computing class numbers. Dirichlet could do this by obtaining a finite expression for $L(1,\chi)$ using *Gauss sums.*

Let χ be a Dirichlet character mod m. We say χ is **even** if $\chi(-1) = 1$ and χ is **odd** if $\chi(-1) = -1$. By multiplicativity, these conditions are equivalent to the conditions $\chi(-n) = \chi(n)$ and $\chi(-n) = -\chi(n)$, respectively, justifying the terminology of even and odd. For $k \in \mathbb{N} \cup \{0\}$, define the k-th **Gauss sum** associated to χ to be

$$\tau_k(\chi) = \sum_{a \in (\mathbb{Z}/m\mathbb{Z})^\times} \chi(a) e^{2\pi i a k/m}.$$

From the definition, notice that the Gauss sums are something like "Fourier coefficients" for χ.

We will assume that χ is **primitive**, which means that χ is nontrivial and (regarded as a character of $(\mathbb{Z}/m\mathbb{Z})^\times$) it does not come from a character of $(\mathbb{Z}/d\mathbb{Z})^\times$ for any proper divisor d of m. (If $d|m$, then any character of $(\mathbb{Z}/d\mathbb{Z})^\times$ gives a character of $(\mathbb{Z}/m\mathbb{Z})^\times$ just by composition with the mod d map $(\mathbb{Z}/m\mathbb{Z})^\times \to (\mathbb{Z}/d\mathbb{Z})^\times$.)

Exercise 3.13. *Determine which Dirichlet characters mod m are primitive for $m = 4, 8, 15$. Find a formula for the number of primitive Dirichlet characters mod m for any $m \in \mathbb{N}$.*

Proposition 3.4.8. *Let χ be a primitive character mod m. Then*

$$L(1,\chi) = \begin{cases} \frac{\pi i \tau_1(\chi)}{m^2} \sum_{k \in (\mathbb{Z}/m\mathbb{Z})^\times} \chi^{-1}(k) k & \chi \text{ odd} \\ -\frac{\tau_1(\chi)}{m} \sum_{k \in (\mathbb{Z}/m\mathbb{Z})^\times} \chi^{-1}(k) \log \sin \frac{k\pi}{m} & \chi \text{ even.} \end{cases}$$

Proof. For $s > 1$, the absolute convergence of the L-series for $L(s,\chi)$ implies we can write

$$L(s,\chi) = \sum_{a \in (\mathbb{Z}/m\mathbb{Z})^\times} \left\{ \chi(a) \sum_{n \in m\mathbb{N}+a} \frac{1}{n^s} \right\}.$$

Note the characters for the additive group $(\mathbb{Z}/m\mathbb{Z})$ are just given by $\omega_k(a) = e^{2\pi i a k/m}$ for $0 \le k \le m-1$. Hence the character orthogonality relations, namely that

$$\sum_{k=0}^{m-1} \omega_k(a) \omega_k^{-1}(n) = \begin{cases} m & a \equiv n \bmod m \\ 0 & \text{else,} \end{cases}$$

tell us we can rewrite the inner sum above as

$$\frac{1}{m} \sum_{n=1}^{\infty} \frac{\sum_{k=0}^{m-1} \omega_k(a) \omega_k^{-1}(n)}{n^s} = \frac{1}{m} \sum_{n=1}^{\infty} \frac{\sum_{k=0}^{m-1} e^{2\pi i (a-n)k/m}}{n^s}.$$

Plugging this in to our first equation gives

$$L(s,\chi) = \frac{1}{m} \sum_{k=0}^{m-1} \left\{ \tau_k(\chi) \sum_{n=1}^{\infty} \frac{\omega^{-nk}}{n^s} \right\},$$

for $s > 1$, where $\omega = e^{2\pi i/m}$. It is easy to see the Dirichlet series $\sum_{n=1}^{\infty} \frac{\omega^{-nk}}{n^s}$ converges for $s > 1$, and as $s \to 1^+$, one can check it approaches $-\log(1 - \omega^{-k})$. (Plugging in $s = 1$ gives the Taylor expansion for $-\log(1-z)$ evaluated at $z = \omega^{-k}$.) This gives

$$L(1,\chi) = -\frac{1}{m} \sum_{k=1}^{m-1} \tau_k(\chi) \log(1 - \omega^{-k}).$$

(Up to here, we do not need that χ is primitive, just nontrivial.)

