Bachelor's project

NTNU
Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Mathematical Sciences

Markus Valås Hagen

Dedekind zeta functions

Bachelor's project in BMAT
Supervisor: Petter Andreas Bergh
May 2021

NTNU
Norwegian University of
Science and Technology

Markus Valås Hagen

Dedekind zeta functions

Bachelor's project in BMAT
Supervisor: Petter Andreas Bergh
May 2021

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Mathematical Sciences

Dedekind zeta functions

Markus Valås Hagen

May 18, 2021

Introduction

In 1847 the French mathematician Gabriel Lamé proposed a proof of Fermat's Last Theorem to the Paris Academy. However the theorem was not established until recently by Wiles et.al. So Lamé was wrong. His blunder was to assume that the cyclotomic integers $\mathbb{Z}[\zeta_p]$ had unique factorization for all primes p, something Kummer had proven was not true even before Lamé proposed his proof.

The ring $\mathbb{Z}[\zeta_p]$ is the *ring of integers* of the cyclotomic extension $\mathbb{Q}(\zeta_p)$ of $\mathbb{Q}$. To any number field K, which we will very soon define, we can associate a ring of integers $\mathcal{O}_K$ which mimics how $\mathbb{Z}$ lies in $\mathbb{Q}$. The failure of unique factorization in such rings is measured by what is called the *ideal class group.* Towards the end of this thesis we will give a formula for calculating the class number - the order of the ideal class group. Although such a ring of integers may lack unique factorization, they have the amazing property that every non-zero ideal factorize uniquely into prime ideals. At the heart of this thesis is the study of $\mathcal{O}_K$.

The organization of this thesis is roughly as follows:

- In chapter 1 we introduce number fields and their rings of integers and prove some basic properties about them.
- In chapter 2 we widen our focus and prove that every *Dedekind domain* has the property of non-zero ideals factorizing uniquely into prime ideals, before specializing to $\mathcal{O}_K$ again as a specific example of a Dedekind domain.
- In chapter 3 we introduce some geometric methods, especially lattice theory, to prove that the ideal class group is finite as well as study the structure of the group of units of $\mathcal{O}_K$. Finally we use the geometric methods introduced to describe how ideals in $\mathcal{O}_K$ are distributed with respect to their ideal norm.
- In chapter 4 we introduce Dedekind's zeta function ζ_K, *the* generalization of Riemann's zeta function to a number field K. We prove the class number formula, which relates the residue of ζ_K at $s = 1$ to various invariants of a number field, one of them being the class number. With the theory of Dedekind zeta functions we give a proof of the non-vanishing of $L(1, \chi)$ for a non-trivial Dirichlet character χ and thus deduce Dirichlet's theorem on primes in arithmetic progressions.

I want to thank my advisor Petter Andreas Bergh for suggesting the topic of this thesis, which I have grown to appreciate a lot. I also want to thank him for the weekly meetings we have had over the past year. They have given me good advices, not only in the work for this thesis, but also for the mathematical journey after finishing my bachelor degree. I finally want to thank Wojtek Wawrów for proof reading the entire thesis.

Contents

Chapter 1

Number fields and rings of integers

In this chapter we introduce number fields, K, and related concepts such as the norm, trace and discriminant. Then we introduce the ring of integers of K, $\mathcal{O}_K$, which consists of all elements in K that satisfy a monic polynomial with integer coefficients. $\mathcal{O}_K$ will be of key interest throughout this whole thesis. One can say that the study of the arithmetic of $\mathcal{O}_K$ is the heart of algebraic number theory.

1.1 Number fields

Definition 1.1.1. A number field $K \subseteq \mathbb{C}$ is a finite (and hence algebraic) field extension of $\mathbb{Q}$.

The elements in K that we will be most interested in are those satisfying a monic polynomial with integer coefficients. We will study those in more detail in chapter 1.2, and for the time being just define algebraic integers as we will need that notion for this section.

Definition 1.1.2. Let $\alpha \in \mathbb{C}$. Then α is said to be an algebraic number if $\exists p(x) \in \mathbb{Q}[x]$, non-zero, such that $p(\alpha) = 0$. α is said to be an algebraic integer if $\exists p(x) \in \mathbb{Z}[x]$ which is monic so that $p(\alpha) = 0$. The set of all algebraic integers will be denoted by $\mathbb{A}$.

Let $n = [K : \mathbb{Q}]$. The primitive element theorem [4, Theorem 16.5.2] tells us that there exists some $\alpha \in K$ so that

$$K = \mathbb{Q}(\alpha) = \left\{a_0 + a_1\alpha + \cdots + a_{n-1}\alpha^{n-1} : a_i \in \mathbb{Q}\right\}.$$

In other words, $\{1, \alpha, \ldots, \alpha^{n-1}\}$, is a $\mathbb{Q}$-basis for K. Throughout this thesis we will be interested in the embeddings of K into $\mathbb{C}$ in various settings. Hence we start by studying those embeddings.

Theorem 1.1.3. *Let $K = \mathbb{Q}(\alpha)$ be a number field, and let $p(x)$ be the minimal polynomial of α over $\mathbb{Q}$ with roots $\alpha_1 = \alpha, \alpha_2, \ldots, \alpha_n$. Then there are n embeddings σ_i of K into $\mathbb{C}$, each defined by $\sigma_i(\alpha) = \alpha_i$.*

Proof. Since $p(x)$ is the minimal polynomial of α over $\mathbb{Q}$, $p(x)$ has distinct roots. An embedding of K into $\mathbb{C}$ must fix $\mathbb{Q}$, and will hence send roots of $p(x)$ to roots of $p(x)$. Hence there are at most n possibilites for where α can be sent by such an embedding. Conversely, as α, α_j are roots of the same irreducible polynomial we have

$$\mathbb{Q}(\alpha) \cong \mathbb{Q}[x]/(p(x)) \cong \mathbb{Q}(\alpha_j)$$

where this isomorphism is simply given by $\alpha \mapsto \alpha_j$. Hence there are exactly n embeddings of K into $\mathbb{C}$. □

Now that we know how the embeddings of $\mathbb{Q}(\alpha)$ into $\mathbb{C}$ behave, we introduce three important concepts that is defined via those embeddings.

Definition 1.1.4. Let K be a number field with $[K:\mathbb{Q}] = n$, and let $\sigma_1, \dots, \sigma_n$ be all the embeddings of K into $\mathbb{C}$. Then for any element $\beta \in K$ we define the **norm** as

$$N(\beta) = \sigma_1(\beta) \cdots \sigma_n(\beta),$$

and the **trace** as

$$T(\alpha) = \sigma_1(\beta) + \cdots + \sigma_n(\beta).$$

For any n-tuple $(\beta_1, \dots, \beta_n) \in K^n$ the **discriminant** is defined as

$$\operatorname{disc}(\beta_1, \dots, \beta_n) = \det \begin{pmatrix} \sigma_1(\beta_1) & \sigma_1(\beta_2) & \cdots & \sigma_1(\beta_n) \\ \sigma_2(\beta_1) & \sigma_2(\beta_2) & \cdots & \sigma_2(\beta_n) \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_n(\beta_1) & \sigma_n(\beta_2) & \cdots & \sigma_n(\beta_n) \end{pmatrix}^2 = |\sigma_i(\beta_j)|^2$$

Since interchanging columns only alter the sign of the determinant, the discriminant is well defined. The last equality in the definition of the discriminant is the notation we will use. It follows readily from the definition that $T(\alpha+\beta) = T(\alpha)+T(\beta)$ and $N(\alpha\beta) = N(\alpha)N(\beta)$ for $\alpha, \beta \in K$. We continue by examining some of the properties of the norm and trace, before we turn to the discriminant. We first need some lemmata about polynomials. Recall the definition of a primitive polynomial $f \in \mathbb{Z}[x]$: f is primitive if the gcd of its coefficient is 1.

Lemma 1.1.5. (Gauss' lemma) *If $f, g \in \mathbb{Z}[x]$ are primitive, so is their product fg.*

A proof of this can be found in [4, Lemma 11.4.2].

Lemma 1.1.6. *Let $f(x) \in \mathbb{Z}[x]$ be monic. Suppose there are monic polynomials $f_1, \dots, f_n \in \mathbb{Q}[x]$ so that $\prod_{i=1}^n f_i(x) = f(x)$. Then $f_i(x) \in \mathbb{Z}[x]$ for all $1 \leq i \leq n$.*

Proof. We first claim that there is a $c_i \in \mathbb{N}$ such that $c_i f_i \in \mathbb{Z}[x]$ is primitive. Write $f_i(x) = \frac{a_0}{b_0} + \frac{a_1}{b_1}x + \cdots + x^k$. There is clearly a least $c_i \in \mathbb{N}_{>0}$ so that $c_i f_i \in \mathbb{Z}[x]$, so suppose this least c_i does not make $c_i f_i$ primitive. Then there is $d > 1$ so that $d \mid c_i \frac{a_j}{b_j}$ for all $1 \leq j \leq k$ so we find $m_j \in \mathbb{N}_{>0}$ such that $dm_j = c_i \frac{a_j}{b_j}$. Then

$$\begin{aligned} c_i f_i &= c_i \frac{dm_0}{c_i} + c_i \frac{dm_1}{c_i} x + \cdots + dm_k x^k \\ &= d\left(m_0 + m_1 x + \cdots + m_k x^k\right) \end{aligned}$$

Since f_i is monic $d \mid c_i$. Since $d > 1$, $\frac{c_i}{d} < c_i$, so this contradicts the minimality of c_i. Hence $c_i f_i$ is primitive for all i, and we thus get by Gauss' lemma that $c_1 f_1 c_2 f_2 \cdots c_n f_n = (c_1 c_2 \cdots c_n)(f_1 \cdots f_n)$ is primitive. This forces $c_1 c_2 \cdots c_n = \pm 1$, so $c_i = \pm 1$, so f_i is primitive. □

Lemma 1.1.7. *An algebraic number α is an algebraic integer if and only if its minimal polynomial over $\mathbb{Q}$ has coefficients in $\mathbb{Z}$.*

Proof. Let $p(x)$ be the minimal polynomial of α over $\mathbb{Q}$. Then it is irreducible over $\mathbb{Q}$ and is monic. If it has coefficients in $\mathbb{Z}$ it is by definition an algebraic integer. Conversely, if α is an algebraic integer, there is a monic polynomial $f(x) \in \mathbb{Z}[x]$ such that $f(\alpha) = 0$. By minimality of $p(x)$, we have in $\mathbb{Q}[x]$ that $p(x) \mid f(x)$, and as both $f(x), p(x)$ are monic, there must be monic $k(x) \in \mathbb{Q}[x]$ such that $p(x)k(x) = f(x)$. Lemma 1.1.6, then gives that $p(x) \in \mathbb{Z}[x]$, as we wanted to show. □

The following lemma, paired up with Galois theory, is very useful when trying to prove that some polynomial has rational coefficients.

Lemma 1.1.8. *Let $f \in \mathbb{C}[x]$ be a polynomial such that $f(q) \in \mathbb{Q}$ $\forall q \in \mathbb{Q}$. Then $f(x) \in \mathbb{Q}[x]$.*

Proof. Let f have degree n. We construct a polynomial $g(x) \in \mathbb{Q}[x]$ of degree n that agrees with f in $n+1$ points. Then it follows that $f - g$, a polynomial of degree n, has $n+1$ roots, and hence has to be 0. Moreover, in this case $f = g$. Now $g(x)$ is simply the Lagrange interpolation of f in the $n+1$ points, $(0, f(0)), (1, f(1)), \ldots, (n, f(n))$:

$$g(x) = \sum_{j=0}^{n} \left(f(j) \prod_{0 \leq m \leq n, m \neq j} \frac{x - m}{j - m} \right)$$

By assumption, $g(x) \in \mathbb{Q}[x]$, which proves the statement. □

Theorem 1.1.9. *Let K be a number field and $\alpha \in K$. Then $N(\alpha), T(\alpha) \in \mathbb{Q}$. Furthermore if α is an algebraic integer then $N(\alpha), T(\alpha) \in \mathbb{Z}$.*

Proof. We will prove the first statement using Galois theory. To that end, let $K = \mathbb{Q}(\theta)$ with $p(x)$ the minimal polynomial of θ. Furthermore, let L be the splitting field of $p(x)$, so that $L/\mathbb{Q}$ is a Galois extension. We introduce the auxiliary polynomial f_α defined as

$$f_\alpha(x) = (x - \sigma_1(\alpha)) \cdots (x - \sigma_n(\alpha))$$

Observe that the constant term in $f_\alpha(x)$ is $(-1)^j N(\alpha)$ and the coefficient in front of x^{n-1} is $-T(\alpha)$ for some integer j. Hence, if we are able to prove that $f_\alpha(x) \in \mathbb{Q}[x]$, the first claim will follow. As $\alpha \in K = \mathbb{Q}(\theta)$, there is a polynomial $r(x) \in \mathbb{Q}[x]$ so that $\alpha = r(\theta)$. By Theorem 1.1.3 it follows that $\sigma_i(\alpha) = \sigma_i(r(\theta)) = r(\sigma_i(\theta)) = r(\theta_i)$ where $\theta_i = \sigma_i(\theta)$ as in Theorem 1.1.3. Now let $\tau \in \mathrm{Gal}(L/\mathbb{Q})$. Then if $x \in \mathbb{Q}$ we have

$$\tau(f_\alpha(x)) = \tau\left(\prod_{i=1}^{n}(x - r(\theta_i))\right) = \prod_{i=1}^{n}(x - \tau(r(\theta_i))) = \prod_{i=1}^{n}(x - r(\tau(\theta_i)))$$

Now τ will send a root of $p(x)$ to a root of $p(x)$ and as it is injective it will actually permute those. As the product runs over all the roots of $p(x)$, this implies that $\tau(f_\alpha(x)) = f_\alpha(x)$ and as τ was arbitrary in $\mathrm{Gal}(L/\mathbb{Q})$ it follows by the fundamental theorem of Galois theory that $f_\alpha(x) \in \mathbb{Q}[x]$. To prove the second claim, we first show that $f_\alpha(t) = (p_\alpha(t))^k$, where $k \in \mathbb{N}$ and p_α is the minimal polynomial of α over $\mathbb{Q}$. Since one of the embeddings of K into $\mathbb{C}$ is identity on K, $f_\alpha(\alpha) = 0$. Hence $p_\alpha \mid f_\alpha$, and we can write $(p_\alpha)^s h = f_\alpha$ where $h \in \mathbb{Q}[x]$ is relatively prime to p_α. Suppose h is not constant. Then there must be some $\sigma_i(\alpha)$ so that $h(\sigma_i(\alpha)) = 0$. Now $\alpha = r(\theta)$ for some $r(x) \in \mathbb{Q}[x]$ so that $h(\sigma_i(\alpha)) = h(r(\theta_i))$. That is θ_i is a root of $h(r(x))$. Let p_θ be the minimal polynomial of θ - then it is also the minimal polynomial of θ_i and so $p_\theta(x) \mid h(r(x))$. Hence $h(r(\theta)) = 0$ as well, and so $h(\alpha) = h(r(\theta)) = 0$, but then h and p_α shares a root, which contradicts that they are relatively prime. Since f_α is monic, $h = 1$. For the final step: let α be an algebraic integer. Then its mimimal polynomial p_α over $\mathbb{Q}$ has coefficients $\mathbb{Z}$, so f_α must also have integer coefficients. This finishes the proof. □

Example 1.1.10. In the field $\mathbb{Q}(i)$, the norm and trace is given by

$$N(\alpha + \beta i) = (\alpha + \beta i)(\alpha - \beta_i) = \alpha^2 + \beta^2 \qquad T(\alpha + \beta i) = 2\alpha$$

We introduce some notation for the sake of space: $[a_{ij}]$ will denote the matrix having a_{ij} in the ith row, jth column. In the same spirit $|a_{ij}|$ will denote the determinant of the same matrix. Discriminant and trace have a close relation.

Theorem 1.1.11. *Let K be a number field, $\alpha_1, \ldots, \alpha_n \in K$. Then*

$$\mathrm{disc}(\alpha_1, \ldots, \alpha_n) = |T(\alpha_i \alpha_j)|$$

Proof. Firstly $[\sigma_i(\alpha_j)]^T[\sigma_i(\alpha_j)] = [\sigma_1(\alpha_i)\sigma_1(\alpha_j) + \cdots + \sigma_n(\alpha_i)\sigma_n(\alpha_j)] = [T(\alpha_i\alpha_j)]$. For a square matrix, we have $\det(A^T) = \det(A)$, and so $|T(\alpha_i\alpha_j)| = |\sigma_i(\alpha_j)|^2 = \mathrm{disc}(\alpha_1, \ldots, \alpha_n)$. □

Corollary 1.1.12. *With the notation above,* $\mathrm{disc}(\alpha_1, \ldots, \alpha_n) \in \mathbb{Q}$*, and if* $\alpha_1, \ldots, \alpha_n \in \mathbb{A}$*, then* $\mathrm{disc}(\alpha_1, \ldots, \alpha_n) \in \mathbb{Z}$.

Proof. By Theorem 1.1.9, $T(\alpha_i\alpha_j) \in \mathbb{Q}$, so $|T(\alpha_i\alpha_j)| \in \mathbb{Q}$. As we will see in Lemma 1.2.4, $\mathbb{A}$ is a ring, thus products and sums of algebraic integers are still algebraic integers. If all $\alpha_1, \dots, \alpha_n \in \mathbb{A}$, then $T(\alpha_i\alpha_j) \in \mathbb{Z}$ for all i, j by Theorem 1.1.9 and so $|T(\alpha_i\alpha_j)| \in \mathbb{Z}$ in that case as well. By Theorem 1.1.11 this is enough. □

We end this section with determining when the discriminant is zero. Before proving this we need a lemma from linear algebra, on the Vandermonde determinant.

