18 The analytic class number formula

The following theorem is usually attributed to Dirichlet, although he originally proved it only for quadratic fields. The formula for the limit in the theorem below was proved by Dedekind, and analytic continuation was proved by Landau. Hecke later showed that, like the Riemann zeta function, the Dedekind zeta function has an analytic continuation to all of $\mathbb{C}$ and satisfies a functional equation, but we won't prove these results here.

Theorem (Analytic Class Number Formula)**.** *Let K be a number field of degree n with r real and s complex places. The Dedekind zeta function $\zeta_K(z)$ extends to a meromorphic function on $\operatorname{Re}(z) > 1 - \frac{1}{n}$ that is holomorphic except for a simple pole at $z = 1$ with residue*

$$\lim_{z\to 1^+}(z-1)\zeta_K(z) = \frac{2^r(2\pi)^s h_K R_K}{w_K|D_K|^{1/2}},$$

where $h_K := \#\operatorname{cl}\mathcal{O}_K$ is the class number, R_K is the regulator, $w_K := \#(\mathcal{O}_K^\times)_{\rm tors}$ is the number of roots of unity in K, and $D_K := \operatorname{disc}\mathcal{O}_K$ is the discriminant.

Recall that the regulator R_K is the covolume of the image of $\mathcal{O}_K^\times$ in the trace-zero hyperplane $\mathbb{R}_0^{r+s}$ under the Log map; see Definition 14.10.

Example 18.1. For $K = \mathbb{Q}$ we already know that

$$\zeta_{\mathbb{Q}}(z) = \zeta(z) = \frac{1}{z-1} + \phi(z)$$

where $\phi(z)$ is holomorphic on $\operatorname{Re} z > 1 - \frac{1}{1} = 0$. The residue of the simple pole at $z = 1$ is

$$\lim_{z\to 1^+}(z-1)\zeta(z) = 1 + (z-1)\phi(z) = 1.$$

In terms of the class number formula, we have $r = 1, s = 0, h_K = 1, R_K = 1, w_K = 2$, and $D_K = 1$ (for the regulator, note that the covolume of a point in $\mathbb{R}^0$ is the determinant of a 0×0 matrix, which is 1). Plugging these values into the theorem gives

$$\lim_{z\to 1^+}(z-1)\zeta_K(z) = \frac{2^1(2\pi)^0\cdot 1\cdot 1}{2\cdot|1|^{1/2}} = 1,$$

as expected.

18.1 Cyclotomic zeta functions and Dirichlet L-series

Before proving the analytic class number formula, let's use it to complete the proof of Dirichlet's theorem on primes in arithmetic progressions that we started in the previous lecture. In order to establish the key claim that the Dirichlet L-series $L(s,\chi)$ does not vanish at $s = 1$ when χ is non-principal, we are going to show that the Dedekind zeta function for the mth cyclotomic field $K = \mathbb{Q}(\zeta_m)$ can be written as a product of Dirichlet L-series

$$\zeta_K(s) = \prod_\chi L(s,\chi),$$

where the product ranges over the primitive Dirichlet characters whose conductor divides m. For the principal character of conductor 1 we have $L(s,\chi) = \zeta(s)$ with a simple pole at $s = 1$, and since $\zeta_K(s)$ also has a simple pole at $s=1$, this implies that none of the $L(s,\chi)$ with χ non-principal can vanish at $s=1$ (by Proposition 17.29, none of them has a pole at $s=1$).

Let ζ_m be a primitive mth root of unity and let $K=\mathbb{Q}(\zeta_m)$ be the mth cyclotomic field. Recall that we have an isomorphism

$$\varphi\colon \operatorname{Gal}(K/\mathbb{Q}) \xrightarrow{\sim} (\mathbb{Z}/m\mathbb{Z})^\times$$

that sends $\sigma\in\operatorname{Gal}(K/\mathbb{Q})$ to the unique $a\in(\mathbb{Z}/m\mathbb{Z})^\times$ for which $\sigma(\zeta_m)=\zeta_m^a$. For primes $p\nmid m$, we have $\varphi(\sigma_p)=p\in(\mathbb{Z}/m\mathbb{Z})^\times$, where $\sigma_p\in\operatorname{Gal}(K/\mathbb{Q})$ is the Frobenius element at p (which is the Frobenius element at any prime $\mathfrak{p}$ of $\mathbb{Q}(\zeta_m)$ above p; these are all conjugate, hence equal, because $\operatorname{Gal}(K/\mathbb{Q})$ is abelian).

Theorem 18.2. *Let $K=\mathbb{Q}(\zeta_m)$ be the mth cyclotomic field. Then*

$$\zeta_K(s)=\prod_\chi L(s,\chi),$$

where χ ranges over the primitive Dirichlet characters of conductor dividing m.

Proof. On the LHS we have

$$\zeta_K(s)=\prod_{\mathfrak{q}}\left(1-\mathrm{N}(\mathfrak{q})^{-s}\right)^{-1}=\prod_p\prod_{\mathfrak{q}|p}((1-\mathrm{N}(\mathfrak{p})^{-s})^{-1},$$

and on the RHS we have

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

It thus suffices to prove that the equality

$$\prod_{\mathfrak{q}|p}\left(1-\mathrm{N}(\mathfrak{q})^{-s}\right)\overset{?}{=}\prod_\chi\left(1-\chi(p)p^{-s}\right)$$

holds for each prime p.