We again use some simple character theory to obtain that

$$\tau_k(\chi) = \begin{cases} \chi^{-1}(k) \tau_1(\chi) & \gcd(k,m) = 1 \\ 0 & \gcd(k,m) > 1. \end{cases}$$

Then

$$L(1,\chi) = -\frac{\tau_1(\chi)}{m} \sum_{k\in(\mathbb{Z}/m\mathbb{Z})^\times} \chi^{-1}(k)\log(1-\omega^{-k}).$$

We can replace k with $-k$ in the sum and pull out a $\chi(-1)$ to get

$$L(1,\chi) = -\frac{\chi(-1)\tau_1(\chi)}{m} \sum_{k\in(\mathbb{Z}/m\mathbb{Z})^\times} \chi^{-1}(k)\log(1-\omega^{k}).$$

Using the fact that

$$\log(1-\omega^k) = \log 2 + \log\sin\frac{k\pi}{m} + \left(\frac{k}{m}-\frac{1}{2}\right)\pi i$$

one can write

$$L(1,\chi) = -\frac{\chi(-1)\tau_1(\chi)}{m} \sum_{k\in(\mathbb{Z}/m\mathbb{Z})^\times} \chi^{-1}(k)\left(\log\sin\frac{k\pi}{m} + \frac{k\pi i}{m}\right).$$

Note that the $\log 2$ terms drop out since $\sum_k \chi^{-1}(k) = 0$ by orthogonality relations. One finishes the proof by checking that $\sum_k \chi^{-1}(k)\log\sin\frac{k\pi}{m} = 0$ if χ is odd and $\sum_k \chi^{-1}(k)k = 0$ if χ is even. □

This immediately gives a more useable version of Dirichlet's class number formula. Though the version we give below comes from one additional simplification—namely, since $h_K > 0$, we only need a formula for the absolute value $|L(1,\chi)|$ to compute h_K. By taking absolute values in the above proposition, one can get rid of the Gauss sum (which has absolute value $\sqrt{m}$). We omit the details of the complete simplification, but in the end, one has the following.

Theorem 3.4.9. (Dirichlet's class number formula, second form) *Let $K = \mathbb{Q}(\sqrt{d})$ be the quadratic field of discriminant Δ. Then*

$$h_K = \begin{cases} \frac{1}{2-\chi_\Delta(2)}\left|\sum_{1\le k<|\Delta|/2}\chi_\Delta(k)\right| & \Delta<0, \Delta\neq -3,-4 \\ \frac{1}{\log\eta}\left|\sum_{1\le k<\Delta/2}\chi_\Delta(k)\log\sin\frac{k\pi}{\Delta}\right| & \Delta>0. \end{cases}$$

Here η denotes the fundamental unit of $\mathcal{O}_K$ when $\Delta > 0$.

Note with this form of the class number formula, we can effectively compute the class number of any quadratic field, and this is much easier than the approach via Minkowski's bound (though in the real quadratic case one needs to determine the fundamental unit). Additionally, since $\chi_\Delta(n) = 0$ unless $\gcd(n,\Delta) = 1$, we may restrict the above sums to k relatively prime to Δ.

Example 3.4.10. *For $d = -2$, i.e., $K = \mathbb{Q}(\sqrt{-2})$, we see $\Delta = -8$,*

$$\chi_\Delta(n) = \begin{cases} 1 & n \text{ odd} \\ 0 & n \text{ even,} \end{cases}$$

so

$$h_K = \frac{1}{2}|\chi_\Delta(1)+\chi_\Delta(3)| = 1.$$

Exercise 3.14. *Using the second form of the class number formula compute h_K where $K = \mathbb{Q}(\sqrt{d})$ for $d = -5, -6, -7, -10, 2$.*

3.5 Postlude: Beyond Dirichlet

In this chapter we have shown (modulo a few details) two landmark results in number theory:

(i) Dirichlet's theorem on primes in arithmetic progressions
(ii) Dirichlet's class number formula for quadratic fields

both of which used Dirichlet L-functions. Subsequent key developments in number theory, generalizing the above, are

(I) prime density theorems
(II) class number formulas for general number fields.