Lemma 1.1.13. *Let R be a commutative ring and $a_1, \dots, a_n \in R$. Then*

$$\begin{vmatrix} 1 & a_1 & \cdots & a_1^{n-1} \\ 1 & a_2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & \cdots & a_n^{n-1} \end{vmatrix} = \prod_{1 \le r < s \le n} (a_s - a_r)$$

Proof. The proof is by induction. The base case $n = 2$ is trivial, so assume the statement holds for $n = k$. Let $f \in R[x]$ be any monic polynomial of degree k. Since elementary column operations do not change the value of the determinant we have

$$\begin{vmatrix} 1 & a_1 & \cdots & a_1^{k} \\ 1 & a_2 & \cdots & a_2^{k} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & a_{k+1} & \cdots & a_{k+1}^{k} \end{vmatrix} = \begin{vmatrix} 1 & a_1 & \cdots & f(a_1) \\ 1 & a_2 & \cdots & f(a_2) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & a_{k+1} & \cdots & f(a_{k+1}) \end{vmatrix}$$

Now we choose $f(x) = (x - a_1)(x - a_2) \cdots (x - a_k)$ which is indeed a monic polynomial in $R[x]$. This yields

$$\begin{vmatrix} 1 & a_1 & \cdots & a_1^{k} \\ 1 & a_2 & \cdots & a_2^{k} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & a_{k+1} & \cdots & a_{k+1}^{k} \end{vmatrix} = \begin{vmatrix} 1 & a_1 & \cdots & 0 \\ 1 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & a_{k+1} & \cdots & f(a_{k+1}) \end{vmatrix} = (-1)^{2k+2} f(a_{k+1}) \prod_{1 \le r < s \le n} (a_s - a_r)$$

□

Theorem 1.1.14. *Let K be a number field and $(\alpha_1, \dots, \alpha_n) \in K^n$. Then* $\operatorname{disc}(\alpha_1, \dots, \alpha_n) =$ 0 *if and only if $\{\alpha_1, \dots, \alpha_n\}$ is a linearly dependent set over $\mathbb{Q}$.*

Proof. Assume first that $\{\alpha_1, \dots, \alpha_n\}$ is a linearly dependent set over $\mathbb{Q}$. Then there exists $a_1, \dots, a_n \in \mathbb{Q}$, not all zero, such that $a_1\alpha_1 + \cdots + a_n\alpha_n = 0$, and hence for any $i = 1, \dots, n$ we have $a_1\sigma_i(\alpha_1) + \cdots + a_n\sigma_i(\alpha_n) = 0$. Thus

$$a_1 \begin{pmatrix} \sigma_1(\alpha_1) \\ \vdots \\ \sigma_n(\alpha_1) \end{pmatrix} + \cdots + a_n \begin{pmatrix} \sigma_1(\alpha_n) \\ \vdots \\ \sigma_n(\alpha_n) \end{pmatrix} = 0$$

so the columns of the matrix $[\sigma_i(\alpha_j)]$ are linearly dependent, and hence the determinant is zero. We prove the converse contrapositively. Let $K = \mathbb{Q}(\theta)$ - then $\{1, \theta, \ldots, \theta^{n-1}\}$ is a basis for K over $\mathbb{Q}$. Another basis is $\{\alpha_1, \ldots, \alpha_n\}$. Hence there is an invertible matrix $[c_{ij}]$ such that $\alpha_i = \sum_{j=0}^{n-1} c_{ij}\theta^j$ and hence $[\sigma_i(\alpha_j)] = [\sigma_i(\theta^j)][c_{ij}]^T$. Taking determinants on both sides and remembering $|c_{ij}| = |c_{ji}|$ yields $\mathrm{disc}(\alpha_1, \ldots, \alpha_n) = |c_{ij}|^2 \mathrm{disc}(1, \theta, \ldots, \theta^{n-1})$. Hence it is enough to prove that $\mathrm{disc}(1, \theta, \ldots, \theta^{n-1}) \neq 0$. But this is the square of a Vandermonde determinant and hence if we denote $\theta_i = \sigma_i(\theta)$, then

$$\mathrm{disc}(1, \theta, \ldots, \theta^{n-1}) = \left(\prod_{1 \leq i < j \leq n} (\theta_j - \theta_i) \right)^2$$

Clearly this product is 0 if and only if $\theta_j = \theta_i$ for $i \neq j$, but K is a seperable extension of $\mathbb{Q}$ and hence any minimal polynomial has no double roots. □

Remark: Let $\{\alpha_1, \ldots, \alpha_n\}$ and $\{\beta_1, \ldots, \beta_n\}$ be two bases for K over $\mathbb{Q}$. Furthermore let $\alpha_i = \sum_{j=1}^{n} c_{ij}\beta_i$. Then a part of the argument over is easily generalized to show that $\mathrm{disc}(\alpha_1, \ldots, \alpha_n) = |c_{ij}|^2 \mathrm{disc}(\beta_1, \ldots, \beta_n)$. We will use this relation later.

1.2 Ring of integers of K

Central to this chapter will be the notion of free abelian groups.

Definition 1.2.1. A free abelian group is a free $\mathbb{Z}$-module.

Lemma 1.2.2. *Let G be a free abelian group of rank n, and let H be a subgroup of G. Then H is free abelian of rank $\leq n$.*

Proof. The proof is by induction. If G is of rank $n = 1$, then $G \cong \mathbb{Z}$. A subgroup of $\mathbb{Z}$ is of the form $n\mathbb{Z}$, which is also free abelian. Hence any subgroup H of G is free abelian of rank at most one. Now assume the result holds for $n - 1$. Let $G \cong \mathbb{Z} \times \cdots \times \mathbb{Z}$ (n times). Define the homomorphism

$$\pi : \mathbb{Z} \times \cdots \times \mathbb{Z} \to \mathbb{Z} \qquad (x_1, \ldots, x_n) \mapsto x_1$$

Then $\ker(\pi) = \mathbb{Z}^{n-1}$. Let H be a subgroup of G (viewed as a subgroup of $\mathbb{Z}^n$). Then $H \cap \ker(\pi)$ is a subgroup of $\mathbb{Z}^{n-1}$ and hence by the induction hypothesis is free abelian of rank $\leq n-1$. Now the image of H, $\pi(H)$, is a subgroup of $\mathbb{Z}$, and hence either infinite cyclic or $\{0\}$. If $\pi(H) = \{0\}$, then $H \subseteq \ker(\pi)$, and hence $H = H \cap \ker(\pi)$, and hence H is free abelian of rank $\leq n - 1$. For the other case, fix some $h \in H$ so that $\pi(h)$ generates $\pi(H)$. We want to prove that $H = \mathbb{Z}h \oplus (H \cap \ker \pi)$. Let $x \in H$, then $\pi(x) = k\pi(h) = \pi(kh)$ for some $k \in \mathbb{Z}$. Then clearly $\pi(x - kh) = \pi(x) - \pi(kh) = 0$, so $x - kh \in H \cap \ker(\pi)$. Hence for any $x \in H$, $x = kh + (x - kh) \in \mathbb{Z}h + (H \cap \ker(\pi))$. If $y \in \mathbb{Z}h \cap (H \cap \ker(\pi))$ then $y = kh \in \ker(\pi)$, that is $0 = \pi(y) = \pi(kh) = k\pi(h)$ and hence $k = 0$ so $y = 0$. As $\mathbb{Z}h$ and $(H \cap \ker(\pi))$ clearly are subgroups of H, this shows that $H = \mathbb{Z}h \oplus (H \cap \ker \pi)$ - a direct sum of two free abelian groups, one of rank 1, the other of rank $\leq n-1$. Hence H is free abelian of rank $\leq n$. □

The next thing we want to establish is that the algebraic integers $\mathbb{A}$ actually form a ring! We start by finding an equivalent definition of algebraic integers, that will be easier to work with.

Lemma 1.2.3. *An element $\alpha \in \mathbb{C}$ is an algebraic integer if and only if $\Gamma_\alpha = (1, \alpha, \alpha^2, \dots)$ is a finitely generated abelian group.*

Proof. If Γ_α is finitely generated, then there exists some $n \in \mathbb{N}$ such that $\alpha^n = a_0 + a_1\alpha + \cdots + a_{n-1}\alpha^{n-1}$ for $a_i \in \mathbb{Z}$, but this is exactly saying that α satisfies some monic polynomial with integer coefficients. Conversely, suppose $\alpha \in \mathbb{C}$ is an algebraic integer. Then for some $n \in \mathbb{N}$ we have $\alpha^n = a_0 + a_1\alpha + \cdots + a_{n-1}\alpha^{n-1}$ where $a_i \in \mathbb{Z}$. We claim that $(1, \alpha, \dots, \alpha^{n-1})$ genereates Γ_α, and want to show this by induction. The base case $m = 0$ is clear. Assume $\alpha^{n+m-1} = b_0 + b_1\alpha + \cdots + b_{n-1}\alpha^{n-1}$, $b_i \in \mathbb{Z}$. Then

$$\begin{aligned}
\alpha^{n+m} &= \alpha\alpha^{n+m-1} \\
&= \alpha(b_0 + b_1\alpha + \cdots + b_{n-1}\alpha^{n-1}) = b_0\alpha + b_1\alpha^2 + \cdots + b_{n-1}\alpha^n \\
&= b_0\alpha + b_1\alpha^2 + \cdots + b_{n-1}(a_0 + a_1\alpha + \cdots + a_{n-1}\alpha^{n-1}) \\
&= b_{n-1}a_0 + (b_0 + b_{n-1}a_1)\alpha + \cdots + b_{n-1}a_{n-1}\alpha^{n-1} \in (1, \alpha, \dots, \alpha^{n-1})
\end{aligned}$$

Which finishes the proof by induction. □

We are now just one lemma away from a very important definition in this thesis.

Lemma 1.2.4. *The algebraic integers, $\mathbb{A}$, form a subring of $\mathbb{C}$.*

Proof. $\mathbb{A}$ is non-empty as $1 \in \mathbb{A}$. What is left to show is that $\forall \alpha, \beta \in \mathbb{A}$, $\alpha - \beta \in \mathbb{A}$ and $\alpha\beta \in \mathbb{A}$. By Lemma 1.2.3 it is enough to show $\Gamma_{\alpha-\beta}$ and $\Gamma_{\alpha\beta}$ are finitely generated. Since α, β are algebraic integers, we know $\Gamma_\alpha, \Gamma_\beta$ is finitely generated, say by $\{1, \alpha, \dots, \alpha^m\}$ and $\{1, \beta, \dots, \beta^k\}$ respectively. Then $\{\alpha^i\beta^j\}_{0\leq i\leq m, 0\leq j\leq k}$ generates a (finitely generated) abelian group where both $\Gamma_{\alpha-\beta}$ and $\Gamma_{\alpha\beta}$ are contained. Since a subgroup of a finitely generated abelian group is finitely generated[1], we are done. □

As an immediate consequence of the lemma, $\mathbb{A} \cap K$ becomes a ring for any number field K.

Definition 1.2.5. The ring of integers of (a number field) K is defined as $\mathbb{A} \cap K$, and is denoted by $\mathcal{O}_K$. In other words, $\mathcal{O}_K$ consists of every element in K that satifies a monic polynomial with integer coefficients.

$\mathcal{O}_K$ has a very nice structure. Most importantly is perhaps the fact that every proper non-zero ideal in $\mathcal{O}_K$ factors into prime ideals uniquely, so that ideals in $\mathcal{O}_K$ behave almost like ordinary numbers. This fact isn't just specific to $\mathcal{O}_K$ - domains where this fact hold are

[1] A proof of this fact can be obtained by a similar argument as in Lemma 1.2.2. For details see [9, Proposition 3.18]

known as Dedekind domains, and we will define those more specifically and study them in more detail in the next section. In the last part of this section we prove three central facts about $\mathcal{O}_K$, which we will see in the next section is the very definition of a Dedekind domain. The first fact we are going to prove is that $\mathcal{O}_K$ is Noetherian. We start with two lemmata.

Lemma 1.2.6. *Let α be an algebraic number. Then there exists $0 \neq m \in \mathbb{Z}$ such that $m\alpha$ is an algebraic integer.*

Proof. As α is an algebraic number, let $f(x) = \frac{a_0}{b_0} + \cdots + \frac{a_{n-1}}{b_{n-1}} x^{n-1} + x^n$ be its minimal polynomial over $\mathbb{Q}$, where without loss of generality, $\gcd(a_i, b_i) = 1$ for each i. Set $m = b_{n-1}^n b_{n-2}^n \cdots b_0^n$. Then

$$\begin{aligned} 0 = m \cdot 0 &= (b_{n-1}^n b_{n-2}^n \cdots b_0^n) \left(\frac{a_0}{b_0} + \cdots + \frac{a_{n-1}}{b_{n-1}} \alpha^{n-1} + \alpha^n \right) \\ &= a_0 b_0^{n-1} b_1^n \cdots b_{n-1}^n + a_1 b_0^{n-1} b_1^{n-2} \cdots b_{n-1}^{n-1} (b_0 b_1 \cdots b_{n-1}\alpha) + \cdots + (b_0 b_1 \cdots b_{n-1} \alpha)^n \\ &= a_0 b_0^{n-1} b_1^n \cdots b_{n-1}^n + a_1 b_0^{n-1} b_1^{n-2} \cdots b_{n-1}^{n-1} (m\alpha) + \cdots + (m\alpha)^n \end{aligned}$$

Hence $m\alpha$ is an algebraic integer. □

Lemma 1.2.7. *Let K be a number field with $[K : \mathbb{Q}] = n$. Then there exists a basis $\{\alpha_1, \ldots, \alpha_n\}$ consisting entirely of algebraic integers for K over $\mathbb{Q}$.*

Proof. Pick a basis for K over $\mathbb{Q}$, say $\{\beta_1, \ldots, \beta_n\}$. By Lemma 1.2.6 we find non-zero integers $m_1, \ldots, m_n$ such that $\alpha_i = m_i \beta_i$ is an algebraic integer. Now, for any $x \in K$:

$$x = a_1 \beta_1 + \cdots + a_n \beta_n = \frac{a_1}{m_1} \alpha_1 + \cdots + \frac{a_n}{m_n} \alpha_n$$

so $\{\alpha_1, \ldots, \alpha_n\}$ spans K over $\mathbb{Q}$. Furthermore if $a_1\alpha_1 + \cdots + a_n\alpha_n = 0$, then $a_1 m_1 \beta_1 + \cdots + a_n m_n \beta_n = 0$, and as $\{\beta_1, \ldots, \beta_n\}$ is a basis, this implies that $a_1 m_1 = \cdots = a_n m_n = 0$, hence $a_1 = \cdots = a_n = 0$ as $m_i \neq 0$. □

Theorem 1.2.8. *$\mathcal{O}_K$ viewed as an additive group is free abelian of rank n.*

Proof. The proof is by contradiction - suppose $\mathcal{O}_K$ is not free abelian of rank n. This is the same as saying that $\mathcal{O}_K$ does not possess any integral basis. Let $\{\omega_1, \ldots, \omega_n\}$ be a basis of algebraic integers for K over $\mathbb{Q}$ (which exists by Lemma 1.2.7) that makes the absolute value of the discriminant minimal. Since this cannot be an integral basis there exists some $\gamma \in \mathcal{O}_K$ such that $\gamma = a_1\omega_1 + \cdots + a_n\omega_n$ where not all $a_i \in \mathbb{Z}$. Let us choose the numbering such that $a_1 \notin \mathbb{Z}$. Then $a_1 = a + r$ where $a \in \mathbb{Z}$, $0 < r < 1$. Let $\psi_1 = \gamma - a\omega_1$ and $\psi_i = \omega_i$ for $i = 2, \ldots, n$. Since $\mathcal{O}_K$ is a ring, $\psi_1, \ldots, \psi_n$ are all in $\mathcal{O}_K$. In fact, we claim that $\{\psi_1, \ldots, \psi_n\}$ is another basis for K over $\mathbb{Q}$. To this end, assume $b_1\psi_1 + \cdots + b_n\psi_n = 0$. This is the same as

$$0 = b_1(\gamma - a\omega_1) + b_2\omega_2 + \cdots + b_n\omega_n = b_1(a_1 - a)\omega_1 + (b_2 + b_1 a_2)\omega_2 + \cdots + (b_n + b_1 a_n)\omega_n$$

Since $\{\omega_1,\dots,\omega_n\}$ is a basis, $b_1(a_1-a)=0$. We cannot have $a_1=a$, because then $r=0$. Hence $b_1=0$, but then we must have $b_2=\dots=b_n=0$ as well. Thus $\{\psi_1,\dots,\psi_n\}$ is a basis for K over $\mathbb{Q}$. By the remark after Theorem 1.1.14 we get the following equality:

$$|\text{disc}(\psi_1,\dots,\psi_n)| = \left|\det\begin{pmatrix} a_1-a & a_2 & a_3 & \cdots & a_n \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}^2\right| |\text{disc}(\omega_1,\dots,\omega_n)|$$

but then as $0<r^2<1$:

$$|\text{disc}(\psi_1,\dots,\psi_n)| = (a_1-a)^2|\text{disc}(\omega_1,\dots,\omega_n)| = r^2|\text{disc}(\omega_1,\dots,\omega_n)| < |\text{disc}(\omega_1,\dots,\omega_n)|$$

which contradicts the minimality of $\{\omega_1,\dots,\omega_n\}$. □

Corollary 1.2.9. *$\mathcal{O}_K$ is Noetherian.*

Proof. Let I be any ideal of $\mathcal{O}_K$. Then I is an additive subgroup of $\mathcal{O}_K$ viewed as an additive group. But by theorem 1.2.6, $\mathcal{O}_K$ is free abelian of rank n, and by Lemma 1.2.2 I must also be free abelian of rank $\leq n$ and hence finitely generated. □

If $\{\gamma_1,\cdots,\gamma_n\}$ and $\{\omega_1,\dots,\omega_n\}$ are two integral bases for $\mathcal{O}_K$, then the change of basis matrix between them consists of integers. It is a well known fact that an invertible matrix with integer entries has determinant ±1. Hence it follows by the remark following Theorem 1.1.14 that the discriminant of those two bases are equal. This justifies the following definition.

Definition 1.2.10. Let K be a number field. We define disc$(\mathcal{O}_K)$ to be the discriminant of some integral basis for $\mathcal{O}_K$. We can generalize the remark after Theorem 1.1.14 to a non-zero ideal I as well, because I is free abelian of rank $\leq n$ and we can restrict the embeddings of K into $\mathbb{C}$ to I. Thus we also define for non-zero ideal I, disc(I) to be the discriminant of some integral basis for I (which is well-defined by the paragraph above).

Lemma 1.2.11. *A finite integral domain D is a field.*

Proof. Take $a\neq0$ in D. Since D is finite, the list $1,a,a^2,a^3,\dots$ repeats itself at some point where $a^n=a^m, n<m$, whence $a^n(1-a^{m-n})=0$ and as $a\neq0$ and D is an integral domain $1=a^{m-n}=aa^{m-n-1}$ so a has an inverse. □

Theorem 1.2.12. *$\mathcal{O}_K$ has Krull dimension 1 - that is every non-zero prime ideal of $\mathcal{O}_K$ is maximal.*

Proof. Let $\mathfrak{p}$ be a non-zero prime ideal of $\mathcal{O}_K$. Let $0 \neq \alpha \in \mathfrak{p}$. Then we have $N(\alpha) = \alpha\sigma_2(\alpha)\cdots\sigma_n(\alpha_n)$, where $\sigma_1(\alpha) = \alpha$ if σ_1 is the identity embeddings of K into $\mathbb{C}$. Hence, let us write $N(\alpha) = \alpha\beta$ with $\beta = \sigma_2(\alpha)\cdots\sigma_n(\alpha)$. By Theorem 1.1.9, $N(\alpha) \in \mathbb{Z}$, and since $\alpha \neq 0$, $N(\alpha) \neq 0$. Since σ sends an algebraic integer to an algebraic integer $\beta \in \mathbb{A}$. We also have $\beta = \frac{N(\alpha)}{\alpha} \in K$ so $\beta \in \mathcal{O}_K$. Since $\mathfrak{p}$ is an ideal, $N(\alpha) = \alpha\beta \in \mathfrak{p}$, so $(N(\alpha)) \subseteq \mathfrak{p}$. From the third isomorphism theorem it follows that

$$\frac{\mathcal{O}_K/(N(\alpha))}{\mathfrak{p}/(N(\alpha))} \cong \mathcal{O}_K/\mathfrak{p}$$

Now as $\mathcal{O}_K \cong \mathbb{Z}^n$, $(N(\alpha)) = k\mathbb{Z}^n$ for some k and hence $|\mathcal{O}_K/(N(\alpha))| = |\mathbb{Z}^n/k\mathbb{Z}^n| = k^n$. This shows that the left hand side of the isomorphism above is finite, and so $\mathcal{O}_K/\mathfrak{p}$ is also finite. But $\mathfrak{p}$ is a prime ideal, so $\mathcal{O}_K/\mathfrak{p}$ is a finite integral domain which we know are all fields. Finally, $\mathcal{O}_K/\mathfrak{p}$ can be a field iff $\mathfrak{p}$ is maximal, which finishes the proof. □

Theorem 1.2.13. *Suppose $\alpha \in \mathbb{C}$ satisfy a monic polynomial with coefficients in $\mathbb{A}$. Then $\alpha \in \mathbb{A}$.*

Proof. Let $a_i \in \mathbb{A}$, not all zero, such that

$$a_0 + a_1\alpha + \cdots + \alpha^n = 0$$

We aim to prove that $\mathbb{Z}[a_0, \dots, a_{n-1}](1, \alpha, \alpha^2, \dots)$ is finitely generated abelian. Clearly $\Gamma_\alpha = (1, \alpha, \alpha^2, \dots)$ is a subgroup of this group, and as a subgroup of a finitely generated abelian group is finitely generated this is enough by Lemma 1.2.3. Because the a_i's are algebraic integers, the Γ_{a_i}'s are finitely generated and hence $\mathbb{Z}[a_0, \dots, a_{n-1}]$ is finitely generated. Now we have that

$$\begin{aligned}
\alpha^{n+1} &= \alpha\alpha^n \\
&= \alpha(-a_0 - a_1\alpha - \cdots - \alpha^{n-1}) = -a_0\alpha - a_1\alpha^2 - \cdots - \alpha^n \\
&= -a_0\alpha - a_1\alpha^2 - \cdots - a_{n-2}\alpha^{n-1} - (a_0 + a_1\alpha + \cdots + a_{n-1}\alpha^{n-1}) \\
&= -a_0 - (a_0 + a_1)\alpha - (a_1 + a_2)\alpha^2 - \cdots - (a_{n-2} + a_{n-1})\alpha^{n-1} \\
&\in \mathbb{Z}[a_0, \dots, a_{n-1}](1, \alpha, \dots, \alpha^{n-1})
\end{aligned}$$

By induction it follows that $\alpha^{n+m} \in \mathbb{Z}[a_0, \dots, a_{n-1}](1, \alpha, \dots, \alpha^{n-1})$ for $m \in \mathbb{N}$. This finishes the proof, by the remarks in the beginning of the proof. □

Chapter 2

Dedekind domains

In this chapter we prove that every non-zero ideal in a Dedekind domain factorizes uniquely into prime ideals. Afterwards we prove some other nice properties of Dedekind domains. First that a Dedekind domain is a PID if and only if it is a UFD. Secondly that ideals are generated by at most two elements. As we will see, the ideal theory of Dedekind domains is very nice: the ideals here behave somewhat like numbers.