Since $K/\mathbb{Q}$ is Galois, for each prime p we have $[K:\mathbb{Q}]=\phi(m)=e_pf_pg_p$, where e_p is the ramification index, f_p is the inertia degree, and g_p is the number of distinct primes $\mathfrak{q}|p$. On the LHS,

$$\prod_{\mathfrak{q}|p}\left(1-\mathrm{N}(\mathfrak{q})^{-s}\right)=\left(1-(p^{f_p})^{-s}\right)^{g_p}=\left(1-(p^{-s})^{f_p}\right)^{g_p}. \tag{1}$$

On the RHS we can ignore factors with $\chi(p)=0$; these occur precisely when p divides the conductor of χ (which never happens if $p\nmid m$ is unramified). Let us write $m=p^km'$ with m' not divisible by p. We may identify the set of primitive Dirichlet characters of conductor dividing m' with the character group of $(\mathbb{Z}/m'\mathbb{Z})^\times$, via Lemma 17.25.

The field $K'=\mathbb{Q}(\zeta_{m'})$ is the maximal extension of $\mathbb{Q}$ in K unramified at p; it has degree $\phi(m')=\phi(m)/e_p=f_pg_p$ (because K'/K is totally ramified at p with degree e_p). Thus

$$\prod_\chi\left(1-\chi(p)p^{-s}\right)=\prod_{\operatorname{cond}(\chi)|m'}\left(1-\chi(p)p^{-s}\right)=\prod_{\alpha^{f_p}=1}\left(1-\alpha p^{-s}\right)^{g_p},$$

since the map $\chi\mapsto\chi(p)$ defines a surjective homomorphism from the character group of $(\mathbb{Z}/m'\mathbb{Z})^\times$ to the group of f_pth roots of unity α, and the kernel of this map has cardinality $\phi(m')/f_p=g_p$.

Over $\mathbb{C}[T]$ we have the identity

$$1 - T^f = \prod_{\alpha \in \mu_{f_p}} (1 - \alpha T),$$

and substituting $T = p^{-s}$ yields

$$\prod_{\chi} \left(1 - \chi(p)p^{-s}\right) = \prod_{\alpha^{f_p}=1} \left(1 - \alpha p^{-s}\right)^{g_p} = \Big(1 - (p^{-s})^{f_p}\Big)^{g_p},$$

which matches the expression in (1) for the LHS as desired. □

Remark 18.3. Theorem 18.2 is sometimes stated in terms of the Dirichlet characters of modulus m, rather than using the primitive Dirichlet characters of conductor dividing m. Both forms of the theorem are equivalent, but using primitive characters as we have here correctly accounts for the Euler factors at primes $p|m$, leading to a prettier formula and a simpler proof. More generally, one can use Dirichlet characters to analyze ramification in any abelian extension of $\mathbb{Q}$ (these are all subfields of cyclotomic fields, but need not be cyclotomic), and for this purpose it is better to use primitive characters.

18.2 Non-vanishing of Dirichlet L-functions with non-principal character

We can now prove the key claim needed to complete our proof of Dirichlet's theorem on primes in arithmetic progressions.

Theorem 18.4. *Let χ be any non-principal Dirichlet character. Then $L(1, \chi) \neq 0$.*

Proof. Let ψ be a non-principal Dirichlet character, say of modulus m. Then ψ is induced by a non-trivial primitive Dirichlet character $\widetilde{\psi}$ of conductor $\widetilde{m}$ dividing m. The L-functions of ψ and $\widetilde{\psi}$ differ at only finitely many Euler factors $(1 - \chi(p)p^{-s})^{-1}$ (corresponding to primes p dividing m but not $\widetilde{m}$), and these factors are all nonzero at $s = 1$ because $p > 1$. Thus we may assume without loss of generality that $\psi = \widetilde{\psi}$ is primitive.

We now consider the order of vanishing at $s = 1$ of both sides of the equality

$$\zeta_K(s) = \prod_{\chi} L(s, \chi),$$

given by Theorem 18.2, where the product ranges over primitive Dirichlet characters of conductor dividing m. By the analytic class number formula (Theorem 18.13), the LHS has a simple pole at $s = 1$, so the same must be true of the RHS. Thus

$$\begin{aligned}
\operatorname{ord}_{s=1} \zeta_K(s) &= \operatorname{ord}_{s=1} \prod_{\chi} L(s, \chi) \\
-1 &= \operatorname{ord}_{s=1} L(s, \mathbb{1}) \prod_{\chi \neq \mathbb{1}_m} L(s, \chi) \\
-1 &= \operatorname{ord}_{s=1} \zeta(s) \prod_{\chi \neq \mathbb{1}} L(s, \chi) \\
-1 &= -1 + \sum_{\chi \neq \mathbb{1}} \operatorname{ord}_{s=1} L(s, \chi)
\end{aligned}$$