Let us first discuss the density theorems. Let $\mathcal{P}$ denote the set of primes in $\mathbb{N}$, and $S \subseteq \mathcal{P}$. We say S has **(natural) density**[‡] ρ if

$$\lim_{x\to\infty} \frac{|\{p \in S : p < x\}|}{|\{p \in \mathcal{P} : p < x\}|} = \rho.$$

Of course any finite set of primes will have density 0, but infinite sets of primes can have density 0 also. Using this notion, one can strengthen Dirichlet's theorem on arithmetic progressions to the following:

Theorem 3.5.1. *Let $m > 1$ and $a \in \mathbb{N}$ such that $\gcd(a, m) = 1$. Then the set of primes $\equiv a \bmod m$ has density $\frac{1}{\phi(m)}$.*

In other words, the number of primes $\equiv a \bmod m$ is the same for any $a \in (\mathbb{Z}/m\mathbb{Z})^\times$. Or put another way, the primes are distributed equally among the invertible congruence classes mod m. For example, there are the same number or primes $\equiv 1 \bmod 4$ as there are $\equiv 3 \bmod 4$ (in the sense of density). This is one example of a statistical regularity that prime numbers satisfy, despite their apparent randomness. Proving this requires significantly more sophisticated analysis of Dirichlet L-functions.

Another way to think about distribution of primes is in terms of how they split in extensions K of $\mathbb{Q}$. For example the above statement about the number of primes $\equiv 1 \bmod 4$ being the same as the number of primes $\equiv 3 \bmod 4$ can be recast as saying the number of primes with split in $\mathbb{Q}(i)$ is the same as the number of primes which are inert in $\mathbb{Q}(i)$. In fact the above theorem is essentially equivalent to the following

Theorem 3.5.2. *Let K be a quadratic field. The set of primes in $\mathcal{P}$ which split in K has density $\frac{1}{2}$, as does the set of primes which remain inert in K.*

In other words, half the primes split in K and half remain prime in K, for any quadratic field K. This is a special case of a very strong theorem, called the *Chebotarev density theorem.* This provides much stronger regularity results about distributions of primes than just considering congruence classes. We haven't defined everything we need at this point to state the complete theorem, but here is a (corollary of a) special case.

Theorem 3.5.3. *Let K be a number field of degree n and S be the set of primes in $\mathcal{P}$ which split completely in K. Then S has density $\rho \geq \frac{1}{n}$. Further $\rho = \frac{1}{n}$ if and only if $K/\mathbb{Q}$ is Galois.*

[‡] There is another weaker notion of density, now called Dirichlet density, that coincides with natural density if the natural density exists, which is essentially what Dirichlet originally considered.

A proof of Chebotarev density involves a more general kind of L-function called an Artin L-function. The basic idea is the following. Even though we have several Dirichlet characters mod m for a given m, the most important Dirichlet characters are those of the form χ_Δ where Δ is the discriminant of some quadratic field K.

On the other hand, if we start with some number field K, there is a natural group associated to it, $\mathrm{Gal}(K/\mathbb{Q})$, and so one can look at the irreducible representations ρ of $\mathrm{Gal}(K/\mathbb{Q})$. Artin defined an L-function $L(s, \rho)$ associated to each such Galois representation ρ in order to study how the primes split in $K/\mathbb{Q}$ (or more generally, an arbitrary extension L/K). This L-function has some L-series expansion as well as an Euler product. If ρ is the trivial character of $\mathrm{Gal}(K/\mathbb{Q})$, then $L(s, \rho)$ is essentially $\zeta_K(s)$ (they have the same Euler product expansions except at a finite number of primes).

In the case where K is quadratic, then $\mathrm{Gal}(K/\mathbb{Q})$ has only 2 irreducible representations, both of dimension 1, i.e., both characters. If ρ is the nontrivial character of $\mathrm{Gal}(K/\mathbb{Q})$, then $L(s, \rho) = L(s, \chi_\Delta)$ where $\Delta = \Delta_K$. If ρ_0 is the trivial character, then $L(s, \rho_0) = L(s, \chi_0)$ where χ_0 is the trivial Dirichlet character mod Δ. Hence the Artin L-functions are a generalization of (at least the most important cases of) Dirichlet L-functions. In this quadratic case we have

$$L(s, \rho_0)L(s, \rho) = L(s, \chi_0)L(s, \chi) \simeq \zeta_K(s)L(s, \chi) = \zeta(s),$$

where $\simeq$ means equal up to a finite number of factors in the Euler product.