2.1 Unique factorization in Dedekind domains

Definition 2.1.1. Let R be an integral domain. Then R is called a Dedekind domain if
(DD1) R is Noetherian.
(DD2) R has Krull dimension 1, that is every non-zero prime ideal is maximal.
(DD3) R is integrally closed in its field of fractions $K = \left\{\frac{\alpha}{\beta} : \alpha, \beta \in R, \beta \neq 0\right\}$ - that is if $\frac{\alpha}{\beta}$ is a root of a monic polynomial with coefficients in K, then $\frac{\alpha}{\beta} \in R$.

Corollary 1.2.9, Theorem 1.2.12 and Theorem 1.2.13 show that the ring of integers of any number field is a Dedekind domain. However this is not the only class of Dedekind domains, another one comes from algebraic geometry: if X is a non-singular affine curve over a field k, the coordinate ring $\Gamma(X/k)$ is a Dedekind domain. Our first goal will be to prove that every non-zero ideal in a Dedekind domain factorizes uniquely into prime ideals. For this we need a few lemmata. Throughout the rest of the chapter, R is always assumed to be a Dedekind domain with field of fractions K.

Lemma 2.1.2. *Let I be a non-zero ideal in R. Then I contains a product of non-zero prime ideals of R.*

Proof. Assume there is a non-zero ideal I that does not contain a product of prime ideals. Then the set

$$\mathcal{P} = \{I \subseteq R : I \text{ ideal}, \nexists P_1, \dots, P_t \text{ non-zero prime ideals such that } P_1 \cdots P_t \subseteq I\}$$

is non-empty. Since R is Noetherian, $\mathcal{P}$ contains a maximal element, call it M. Since $M \in \mathcal{P}$, M cannot be prime. Hence we can find $rs \in M$ such that $r, s \notin M$. The ideals $M + (r)$ and $M + (s)$ are strict supersets of M, and hence by maximality of M cannot be in $\mathcal{P}$. Thus we can find non-zero prime ideals $P_1, \dots, P_t, Q_1, \dots, Q_\ell$ such that $P_1 \cdots P_t \subseteq M + (r)$ and $Q_1 \cdots Q_\ell \subseteq M + (s)$. But as M is an ideal and $(rs) \subseteq M$ we get

$$P_1 \cdots P_t Q_1 \cdots Q_\ell \subseteq (M + (r))(M + (s)) \subseteq M^2 + M(s) + M(r) + (rs) \subseteq M$$

which contradicts that $M \in \mathcal{P}$. □

Lemma 2.1.3. *Let A be a proper ideal of R. Then there is an element $\gamma \in K \backslash R$ such that $\gamma A \subseteq R$.*

Proof. Let $a \in A$ be non-zero. By Lemma 2.1.2, $P_1 \cdots P_t \subseteq (a)$ for some non-zero prime ideals $P_1, \dots, P_t$ in R. Choose t such that it is minimal. Every proper ideal is contained in a maximal ideal (this follows from Zorn's lemma, see [3, Corollary 1.4]). So fix a maximal ideal P (and hence prime ideal) such that $(a) \subseteq P$. Hence P contains the product $P_1 \cdots P_t$, and because it is a prime ideal $P \supseteq P_i$ for some i, say $i = 1$. However, R is a Dedekind domain so DD2 implies that P_1 is a maximal ideal as well, and hence $P = P_1$. Since t was minimal $\exists b \in P_2 \cdots P_t \backslash (a)$. Now let $\gamma = b/a$. Then $\gamma \notin R$, because if it was then $b = \frac{b}{a} \cdot a \in (a)$. Now take $c \in A$. Since $c \in A \subseteq P = P_1$, $cb \in P_1 \cdots P_t \subseteq (a)$. Hence $cb = ar$ for some $r \in R$. That is $\gamma c = \frac{b}{a} c = r \in R$. □

Now we are ready to prove a central theorem: a non-zero ideal is just a multiplication away with some other ideal from being principal. From this theorem the prime factorization of ideals will follow.

Theorem 2.1.4. *Let I be a non-zero ideal in R. Then there exists a non-zero ideal J such that $IJ = (\alpha)$ for some non-zero $\alpha \in I$.*

Proof. We show that $J = \{\beta \in R : \beta I \subseteq (\alpha)\}$ works for any non-zero $\alpha \in I$. First of all we have to show that J is a non-zero ideal. To that end let $x, y \in J, r \in R$. Since $xI \subseteq (\alpha)$ and $yI \subseteq (\alpha)$ we get $(x - y)I \subseteq (\alpha)$. Furthermore $(rx)I = r(xI) \subseteq r(\alpha) \subseteq (\alpha)$. Hence J is an ideal, and non-zero since $\alpha \in J$. If now $z \in I, w \in J$, then $zw \in (\alpha)$, so $IJ \subseteq (\alpha)$. We now proceed to show that this actually is an equality. To this end, define $A = \frac{1}{\alpha} IJ$. Since IJ is an ideal, and $IJ \subseteq (\alpha)$, it follows that A is an ideal. If $A = R$ then $(\alpha) = IJ$, in which case we are done. We now complete the proof by showing that A being a proper ideal yields a contradiction. If A is a proper ideal, then by Lemma 2.1.3, there exists $\gamma \in K \backslash R$ such that $\gamma A \subseteq R$. Since $\alpha \in I$, the ideal J is contained in A. It follows that $\gamma J \subseteq \gamma A \subseteq R$. Now $(\gamma J)I = \gamma JI = \gamma \alpha A = \alpha(\gamma A) \subseteq \alpha R = (\alpha)$, so by definition of J we have $\gamma J \subseteq J$. By DD1 we can find a generating set $\alpha_1, \dots, \alpha_m$ for J, and so $\gamma \alpha_i = a_{i1}\alpha_1 + \cdots + a_{im}\alpha_m$ for $a_{im} \in R$.

This gives us the following system of linear equations:

$$\gamma \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mm} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix}$$

Let us call the matrix in the system of equations above for M. Since not all α_i can be zero, the above shows that $(M - \gamma I_m)\mathbf{x} = \mathbf{0}$ not only has the trivial solution. We then know from linear algebra that $\det(M - \gamma I_m) = 0$. Expanding this gives a monic polynomial in $R[x]$ where γ is a root. By DD3, $\gamma \in R$ which is a contradiction as $\gamma \in K \backslash R$. □

From this we get three important corollaries: two of them says that (non-zero) ideals in Dedekind domains have the cancellation property and that divisibility of ideals and inclusion is equivalent. The last tells ut that the ideal classes of a Dedekind domain form an abelian group (!). This group is called the ideal class group and we will study it more in the next chapter. As we will see, in some sense the order of the ideal class group, called the class number, measures how far away a Dedekind domain is from being a unique factorization domain. The ideal class group will be our main focus after this chapter and throughout the thesis. In fact we will establish a formula for calculating the class number, via the theory of Dedekind zeta functions. We postpone this third corollary to the next chapter.

Corollary 2.1.5. *Let A, B, C be non-zero ideals in R. Then $AB = AC$ implies $B = C$.*

Proof. By Theorem 2.1.4 we find an ideal J such that $AJ = (\alpha)$, $\alpha \neq 0$. Multiplying $AB = AC$ with J gives us $(\alpha)B = (\alpha)C$, so for any $b \in B$, there is a $c \in C$ and an $r \in R$ we have $\alpha b = r\alpha c$, so $b = rc \in C$, that is $B \subseteq C$. That $C \subseteq B$ is proven analogously. □

Corollary 2.1.6. *Let A, B be non-zero ideals in R. Then $A \mid B \iff A \supseteq B$.*

Proof. If $A \mid B$, there exists a non-zero ideal C in R such that $A \supseteq AC = B$. Conversely, we can find an ideal J and a non-zero element α such that $AJ = (\alpha)$ (Theorem 2.1.4). Now if $C = \frac{1}{\alpha}JB$ then

$$AC = A\frac{1}{\alpha}JB = \frac{1}{\alpha}(\alpha)B = B$$

What is left to prove is that C is an ideal in R. If we are able to prove $\frac{1}{\alpha}JB \subseteq R$ this will follow since JB is an ideal. Now $\frac{1}{\alpha}JB \subseteq \frac{1}{\alpha}JA = \frac{1}{\alpha}(\alpha) = R$. □

Now we are finally ready to prove on of the big theorems of this chapter:

Theorem 2.1.7. *Every proper non-zero ideal in a Dedekind domain R factorizes uniquely into prime ideals.*

Proof. We first show existence. Suppose for a contradiction that

$$\mathcal{P} = \{I \text{ non-zero proper ideal in } R : I \text{ does not factorize into prime ideals}\} \neq \emptyset$$

By DD1, $\mathcal{P}$ has a maximal element M. Again as in Lemma 2.1.3, we can fix a maximal (and hence prime) ideal P of R such that $M \subsetneq P$. By Corollary 2.1.6 there is an ideal C such that $PC = M$. Now C strictly contains M, because if $C = M$ then $PM = PC = M = RM$, which by Corollary 2.1.5 would imply $P = R$. That the inclusion is strict implies that $C = P_1 \cdots P_t$ for prime ideals P_i, but then $M = PC = PP_1 \cdots P_t$, which contradicts that $M \in \mathcal{P}$. This shows existence. For the uniqueness part, let I be any non-zero ideal in R, and suppose

$$I = P_1 \cdots P_t = Q_1 \cdots Q_\ell$$

for prime ideals P_i, Q_j say with $t \geq \ell$. Then $P_1 \mid Q_1 \cdots Q_\ell$ so $P_1 \supseteq Q_1 \cdots Q_\ell$ by Corollary 2.1.6. Since P_1 is prime, it contains some Q_i. After rearrangement, say $Q_1 \subseteq P_1$. Since Q_1 is prime and non-zero, DD2 gives that Q_1 is maximal and hence $Q_1 = P_1$. By Corollary 2.1.5 it follows that $P_2 \cdots P_t = Q_2 \cdots Q_\ell$. If $t = \ell$ we get the uniqueness and we are done. If $t > \ell$ we have $P_{\ell+1} \cdots P_t = R$, so $P_{\ell+1} \mid R$, that is $P_{\ell+1} \supseteq R$, which contradicts $P_{\ell+1}$ being prime. □

Theorem 2.1.4 told us that for every non-zero ideal I, there exists an ideal J such that IJ is a non-zero principal ideal. Another way to interpret this is that a Dedekind domain R is not too far away from being principal. We can strengthen this even more: every ideal in a Dedekind domain is generated by at most two elements. To prove this we start by generalizing gcd and lcm from $\mathbb{Z}$ to R. Remember in $\mathbb{Z}$ that gcd is the greatest common divisor of two integers, so the greatest common divisor of two ideals should be the greatest ideal that divides both. From Corollary 2.1.6 we know $A \mid B \iff A \supseteq B$. So if D is the greatest common divisor of I and J we have $D \supseteq I$ and $D \supseteq J$. Furthermore if $C \mid I$, $C \mid J$ as well, then $C \mid D$ and so $C \supseteq D$. From this we see that the greatest common divisor of I and J should be the smallest ideal that contains both.

Definition 2.1.8. Let I, J be non-zero ideals of R. We then define

$$\gcd(I, J) = I + J \qquad \operatorname{lcm}(I, J) = I \cap J$$

Lemma 2.1.9. *If I, J are ideals in a Dedekind domain R such that $1 \in I + J$, then $1 \in I^m + J^n$ for all m, n.*

Proof. Since $1 \in I + J$, write $1 = \alpha + \beta, \alpha \in I, \beta \in J$. By the binomial theorem

$$\begin{aligned} 1 = (\alpha + \beta)^{m+n} &= \sum_{i=0}^{m+n} \binom{m+n}{i} \alpha^i \beta^{m+n-i} \\ &= \beta^n \sum_{i=0}^{m} \binom{m+n}{i} \alpha^i \beta^{m-i} + \alpha^m \sum_{i=m+1}^{m+n} \binom{m+n}{i} \alpha^{i-m} \beta^{m+n-i} \end{aligned}$$

Because $\beta^n \in J^n$ and $\alpha^m \in I^m$, the absorbance property of ideals finishes the proof. □

Theorem 2.1.10. *Let I be any non-zero ideal in a Dedekind domain R, and let α be any non-zero element of I. Then there exists an element $\beta \in I$ such that $I = (\alpha, \beta)$.*

Proof. Suppose that we find a $\beta \in R$ such that $I = \gcd((\alpha), (\beta))$. Then $I = (\alpha)+(\beta) = (\alpha, \beta)$ by the definition above. In that case $I = (\alpha, \beta) \supseteq (\beta)$ and so $\beta \in I$, which is what we want. With this strategy in mind, start by factorizing I into primes:

$$I = P_1^{n_1} \cdots P_t^{n_t}$$

such that all the P_i are distinct. Since $(\alpha) \subseteq I$, $P_1^{n_1} \cdots P_t^{n_t} = I \mid (\alpha)$ and so (α) is divisible by all the $P_i^{n_i}$. There may be more divisors: denote those by $Q_1, \ldots, Q_\ell$. We must construct a β such that none of these Q_i divides (β) and such that $P_i^{n_i}$ is the exact power of P_i that divides (β). In other words we want to find β such that

$$\beta \in \bigcap_{i=1}^{t} (P_i^{n_i} \backslash P_i^{n_i+1}) \cap \bigcap_{j=1}^{\ell} (R \backslash Q_j)$$

We now show such β exist, which would finish the proof. By unique factorization we can fix non-zero elements $\beta_i \in P_i^{n_i} \backslash P_i^{n_i+1}$. Consider the congruences

$$\begin{aligned} x &\equiv \beta_i \pmod{P_i^{n_i+1}} \qquad & i = 1, \ldots, t \\ x &\equiv 1 \pmod{Q_j} \qquad & j = 1, \ldots, \ell \end{aligned}$$

The ideals $P_1^{n_1}, \ldots, P_t^{n_t}, Q_1, \cdots, Q_\ell$ are all pairwise relatively prime by Lemma 2.1.9 and thus the Chinese Remainder Theorem gives a solution x to the congruences over. This solution satisfy $x \in P_i^{n_i}$ but since β_i is non-zero $x \notin P_i^{n_i+1}$. Also such a solution will by definition satisfy $x \notin Q_j$. This solution x is exactly the β we are seeking. □

Recall that every PID is a UFD, but the converse is not in general true. However, for Dedekind domains the converse is true.

Theorem 2.1.11. *A Dedekind domain R is a UFD if and only if it is a PID.*

Proof. Assume R is a UFD. Let I be a non-zero ideal of R. By Theorem 2.1.4 we find an ideal J such that $IJ = (\alpha), \alpha \neq 0$. Hence $I \mid (\alpha)$. Since R is a UFD, α factorizes into prime *elements* in R. Any such prime divisor p will generate a prime ideal (p). Hence I divides a product of principal ideals whom are all prime. It follows that I itself is a product of principal prime ideals, and hence a principal ideal. □

2.2 The Ideal Norm

We now go back to the special case where our Dedekind domain R is the ring of integers $\mathcal{O}_K$ for a number field K. There is also a notion of norm for ideals, defined as $\|I\| = |R/I|$. While this can be defined in any Dedekind domain (or any ring whatsoever), we get some very nice properties of the ideal norm when restricting to $\mathcal{O}_K$.

Theorem 2.2.1. *Let I, J be non-zero ideals in $\mathcal{O}_K$. Then $\|I\| < \infty$ and $\|IJ\| = \|I\|\|J\|$.*

Proof. The finiteness of the norm follows from the exact same argument as in the proof of Theorem 1.2.12. For multiplicativeness, we first prove that $\|P^m\| = \|P\|^m$ for P a prime ideal. Since $\mathcal{O}_K \supseteq P \supseteq P^2 \supseteq \cdots$ we get by the third isomorphism theorem that

$$\begin{aligned}
(\mathcal{O}_K/P^m)/(P^{m-1}/P^m) &\cong \mathcal{O}_K/P^{m-1} \\
(\mathcal{O}_K/P^{m-1})/(P^{m-2}/P^{m-1}) &\cong \mathcal{O}_K/P^{m-2} \\
&\vdots \\
(\mathcal{O}_K/P^2)/(P^1/P^2) &\cong \mathcal{O}_K/P^1
\end{aligned}$$

Considering orders we get

$$\begin{aligned}
\|P^m\| &= |P^{m-1}/P^m|\|P^{m-1}\| \\
&= |P^{m-1}/P^m||P^{m-2}/P^{m-1}|\|P^{m-2}\| \\
&= \cdots = |P^{m-1}/P^m| \cdots |P^0/P^1|
\end{aligned}$$

Hence to prove $\|P^m\| = \|P\|^m$ it is enough to show that $|P^k/P^{k+1}| = \|P\|$ for all $k = 0, \ldots, m-1$. Here $P^0 = \mathcal{O}_K$.