Each $\chi \neq \mathbb{1}$ in the sum is necessarily non-principal (since it is primitive), and we proved in Proposition 17.29 that for non-principal χ the Dirichlet L-series $L(s,\chi)$ has an analytic continuation to $\operatorname{Re} s > 0$; in particular, it does not have a pole at $s = 1$, so $\operatorname{ord}_{s=1} L(s,\chi) \geq 0$. But then

$$\sum_{\chi \neq \mathbb{1}} \operatorname{ord}_{s=1} L(s,\chi) = 0$$

can only hold if every term in the sum is zero. So $L(1,\chi) \neq 0$ for every non-trivial primitive Dirichlet character χ of conductor dividing m, and in particular for $\psi = \widetilde{\psi}$. □

18.3 Preparation for proving the analytic class number formula

Recall that in §14.2 of Lecture 14 we defined the locally compact group

$$K_{\mathbb{R}}^{\times} := \prod_{\text{real } v|\infty} \mathbb{R}^{\times} \times \prod_{\text{complex } v|\infty} \mathbb{C}^{\times},$$

whose multiplication is defined component-wise, with the subspace topology inherited from $K_{\mathbb{R}} := K \otimes_{\mathbb{Q}} \mathbb{R} \simeq \mathbb{R}^n$ (note that $\mathbb{R}^r \times \mathbb{C}^s \simeq \mathbb{R}^n$ as $\mathbb{R}$-vector spaces, and even though they are not isomorphic as $\mathbb{R}$-algebras, multiplication is continuous, so we do get a locally compact topological group). We have a natural embedding

$$\begin{aligned} K^{\times} &\hookrightarrow K_{\mathbb{R}}^{\times} \\ x &\mapsto (x_v), \end{aligned}$$

where v ranges over the $r+s$ archimedean places of K; this allows us to view $K^{\times}$ as a subgroup of $K_{\mathbb{R}}^{\times}$ that contains the nonzero elements of $\mathcal{O}_K$. We then defined the log map

$$\begin{aligned} \operatorname{Log}: K_{\mathbb{R}}^{\times} &\to \mathbb{R}^{r+s} \\ (x_v) &\mapsto (\log \|x_v\|_v), \end{aligned}$$

and proved that we have an exact sequence of abelian groups

$$0 \longrightarrow (\mathcal{O}_K^{\times})_{\text{tors}} \longrightarrow \mathcal{O}_K^{\times} \xrightarrow{\text{Log}} \Lambda_K \to 0,$$

in which Λ_K is a lattice in the trace-zero hyperplane $\mathbb{R}_0^{r+s} := \{x \in \mathbb{R}^{r+s} : T(x) = 0\}$ (the trace $T(x)$ is just the sum of the coordinates of x). We then defined the regulator R_K as the covolume of Λ_K in $\mathbb{R}_0^{r+s}$ (see Definition 14.10), where we endow $\mathbb{R}_0^{r+s}$ with the Euclidean measure induced by any coordinate projection $\mathbb{R}^{r+s} \to \mathbb{R}^{r+s-1}$.

18.3.1 Lipschitz parametrizability

To prove the analytic class number formula we need an asymptotic estimate of the number of nonzero $\mathcal{O}_K$-ideals I with norm $N(I)$ bounded by some real t that we will let tend to infinity; this is needed to understand the behavior of $\zeta_K(z) = \sum_I N(I)^{-s}$ as $z \to 1^+$. Our strategy for doing so is to count points in the image of $\mathcal{O}_K$ under the Log map that lie inside a suitably chosen region S of $\mathbb{R}^{r+s}$ that we will than scale by t. In order to bound this count as a function of t we need a condition on S that ensures that this count grows smoothly with t; this is not guaranteed for arbitrary S, we need S to have a "reasonable" shape, one that is not too convoluted. A sufficient condition for this is *Lipschitz parametrizability.*

Definition 18.5. Let X and Y be metric spaces. A function $f : X \to Y$ is *Lipschitz* if there exists $c > 0$ such that for all $x_1, x_2 \in X$

$$d(f(x_1), f(x_2)) \leq cd(x_1, x_2).$$

Every Lipschitz function is continuous, in fact, uniformly continuous, but the converse need not hold. For example, the function $f(x) = \sqrt{x}$ on $[0, 1]$ is uniformly continuous but not Lipschitz, since $|\sqrt{1/n} - 0|/|1/n - 0| = \sqrt{n}$ is unbounded as $1/n \to 0$.

Definition 18.6. A set $B \subseteq \mathbb{R}^n$ is *d-Lipschitz parametrizable* if it is the union of the images of a finite number of Lipschitz maps $[0, 1]^d \to B$.

Lemma 18.7. *Let $S \subseteq \mathbb{R}^n$ be a set whose boundary $\partial S := \overline{S} - S^0$ is $(n - 1)$-Lipschitz parametrizable. Then*

$$\#(tS \cap \mathbb{Z}^n) = \mu(S)t^n + O(t^{n-1}),$$

as $t \to \infty$, where μ is the usual Lebesgue measure on $\mathbb{R}^n$.