Similarly in the case of K a general number field, we have

$$\prod_{\rho} L(s, \rho) \simeq \zeta_K(s) \prod_{\rho \text{ nontrivial}} L(s, \rho) = \zeta(s),$$

where ρ runs over all irreducible representations of $\mathrm{Gal}(K/\mathbb{Q})$. One has the following generalization of Dirichlet's class number formula.

Theorem 3.5.4. (General class number formula) *Let K be a number field and $\{\rho\}$ be the set of irreducible representations of* $\mathrm{Gal}(K/\mathbb{Q})$. *Then*

$$\prod_{\rho \text{ nontrivial}} L(1, \rho) = \lim_{s \to 1^+} \frac{\zeta_K(s)}{\zeta(s)} = \frac{2^{r_1}(2\pi)^{r_2} R}{w\sqrt{|\Delta_K|}} h_K,$$

where r_1 (resp. r_2) is the number of real (resp. complex) embeddings of K, w is the number of roots of unity in K, and R is the regulator *of K.*

The regulator is basically the volume of a certain lattice which comes up in Dirichlet's units theorem. This provides a way to compute and study class numbers for number fields of degree greater than 2. In the case $K = \mathbb{Q}(\zeta_p)$ where ζ_p is a primitive p-th root of unity, the class number formula provides a way to determine which primes are *regular*, i.e., for which p is $\gcd(p, h_K) = 1$. The significance of this is that Kummer was able to prove Fermat's Last Theorem in the case of regular prime exponents.

In fact, Wiles proved Fermat's Last Theorem by (more or less) proving the Taniyama–Shimura Conjecture, which is itself a statement about L-functions. To certain curves (irreducible smooth cubic curves) called *elliptic curves E* one associates an L-function $L(s, E)$. (It can be given by an Euler product where the factor corresponding to the prime p is given by the number of points on E mod p.) On the other hand, one also has L-functions $L(s, f)$ attached to *modular forms f*, which are certain periodic functions on the upper half plane. (The Euler product factors are given by the p-th Fourier coefficients in a certain Fourier expansion.) The Taniyama–Shimura Conjecture is that every elliptic curve E corresponds to a modular form f in the sense that $L(s, E) = L(s, f)$. (Going from modular forms to elliptic curves is easier, and is known by work of Eichler and Shimura.)

While the whole story is rather involved, let me just say that it is difficult to directly compare elliptic curves and modular forms, but one can associate to both of these objects certain 2-dimensional Galois representations. Thus to prove Taniyama–Shimura, one wants to show the Galois representations coming from elliptic curves are contained in the set of Galois representations coming from modular forms. In a herculian endeavor, Wiles reduced (a sufficient part of) the Taniyama–Shimura conjecture to a theorem of Langlands and Tunnell (also a very difficult result) which tells us every (odd) 2-dimensional complex Galois representation $\rho : \mathrm{Gal}(K/\mathbb{Q}) \to \mathrm{GL}_2(\mathbb{C})$ with solvable image corresponds to a modular form.

Let me emphasize 2 points:

1) L-functions provide a practical (though not typically easy) way to compare objects of different types: geometric objects (curves and varieties), algebraic/arithmetic objects (number fields/Galois representations) and analytic objects (modular/automorphic forms).

2) Galois representations and/or automorphic forms/automorphic representations provide a general framework for studying many number theoretic problems. For example, if K is an abelian extension of $\mathbb{Q}$, i.e., $K/\mathbb{Q}$ is Galois with abelian Galois group, then all representations of $\mathrm{Gal}(K/\mathbb{Q})$ are 1-dimensional. Hence the theory of 1-dimensional Galois representations provides a way to study abelian extensions of $\mathbb{Q}$, and contains the case of Dirichlet characters and everything we did in this chapter. The fact that 1-dimensional Galois representations correspond to "1-dimensional" automorphic forms is essentially *class field theory*, which provides a way to understand the abelian extensions of a number field, and is the crowning achievement of classical algebraic number theory.

In this context, we can view the Taniyama-Shimura conjecture as a sort of "2-dimensional" analogue of class field theory. A large amount of present work in modern number theory is studying higher-dimensional analogues of class field theory (often called non-abelian class field theory). In fact, this is essentially the area of research Ralf Schmidt, Ameya Pitale and I specialize in (and it is related to Alan Roche and Tomasz Prezbinda's research as well). For example, in my thesis I proved that certain 4-dimensional Galois representations with solvable image correspond to automorphic forms.

We will give a brief introduction to class field theory and higher-dimensional analogues in Part III.