Let $\alpha \in P^k \backslash P^{k+1}$ (that such an α exists follows from unique factorization). Let us now consider those ideals as additive groups. We get a canonical isomorphism $R/P \to \alpha R/\alpha P$. Since $\alpha \in P^k$ we get an inclusion $\alpha R \hookrightarrow P^k$. This induces a homomorphism to the quotient $\psi : \alpha R \to P^k/P^{k+1}$. Clearly $\ker(\psi) = P^{k+1} \cap \alpha R$ and $\operatorname{Im}(\psi) = (\alpha R + P^{k+1})/P^{k+1}$. Now $\alpha R + P^{k+1} = \gcd(\alpha R, P^{k+1})$. Since $\alpha R \subseteq P^k$ it follows that $P^k \mid \alpha R$. However as $\alpha \in P^k \backslash P^{k+1}$ we also have $\alpha R \not\subseteq P^{k+1}$ so it follows that $\gcd(\alpha R, P^{k+1}) = P^k$. Now we continue with working out $P^{k+1} \cap \alpha R$. First observe that $P^{k+1} \cap \alpha R \subseteq P^k \cap \alpha P \subseteq \alpha P$ because $\alpha P \subseteq P^k$. Conversely if $x \in \alpha P \subseteq \alpha R$ then $\alpha \in P^{k+1}$ because $\alpha \in P^k$. We conclude that $P^{k+1} \cap \alpha R = \alpha P$. Piecing all this together and using the first isomorphism theorem we get an isomorphism $\alpha R/\alpha P \to P^k/P^{k+1}$, so we have our desired isomorphism $R/P \to P^k/P^{k+1}$. This proves $\|P^m\| = \|P\|^m$.

Now let T be any non-zero ideal and P a non-zero prime ideal such that P does not divide T. Then P^m and T are relatively prime for any exponent m by Lemma 2.1.9, because P and T are relatively prime. The Chinese Remainder Theorem then gives us an isomorphism $R/P^m \times R/T \to R/P^m T$ from which by taking norms, and using what we proved earlier yield $\|P\|^m\|T\| = \|P^m\|\|T\| = \|P^m T\|$. Using the last equation repeatedly finishes the proof, because $I = P_1^{e_1} \cdots P_t^{e_t}$, $J = Q_1^{r_1} \cdots Q_s^{r_s}$ for some primes P_i and Q_j. □

In the case where our ideal is principal the norm is easily calculated. To see this we first need a few more results. We first state a theorem that we will need.

Theorem 2.2.2. *Let G be a free abelian group with finite rank n and H a non-zero subgroup. Then there exists a basis $\{\beta_1, \ldots, \beta_n\}$ of G, an integer $r \geq 1$ and positive integers $d_1 \mid \cdots \mid d_r$ such that $\{d_1\beta_1, \ldots, d_r\beta_r\}$ is a basis for H.*

For a proof see [5, Theorem II.1.6]. If I is a non-zero ideal of $\mathcal{O}_K$ we know that $\|I\| < \infty$, that is $|\mathcal{O}_K/I| < \infty$. As I is also free abelian of rank $\leq n$, it follows from $|\mathcal{O}_K/I| < \infty$ that

I needs to be free abelian of rank n. The discriminant of I (see Definition 1.2.10) is closely related to $\operatorname{disc}(\mathcal{O}_K)$:

Theorem 2.2.3. *Let I be a non-zero ideal of $\mathcal{O}_K$ where $[K : \mathbb{Q}] = n$. Then I is free abelian of rank n and $\operatorname{disc}(I) = |\mathcal{O}_K/I|^2 \operatorname{disc}(\mathcal{O}_K)$.*

Proof. We have already proven the first claim. By Theorem 2.2.2, fix a basis $\{\omega_1, \ldots, \omega_n\}$ for $\mathcal{O}_K$ such that $\{d_1\omega_1, \ldots, d_n\omega_n\}$ is a basis for I for some positive integers $d_1, \ldots, d_n$. As mentioned in Definition 1.2.10, the remark after Theorem 1.1.14 can be used in this case as well. Now every basis element is just multiplied by a positive integer. The determinant in that formula then becomes $(d_1 \cdots d_k)^2$. Now look at the homomorphism

$$\varphi : \mathcal{O}_K \to \mathbb{Z}/d_1\mathbb{Z} \times \cdots \times \mathbb{Z}/d_k\mathbb{Z} \qquad a_1\omega_1 + \cdots + a_n\omega_n \mapsto (a_1, \ldots, a_n)$$

which is clearly surjective and has kernel I. Hence $|G/I| = d_1 \cdots d_k$ by the first isomorphism theorem, which finishes the proof. □

Corollary 2.2.4. *Let $a \in \mathcal{O}_K$ be non-zero. Then $\|(a)\| = |N(a)|$.*

Proof. A basis for (a) is $\{a\omega_1, \ldots, a\omega_n\}$ where $\{\omega_1, \ldots, \omega_n\}$ is a basis for $\mathcal{O}_K$. Hence $\operatorname{disc}(a\omega_1, \ldots, a\omega_n) = \|(a)\|^2 \operatorname{disc}(\omega_1, \ldots, \omega_n)$ from Theorem 2.2.3. From basic determinant rules we get that

$$\operatorname{disc}(a\omega_1, \ldots, a\omega_n) = |N(a)|^2 \operatorname{disc}(\omega_1, \ldots, \omega_n)$$

The proof follows by combining this with the previous formula. □

There are some relations between prime numbers and prime ideals via the norm as well.

Theorem 2.2.5. *Let I be a non-zero ideal of $\mathcal{O}_K$. Then $\|I\| \in I$. Furthermore if $\|I\|$ is prime, I is also prime. Conversely, if I is prime, then $\|I\| = p^m$ for some exponent $m \leq [K : \mathbb{Q}]$. Finally, given a positive integer t, there are at most finitely many non-zero ideals I, satisfying $\|I\| = t$.*

Proof. Consider $\mathcal{O}_K$ as an additive group. For any $x \in \mathcal{O}_K$, $x\|I\|$ is zero in $\mathcal{O}_K/I$ by definition of $\|I\|$. That is $x\|I\| \in I$. Now let $x = 1$.

For the next statement, start by uniquely factorizing: $I = P_1^{e_1} \cdots P_t^{e_t}$. Taking norms we obtain $\|I\| = \|P_1\|^{e_1} \cdots \|P_t\|^{e_t}$. Since $\|I\|$ is prime then necessarily all but one e_i are 0, and for this non-zero e_i, $e_i = 1$. Hence we have $I = P_i$. Now assume I is prime. Since $\|I\|$ is an integer we have some unique factorization $\|I\| = p_1^{m_1} \cdots p_t^{m_t}$. By the first statement of the theorem, $I \mid (\|I\|) = (p_1)^{m_1} \cdots (p_t)^{m_t}$. It follows that $I \mid (p_i)^{m_i}$ for at least one i. Since I is prime, we get that $I \mid (p_i)$. If there was another $(p_j)^{m_j}$ such that $I \mid (p_j)^{m_j}$, then again since I is prime $I \mid (p_j)$. Since $(p_i), (p_j)$ are coprime, we can find $a, b \in \mathbb{Z}$ such that $ap_i + bp_j = 1$, thus $(p_i) + (p_j) = \mathcal{O}_K$. Then as $I \mid (p_i) \iff I \supseteq (p_i)$, we get $\mathcal{O}_K = (p_i) + (p_j) \subseteq I$, so $I = \mathcal{O}_K$, but I is prime and cannot be the whole ring. Hence we have a contradiction, and so $I \mid (p_i)$ for exactly one i. Taking norms we get $\|I\| = \|(p_i)\| = |N(p_i)|$. p_i being a prime in $\mathbb{Z}$ is fixed by all embeddings of K into $\mathbb{C}$, and hence $|N(p_i)| = p_i^n$. It follows that $\|I\| = p_i^m$ for $m \leq n$.

For the last part of the theorem, consider such a positive integer t. By unique factorization, (t) has finitely many divisors, or equivalently: (t) belongs to only a finite number of ideals $\mathcal{O}_K$, and hence t only belongs to a finite number of ideals of $\mathcal{O}_K$ (if t is belongs to some ideal, then necessarily (t) also must belong there). Let I be an ideal such that $\|I\| = t$. By the first statement of this theorem, $t = \|I\| \in I$. From what we just proved, there can be at most finitely many I where t belongs. In other words, there are at most finitely many ideals I satisfying $\|I\| = t$. □

Chapter 3

Geometry of numbers

In this chapter we interpret number fields geometrically via lattices. More specifically we will use the embeddings of a number field K into $\mathbb{C}$, to map $\mathcal{O}_K$ onto a lattice $\mathbb{R}^n$ for some n. We also introduce two groups: the ideal class group and the unit group. Roughly speaking, the ideal class group measures how much unique factorization of elements fail in $\mathcal{O}_K$. We will show that this group is finite and give a nice bound on its order. The unit group consists of the units of $\mathcal{O}_K$. We will give a explicit description of it. Finally we will use these geometric ideas to study the distribution of ideals in $\mathcal{O}_K$. This will be important in the next chapter where we (finally!) introduce Dedekind Zeta functions.

3.1 The ideal class group

Given any Dedekind domain R, we have an equivalence relation $\sim$ defined for non-zero ideals of R:

$$I \sim J \iff \alpha I = \beta J \text{ for some non-zero } \alpha, \beta \in R$$

Definition 3.1.1. The ideal class group consists of the (non-zero) ideal classes of R with respect to the equivalence relation above, with ideal multiplication as the operation.

Lemma 3.1.2. *The ideal class group is indeed a group (with multiplication of ideals).*

Proof. Let C_1, C_2 be two equivalence classes of ideals and choose representatives $I, I' \in C_1$, $J, J' \in C_2$. We want to show that $IJ \sim I'J'$ so that the multiplication in this group is well-defined. Since $I \sim I'$ we have $\alpha I = \beta I'$ and since $J \sim J'$ we have $\gamma J = \delta J'$ where $\alpha, \beta, \gamma, \delta$ are non-zero. Combining this we get $\alpha\gamma IJ = \beta\delta I'J'$, so $IJ \sim I'J'$. Associativity is clear, and the identity is the class of R. What we need to show is that every ideal class has an inverse. To this end let C be an ideal class and pick a representative I. From Theorem 2.1.4 there is some ideal J such that $1(IJ) = IJ = (\alpha) = \alpha R$ for some non-zero $\alpha \in I$. That is $IJ \sim R$. □

Let $(\alpha), (\beta)$ be any two non-zero principal ideals of a Dedekind domain R. Then $\beta(\alpha) = \alpha(\beta)$ so $(\alpha) \sim \beta$. That is, any two non-zero principal ideals are equivalent. Suppose now that all non-zero ideal of R are equivalent and let I be some non-zero ideals of R. Then we have $\alpha I = \beta R$ for some $\alpha, \beta \in R$. Then there is some $\eta \in I$ such that $\alpha\eta = \beta$, so $\frac{\beta}{\alpha} = \eta \in I \subseteq R$ and hence $I = \frac{\beta}{\alpha}R$ is principal. This argument shows that the class group of R is trivial if and only if R is a PID. From Theorem 2.1.11 we then get

Theorem 3.1.3. *A Dedekind domain R is a UFD if and only if its class group is trivial.*

Proof. This follows from Theorem 2.1.11 and the paragraph above. □

This justifies the idea that the ideal class group measures how far a Dedekind domain is from having unique factorization. Our next goal is to prove that in the case $R = \mathcal{O}_K$, the ideal class group is indeed finite. The idea of the proof is to find a common element α of any non-zero ideal I of $\mathcal{O}_K$ with a certain bound on its norm. This will give us finiteness of the ideal class group. The proof we will give is by no means the easiest, but one of its strengths is that it gives a better bound on the order than more elementary methods. For a more elementary proof see [8, Theorem 35 & Corollary 2].

To get us started, fix a number field K with degree n over $\mathbb{Q}$. Any embedding of K into $\mathbb{C}$ is either purely real (that is its image is contained in $\mathbb{R}$) or has purely imaginary elements in its image as well. In the latter case, such an embedding τ has a conjugate $\overline{\tau}$ which also is an embedding of K into $\mathbb{C}$. We thus list up our embeddings as

$$\sigma_1, \dots, \sigma_r, \tau_1, \overline{\tau_1}, \dots, \tau_s, \overline{\tau_s}$$

where $\sigma_1, \dots, \sigma_r$ are the ones whose image is contained in $\mathbb{R}$. Theorem 1.1.3 tells us that $r + 2s = n$.

Definition 3.1.4. Let K be a number field with $[K : \mathbb{Q}] = n$ and with the notation above, so that $r + 2s = n$. Then we define the following mapping $K \longrightarrow \mathbb{R}^n$:

$$\mathrm{Lat}_K : K \longrightarrow \mathbb{R}^n \qquad \alpha \mapsto (\sigma_1(\alpha), \dots, \sigma_r(\alpha), \mathrm{Re}(\tau_1(\alpha)), \mathrm{Im}(\tau_1(\alpha)), \dots, \mathrm{Re}(\tau_s(\alpha)), \mathrm{Im}(\tau_s(\alpha))$$

where Re and Im denote real part and imaginary part respectively.

The name of the function Lat_K comes from the fact that it maps $\mathcal{O}_K$ onto a lattice in $\mathbb{R}^n$.

Definition 3.1.5. A lattice Λ in $\mathbb{R}^n$ is a subgroup of the additive group $\mathbb{R}^n$, so that Λ has a $\mathbb{Z}$-basis which spans an additive group isomorphic to $\mathbb{R}^r$ for some $0 \leq r \leq n$. If $r = n$ we say Λ has full rank. A fundamental parallellotope for Λ is a set of the form

$$\left\{ \sum_{i=1}^{r} a_i v_i : a_i \in [0, 1) \right\}$$

where $\{v_1, \dots, v_r\}$ is any $\mathbb{Z}$-basis for Λ.

Before we continue we need to agree on precisely what we mean by an n-dimensional volume.

Definition 3.1.6. For any Lebesgue-measureable subset $X \subseteq \mathbb{R}^n$ we define the n-dimensional volume of X, $\mathrm{vol}(X)$, to be the Lebesgue measure of X.

The Lebesgue measure coincides with any reasonable intuitive way of thinking of n-dimensional volume. This is rather vague, but we will not explain this any further. From linear algebra we know that the n-dimensional volume of a parallellotope as described above is the absolute value of the determinant whose rows are $v_1, \dots, v_n$. Since any change of basis matrix between two $\mathbb{Z}$-bases has determinant ± 1, it follows that the volume is well-defined.

Definition 3.1.7. Let Λ be a lattice of full rank in $\mathbb{R}^n$. By $\mathrm{vol}(\mathbb{R}^n/\Lambda)$ we mean the volume of a fundamental parallellotope of Λ as defined above.

Theorem 3.1.8. *Let K be a number field with $[K : \mathbb{Q}] = n = r + 2s$. Denote the image $\mathrm{Lat}_K(\mathcal{O}_K)$ by $\Lambda_{\mathcal{O}_K}$. Then $\Lambda_{\mathcal{O}_K}$ is an n-dimensional lattice in $\mathbb{R}^n$ with*

$$\mathrm{vol}(\mathbb{R}^n/\Lambda_{\mathcal{O}_K}) = \frac{1}{2^s}\sqrt{|\mathrm{disc}(\mathcal{O}_K)|}$$

Proof. If we view Lat_K as an additive homomorphism $K^+ \to \mathbb{R}^n$ then it follows from the definition of Lat_K that $\ker(\mathrm{Lat}_K)$ is trivial. Hence Lat_K is an embedding of K^+ into $\mathbb{R}^n$. Now, let us fix an integral basis $\{\alpha_1, \dots, \alpha_n\}$ for $\mathcal{O}_K$. Then $\{\mathrm{Lat}_K(\alpha_1), \dots, \mathrm{Lat}_K(\alpha_n)\}$ generates $\Lambda_{\mathcal{O}_K}$ over $\mathbb{R}^n$. It will follow that $\Lambda_{\mathcal{O}_K}$ is a n-dimensional lattice in $\mathbb{R}^n$ if $\{\mathrm{Lat}_K(\alpha_1), \dots, \mathrm{Lat}_K(\alpha_n)\}$ is linearly independent over $\mathbb{R}$. To this end assume

$$a_1\mathrm{Lat}_K(\alpha_1) + \cdots + a_n\mathrm{Lat}_K(\alpha_n) = 0$$

for $a_i \in \mathbb{R}$. From this we get a system of equations:

$$\begin{aligned}
0 &= a_1\sigma_1(\alpha_1) + \cdots + a_n\sigma_1(\alpha_n) \\
&\vdots \\
0 &= a_1\sigma_r(\alpha_1) + \cdots + a_n\sigma_r(\alpha_n) \\
0 &= a_1\mathrm{Re}(\tau_1(\alpha_1)) + \cdots + a_n\mathrm{Re}(\tau_1(\alpha_n)) \\
0 &= a_1\mathrm{Im}(\tau_1(\alpha_1)) + \cdots + a_n\mathrm{Im}(\tau_1(\alpha_n)) \\
&\vdots \\
0 &= a_1\mathrm{Re}(\tau_s(\alpha_1)) + \cdots + a_n\mathrm{Re}(\tau_s(\alpha_n)) \\
0 &= a_1\mathrm{Im}(\tau_s(\alpha_1)) + \cdots + a_n\mathrm{Im}(\tau_s(\alpha_n))
\end{aligned}$$

Thus it will follow that $a_1 = a_2 = \cdots = a_n = 0$ if the determinant of the following matrix is non-zero:

$$\begin{pmatrix}
\sigma_1(\alpha_1) & \cdots & \sigma_1(\alpha_n) \\
\vdots & \ddots & \vdots \\
\sigma_r(\alpha_1) & \cdots & \sigma_r(\alpha_n) \\
\mathrm{Re}(\tau_1(\alpha_1)) & \cdots & \mathrm{Re}(\tau_1(\alpha_n)) \\
\mathrm{Im}(\tau_1(\alpha_1)) & \cdots & \mathrm{Im}(\tau_1(\alpha_n)) \\
\vdots & \ddots & \vdots \\
\mathrm{Re}(\tau_s(\alpha_1)) & \cdots & \mathrm{Re}(\tau_s(\alpha_n)) \\
\mathrm{Im}(\tau_s(\alpha_1)) & \cdots & \mathrm{Im}(\tau_s(\alpha_n))
\end{pmatrix}
=
\begin{pmatrix}
\sigma_1(\alpha_1) & \cdots & \sigma_1(\alpha_n) \\
\vdots & \ddots & \vdots \\
\sigma_r(\alpha_1) & \cdots & \sigma_r(\alpha_n) \\
\frac{\tau_1(\alpha_1)+\overline{\tau_1}(\alpha_1)}{2} & \cdots & \frac{\tau_1(\alpha_n)+\overline{\tau_1}(\alpha_n)}{2} \\
\frac{\tau_1(\alpha_1)-\overline{\tau_1}(\alpha_1)}{2} & \cdots & \frac{\tau_1(\alpha_n)-\overline{\tau_1}(\alpha_n)}{2} \\
\vdots & \ddots & \vdots \\
\frac{\tau_s(\alpha_1)+\overline{\tau_s}(\alpha_1)}{2} & \cdots & \frac{\tau_s(\alpha_n)+\overline{\tau_s}(\alpha_n)}{2} \\
\frac{\tau_s(\alpha_1)-\overline{\tau_s}(\alpha_1)}{2} & \cdots & \frac{\tau_s(\alpha_n)-\overline{\tau_s}(\alpha_n)}{2}
\end{pmatrix}$$

Elementary row operations on this matrix shows that the determinant of the matrix is

$$\frac{1}{(-2i)^s}\sqrt{|\mathrm{disc}(\mathcal{O}_K)|}$$

From Theorem 1.1.14 we know that $\mathrm{disc}(\mathcal{O}_K)$ is non-zero so we indeed get $a_1 = \cdots = a_n = 0$. The last claim of the theorem also follows since the absolute value of the determinant we calculated is indeed the volume of a fundamental parallellotope of $\Lambda_{\mathcal{O}_K}$. □

Now that we know that $\mathcal{O}_K$ corresponds to a certain lattice $\Lambda_{\mathcal{O}_K}$, we want to also establish a link between non-zero ideals of $\mathcal{O}_K$ and n-dimensional sublattices of $\Lambda_{\mathcal{O}_K}$.