Proof. We can partition $\mathbb{R}^n$ as the disjoint union of half-open cubes of the form

$$C(a_1, \ldots, a_n) = \{(x_1, \ldots, x_n) \in \mathbb{R}_n : x_i \in [a_i, a_i + 1)\},$$

with $a_1, \ldots, a_n \in \mathbb{Z}$. Let $\mathcal{C}$ be the set of all such half-open cubes C. For each $t > 0$ define

$$B_0(t) := \#\{C \in \mathcal{C} : C \subseteq tS\},$$
$$B_1(t) := \#\{C \in \mathcal{C} : C \cap tS\}.$$

For every $t > 0$ we have

$$B_0(t) \leq \#(tS \cap \mathbb{Z}^n) \leq B_1(t).$$

We can bound $B_1(t) - B_0(t)$ by noting that each $C(a_1, \ldots, a_n)$ counted by this difference has $(a_1, \ldots, a_n)$ within a distance $\sqrt{n} = O(1)$ of a point in ∂tS.

Let $\tau = \lfloor t \rfloor$. Let $f_1, \ldots, f_m$ be Lipschitz functions $[0, 1]^{n-1} \to \partial S$ whose images cover ∂S. There is an absolute constant c (independent of τ) such that every point in ∂S is within a distance $c/\tau = O(1/t)$ of a point in the set

$$\mathcal{P} = \left\{ f_i\left(\frac{a_1}{\tau}, \ldots, \frac{a_{n-1}}{\tau}\right) : 1 \leq i \leq m, a_1, \ldots, a_{n-1} \in [0, \tau) \cap \mathbb{Z} \right\},$$

which has cardinality $m\tau^{n-1} = O(t^{n-1})$. It follows that every point of ∂tS is within a distance $O(1)$ of one of the $O(t^{n-1})$ points tP with $P \in \mathcal{P}$. The number of integer lattice points within a distance $\sqrt{n}$ of a point in ∂tS is thus also $O(t^{n-1})$, and therefore

$$B_1(t) - B_0(t) = O(t^{n-1}).$$

We now note that $B_0(T) \leq \mu(tS) \leq B_1(T)$ and $\mu(tS) = t^n\mu(S)$; the lemma follows. □

Corollary 18.8. *Let Λ be a Lattice in $\mathbb{R}^n$ and let $S \subseteq \mathbb{R}^n$ be a set whose boundary is $(n - 1)$-Lipschitz parametrizable. Then*

$$\#(tS \cap \Lambda) = \frac{\mu(S)}{\operatorname{covol}(\Lambda)}t^n + O(t^{n-1}).$$

Proof. The case $\Lambda \subseteq \mathbb{Z}^n$ is clear. If the corollary holds for $s\Lambda$, for some $s > 0$, then it also holds for Λ, since $tS \cap s\Lambda = (t/s)S \cap \Lambda$.

For any lattice Λ, we can choose $s > 0$ so that $s\Lambda$ is arbitrarily close to an integer lattice (take s to be the LCM of the denominators of rational approximations of the coordinates of a basis for Λ); the corollary follows. □

18.3.2 Counting algebraic integers of bounded norm

By Dirichlet's unit theorem (Theorem 14.8), we can write

$$\mathcal{O}_K^\times = U \times (\mathcal{O}_K^\times)_{\mathrm{tors}},$$

where $U \subseteq \mathcal{O}_K^\times$ is free of rank $r+s-1$ (the subgroup U is not uniquely determined, but let us fix a choice). In order to understand the behavior of the Dedekind zeta function

$$\zeta_K(z) = \sum_I N(I)^{-z}$$

as $z \to 1^+$, we want to estimate the quantity

$$\#\{I : N(I) \le t\},$$

where I ranges over nonzero ideals of $\mathcal{O}_K$, as $t \to \infty$.

As a first step in this direction, let us try to count the set

$$\{\text{nonzero principal ideals } I \subseteq \mathcal{O}_K : N(I) \le t\}.$$

Equivalently, we want to count

$$\{\alpha \in \mathcal{O}_K - \{0\} : |N(\alpha)| \le t\}/\mathcal{O}_K^\times,$$

which is equivalent to

$$\left(K_{\mathbb{R},\le t}^\times \cap \mathcal{O}_K\right)/\mathcal{O}_K^\times,$$

where we are viewing $\mathcal{O}_K - \{0\}$ as a subset of $K_\mathbb{R}^\times$ containing the subgroup $\mathcal{O}_K^\times$ and

$$K_{\mathbb{R},\le t}^\times := \{x \in K_\mathbb{R}^\times : |N(x)| \le t\}.$$

Recall that for $x = (x_v) \in K_\mathbb{R}^\times$ the norm map $\mathrm{N}\colon K_\mathbb{R}^\times \to \mathbb{R}^\times$ is defined by

$$\mathrm{N}(x) := \prod_{v \text{ real}} x_v \prod_{v \text{ complex}} x_v \bar{x}_v,$$

and satisfies $\mathrm{T}(\mathrm{Log}(x)) = \log|\mathrm{N}(x)|$ for all $x \in K_\mathbb{R}^\times$. To simplify matters, let us replace $\mathcal{O}_K^\times$ with the free group U; we then have a w_K–to–1 map

$$(K_{\mathbb{R},\le t}^\times \cap \mathcal{O}_K)/U \longrightarrow \left(K_{\mathbb{R},\le t}^\times \cap \mathcal{O}_K\right)/\mathcal{O}_K^\times.$$

If F is a fundamental region for $K_\mathbb{R}^\times/U$, it suffices to consider

$$F_{\le t} \cap \mathcal{O}_K,$$

where $F_{\le t} := \{x \in F : \mathrm{N}(x) \le t\}$. Note that F is not compact, but $F_{\le t}$ is, and $\mathcal{O}_K - \{0\}$ is discrete as a subset of $K_\mathbb{R}^\times$, so $F_{\le t} \cap \mathcal{O}_K$ is a finite set; we want to understand how its cardinality grows as $t \to \infty$.