Corollary 3.1.9. *A non-zero ideal I of $\mathcal{O}_K$ corresponds to an n-dimensional lattice Λ_I of $\Lambda_{\mathcal{O}_K}$ with*

$$\mathrm{vol}(\mathbb{R}^n/\Lambda_I) = \frac{1}{2^s}\sqrt{|\mathrm{disc}(\mathcal{O}_K)|}\|I\|$$

Proof. Following the proof of Theorem 3.1.8, we see that I is a n-dimensional sublattice Λ_I of $\Lambda_{\mathcal{O}_K}$ as I is also free abelian of rank n (Theorem 2.2.3). Because $\Lambda_{\mathcal{O}_K}$ is free abelian of rank n, let us by Theorem 2.2.2 fix a basis $\{\lambda_1, \ldots, \lambda_n\}$ of $\Lambda_{\mathcal{O}_K}$ such that for positive integers $d_1, \ldots, d_n$, $\{d_1\lambda_1, \ldots, d_n\lambda_n\}$ is a basis for Λ_I. The homomorphism

$$\varphi : \Lambda_{\mathcal{O}_K} \to \mathbb{Z}/d_1\mathbb{Z} \times \cdots \times \mathbb{Z}/d_n\mathbb{Z}$$

has kernel Λ_I. Hence $|\Lambda_{\mathcal{O}_K}/\Lambda_I| = d_1 \cdots d_k$. We also have

$$\mathrm{vol}(\mathbb{R}^n/\Lambda_I) = |\det(d_1\lambda_1, \ldots, d_n\lambda_n)| = d_1 \cdots d_k |\det(\lambda_1, \ldots, \lambda_n)| = |\Lambda_{\mathcal{O}_K}/\Lambda_I| \mathrm{vol}(\mathbb{R}^n/\Lambda_{\mathcal{O}_K})$$

We must have $|\mathcal{O}_K/I| = |\Lambda_{\mathcal{O}_K}/\Lambda_I|$ and hence by the formula from the previous theorem we get

$$\mathrm{vol}(\mathbb{R}^n/\Lambda_I) = |\mathcal{O}_K/I|\frac{1}{2^s}\sqrt{|\mathrm{disc}(\mathcal{O}_K)|} = \frac{1}{2^s}\sqrt{|\mathrm{disc}(\mathcal{O}_K)|}\|I\|$$

□

Now we transfer the norm we have defined on K to a norm on $\mathbb{R}^n$. Let us denote the usual norm on K as defined earlier as N^K. As usual $[K : \mathbb{Q}] = n = r + 2s$. For a vector $(x_1, \ldots, x_n) \in \mathbb{R}^n$ we define the norm as

$$N(x) = x_1 \cdots x_r(x_{r+1}^2 + x_{r+2}^2) \cdots (x_{n-1}^2 + x_n^2)$$

If $\alpha \in K$ then

$$\begin{aligned} N^K(\alpha) &= \sigma_1(\alpha) \cdots \sigma_r(\alpha)\tau_1(\alpha)\overline{\tau_1}(\alpha) \cdots \tau_1(\alpha)\overline{\tau_1}(\alpha) \\ &= \sigma_1(\alpha) \cdots \sigma_n(\alpha)((\mathrm{Re}(\tau_1(\alpha)))^2 + \mathrm{Im}(\tau_1(\alpha))^2) \cdots ((\mathrm{Re}(\tau_s(\alpha)))^2 + (\mathrm{Im}(\tau_s(\alpha)))^2) \end{aligned}$$

so $N|_K = N^K$. This justifies the notation N for this special norm. Our goal is to show that any n-dimensional lattice in $\mathbb{R}^n$ has some non-zero point x with a certain bound, from which the finiteness of the class group will follow. Interestingly, but perhaps not so surprising by now, it will also give us a lower bound for the discriminant. The existence of such a point will follow from a classical result due to Minkowski. Before we proceed we remark the three following properties of the Lebesgue measure, and hence our notion of volume in this specific setting.

1. The Lebesgue measure is countably additive: if $E_1, E_2, \ldots$ are pairwise disjoint Lebesgue-measureable sets then
$$\mathrm{vol}\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mathrm{vol}(E_i)$$
2. The Lebesgue measure is translation-invariant: if $E \subseteq \mathbb{R}^n$ is Lebesgue-measureable then for any $x \in \mathbb{R}^n$ we have $\mathrm{vol}(E + x) = \mathrm{vol}(E)$.
3. If $A, B \subseteq \mathbb{R}^n$ are Lebesgue measureable and $A \subseteq B$ then $\mathrm{vol}(A) \leq \mathrm{vol}(B)$.

We also need two definitons.

Definition 3.1.10. Let E be some subset of $\mathbb{R}^n$.

1. E is said to be convex if for any $x, y \in E$ the line segment joining x, y is contained in E.
2. E is said to be centrally symmetric if $x \in E$ implies $-x \in E$.

Theorem 3.1.11. (Minkowski's theorem)
Let Λ be a n-dimensional lattice in $\mathbb{R}^n$ and suppose that E is a convex, Lebesgue-measureable, centrally symmetric subset of $\mathbb{R}^n$ such that
$$\mathrm{vol}(E) > 2^n \mathrm{vol}(\mathbb{R}^n/\Lambda)$$
Then there is some non-zero x such that $x \in E \cap \Lambda$. If E is compact, then the inequality assumption can be weakened to $\mathrm{vol}(E) \geq 2^n \mathrm{vol}(\mathbb{R}^n/\Lambda)$

Proof. Let $\mathcal{F}$ be a fundamental parallellotope of Λ. By definition, $\mathbb{R}^n$ is a disjoint union of translations of $\mathcal{F}$: $x + \mathcal{F}$, $x \in \Lambda$. More specifically, any subset of $\mathbb{R}^n$ is contained in a disjoint union of translations of $\mathcal{F}$. Hence we get
$$\frac{1}{2}E = \bigsqcup_{x \in \Lambda} \left(\left(\frac{1}{2}E\right) \cap (x + \mathcal{F}) \right)$$
where $tE = \{te : e \in E\}$ with $t \in \mathbb{R}$ and where $\sqcup$ denotes disjoint union. By assumption we get
$$\mathrm{vol}(\mathcal{F}) < \frac{1}{2^n}\mathrm{vol}(E) = \mathrm{vol}\left(\frac{1}{2}E\right)$$
where the latter equality follows from the fact that E is a subset of $\mathbb{R}^n$, so scaling it down by a half scales it volume by $\frac{1}{2^n}$. By countable additivity and translation invariance of the Lebesgue measure we get that
$$\begin{aligned} \mathrm{vol}\left(\bigsqcup_{x \in \Lambda} \left(\left(\frac{1}{2}E\right) \cap (x + \mathcal{F}) \right) \right) &= \sum_{x \in \Lambda} \mathrm{vol}\left(\left(\frac{1}{2}E\right) \cap (x + \mathcal{F}) \right) \\ &= \sum_{x \in \Lambda} \mathrm{vol}\left(\left(\left(\frac{1}{2}E\right) - x \right) \cap \mathcal{F} \right) \end{aligned}$$

and hence

$$\mathrm{vol}(\mathcal{F}) < \sum_{x\in\Lambda} \mathrm{vol}\left(\left(\left(\frac{1}{2}E\right) - x\right) \cap \mathcal{F}\right)$$

This strict inequality shows that not all sets $\left(\left(\frac{1}{2}E\right) - x\right) \cap \mathcal{F}$ can be disjoint, because if they were the latter volume sum would be bounded above by $\mathrm{vol}(\mathcal{F})$, which is absurd. Hence we may fix distinct $x, y \in \Lambda$ such that

$$\left(\left(\frac{1}{2}E\right) - x\right) \cap \left(\left(\frac{1}{2}E\right) - y\right) \neq \emptyset$$

Then $x - y \neq 0$ and $\frac{1}{2}e - x = \frac{1}{2}e' - y$ for some $e, e' \in E$, that is $x - y = \frac{1}{2}e - \frac{1}{2}e'$. Since E is centrally symmetric, $-e' \in E$. Since E is convex, the line segment between e and $-e'$ is in E, that is $(1-t)e - te' \in E$ for $0 \leq t \leq 1$. Hence with $t = \frac{1}{2}$ we see that $\frac{1}{2}e - \frac{1}{2}e' \in E$. This finishes the proof in the case that E is not necessarily compact. Now suppose E is compact and that the inequality is weakened to $\geq$. It is a classical result in general topology that compact is equivalent to closed and bounded for subsets of $\mathbb{R}^n$. We have

$$\mathrm{vol}\left(\left(1 + \frac{1}{m}\right) E\right) = \left(1 + \frac{1}{m}\right)^n \mathrm{vol}(E) > \mathrm{vol}(E) \geq 2^n \mathrm{vol}(\mathbb{R}^n/\Lambda)$$

for $m = 1, 2, \cdots$, so the part we proved earlier implies there is a non-zero

$$x_m \in \left(\left(1 + \frac{1}{m}\right) E\right) \cap \Lambda$$

Now we consider the sequence $\{x_m\}_{m\in\mathbb{Z}_{\geq 1}}$ made of such x_m. Since E is bounded, $2E$ is also bounded, and $\{x_m\}_m \subseteq 2E$ so the sequence $\{x_m\}_m$ is bounded. Since all $x_m \in \Lambda$ and $\{x_m\}_m$ is bounded, there cannot be more than finitely many distinct elements x_m in the sequence (The intersection of a ball with finite radius and a lattice is a finite set. For a proof of this see [10, Theorem 6.1]). Hence we may fix an element $x^* \in \{x_m\}_m$ so that x^* is in $\left(1 + \frac{1}{m}\right) E$ for infinitely many m. Hence $x^* \in \overline{E}$, but E is closed so $\overline{E} = E$, and $x^* \in E$, so $x^* \in E \cap \Lambda$. □

Now we only need one more lemma before the main theorem of this section. Before this lemma we state a well-known inequality in mathematics, the so called Arithmetic Mean-Geometric Mean inequality (often abbreviated AM-GM).

Theorem 3.1.12. (The AM-GM Inequality)
Let $x_1, \ldots, x_n$ be non-negative real numbers. Then

$$\frac{x_1 + \cdots + x_n}{n} \geq \sqrt[n]{x_1 \cdots x_n}$$

Proof. This elegant proof is due to Pólya. Let $f(x) = e^{x-1} - x$. Then $f''(x) = e^{x-1} > 0$ so f is convex and hence $x \geq e^{x-1}$. Let $A = \frac{x_1 + \cdots + x_n}{n}$. Then the inequality gives

$$A^{-n} x_1 \cdots x_n = \frac{x_1}{A} \cdots \frac{x_n}{A} \geq e^{x_1/A-1} \cdots e^{x_n/A-1} = e^{\frac{x_1+\cdots+x_n}{A} - n} = e^{n-n} = 1$$

Rearranging we get the desired inequality. □

Lemma 3.1.13. *There is a compact, convex, centrally symmetric set* $A \subseteq \mathbb{R}^n$, $n = r + 2s$, *with* $\operatorname{vol}(A) = \frac{n^n}{n!} 2^r \left(\frac{\pi}{2}\right)^s$ *and such that* $a \in A \implies |N(a)| \leq 1$.

Proof. We will prove that the following A work:

$$A = \left\{ (x_1, \ldots, x_n) \in \mathbb{R}^n : |x_1| + \cdots + |x_r| + 2\left(\sqrt{x_{r+1}^2 + x_{r+2}^2} + \cdots + \sqrt{x_{n-1}^2 + x_n^2} \right) \leq n \right\}$$

Suppose $a = (x_1, \ldots, x_n) \in A$. Then by Theorem 3.1.12

$$\begin{aligned} |N(a)|^{1/n} &= \sqrt[n]{|x_1| \cdots |x_r| \sqrt{x_{r+1}^2 + x_{r+2}^2} \sqrt{x_{r+1}^2 + x_{r+2}^2} \cdots \sqrt{x_{n-1}^2 + x_n^2} \sqrt{x_{n-1}^2 + x_n^2}} \\ &\leq \frac{|x_1| + \cdots + |x_r| + \sqrt{x_{r+1}^2 + x_{r+2}^2} + \sqrt{x_{r+1}^2 + x_{r+2}^2} + \ldots \sqrt{x_{n-1}^2 + x_n^2} + \sqrt{x_{n-1}^2 + x_n^2}}{n} \\ &\leq \frac{n}{n} = 1 \end{aligned}$$

so $N(a) \leq 1$. If $a \in A$, then we easily see that also $-a \in A$. A is clearly bounded. That the inequality in the definition of A is non-strict shows A is closed. Hence A is compact.

Now we show that A is convex. Suppose $x, y \in A$, $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n)$. Then we first claim $\frac{x+y}{2} \in A$. By the normal triangle inequality:

$$\left| \frac{x_1 + y_1}{2} \right| + \cdots + \left| \frac{x_r + y_r}{2} \right| \leq \frac{1}{2} \left(|x_1| + \cdots + |x_r| + |y_1| + \cdots + |y_r| \right)$$

By the triangle inequality in $\mathbb{R}^2$:

$$\begin{aligned} &2\left(\sqrt{\left(\frac{x_{r+1} + y_{r+1}}{2}\right)^2 + \left(\frac{x_{r+2} + y_{r+2}}{2}\right)^2} + \cdots + \sqrt{\left(\frac{x_{n-1} + y_{n-1}}{2}\right)^2 + \left(\frac{x_n + y_n}{2}\right)^2} \right) \\ &\leq 2\left(\sqrt{\frac{x_{r+1}^2}{2^2} + \frac{x_{r+2}^2}{2^2}} + \sqrt{\frac{y_{r+1}^2}{2^2} + \frac{y_{r+2}^2}{2^2}} + \cdots + \sqrt{\frac{x_{n-1}^2}{2^2} + \frac{x_n^2}{2^2}} + \sqrt{\frac{y_{n-1}^2}{2^2} + \frac{y_n^2}{2^2}} \right) \\ &= \sqrt{x_{r+1}^2 + x_{r+2}^2} + \cdots + \sqrt{x_{n-1}^2 + x_n^2} + \sqrt{y_{r+1}^2 + y_{r+2}^2} + \cdots + \sqrt{y_{n-1}^2 + y_n^2} \end{aligned}$$

Combining those two inequalities and using that $x, y \in A$ we get that $\frac{x+y}{2} \in A$. Let $x, y \in A$. Since A is closed under taking midpoints, the set $\mathcal{L}_{x,y} = \{(1-t)x + ty : 0 \leq t \leq 1\}$ has a dense subset M that is contained in A. Since A is closed it follows that A must contain the closure of M, which is indeed $\mathcal{L}_{x,y}$. Hence A is convex.

What is left to prove is that $\operatorname{vol}(A) = \frac{n^n}{n!} 2^r \left(\frac{\pi}{2}\right)^s$. To show this we integrate over A to get the volume. By $V_{r,s}(t)$ we mean the volume of the subset of $\mathbb{R}^n$, $n = r + 2s$, defined by

$$|x_1| + \cdots + |x_r| + 2\left(\sqrt{x_{r+1}^2 + x_{r+2}^2} + \cdots + \sqrt{x_{n-1}^2 + x_n^2} \right) \leq t$$

Since we are dealing with n-dimensional volumes we have $V_{r,s}(t) = t^n V_{r,s}(1)$. Thus we only have to calculate $V_{r,s}(1)$. If $r = 0$, then $s > 0$ and we only have to calculate $V_{0,s}(1)$. Either way it actually reduces down to this (but not neccesarily with $s > 0$). To see this, suppose $r > 0$. Then

$$V_{r,s}(1) = 2\int_0^1 V_{r-1,s}(1-x)\,\mathrm{d}x = 2V_{r-1,s}(1)\int_0^1 (1-x)^{r-1+2s}\,\mathrm{d}x = \frac{2}{r+2s}V_{r-1,s}(1)$$

If we use this repeatedly we can reduce $V_{r,s}(1)$ to

$$V_{r,s}(1) = \frac{2^r}{(r+2s)(r+2s-1)\cdots(2s+1)}V_{0,s}(1)$$

where $V_{0,0}(1) = 1$. So it is enough to calculate $V_{0,s}(1)$ for $s > 0$ as claimed. We start similarly as for the reduction procedure from $V_{r,s}(1)$ to $V_{r-1,s}(1)$:

$$V_{0,s}(1) = \iint_{x^2+y^2\leq\frac{1}{4}} V_{0,s-1}(1-2\sqrt{x^2+y^2})\,\mathrm{d}x\mathrm{d}y$$

Then changing to polar coordinates gives

$$V_{0,s}(1) = \int_0^{2\pi}\int_0^{1/2} V_{0,s-1}(1-2\rho)\rho\mathrm{d}\rho\,\mathrm{d}\theta = 2\pi V_{0,s-1}(1)\int_0^{1/2}(1-2\rho)^{2(s-1)}\rho\,\mathrm{d}\rho$$

The last integral is easily calculated by a subsitution $u = 1-2\rho$ and we finally get

$$V_{0,s}(1) = V_{0,s-1}(1)\frac{\pi}{2}\frac{1}{2s(2s-1)}$$

Using this repeatedly we get a similar formula as above, namely

$$V_{0,s}(1) = \left(\frac{\pi}{2}\right)^s\frac{1}{(2s)!}$$

and piecing everything together we get

$$V_{r,s}(1) = \frac{2^r}{(r+2s)(r+2s-1)\cdots(2s+1)}V_{0,s}(1) = \left(\frac{\pi}{2}\right)^s\frac{2^r}{(r+2s)!}$$

and thus finally we get

$$\mathrm{vol}(A) = (r+2s)^{r+2s}V_{r,s}(1) = \frac{n^n}{n!}2^r\left(\frac{\pi}{2}\right)^s$$

which finishes the proof. □

Theorem 3.1.14. *Every n-dimensional lattice Λ in $\mathbb{R}^n$, where $n = r+2s$, contains some non-zero point x with*

$$|N(x)| \leq \frac{n!}{n^n}\left(\frac{8}{\pi}\right)^s \mathrm{vol}(\mathbb{R}^n/\Lambda)$$

Proof. By Lemma 3.1.13, there is some subset A of $\mathbb{R}^n$ that is compact, convex and centrally symmetric with $\mathrm{vol}(A) = \frac{n^n}{n!} 2^r \left(\frac{\pi}{2}\right)^s$ and the property $a \in A \implies |N(a)| \leq 1$. Let

$$t = \sqrt[n]{\frac{2^n}{\mathrm{vol}(A)} \mathrm{vol}(\mathbb{R}^n/\Lambda)}$$

Then tA is convex, centrally symmetric and compact. Furthermore

$$\mathrm{vol}(tA) = t^n \mathrm{vol}(A) = \frac{2^n}{\mathrm{vol}(A)} \mathrm{vol}(\mathbb{R}^n/\Lambda)\mathrm{vol}(A) = 2^n \mathrm{vol}(\mathbb{R}^n/\Lambda)$$

It follows from Minkowski's theorem that there is a non-zero $x \in tA \cap \Lambda$. Hence $x = ta$ for some $a \in A$ and because $|N(a)| \leq 1$:

$$|N(x)| = |N(ta)| \leq t^n |N(a)| \leq t^n = \frac{2^n}{\mathrm{vol}(A)} \mathrm{vol}(\mathbb{R}^n/\Lambda)$$

Inserting the value of $\mathrm{vol}(A)$ gives

$$|N(x)| \leq \frac{n! 2^s 2^n}{2^r n^n \pi^s} \mathrm{vol}(\mathbb{R}^n/\Lambda) = \frac{n!}{n^n} \left(\frac{8}{\pi}\right)^s \mathrm{vol}(\mathbb{R}^n/\Lambda)$$

□

Corollary 3.1.15. *Let K be a number field with $[K : \mathbb{Q}] = n = r + 2s$. Then every non-zero ideal I of $\mathcal{O}_K$ has a non-zero element α such that*

$$|N(\alpha)| \leq \frac{n!}{n^n} \left(\frac{4}{\pi}\right)^s \sqrt{|\mathrm{disc}(\mathcal{O}_K)|} \|I\|$$

Furthermore, every ideal class of $\mathcal{O}_K$ (under the equivalence relation described in the start of this section) contains an ideal J such that

$$\|J\| \leq \frac{n!}{n^n} \left(\frac{4}{\pi}\right)^s \sqrt{|\mathrm{disc}(\mathcal{O}_K)|}$$

As a consequence, the ideal class group of $\mathcal{O}_K$ is finite.

Proof. From Theorem 3.1.14, the lattice Λ_I has a non-zero point x satisfying the bound in the referenced theorem. Then $x = N(\alpha)$, $0 \neq \alpha \in I$. Using the bound given from the theorem and also Theorem 3.1.8 we get the bound

$$|N(\alpha)| \leq \frac{n!}{n^n} \left(\frac{4}{\pi}\right)^s \sqrt{|\mathrm{disc}(\mathcal{O}_K)|} \|I\|$$

Take any ideal class $\mathcal{C}$ in the ideal class group. Since the ideal class group is indeed a group, there is an inverse class $\mathcal{C}^{-1}$. Let $I \in \mathcal{C}^{-1}$. We know from Theorem 2.1.4 that for any non-zero $\alpha \in I$, there is a non-zero ideal J such that $IJ = (\alpha)$. Then $IJ \sim \mathcal{O}_K$, so $J \in \mathcal{C}$. Take $\alpha \in I$ to be the same as the α in the first part of this proof. It then follows that

$$\|J\| = \frac{\|(\alpha)\|}{\|I\|} = \frac{1}{\|I\|}|N(\alpha)| \leq \frac{n!}{n^n}\left(\frac{4}{\pi}\right)^s \sqrt{|\mathrm{disc}(\mathcal{O}_K)|} =: \lambda$$

For the last part of the statement, let λ^* be the least positive integer such that $\lambda^* \geq \lambda$. For each $i = 1, 2, \dots, \lambda^*$, there are at most finitely many ideals J satisfying $\|J\| \leq i$ by Theorem 2.2.5. Hence there can be at most finitely many ideals satisfying $\|J\| \leq \lambda$. But every ideal class has an ideal that is equivalent to some ideal with norm $\leq \lambda$. Thus there cannot be more than finitely many ideal classes. This proves the finiteness of the ideal class group. □

3.2 The unit group

The goal of this section is to prove the unit theorem due to Dirichlet. The geometric ideas from the previous section give us a quite explicit description of the group of units of $\mathcal{O}_K$. Again we use our embedding of K into a lattice and Minkowski's theorem to get the existence of certain algebraic numbers satisfying specific inequalities. This time however we introduce another space as well: the logarithmic space.

Definition 3.2.1. Let K be a number field with degree $n = r+2s$ over $\mathbb{Q}$. For $(x_1, \dots, x_n) \in \Lambda_{\mathcal{O}_K}\backslash\{0\}$, define $\log : \Lambda_{\mathcal{O}_K}\backslash\{0\} \to \mathbb{R}^{r+s}$ by

$$(x_1, \dots, x_n) \mapsto (\log|x_1|, \dots, \log|x_r|, \log(x_{r+1}^2 + x_{r+2}^2), \dots, \log(x_{n-1}^2 + x_n^2))$$

Here the log in the coordinates is just the normal (natural) logarithm. This is well-defined: first we show that $|x_1|, \dots, |x_r|$ are all strictly positive. Recalling the definition of Lat_K we see that if one of these are zero, then all $x_1, \dots, x_r$ are zero and also all $x_{r+1}^2 + x_{r+2}^2, \dots, x_{n-1}^2 + x_n^2$. This is because all the embeddings in Lat_K have trivial kernels. Similarly if one of the $x_{t+1}^2 + x_{t+2}^2$ are zero, then this can happen if and only if $x_{t+1} = x_{t+2} = 0$, which again implies that all of $x_1, \dots, x_n, x_{r+1}^2 + x_{r+2}^2, \dots, x_{n-1}^2 + x_n^2$ are zero. Hence the log function is well-defined. If U is the group of units of $\mathcal{O}_K$ we have the following sequence:

$$U \longrightarrow \mathcal{O}_K\backslash\{0\} \xrightarrow{\mathrm{Lat}_K} \Lambda_{\mathcal{O}_K}\backslash\{0\} \xrightarrow{\log} \mathbb{R}^{r+s}$$

From now on we will denote the composition $\log \circ \mathrm{Lat}_K$ by log and $\mathbb{R}^{r+s}$ as the logarithmic space. Given this slight redefinition of log we prove some nice properties.

Lemma 3.2.2. *Let K be a number field with degree $n = r + 2s$ over $\mathbb{Q}$. Let U be the group of units of $\mathcal{O}_K$. Then* log *is a group homomorphism $U \to \mathbb{R}^{r+s}$. Furthermore, its image is contained in a hyperplane $H \subset \mathbb{R}^{r+s}$ defined by $y_1 + \cdots + y_{r+s} = 0$. Finally, any bounded set in $\mathbb{R}^{r+s}$ has a finite preimage (by* log*) in U.*

Proof. If σ_i is a purely real embedding then

$$\log|\sigma_i(ab)| = \log|\sigma_i(a)\sigma_i(b)| = \log|\sigma_i(a)| + \log|\sigma_i(b)|$$

If τ_i is an embedding that is not purely real then observe that $(\mathrm{Re}(\tau_i(ab)))^2 + (\mathrm{Im}(\tau_i(ab))^2 = |\tau_i(ab)|^2$ and hence

$$\begin{aligned}&\log((\mathrm{Re}(\tau_i(ab)))^2 + (\mathrm{Im}(\tau_i(ab)))^2)\\&= \log((\mathrm{Re}(\tau_i(a)))^2 + (\mathrm{Im}(\tau_i(a)))^2) + (\mathrm{Re}(\tau_i(b)))^2 + (\mathrm{Im}(\tau_i(b)))^2\end{aligned}$$

These two equalities show that log is a group homomorphism as we get $\log(\alpha\beta) = \log(\alpha) + \log(\beta)$ for $\alpha, \beta \in \mathcal{O}_K\backslash\{0\}$. Now suppose $\alpha \in U$. Then there is some $\beta \in U$ such that $\alpha\beta = 1$. We then have $N(\alpha)N(\beta) = N(\alpha\beta) = N(1) = 1$. Since α, β are algebraic integers, the norm is a non-zero integer (see Theorem 1.1.9), hence $N(\alpha) = \pm 1$. Let $\mathrm{Lat}_K(\alpha) = (x_1, \ldots, x_r)$. Using that $\log(ab) = \log(a) + \log(b)$ for $a, b \in \mathbb{R}^+$ we get

$$\log|x_1| + \cdots + \log|x_r| + \log(x_{r+1}^2 + x_{r+2}^2) + \cdots + \log(x_{n-1}^2 + x_n^2) = \log|N(\alpha)| = 0$$

which proves the second claim. Let M be any bounded subset of $\mathbb{R}^{r+s}$. Then we can find a polydisc $\Delta^n(\mathbf{x}; \xi) = \{(x_1, \ldots, x_{r+s}) : |x_1| < \xi, \ldots, |x_{r+s}| < \xi\}$ which contains M. Let us first consider $\log|x| : \mathbb{R}\backslash\{0\} \to \mathbb{R}$, and moreover the preimage of $\mathbb{B}^1(\xi) = \{x \in \mathbb{R} : |x| < \xi\}$:

$$\log^{-1}|\mathbb{B}^1(\xi)| = \{x \in \mathbb{R} : \log|x| < \xi\} = \{x \in \mathbb{R} : |x| < e^\xi\} = \mathbb{B}^1(e^\xi)$$

Next let us consider the logarithm of the form $\log(x_{t+1}^2 + x_{t+2}^2)$, which we for simplicity name ϱ. Then we have

$$\varrho^{-1}(\mathbb{B}^1(\xi)) = \{x \in \mathbb{R}^2 : \log(x_{t+1}^2 + x_{t+2}^2) < \xi\} = \{x \in \mathbb{R}^2 : x_{t+1}^2 + x_{t+2}^2 < e^\xi\} = \mathbb{B}^2(e^{\xi/2})$$

where $\mathbb{B}^2(\xi)$ is the 2-dimensional ball with radius ξ. It follows that the preimage of the logarithm in our lattice (that is we don't consider the composition $\log \mathrm{Lat}_K$ but only log) is such that

$$\log^{-1}(M) \subseteq \log^{-1}(\Delta^n(\mathbf{x}; \xi)) \subseteq \mathbb{B}^1(e^\xi) \times \cdots \times \mathbb{B}^1(e^\xi) \times \mathbb{B}^2(e^{\xi/2}) \times \cdots \times \mathbb{B}^2(e^{\xi/2})$$

which shows the preimage of M is also bounded. However a bounded set in a lattice can clearly only contain finitely many points, and so the preimage of M in the lattice is finite. As Lat_K is injective it follows that the preimage of M by the composition $\log \mathrm{Lat}_K$ is finite. This proves the last claim. □

Observe that the third point in the lemma actually implies that any bounded subset of $\log(U)$ is finite: for let M be a bounded subset of $\log(U)$ and suppose that M is not finite. Then since log is surjective onto $\log(U)$ and $M \subseteq \log(U)$, this forces the preimage $\log^{-1}(M)$ to be infinite, but this is a contradiction to the lemma we just proved. This is an important observation because it acutally implies that $\log(U)$ is a lattice:

Lemma 3.2.3. *Let G be a subgroup of $\mathbb{R}^m$ such that every bounded subset of G is finite. Then G is a lattice.*

Proof. G contains a lattice $\{0\}$, so there exists a (not necessarily unique) lattice Λ of maximal dimension, say d, contained in G. Let $\{v_1, \ldots, v_d\}$ be a $\mathbb{Z}$-basis for Λ. Let $v \in G$. Then $\{v, v_1, \ldots, v_d\}$ must be a linearly dependent set in $\mathbb{R}^m$ since otherwise this would contradict the maximality of Λ. Hence G is contained in the subspace of $\mathbb{R}^m$ generated by Λ. Now we claim that G/Λ is finite. To see this fix a fundamental parallellotope $\mathcal{F}$ for Λ. Let $v+\Lambda, v \in G$ be a a coset. Since G is contained in the subspace of $\mathbb{R}^m$ generated by G, there are $a_i \in \mathbb{R}$ such that

$$g = a_1 v_1 + \cdots + a_d v_d$$

For each a_i let us write $a_i = n_i + r_i$ where $n_i \in \mathbb{Z}$ and $0 \leq r_i < 1$. Then $r_1 v_1 + \cdots + r_d v_d \in (v+\Lambda) \cap \mathcal{F}$. This shows that every coset has a representative in $\mathcal{F}$. Now $\mathcal{F} \cap G$ is a bounded subset of G and hence is finite. As $\mathcal{F} \cap G$ contains a representative from each coset $v+\Lambda$, this shows G/Λ is finite. From this it is easy to see that $|G/\Lambda| G \subseteq \Lambda$, because if we pick any $g \in G$, then $|G/\Lambda| g \equiv 0 \pmod{\Lambda}$, that is $|G/\Lambda| g \in \Lambda$. Let us write $r = |G/\Lambda|$. By Lemma 1.2.2 we get that rG is a free abelian group of rank $\leq d$. From this we can conclude G is actually free abelian of rank d. Fix a basis $\{rw_1, \ldots, rw_k\}$ for rG. Then $\{w_1, \ldots, w_k\}$ is a basis for G, so G is free abelian of rank $\leq d$. However $\Lambda \subseteq G$, so the rank has to be d. We now have all the ingredients to show that G is a lattice. Namely, fix a $\mathbb{Z}$-basis $\{w_1, \ldots, w_d\}$ for G. This must be $\mathbb{R}$-independent because G contains a basis for Λ of cardinality d which is $\mathbb{R}$-independent. Hence G is a lattice and we are done. □

Before we prove the unit theorem we need two lemmata.

Lemma 3.2.4. *Let K be a number field with group of units U, and $[K:\mathbb{Q}] = n = r+2s$. Let k be an integer such that $1 \leq k \leq r+s$. Then there is some $\mu \in U$ with the property that if $\log(\mu) = (y_1, \ldots, y_{r+s})$, then $y_i < 0$ for all $i \neq k$.*

Proof. We first prove that for any non-zero $\alpha \in \mathcal{O}_K$ there is some $\beta \in \mathcal{O}_K$ such that

$$|N(\beta)| \leq \left(\frac{2}{\pi}\right)^s \sqrt{|\mathrm{disc}(\mathcal{O}_K)|}$$

and such that if $\log(\alpha) = (a_1, \ldots, a_{r+s}), \log(\beta) = (b_1, \ldots, b_{r+s})$, then $b_i < a_i$ for all $i \neq k$. We achieve this by Minkowski's theorem. Let

$$E = \{(x_1, \ldots, x_{r+s}) \in \mathbb{R}^{r+s} : |x_1| \leq c_1, \ldots |x_r| \leq c_r, x_{r+1}^2 + x_{r+2}^2 \leq c_{r+1}, \ldots, x_{n-1}^2 + x_n^2 \leq c_{r+s}\}$$

where $0 < c_i < e^{a_i}$ for all $i \neq k$ and

$$c_1 \cdots c_{r+s} = \left(\frac{2}{\pi}\right)^s \sqrt{|\mathrm{disc}(\mathcal{O}_K)|}$$

Since c_k is not bounded we can meet this equality requirement. We need to check that E satisfies the assumptions of Minkowski's theorem. It is clear that E is centrally symmetric. For the convexity: observe that E is a cartesian product of convex sets (intervals and discs). That the inequalities are non-strict implies that E is closed (E is a cartesian product of closed lines and discs), and E is clearly bounded. The volume is easily calculated:

$$\mathrm{vol}(E) = (2c_1)\cdots(2c_r)(\pi c_{r+1})\cdots(\pi c_{r+s}) = 2^{r+s}\sqrt{|\mathrm{disc}(\mathcal{O}_K)|} = 2^n \mathrm{vol}(\mathbb{R}^n/\Lambda_{\mathcal{O}_K})$$

where the last equality is due to Theorem 3.1.8. Then Minkowski's theorem gives us a non-zero $\beta \in \mathcal{O}_K$ such that

$$|N(\beta)| \leq c_1 \cdots c_{r+s} = \left(\frac{2}{\pi}\right)^s \sqrt{|\mathrm{disc}(\mathcal{O}_K)|}$$

Furthermore if $\mathrm{Lat}_K(\beta) = (w_1, \ldots, w_{r+s})$, then for $i = 1, \ldots, r$, $i \neq k$ we have $|w_i| \leq c_i < e^{a_i}$ and for $i = r+1, \ldots, r+s-1$, $i \neq k$, we have $w_i^2 + w_{i+1}^2 \leq c_i < e^{a_i}$. Taking log gives

$$\log|w_i| < a_i \qquad \log(w_i^2 + w_{i+1}^2) < a_i$$

so if $\log(\beta) = (b_1, \ldots, b_{r+s})$ we get $b_i < a_i$ for all $i \neq k$, as desired. Now we can prove the statement in the lemma. If we start with any non-zero α_1, we can use what we just proved to get a sequence $\alpha_1, \alpha_2, \alpha_3, \dots \in \mathcal{O}_K$ such that for each $i \neq k$, and for each $j \geq 1$, the ith coordinate of $\log(\alpha_j)$ is larger than the ith coordinate of $\log(\alpha_{j+1})$. Furthemore each α_j is such that

$$\|(\alpha_j)\| = |N(\alpha_j)| \leq \left(\frac{2}{\pi}\right)^s \sqrt{|\mathrm{disc}(\mathcal{O}_K)|}$$

The quantity on the right is finite, and hence, just like the proof in Theorem 3.1.15, this implies that there can be at most finitely many such ideals (α_j). Hence fix some ℓ, t such that $(\alpha_\ell) = (\alpha_t)$ with $t > \ell$. Then we have $\mu, \nu \in \mathcal{O}_K$ such that $\alpha_\ell = \nu\alpha_t$, $\alpha_t = \mu\alpha_\ell$. Hence $\alpha_\ell = \nu\alpha_t = \mu\nu\alpha_\ell$, so $\mu\nu = 1$ since $\mathcal{O}_K$ is an integral domain. Hence μ is a unit: $\mu \in U$ and $\mu = \dfrac{\alpha_t}{\alpha_\ell}$. It makes sense to talk about the composition $\log \circ \mathrm{Lat}_K$ for all of $K \backslash \{0\}$ as well (log is still well-defined), and $\log(\alpha\beta) = \log(\alpha) + \log(\beta)$ still. Taking logs we obtain

$$\log(\mu) = \log\left(\frac{\alpha_t}{\alpha_\ell}\right) = \log(\alpha_t) + \log\left(\frac{1}{\alpha_\ell}\right) = \log(\alpha_t) - \log(\alpha_\ell)$$

It follows that $\log(\mu) = (y_1, \ldots, y_{r+s})$ is such that $y_i < 0$ for all $i \neq k$ because $t > \ell$. □

The next lemma is pure linear algebra

Lemma 3.2.5. *Let $M = (m_{ij})$ be an $\ell \times \ell$-matrix with entries in $\mathbb{R}$, such that the diagonal is positive: $m_{ii} > 0$ for all i, and everything else is negative: $m_{ij} < 0$ for all $i \neq j$. Furthermore suppose each row sums to 0. Then the rank of M is $\ell - 1$.*

Proof. Let the first $\ell - 1$ columns of M be $v_1, \ldots, v_{\ell-1}$ and suppose $t_1 v_1 + \cdots + t_{\ell-1} v_{\ell-1} = 0$, $t_j \in \mathbb{R}$. Suppose not all t_j are zero. Divide the expression by the largest t_k. Then we can without loss of generality assume $t_k = 1$ and all other $t_j \leq 1$. Now we calculate:

$$0 = \sum_{j=1}^{\ell-1} t_j m_{kj} = m_{kk} + \sum_{j=1, j\neq k}^{\ell-1} t_j m_{kj} \geq \sum_{j=1}^{\ell-1} m_{kj} > \sum_{j=1}^{\ell} m_{kj} = 0$$

The first inequality comes from the fact that all m_{kj}, $k \neq j$, are negative, the last inequality comes from the fact that they are strictly negative. The inequality $0 < 0$ gives a contradiction, so all t_j are zero. This shows that the $\mathrm{rank}(M) \geq \ell - 1$. By assumption we have $0 = m_{i1} + \cdots + m_{i\ell}$ so $m_{i\ell} = -m_{i1} + \cdots + m_{i(\ell-1)}$ from which it follows $v_\ell = -v_1 - \cdots - v_{\ell-1}$. Thus $\mathrm{rank}(M) \neq \ell$ which finishes the proof. □

We are now ready to prove the unit theorem.

Theorem 3.2.6. (Dirichlet's Unit Theorem)
Let K be a number field with degree n over $\mathbb{Q}$ and $n = r + 2s$. Let U be the group of units of $\mathcal{O}_K$. Then $U \cong W \times V$ where W is a finite cyclic group that consists of the roots of unity of K, and V is a free abelian group of rank $r + s - 1$.