In order to explicitly construct F we define the map

$$\begin{aligned}\sigma\colon K_\mathbb{R}^\times &\twoheadrightarrow K_{\mathbb{R},1}^\times\\ x &\mapsto x|\mathrm{N}(x)|^{-1/n}\end{aligned}$$

which rescales each $x \in K_{\mathbb{R}}^{\times}$ so that it has norm 1. The image of $K_{\mathbb{R},1}^{\times}$ under the Log map is precisely the trace-zero hyperplane $\mathbb{R}_0^{r+s}$ in $\mathbb{R}^{r+s}$, in which $\operatorname{Log} U = \Lambda_K$ is a lattice. If we pick a fundamental domain R for the lattice Λ_K in $\mathbb{R}_0^{r+s}$ then

$$F := \sigma^{-1}\left(\operatorname{Log}^{-1}(R)\right)$$

is a fundamental region for $K_{\mathbb{R}}^{\times}/U$. Note that $tF_{\le 1} = F_{\le t^n}$, so $F_{\le t} = t^{1/n}F_{\le 1}$.

Recall that the map $\operatorname{Log}\colon K_{\mathbb{R}}^{\times} \to \mathbb{R}^{r+s}$ satisfies

$$(x_1, \ldots, x_r, z_1 \ldots, z_s) \mapsto (\log|x_1|, \ldots, \log|x_r|, 2\log|z_1|, \ldots, 2\log|z_s|),$$

where $x_1, \ldots, x_r \in \mathbb{R}^{\times}$ and $z_1, \ldots, z_s \in \mathbb{C}^{\times}$ and $|\ |$ denotes the usual absolute value in $\mathbb{R}$ and $\mathbb{C}$. The kernel of the Log map is $\{\pm 1\}^r \times \mathrm{U}(1)^s$, where $\mathrm{U}(1) = \{z : |z| = 1\}$ is the unit circle in $\mathbb{C}$. We thus have a continuous isomorphism of locally compact groups

$$K_{\mathbb{R}}^{\times} = (\mathbb{R}^{\times})^r \times (\mathbb{C}^{\times})^s \xrightarrow{\sim} \mathbb{R}^{r+s} \times \{\pm 1\}^r \times \mathrm{U}(1)^s, \tag{2}$$

where the map to $\mathbb{R}^{r+s}$ is the Log map, the map to $\{\pm 1\}^r$ is the vector of signs of the r real components, and the map to $\mathrm{U}(1)^s$ is the vector of radial projections to $\mathrm{U}(1)$ of the s complex components.

The set $F_{\le 1} = F_{<1}$ consists of 2^r connected components, one for each element of $\{\pm 1\}$. We can parametrize each of these component using n real parameters as follows:

- $r + s - 1$ parameters in $[0, 1)$ that encode a point in R as an $\mathbb{R}$-linear combination of $\operatorname{Log}(\epsilon_1), \ldots, \operatorname{Log}(\epsilon_{r+s-1})$, where $\epsilon_1, \ldots, \epsilon_{r+s-1}$ are a basis for U;
- s parameters in $[0, 1)$ that encode an element of $\mathrm{U}(1)^s$;
- a parameter in $(0, 1]$ that encodes the nth-root of the norm.

These parameterizations define a continuously differentiable bijection from the set

$$C = [0, 1)^{n-1} \times (0, 1] \simeq [0, 1)^n \subseteq [0, 1]^n$$

to each of the 2^r disjoint components of F; it can be written out explicitly in terms of exponentials and the identity function. The boundary ∂C of the unit cube is clearly $(n-1)$-Lipschitz parametrizable, so $\partial F_{\le 1}$ is $(n-1)$-Lipschitz parameterizable.

Applying Corollary 18.8 to the set $S = F_{\le 1}$ with t replaced by $t^{1/n}$ and recalling that $F_{\le t} = t^{1/n}F_{\le 1}$ yields the asymptotic bound

$$\#(F_{\le t} \cap \mathcal{O}_K) = \frac{\mu(F_{\le 1})}{\operatorname{covol}(\mathcal{O}_K)}(t^{1/n})^n + O\big((t^{1/n})^{n-1}\big) = \left(\frac{\mu(F_{\le 1})}{|\operatorname{disc}\mathcal{O}_K|^{1/2}}\right)t + O\big(t^{1-1/n}\big), \tag{3}$$

so the number of elements of $\mathcal{O}_K$ in $F_{\le t}$ grows linearly with t.