Proof. By Lemma 3.2.2, the preimage of $\{0\}$ by log in U is finite. That is, the kernel of log is finite and hence every element ζ in the kernel has finite order: $\zeta^t = 1$. Hence the kernel consist of roots of unity. Conversely, if ζ is a root of unity in K, say $\zeta^\ell = 1$, then $0 = \log(\zeta^\ell) = \ell \log(\zeta)$ and hence $\log(\zeta) = 0$ so $\zeta \in \ker \log$. It is a classical result from field theory that all finite subgroups of F^* for any field F are cyclic. The roots of unity in K is as shown above a finite subgroup of K^*, and hence cyclic. Continuing, the remark leading up to Lemma 3.2.3, as well as the lemma itself, implies that $\log(U)$ is a lattice. We also know from Lemma 3.2.2 that it is contained in a hyperplane $H \subseteq \mathbb{R}^{r+s}$, and hence the rank of $\log(U)$ is at most $r+s-1$. As we have just seen, $\log(U)$ is free abelian of rank $d \leq r+s-1$, so fix $u_1, \ldots, u_d \in U$, such that $\log(u_1), \ldots, \log(u_d)$ is a $\mathbb{Z}$-basis for $\log(U)$. Now we claim that the V in the theorem is the subgroup of U generated multiplicatively by $u_1, \ldots, u_d$. We first prove that $u_1, \ldots, u_d$ generate V freely. To that end, assume $u_1^{k_1} \cdots u_d^{k_d} = 1$. Taking log we get

$$k_1 \log(u_1) + \cdots + k_d \log(u_d) = 0$$

which implies $k_1 = \cdots = k_d = 0$ as $\{\log(u_1), \ldots, \log(u_d)\}$ constitutes a $\mathbb{Z}$-basis of $\log(U)$. This shows V is free abelian of rank d. Now we claim that indeed $U \cong W \times V$. To show this, we first show $W \cap V = \{1\}$. Let $x = u_1^{k_1} \cdots u_d^{k_d} \in V \cap W$. Then there is some non-zero integer ℓ such that $x^\ell = 1$. Taking log we obtain

$$\ell k_1 \log(u_1) + \cdots + \ell k_d \log(u_d) = 0$$

which by linear independence in $\log(U)$, implies $\ell k_1 = \cdots = \ell k_d = 0$, and hence $k_1 = \cdots = k_d = 0$. It follows that $x = 1$. Now let $x \in U$. Then $\log(x) = a_1 \log(u_1) + \cdots + a_d \log(u_d) = \log(u_1^{a_1} \cdots u_d^{a_d})$, which goes to show that

$$\log(x u_1^{-a_1} \cdots u_d^{-a_d}) = 1 \implies x u_1^{-a_1} \cdots u_d^{-a_d} \in \ker \log$$

so there is some $z \in \ker \log = W$ such that $x = z u_1^{a_1} \cdots u_d^{a_d} \in WV$. Conversely $VW \subseteq U$, as both V and W are multiplicative subgroups of U. Now define $\varphi : W \times V \to WV$ by $(\zeta, u_1^{a_1} \cdots u_d^{a_d}) \mapsto \zeta u_1^{a_1} \cdots u_d^{a_d}$. Since $u_1, \ldots, u_d$ generate V freely this is well-defined. Since any $x \in U$ can be written as a product in WV it follows that φ is surjective. Suppose $\zeta u_1^{a_1} \cdots u_d^{a_d} = 1$. Then $\zeta u_1^{a_1} \cdots u_d^{a_d} \in W \cap V$. Multiplying, it follows that $\zeta \in V$ and $u_1^{a_1} \cdots u_d^{a_d} \in W$. This implies $\zeta = u_1^{a_1} \cdots u_d^{a_d} = 1$. Hence we have a trivial kernel, and φ is injective, and hence furthermore an isomorphism: $U = WV \cong W \times V$. The only thing that is left to prove is that $d = r+s-1$. We do this the completely obvious way: produce $r+s-1$ elements in U whose images in $\log(U)$ are linearly independent. Lemma 3.2.4 gives us the existence of $\mu_1, \ldots, \mu_{r+s}$ such that all coordinates of $\log(u_j)$ are negative excect the jth row. Since $\log(u_j)$ is contained in a hyperplane H defined by $y_1 + \cdots + y_{r+s} = 0$, this forces the jth coordinate to be positive. If we now form the $(r+s) \times (r+s)$-matrix $\mathcal{M}$ with $\log(\mu_j)$

as its jth row, we get a matrix where the diagonal is strictly positive, and strictly negative elsewhere. Furthermore each row sums to zero since $\log(U)$ is contained in the hyperplane H. By Lemma 3.2.5 we see that $\mathcal{M}$ has rank $r+s-1$. Hence $\{\log(\mu_1), \ldots, \log(\mu_{r+s})\}$ has $r+s-1$ linearly independent vectors. This finishes the proof. □

3.3 Distribution of ideals in $\mathcal{O}_K$

We now give a rather explicit description of the distribution of ideals in $\mathcal{O}_K$ with respect to their ideal norm. From this distribution the class number formula will follow without too much work. For each real number $t \geq 0$ we define

$$i(t) = \#\{I | I \text{ ideal in } \mathcal{O}_K, \|I\| \leq t\}$$

i.e. the number of ideals in $\mathcal{O}_K$ with norm at most t. Given some ideal class C we define

$$i_C(t) = \#\{I | I \text{ ideal in } \mathcal{O}_K, I \in C, \|I\| \leq t\}$$

i.e. the number of ideals in C with norm at most t. Summing over all ideal classes we get

$$i(t) = \sum_C i_C(t)$$

Since the ideal class group is finite (Corollary 3.1.15) this sum is finite. We will show that $i_C(t) = \kappa t + O\left(t^{1-\frac{1}{n}}\right)$ for any ideal class C, where $n = [K : \mathbb{Q}]$ and κ is some constant independent of C. Since the sum is finite, it will then follow that

$$i(t) = h\kappa t + O\left(t^{1-\frac{1}{n}}\right)$$

where h is the class number (the order of the ideal class group). We will also determine the constant κ. It turns out that this constant κ encodes many invariants of a number field:

$$\kappa = \frac{2^{r+s}\pi^s \operatorname{reg}(O_K)}{w\sqrt{|\operatorname{disc}(\mathcal{O}_K)|}}$$

where as before r is the number of real embeddings of K into $\mathbb{C}$ and s is half the number of non-real embeddings K in $\mathbb{C}$. Furthermore w is the number of roots of unity in K and finally $\operatorname{reg}(\mathcal{O}_K)$ is the regulator of $\mathcal{O}_K$. We will come more back to this, but say for now that it, loosely speaking, measures the density of the units: a small regulator means many units. We will need two lemmata for the distribution theorem.

Lemma 3.3.1. *Let $f : G \to G'$ be a homomorphism of abelian groups and let S be a subgroup of G which is sent isomorphically to a subgroup S' of G' by f. If D' is a set of coset representatives for S' in G', then $D = f^{-1}(D')$ is a set of coset representatives for S in G.*

Proof. First take $x, y \in D$ such that $x \neq y$ and assume $x \equiv y \pmod{S}$. Then $x - y \in S$ so $f(x-y) \in S'$, in other words $f(x) \equiv f(y) \pmod{S'}$. Since $f|_S$ is an isomorphism we have $f(x) - f(y) = f(x-y) \neq 0$ since $x \neq y$. But $f(x), f(y) \in D'$, so they cannot be not equal while congruent modulo S' since D' is a set of coset representative of S' in G'. This shows that the elements in D represent different cosets of S in G. To see that we hit all, let z be a coset represenative in G/S. Then $f(z) \equiv w \pmod{S'}$ for some $w \in D'$. Since $f|_S$ is an isomorphism, there is $\widetilde{w} \in S$ such that $f(\widetilde{w}) = f(z) - w$ so $w = f(z - \widetilde{w})$. Hence $z - \widetilde{w} \in D$. Since $\widetilde{w} \in S$, $z - \widetilde{w} \equiv z \pmod{S}$ this gives us the desired coset representative. □

For the next lemma we need a definition.

Definition 3.3.2. A bounded set $B \subseteq \mathbb{R}^n$ is said to have $(n-1)$-Lipschitz parameterizable boundary if its boundary ∂B is contained in the union of the images of finitely many Lipschitz functions $f : [0,1]^{n-1} \to \mathbb{R}^n$. That f is Lipschitz in this case means that there is a constant C such that

$$\left| \frac{f(x) - f(y)}{x - y} \right| \leq C \qquad x, y \in [0,1]^{n-1}$$

Lemma 3.3.3. *Let Λ be any n-dimensional lattice in $\mathbb{R}^n$ and B any bounded subset of $\mathbb{R}^n$. If B is $(n-1)$-Lipschitz parameterizable then*

$$|\Lambda \cap aB| = \frac{\mathrm{vol}(B)}{\mathrm{vol}(\mathbb{R}^n/\Lambda)} a^n + O(a^{n-1})$$

A proof of Lemma 3.3.3 can be found in [8, Lemma 6.2]. The idea is to first reducing the problem to $\mathbb{Z}^n$ and then consider translates of n-cubes $[0,1]^n$.

Theorem 3.3.4. *Let K be a number field, $n = [K : \mathbb{Q}]$. For any ideal class C,*

$$i_C(t) = \kappa t + O\left(t^{1-\frac{1}{n}}\right)$$

where κ is the aforementioned constant.

Proof. This proof is rather long. We start by summarizing what we are going to do, to try to avoid getting completely lost in the details. Start by fixing an ideal $J \in C^{-1}$ (the inverse ideal class). Then we claim there is a bijection

$$\begin{aligned} \{\text{ideals } I \in C \text{ with } \|I\| \leq t\} &\longleftrightarrow \{\text{principal ideals } (\alpha) \subseteq J \text{ with } \|(\alpha)\| \leq t\|J\|\} \\ I &\longmapsto IJ \end{aligned}$$

This reduces the counting of ideals with norm at most t to counting principal ideals in J with certain bounded norm. Now counting principal ideals is almost like counting elements apart from the fact that α is determined from (α) up to a unit factor. It would be nice if we could ignore this problem. To do this we construct a subset of $\mathcal{O}_K$ in which no two elements differ by a unit factor and every non-zero element of $\mathcal{O}_K$ has a unit multiple. After we have passed to this set, we can simply count elements of J there. This set is exactly a set of coset representatives for the unit group in the multiplicative semigroup $\mathcal{O}_K\backslash\{0\}$. Then we pass to a lattice via the log-mapping and show that the problem reduces to calculating the volume of a certain domain in $\mathbb{R}^n$. Then we count another way and compare to get a formula for $i_C(t)$. How this is done more specifically will get more clear later in the proof.

Let us get going: first we show that the map above is well-defined. Since $IJ \sim \mathcal{O}_K$ there are non-zero $\gamma, \delta \in \mathcal{O}_K$ such that $\gamma IJ = \delta\mathcal{O}_K$. Hence there is non-zero $\eta \in IJ$ such that $\gamma\eta = \delta$. Thus $\frac{\delta}{\gamma} = \eta \in IJ \subseteq \mathcal{O}_K$. Hence $IJ = \left(\frac{\delta}{\gamma}\right)$ is a principal ideal and is contained in J as $IJ \subseteq J$. Finally $\left\|\left(\frac{\delta}{\gamma}\right)\right\| = \|IJ\| \leq t\|J\|$. This shows well-definedness. If $IJ = I'J$ then $I = I'$ by Corollary 2.1.5, so the map is injective. Finally let $(x) \subseteq J$. Then by Corollary 2.1.6, $J \mid (x)$, that is there is an ideal I so that $IJ = (x)$. Since $(x) \sim \mathcal{O}_K$, $I \in C$ and finally $\|I\|\|J\| = \|IJ\| = \|(x)\| \leq t\|J\|$ so $\|I\| \leq t$. This shows surjectivity, proving the bijection.

Now we move on to constructing the set of coset representatives mentioned above. Actually it will be sufficient to construct a set of coset representatives for a free abelian subgroup $V \subseteq U$ of rank $r+s-1$. Such a V exists by Dirichlet's Unit Theorem: recall $U \cong V \times W$ where W is the group of roots of unity in K. Such a V is not unique so let us fix one such V. We will see why this is sufficient soon: it comes down to the fact that we can easily account for the roots of unity. We have the following sequence of maps:

$$V \hookrightarrow U \hookrightarrow \mathcal{O}_K\backslash\{0\} \xrightarrow{\log} \mathbb{R}^{r+s}$$

where we by log really mean $\log \circ \mathrm{Lat}_K$ as mentioned earlier. In the proof of Dirichlet's Unit Theorem we saw that $\log(U)$ is a lattice, which we from now on denote Λ_U. From the same proof we also saw that W is the kernel of the log-map. Since $U \cong V \times W$ it thus follows that the restriction of

$$U \hookrightarrow \mathcal{O}_K\backslash\{0\} \xrightarrow{\log} \mathbb{R}^{r+s}$$

to V is an isomorphism onto Λ_U. Let us now change our perspective slightly. Before we viewed $\Lambda_{\mathcal{O}_K}$ as a subset of $\mathbb{R}^{r+2s}$ - now we instead think of $\Lambda_{\mathcal{O}_K}$ as a subset of $(\mathbb{R})^r \times (\mathbb{C})^s$ and hence $\Lambda_{\mathcal{O}_K}\backslash\{0\}$ as a subset of $(\mathbb{R}^*)^r \times (\mathbb{C}^*)^s$. This requires a slight redefinition of the log-map as well where we identify $\mathbb{R}^2$ with $\mathbb{C}$:

$$z = a + bi \mapsto \log(a^2 + b^2) = \log(|z|^2) = 2\log(|z|)$$

so in other words we now have the map

$$(x_1, \ldots, x_r, z_1, \ldots, z_s) \overset{\log}{\longmapsto} (\log|x_1|, \ldots, \log|x_r|, 2\log|z_1|, \ldots, 2\log|z_s|)$$

and the map Lat_K is now

$$\alpha \mapsto (\sigma_1(\alpha), \ldots, \sigma_r(\alpha), \tau_1(\alpha), \ldots, \tau_s(\alpha))$$

where the σ_i are the real embeddings, and $\tau_i, \overline{\tau_i}$ are the non-real embeddings with their complex conjugates as usual. The restriction of Lat_K to $\mathcal{O}_K\backslash\{0\}$ defines a multiplicative embedding, and so V maps isomorphically onto a subgroup V' of $(\mathbb{R}^*)^r \times (\mathbb{C}^*)^s$. Thus we can obtain a set of coset representatives for V in $\mathcal{O}_K \backslash \{0\}$ from a set of coset representatives for V' in $(\mathbb{R}^*)^r \times (\mathbb{C}^*)^s$. This is called a fundamental domain for V'. Since Lat_K takes ideals in $\mathcal{O}_K$ to sublattices Λ_I of $\Lambda_{\mathcal{O}_K}$, counting members of the ideal J in the desired subset of coset representatives of V is the same as counting elements of Λ_J in the fundamental domain for V'. Let us translate the norm criterion $\|(\alpha)\| \leq t\|J\|$ to this setting: recall that we had a norm on $\mathbb{R}^n$, namely N^K, that was the same as N when restricted to the image of K. Now with our change from $\mathbb{R}^{r+2s}$ to $\mathbb{R}^r \times \mathbb{C}^s$ this norm is defined by

$$N(x_1, \dots, x_r, z_1, \dots, z_s) = x_1 \cdots x_r |z_1|^2 \cdots |z_s|^2$$

By Lemma 2.2.4 we get

$$\|(\alpha)\| = |N(\alpha)| = |\sigma_1(\alpha) \cdots \sigma_r(\alpha) \underbrace{\tau_1(\alpha)\overline{\tau_1}(\alpha)}_{|\tau_1(\alpha)|^2} \cdots \underbrace{\tau_s(\alpha)\overline{\tau_s}(\alpha)}_{|\tau_s(\alpha)|^2}| = |N(\mathrm{Lat}_K(\alpha))|$$

Hence if we let x denote the image of α in $(\mathbb{R}^*)^r \times (\mathbb{C}^*)^s$ then $\|(\alpha)\| \leq t\|J\|$ is equivalent to $|N(x)| \leq t\|J\|$. Thus what we have to do is find a set D of coset representatives for V' in $(\mathbb{R}^*)^r \times (\mathbb{C}^*)^s$ and count elements $x \in \Lambda_J \cap D$ with $|N(x)| \leq t\|J\|$. The number of such x is almost the number of principal ideals $(\alpha) \subseteq J$ with $\|(\alpha)\| \leq t\|J\|$, but remember that we are counting inside a set of coset representatives for V - not for the whole of U. Thus each ideal (α) is counted $|W|$ times. From the bijection in the very beginning it follows that the number of $x \in \Lambda_J \cap D$ with $|N(x)| \leq t\|J\|$ is $|W| i_C(t)$. With the help of Lemma 3.3.3 we will count the number of $x \in \Lambda_J \cap D$ another way and from that deduce a value of $i_C(t)$. However there is an unknown volume in that expression that we need to calculate.