Our next task is compute $\mu(F_{\le 1})$. To do this we will use the isomorphism in (2) to make a change of coordinates, and we need to understand how this affects the Haar measure μ on $K_{\mathbb{R}}^{\times} \subseteq K_{\mathbb{R}}$ (normalized as in §13.2 using the canonical inner product). For each factor $\mathbb{R}^{\times}$ of $K_{\mathbb{R}}^{\times} = (\mathbb{R}^{\times})^r \times (\mathbb{C}^{\times})^s$ we have

$$\begin{aligned} \mathbb{R}^{\times} &\to \mathbb{R} \times \{\pm 1\} \\ x &\mapsto (\log|x|, \operatorname{sgn} x) \\ \pm e^{\ell} &\leftarrow\!\shortmid (\ell, \pm 1) \\ dx &\mapsto e^{\ell} d\ell \mu_{\{\pm 1\}}, \end{aligned}$$

where dx and $d\ell$ denote the standard Lebesgue measures and $\mu_{\{\pm 1\}}$ is just the counting measure on the discrete set $\{\pm 1\}$. For each factor $\mathbb{C}^\times$,

$$\begin{aligned}
\mathbb{C}^\times &\to \mathbb{C} \times [0, 2\pi)\\
z &\mapsto (2\log|z|, \arg z)\\
e^{\ell/2} &\leftarrow\!\shortmid (\ell, \theta)\\
2dA &\mapsto 2e^{\ell/2}d(e^{\ell/2})d\theta = e^\ell d\ell d\theta,
\end{aligned}$$

where dA is the standard Lebesgue measure on $\mathbb{C}$ (so $2dA$ is the measure on $\mathbb{C}^\times$ as a component of $K_\mathbb{R}^\times$ under the Haar measure μ on $K_\mathbb{R}^\times \subseteq K_\mathbb{R}$), and $d\ell$ and $d\theta$ are the usual Lebesuge measures on $\mathbb{R}$ and $[0, 2\pi)$, respectively. We therefore have

$$\begin{aligned}
K_\mathbb{R}^\times &\xrightarrow{\sim} \mathbb{R}^{r+s} \times \{\pm 1\}^r \times [0, 2\pi)^s\\
\mu &\mapsto e^{\mathrm{T}(x)}\mu_{\mathbb{R}^{r+s}}\mu_{\{\pm 1\}}^r\mu_{[0,2\pi)}^s
\end{aligned}$$

We now make one further change of coordinates

$$\begin{aligned}
\mathbb{R}^{r+s} &\to \mathbb{R}^{r+s-1} \times \mathbb{R}\\
x = (x_1, \ldots, x_{r+s}) &\mapsto (x_1, \ldots, x_{r+s-1}, y := \mathrm{T}(x))\\
e^{\mathrm{T}(x)}\mu_{\mathbb{R}^{r+s}} &\mapsto e^y\mu_{\mathbb{R}^{r+s-1}}dy
\end{aligned}$$

The map $\pi\colon \mathbb{R}^{r+s} \to \mathbb{R}^{r+s-1}$ is just the coordinate projection, and the measure of $\pi(R)$ in $\mathbb{R}^{r+s-1}$ is, by definition, the regulator R_K (see Definition 14.10).

The Log map gives us a bijection

$$\begin{aligned}
F_{\leq 1} &\xrightarrow{\sim} R + (-\infty, 0]\left(\frac{1}{n}, \ldots, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{2}{n}\right),\\
x = |N(x)|^{1/n}\sigma(x) &\mapsto \log\sigma(x) + \log|\mathrm{N}(x)|\left(\frac{1}{n}, \ldots, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{2}{n}\right).
\end{aligned}$$

Thus the coordinate $y \in (-\infty, 0]$ is given by $y = \mathrm{T}(\operatorname{Log} x) = \log|\mathrm{N}(x)|$, and we can view $F_{\leq 1}$ as a union of cosets of $\operatorname{Log}^{-1}(R)$ parameterized by $e^y = |\mathrm{N}(x)| \in (0, 1]$.

Under our change of coordinates we thus have

$$\begin{aligned}
K_\mathbb{R}^\times &\xrightarrow{\sim} \mathbb{R}^{r+s-1} \times \mathbb{R} \times \{\pm 1\}^r \times [0, 2\pi)^s\\
F_{\leq 1} &\to \pi(R) \times (-\infty, 0] \times \{\pm 1\}^r \times [0, 2\pi)^s
\end{aligned}$$

Since $R_K = \mu_{\mathbb{R}^{r+s-1}}(\pi(R))$, we have

$$\begin{aligned}
\mu(F_{\leq 1}) &= \int_{-\infty}^0 e^y R_K 2^r(2\pi)^s dy\\
&= 2^r(2\pi)^s R_K.
\end{aligned}$$

Plugging this into (3) yields

$$\#(F_{\leq t} \cap \mathcal{O}_K) = \left(\frac{2^r(2\pi)^s R_K}{|\operatorname{disc}\mathcal{O}_K|^{1/2}}\right)t + O\big(t^{1-1/n}\big). \tag{4}$$

18.4 Proof of the analytic class number formula

We are now ready to prove the analytic class number formula. Our main tool is the following theorem, which uses our analysis in the previous section to give a precise asymptotic estimate on the number of ideals of bounded norm.

Theorem 18.9. *Let K be a number field of degree $n = r + 2s$ with r real and s complex places. As $t \to \infty$, the number of nonzero ideals $I \subseteq \mathcal{O}_K$ of absolute norm $N(I) \le t$ is*

$$\left(\frac{2^r(2\pi)^s h_K R_K}{w_k|D_K|^{1/2}}\right) t + O(t^{1-1/n}),$$

where $h_K = \#\operatorname{cl}\mathcal{O}_K$ is the class number, $R_K := \operatorname{covol}(\mathcal{O}_K^\times)$ is the regulator, $w_K : \#(\mathcal{O}_K)^\times_{\rm tors}$ is the number of roots of unity in K, and $D_K := \operatorname{disc}\mathcal{O}_K$ is the discriminant.

Proof. In order to count nonzero ideals $I \subseteq \mathcal{O}_K$ of norm $N(I) \le t$ we will group them by ideal class. For the trivial class, we just need to count nonzero principal ideals (α), equivalently, the number of nonzero $\alpha \in \mathcal{O}_K$ with $\mathrm{N}(\alpha) \le t$, modulo the unit group $\mathcal{O}_K^\times$. Dividing (4) by w_K to account for the w_K-to-1 map

$$F_{\le t} \cap \mathcal{O}_K \longrightarrow (K^\times_{\mathbb{R},\le t} \cap \mathcal{O}_K)/\mathcal{O}_K^\times$$

we obtain

$$\#\{(\alpha) \subseteq \mathcal{O}_K : \mathrm{N}(\alpha) \le t\} = \left(\frac{2^r(2\pi)^s R_K}{w_K|D_K|^{1/2}}\right) t + O(t^{1-1/n}).$$

To complete the proof we just need to show that we get the same answer for every ideal class; equivalently, that the nonzero ideals I of norm $N(I) \le t$ are equidistributed among ideal classes, as $t \to \infty$.

Let us fix an ideal class $c = [I_c]$, with $I_c \subseteq \mathcal{O}_K$ a nonzero (integral) ideal (recall that every ideal class contains an integral ideal, see Theorem 13.18). Multiplication by I_c defies a bijection

$$\begin{aligned}\{\text{ideals } I \in [I_c^{-1}] : N(I) \le t\} &\xrightarrow{\times I_c} \{\text{nonzero principal ideals } J \subseteq I_c : N(J) \le tN(I_c)\}\\ &\longleftrightarrow \{\text{nonzero } \alpha \in I_c : |N(\alpha)| \le tN(I_c)\}/\mathcal{O}_K^\times.\end{aligned}$$

Let S_c denote the RHS. Applying the exact same argument as in the case $I_c = \mathcal{O}_K$, we have

$$\begin{aligned}\#S_c &= \left(\frac{2^r(2\pi)^s R_K}{w_k \operatorname{covol}(I_c)}\right) tN(I_c) + O(t^{1-1/n})\\ &= \left(\frac{2^r(2\pi)^s R_K}{w_k \operatorname{covol}(\mathcal{O}_K)N(I_c)}\right) tN(I_c) + O(t^{1-1/n})\\ &= \left(\frac{2^r(2\pi)^s R_K}{w_k|D_K|^{1/2}}\right) t + O(t^{1-1/n}),\end{aligned}$$

which does not depend on the ideal class c. Summing over ideal classes then yields

$$\#\{\text{nonzero ideals } I \subseteq \mathcal{O}_K : N(I) \le t\} = \sum_{c\in\operatorname{cl}(\mathcal{O}_K)} \#S_c = \left(\frac{2^r(2\pi)^s h_K R_K}{w_K|D_K|^{1/2}}\right) t + O(t^{1-1/n}),$$

as claimed. □

To derive the analytic class number formula from Theorem 18.9 we need a couple of easy lemmas from complex analysis.

Lemma 18.10. *Let $a_1, a_2, \ldots$ be a sequence of complex numbers and let σ be a real number. Suppose that*

$$a_1 + \cdots + a_t = O(t^\sigma) \qquad (\text{as } t \to \infty).$$

Then the Dirichlet series $\sum a_n n^{-s}$ converges to a holomorphic function for $\operatorname{Re} s > \sigma$.

Proof. Let $A(x) := \sum_{0<n\le x} a_n$. Writing the Dirichlet sum as a Stieltjes integral (apply Corollary 17.36 with $f(n) = n^{-s}$ and $g(n) = a_n$), for $\operatorname{Re}(s) > \sigma$ we have

$$\begin{aligned}\sum_{n=1}^{\infty} a_n n^{-s} &= \int_{1^-}^{\infty} x^{-s}\, dA(x) \\ &= \frac{A(x)}{x^s}\Big|_{1^-}^{\infty} - \int_{1^-}^{\infty} A(x)\, dx^{-s} \\ &= (0-0) - \int_{1^-}^{\infty} A(x)(-sx^{-s-1})\, dx \\ &= s\int_{1^-}^{\infty} \frac{A(x)}{x^{s+1}}\, dx.\end{aligned}$$

Note that we used $|A(x)| = O(x^\sigma)$ and $\operatorname{Re}(s) > \sigma$ to conclude that $\lim_{x\to\infty} A(x)/x^s = 0$. For any $\epsilon > 0$, the integral on the RHS converges uniformly on $\operatorname{Re}(s) \ge \sigma + \epsilon$, thus the sum converges to a holomorphic function on $\operatorname{Re}(s) \ge \sigma + \epsilon$ for all $\epsilon > 0$, hence on $\operatorname{Re}(s) > \sigma$. □

Remark 18.11. The lemma gives us an *abscissa of convergence* σ for the Dirichlet series $\sum a_n n^{-s}$; this is analogous to the radius of convergence of a power series.

Lemma 18.12. *Let $a_1, a_2, \ldots$ be a sequence of complex numbers that satisfies*

$$a_1 + \cdots + a_t = \rho t + O(t^\sigma) \qquad (\text{as } t \to \infty)$$

for some $\sigma \in [0,1)$ and $\rho \in \mathbb{C}^\times$. Then the Dirichlet series $\sum a_n n^{-s}$ converges on $\operatorname{Re}(s) > 1$ and has a meromorphic continuation to $\operatorname{Re}(s) > \sigma$ that is holomorphic except for a simple pole at $s = 1$ with residue ρ.

Proof. Define $b_n := a_n - \rho$. Then $b_1 + \cdots + b_t = O(t^\sigma)$ and

$$\sum a_n n^{-s} = \rho \sum n^{-s} + \sum b_n n^{-s} = \rho\,\zeta(s) + \sum b_n n^{-s}.$$

We have already proved that the Riemann zeta function $\zeta(s)$ is holomorphic on $\operatorname{Re}(s) > 1$ and has a meromorphic continuation to $\operatorname{Re}(s) > 0$ that is holomorphic except for a simple pole at 1 with residue 1. By the previous lemma, $\sum b_n n^{-s}$ is holomorphic on $\operatorname{Re}(s) > \sigma$, and since $\sigma < 1$ it is holomorphic at $s = 1$. So the entire RHS has a meromorphic continuation to $\operatorname{Re}(s) > \sigma$ that is holomorphic except for the simple pole at 1 coming from $\zeta(s)$, and the residue at $s = 1$ is $\rho \cdot 1 + 0 = \rho$. □

We are now ready to prove the analytic class number formula.

Theorem 18.13 (Analytic Class Number Formula)**.** *Let K be a number field of degree n with r real and s complex places. The Dedekind zeta function $\zeta_K(z)$ extends to a meromorphic function on $\operatorname{Re}(z) > 1 - \frac{1}{n}$ that is holomorphic except for a simple pole at $z = 1$ with residue*

$$\lim_{z\to 1^+}(z-1)\zeta_K(z) = \rho_K := \frac{2^r(2\pi)^s h_K R_K}{w_K|D_K|^{1/2}},$$

where $h_K := \#\operatorname{cl}\mathcal{O}_K$ is the class number, R_K is the regulator, $w_K := \#(\mathcal{O}_K^\times)_{\mathrm{tors}}$ is the number of roots of unity in K, and $D_K := \operatorname{disc}\mathcal{O}_K$ is the discriminant.

Proof. We have

$$\zeta_K(z) = \sum_I N(I)^{-s} = \sum_{m\geq 1} a_m m^{-s},$$

where I ranges over nonzero ideals of $\mathcal{O}_K$, and $a_m := \#\{I : N(I) = m\}$. By Theorem 18.9,

$$a_1 + \cdots + a_t = \#\{I : N(I) \leq t\} = \rho t + O(t^{1-1/n}) \qquad (\text{as } t \to \infty)$$

Applying Lemma 18.12 with $\sigma = 1 - 1/n$, we see that $\zeta_K(z) = \sum a_m m^{-s}$ extends to a meromorphic function on $\operatorname{Re}(z) > 1 - 1/n$ that is holomorphic except for a simple pole at $z = 1$ with residue ρ_K. □

Remark 18.14. As previously noted, Hecke proved that $\zeta_K(z)$ extends to a meromorphic function on $\mathbb{C}$ with no poles other than the simple pole at $z = 1$, and it satisfies a functional equation. If we define the *gamma factors*

$$\Gamma_{\mathbb{R}}(z) := \pi^{-z/2}\Gamma\left(\tfrac{z}{2}\right), \qquad \text{and} \qquad \Gamma_{\mathbb{C}}(z) := (2\pi)^{-z}\Gamma(z),$$

and define the *completed zeta function*

$$\xi_K(z) := |D_K|^{z/2}\Gamma_{\mathbb{R}}(z)^r\Gamma_{\mathbb{C}}(z)^s\zeta_K(z),$$

where r and s are the number of real and complex places of K, respectively, then $\xi_K(z)$ is holomorphic except for simple poles at $z = 0, 1$ and satisfies the *functional equation*

$$\xi_K(z) = \xi_K(1-z).$$

In the case $K = \mathbb{Q}$, we have $r = 1$ and $s = 0$, so

$$\xi_{\mathbb{Q}}(z) = \Gamma_{\mathbb{R}}(z)\zeta(z) = \pi^{z/2}\Gamma(\tfrac{z}{2})\zeta_{\mathbb{Q}}(z),$$

which is precisely the completed zeta function $Z(z)$ we defined for the Riemann zeta function $\zeta(z) = \zeta_{\mathbb{Q}}$ in Lecture 16.