First however, let us construct the set D of coset represenatives for V' in $(\mathbb{R}^*)^r \times (\mathbb{C}^*)^s$. To do so we instead construct a set D' of coset representatives for Λ_U in $\mathbb{R}^{r+s}$, then apply Lemma 3.3.1 to the homomorphism $\log : (\mathbb{R}^*)^r \times (\mathbb{C}^*)^s \to \mathbb{R}^{r+s}$ remembering that $\log(V') \cong \Lambda_U$. From now on, fix a fundamental parallellotope $\mathcal{F}$ for the lattice Λ_U. Recall that Λ_U lies inside a hyperplane H defined by $y_1 + \cdots + y_{r+s} = 0$ (Lemma 3.2.2). By Dirichlet's Unit Theorem, Λ_U is a full lattice in H. We define its volume by using volume on H induced by the inner product. Let L be any line through the origin, but not contained in H. Then $\mathcal{F} \oplus L$ is a set of coset representatives for Λ_U in $\mathbb{R}^{r+s}$, and hence we let $D' = \mathcal{F} \oplus L$. We want to choose a “good” L. Another way to describe D' is by fixing a vector $v \in \mathbb{R}^{r+s} \backslash H$ and then let $D' = \mathcal{F} \oplus \mathbb{R}v$. Then by Lemma 3.3.1 the set

$$D = \{x \in (\mathbb{R}^*)^r \times (\mathbb{C}^*)^s : \log x \in \mathcal{F} \oplus \mathbb{R}v\}$$

is a fundamental domain for V' in $(\mathbb{R}^*)^r \times (\mathbb{C}^*)^s$. We will choose $v = (1, \dots, 1, 2, \dots, 2)$ where there are r ones and s twos. We will denote this vector $(1, \dots, 2)$ from now on. This is a good choice as then D becomes homogenenous: $D = aD$ for all $a \neq 0$. With the set of coset represenatives in place, let us count the elements in $\Lambda_J \cap D$ with $|N(x)| \leq t\|J\|$ in another way as promised. To do so we introduce the notation

$$D_a = \{x \in D : |N(x)| \leq a\}$$

Then $D_a = \sqrt[n]{a}D_1$. If we assume for a moment that D_1 is bounded and $(n-1)$-Lipschitz parameterizable and think of $(\mathbb{R}^*)^r \times (\mathbb{C}^*)^s$ the obvious way, then it follows from Lemma 3.3.3 that

$$\left|\Lambda_J \cap \sqrt[n]{t\|J\|}D_1\right| = \frac{\mathrm{vol}(D_1)}{\mathrm{vol}(\mathbb{R}^n/\Lambda_J)}t\|J\| + O\left(t^{1-\frac{1}{n}}\right)$$

Using Corollary 3.1.9 we then get

$$\left|\Lambda_J \cap \sqrt[n]{t\|J\|}D_1\right| = \frac{2^s\mathrm{vol}(D_1)}{\sqrt{|\mathrm{disc}(\mathcal{O}_K)|}}t + O\left(t^{1-\frac{1}{n}}\right)$$

and hence comparing with our count of $|\Lambda_J \cap D|$ earlier we arrive at

$$i_C(t) = \frac{2^s\mathrm{vol}(D_1)}{|W|\sqrt{|\mathrm{disc}(\mathcal{O}_K)|}}t + O\left(t^{1-\frac{1}{n}}\right)$$

which was what we wanted to prove, up to the calculation of $\mathrm{vol}(D_1)$. What we have left to prove is that D_1 is bounded and $(n-1)$-Lipschitz parameterizable and a calculation of $\mathrm{vol}(D_1)$. We start with the first. That $|N(x)| \leq 1$ is the same as $|x_1 \cdots x_r z_1^2 \cdots z_s^2| \leq 1$. Taking the logarithm we get

$$\log|x_1| + \cdots + \log|x_r| + 2\log|z_1| + \cdots + 2\log|z_s| \leq 0$$

and so

$$x \in D_1 \iff \log x \in \mathcal{F} \oplus (-\infty, 0](1, \ldots, 2)$$

Since $\mathcal{F}$ is bounded it follows from the definition of the log-map that the preimage of log by $\mathcal{F} \oplus (-\infty, 0](1, \ldots, 2)$ is bounded. Hence D_1 is bounded. To show that ∂D_1 is $(n-1)$-Lipschitz parameterizable we reduce to the subset

$$D_1^+ = D_1 \cap \{(x_1, \ldots, x_r, z_1, \ldots, z_s) : x_1, \ldots, x_r \geq 0\}$$

By how we have defined log we have that ∂D_1 is $(n-1)$-Lipschitz parameterizable if and only if ∂D_1^+ is $(n-1)$-Lipschitz parameterizable and furthermore $\mathrm{vol}(D_1) = 2^r\mathrm{vol}(D_1^+)$. If we have some $\mathbb{Z}$-basis $\{v_1, \ldots, v_{r+s-1}\}$ for Λ_U, then we write

$$\mathcal{F} = \left\{\sum_{k=1}^{r+s-1} t_k v_k : 0 \leq t_k < 1\right\}$$

For each k we denote

$$v_k = (v_k^{(1)}, \ldots, v_k^{(r+s)})$$

Recall that $x \in D_1 \iff \log x \in \mathcal{F} \oplus (-\infty, 0](1, \ldots, 2)$ so we can characterize a point

$(x_1, \ldots, x_r, z_1, \ldots, z_s) \in D_1^+$ by

$$\begin{aligned}
\log(x_1) &= \sum_{k=1}^{r+s-1} t_k v_k^{(1)} + u \\
&\vdots \\
\log(x_r) &= \sum_{k=1}^{r+s-1} t_k v_k^{(r)} + u \\
2\log|z_1| &= \sum_{k=1}^{r+s-1} t_k v_k^{(r+1)} + 2u \\
&\vdots \\
2\log|z_s| &= \sum_{k=1}^{r+s-1} t_k v_k^{(r+s)} + 2u
\end{aligned}$$

where the x_j are positive, $t_k \in [0,1)$ and finally $u \in (-\infty, 0]$. Define $t_{r+s} = e^u$. Let us write each z_j in polar coordinates: $z_j = \rho_j e^{i\theta_j}$. Then we get a new characterization of $(x_1, \ldots, x_r, \rho_1 e^{i\theta_1}, \ldots, \rho_s e^{i\theta_s}) \in D_1^+$ by applying the exponential function:

$$\begin{aligned}
x_j &= t_{r+s} e^{\sum_{k=1}^{r+s-1} t_k v_k^{(j)}} \\
\rho_j &= t_{r+s} e^{\frac{1}{2}\sum_{k=1}^{r+s-1} t_k v_k^{(r+j)}} \\
\theta_j &= 2\pi t_{r+s+j}
\end{aligned}$$

where $t_{r+s} \in (0,1]$ and all other $t_k \in [0,1)$. Hence all the t_k parameterize D_1^+ by the equation above: that is we have a parameterization of D_1^+ by a half-open n-cube. If we let the t_k take on their boundary value, then we get a parameterization of $\overline{D_1^+}$. To see this, let us first give our function a name

$$f : [0,1]^n \to \mathbb{R}^r \times \mathbb{C}^s \qquad (t_1, \ldots, t_n) \mapsto (x_1, \ldots, x_r, \rho_1 e^{i\theta_1}, \ldots, \rho_s e^{i\theta_s})$$

Then we claim $f([0,1]^n) = \overline{D_1^+}$. We observe that f is continuous, so as $[0,1]^n$ is compact, $f([0,1]^n)$ is compact. Of course $D_1^+ \subseteq f([0,1]^n)$. As $f([0,1]^n)$ is compact, it is closed, so $\overline{D_1^+} \subseteq f([0,1]^n)$. But the half-open n-cube is dense in $[0,1]^n$, so D_1^+ must be dense in $f([0,1]^n)$, so we conclude that $\overline{D_1^+} = f([0,1]^n)$.

Let us write $\overline{D_1^+} = \text{Interior}(D_1^+) \sqcup \partial D_1^+ = I \sqcup B$ where $\sqcup$ denotes disjoint union as before. The boundary of the n-cube is the union of $2n$ $(n-1)$-cubes. Now, assume that we knew $f((0,1)^n) \subseteq I$. Then as $f([0,1]^n) = \overline{D_1^+}$ we would necessarily have $B \subseteq f(\partial[0,1]^n)$, so in other words B would be covered by the images of $2n$ maps from $(n-1)$-cubes. If we were able to show f is Lipschitz as well, then all those maps would be Lipschitz and this would all together prove that ∂D_1^+ is $(n-1)$-Lipschitz parameterizable. Hence we aim to prove

$f((0,1)^n) = I$ and that f is Lipschitz.

Let us first show $f((0,1)^n) \subseteq I$. To do this we prove that f is an open map, that is it maps open sets to open sets. Then necessarily $(0,1)^n$ needs to be mapped into I. To prove this we show that f is the composition of four open maps

$$(0,1)^n \xrightarrow{f_1} \mathbb{R}^n \xrightarrow{f_2} \mathbb{R}^n \xrightarrow{f_3} \mathbb{R}^r \times (0,\infty)^s \times \mathbb{R}^s \xrightarrow{f_4} \mathbb{R}^r \times \mathbb{C}^s$$

where the maps are defined as

- $f_1(t_1,\dots,t_n) = (t_1,\dots,\log(t_{r+s}),\dots,t_n)$ (where log is only in the $(r+s)$th-entry)
- $f_2(u_1,\dots,u_n) = (u_1,\dots,u_n)\mathcal{U}$ where $\mathcal{U}$ is the $n \times n$-matrix defined by

$$\mathcal{U} = \begin{pmatrix} v_1 & & & \\ \vdots & & 0 & \\ v_{r+s-1} & & & \\ 1\dots 2 & & & \\ & 1 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ & 0 & 0 & 1 \end{pmatrix}$$

  Here $1\dots 2$ denote the vector $(1,\dots,2)$ described earlier in the proof.

- $f_3(\xi_1,\dots,\xi_{r+2s}) = (e^{\xi_1},\dots,e^{\xi_r},\frac{1}{2}e^{\xi_{r+1}},\dots,\frac{1}{2}e^{\xi_{r+s}},2\pi\xi_{r+s+1},\dots,2\pi\xi_{r+2s})$
- $f_4(x_1,\dots,x_r,\rho_1,\dots,\rho_s,\theta_1,\dots,\theta_s) = (x_1,\dots,x_r,\rho_1 e^{i\theta_1},\dots,\rho_s e^{i\theta_s})$

We see by direct calculation that $f = f_4 f_3 f_2 f_1$. The map f_1 is open as it is the identity in all coordinates except one, where it is the logarithm, hence it sends a product of open intervals to a product of open intervals. The same argument also goes through for f_3 and f_4. To prove f_2 is open, we prove that $\mathcal{U}$ defines a linear transformation T of full rank. Then T^{-1} exists and since any linear transformation from $\mathbb{R}^n$ to another vector space is continuous it follows that T^{-1} is continuous, or in other words the preimage of an open set by T^{-1} is open. Hence $T(U) = (T^{-1})^{-1}(U)$ is open if U is open. To see that $\mathcal{U}$ has full rank we first observe that by definition $v_1,\dots,v_{r+s-1}$ and $(1\dots 2)$ are linearly indepedent in $\mathbb{R}^{r+s}$. Clearly the other row vectors in the matrix are linearly independent, with those vectors as well. This shows that $f((0,1)^n) \subseteq I$.

From the definition of f we note that all partial derivatives of f exist and are continuous. Since we are working on a (bounded) closed interval, the continuity of the partial derivatives imply that they are bounded. By passing from $\mathbb{C}$ to $\mathbb{R}^2$ and using the mean value theorem for $\mathbb{R}^d$ it follows that f is Lipschitz.

With this proven we are only left with the calculation of κ. As we have shown earlier we only need to calculate $\mathrm{vol}(D_1^+)$ for this. Our starting point is the integral definition of volume, remembering we are using polar coordinates:

$$\mathrm{vol}(D_1^+) = \int_{D_1^+} \rho_1 \cdots \rho_s \, \mathrm{d}x_1 \cdots \mathrm{d}x_r \mathrm{d}\rho_1 \cdots \mathrm{d}\rho_s \cdots \mathrm{d}\theta_1 \cdots \mathrm{d}\theta_s$$

We now substitute accordingly to how we defined the parameterization of $\overline{D_1^+}$ by $[0,1]^n$ (it will not make any changes in the integral if we consider the parameterization by a half-open n-cube or *the* n-cube). We calculate the Jacobian of this substitution: denote $x_1, \ldots, x_s, \rho_1, \ldots, \rho_s, \theta_1, \ldots, \theta_s$ by $w_1, \ldots, w_n$. Then for $k < r+s$ and then $k = r+s$ we have

$$\frac{\partial w_j}{\partial t_k} = \begin{cases} v_k^{(j)} w_j & j \leq r \\ \frac{1}{2} v_k^{(j)} w_j & r < j \leq r+s \\ 0 & j > r+s \end{cases} \qquad \frac{\partial w_j}{\partial t_{r+s}} = \begin{cases} \dfrac{w_j}{t_{r+s}} & j \leq r+s \\ 0 & j > r+s \end{cases}$$

If $k > r+s$ we have

$$\frac{\partial w_j}{\partial t_k} = \begin{cases} 2\pi & j = k \\ 0 & j \neq k \end{cases}$$

Hence the Jacobian matrix $J(t_1, \ldots, t_n)$ is

$$J(t_1, \ldots, t_n) = \begin{pmatrix} v_1^{(1)} w_1 & \cdots & v_1^{(r)} w_r & \frac{1}{2} v_1^{(r+1)} w_{r+1} & \cdots & \frac{1}{2} v_1^{(r+s)} w_{r+s} & \cdots & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ v_{r+s-1}^{(1)} w_1 & \cdots & v_{r+s-1}^{(r)} w_r & \frac{1}{2} v_{r+s-1}^{(r+1)} w_{r+1} & \cdots & \frac{1}{2} v_{r+s-1}^{(r+s)} w_{r+s} & \cdots & 0 & \cdots \\ \frac{w_1}{t_{r+s}} & & & \cdots & & \frac{w_{r+s}}{t_{r+s}} & \cdots & 0 & \cdots \\ & & & & & & 2\pi & & \\ & & & 0 & & & & \ddots & \\ & & & & & & & & 2\pi \end{pmatrix}$$

Taking the absolute value of the determinant and using determinant rules we get

$$|\det(J(t_1, \ldots, t_n))| = (2\pi)^s w_1 \cdots w_{r+s} \frac{1}{2^s} |\det(\mathcal{U})| = \frac{\pi^s x_1 \cdots x_r \rho_1 \cdots \rho_s}{t_{r+s}} |\det(\mathcal{U})|$$

Now

$$x_1 \cdots x_r \rho_1^2 \cdots \rho_s^2 = t_{r+s}^n \exp\left(\sum_{k=1}^{r+s-1} t_k \left(v_k^{(1)} + \cdots + v_k^{(r+s)} \right) \right) = t_{r+s}^n$$

where we have used that $v_k^{(1)} + \cdots + v_k^{(r+s)} = 0$. This is true as each $v_k \in \Lambda_U$ which is contained in a hyperplane H defined by $y_1 + \cdots + y_{r+s} = 0$. Hence the integral reduces to

$$\mathrm{vol}(D_1^+) = \pi^s |\det(\mathcal{U})| \int_{[0,1]^n} t_{r+s}^{n-1} \, \mathrm{d}t_1 \cdots \mathrm{d}t_n = \pi^s |\det(\mathcal{U})| \frac{1}{n}$$

We define $\mathrm{reg}(\mathcal{O}_K) = \frac{1}{n} |\det(\mathcal{U})|$. Of course we have to show that this is well-defined (indepedent of choice of basis of Λ_U). Putting everything we have done together we get

$$i_C(t) = \frac{\pi^s 2^{s+r} \mathrm{reg}(\mathcal{O}_K)}{\sqrt{|\mathrm{disc}\mathcal{O}_K|}} t + O\left(t^{1-\frac{1}{n}}\right)$$

which finishes the proof modulo the well-definedness of $\mathrm{reg}(\mathcal{O}_K)$. □

We end by showing that $\mathrm{reg}(\mathcal{O}_K)$ is well-defined and justify what we said earlier about the regulator meausuring the density of units. Recall that we have a volume on H induced by the inner product. When we write $\mathrm{vol}(H/\Lambda_U)$ in the theorem and its proof its the volume of Λ_U in H we mean.

Theorem 3.3.5. *The definition of* $\mathrm{reg}(\mathcal{O}_K)$ *is independent of the choice of a* $\mathbb{Z}$*-basis of* Λ_U *and furthermore*

$$\mathrm{reg}(\mathcal{O}_K) = \frac{1}{\sqrt{r+s}}\mathrm{vol}(H/\Lambda_U)$$

Proof. Fix a $\mathbb{Z}$-basis $v_1, \dots, v_{r+s-1}$ for Λ_U. We first observe that the (absolute value of the) determinant of $\mathcal{U}$ (from the proof of Theorem 3.3.4) is the same as the determinant of the $(r+s) \times (r+s)$-matrix having rows $v_1, \dots, v_{r+s-1}, (1 \dots 2)$. Let us call this matrix $\widetilde{\mathcal{U}}$. This matrix has coordinate sum 0 for all rows, except the last. Let $\mathcal{V}$ denote any matrix that is obtained by changing the last row of $\widetilde{\mathcal{U}}$ with any other vector that has the same coordinate sum. We claim that $\det(\mathcal{V}) = \det(\widetilde{\mathcal{U}})$. To see this, we observe that $\det(\widetilde{\mathcal{U}}) - \det(\mathcal{V})$ is the determinant of the matrix having the same rows as $\widetilde{\mathcal{U}}$ in the first $r+s-1$ rows and the difference between the old and the new vector in the last row. Adding together all the columns of this matrix we see that we get the zero vector. Hence the vector with 1's everywhere is an eigenvector to this matrix associated to the eigenvalue 0, but then this matrix can't be invertible, hence the determinant is 0, or in other words $\det(\widetilde{\mathcal{U}}) = \det(\mathcal{V})$. The coordinate sum of the last row in $\widetilde{\mathcal{U}}$ is n, so if we change the last row with the vector with $\frac{n}{r+s}$, we get the same result, and hence the same determinant by the claim over.

Now let Λ be the lattice spanned by $v_1, \dots, v_{r+s-1}, \left(\frac{n}{r+s}, \cdots, \frac{n}{r+s}\right)$. Since $v_1, \dots, v_{r+s-1}$ all are in the hyperplane H defined by $y_1 + \cdots + y_{r+s} = 0$, they are orthogonal to the last vector $\left(\frac{n}{r+s}, \cdots, \frac{n}{r+s}\right)$. Hence

$$\mathrm{vol}(\mathbb{R}^{r+s}/\Lambda) = \mathrm{length}\left(\frac{n}{r+s}, \cdots, \frac{n}{r+s}\right) \cdot \mathrm{vol}(H/\Lambda_U) = n\frac{1}{\sqrt{r+s}}\mathrm{vol}(H/\Lambda_U)$$

Note that by definition, $\mathrm{vol}(\mathbb{R}^{r+s}/\Lambda) = n\mathrm{reg}(\mathcal{O}_K)$ and hence the claim for the formula of $\mathrm{reg}(\mathcal{O}_K)$ follows. Finally this also shows that $\mathrm{reg}(\mathcal{O}_K)$ is indepedent of $\mathbb{Z}$-basis, since the volume of the fundamental parallellotope of the lattice Λ_U in H is independent of the $\mathbb{Z}$-basis we choose (see the discussion after Definition 3.1.6). This finishes the proof.

□

Chapter 4

Dedekind zeta functions

In this chapter we introduce the Dedekind zeta function ζ_K of a number field K. We prove that it defines a holomorphic function in the half plane $\mathrm{Re}(z) > 1$, and then give an analytic continuation for $\mathrm{Re}(z) > 1 - \frac{1}{n}$ where $n = [K : \mathbb{Q}]$. The latter as well as the class number formula will follow almost immediately from Theorem 3.3.4. Finally we prove that the Dedekind zeta function of $\mathbb{Q}(\zeta_m)$ for ζ_m a primitive m-th root of unity factorizes, more or less, into Dirichlet L-functions, and from this deduce the non-vanishing of Dirichlet L-functions of non-principal Dirichlet characters. We then show how this implies Dirichlet's theorem.

4.1 The Class Number Formula

From now, all earlier use of r, s, n to denote the number of real and complex embeddings, as well as the dimension of a number field K over $\mathbb{Q}$, stop. Following tradition, we will denote a complex number by $s = \sigma + it$. We will assume the reader knows some basic complex analysis. We need one lemma which is an easy consequence of Morera's theorem in complex analysis:

Lemma 4.1.1. *Let $\{f_n : \Omega \to \mathbb{C}\}_n$ be a sequence of holomorphic functions that converge to a function $f : \Omega \to \mathbb{C}$. Suppose that the convergence is uniform on every compact subset of Ω. Then f is analytic in Ω.*

For a proof, see [2, Theorem 5.1]. Before we define the Dedekind zeta function we need a general lemma on Dirichlet series.

Lemma 4.1.2. *Let $\{a_n\}$ be a sequence of complex numbers and suppose that $\sum_{n \leq t} a_n = O(t^r)$ for some $r > 0$ and all $t \geq 1$. Then*

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

converges for all s with $\sigma > r$ and defines a holomorphic function on that half-plane.

Proof. By Lemma 4.1.1, it is enough to show convergence of the sum for all s with $\sigma > r$, and that the convergence is uniform on all compact subsets of that half-plane. Define $A_k = \sum_{n=1}^{k} a_n$. Then by partial summation we have

$$\sum_{n=m}^{M} \frac{a_n}{n^s} = \frac{A_M}{M^s} - \frac{A_{m-1}}{m^s} + \sum_{n=m}^{M-1} A_n \left(\frac{1}{n^s} - \frac{1}{(n+1)^s} \right)$$

By assumption $|A_k| \leq Bk^r$ for some constant $B > 0$. Furthermore $n^{-s} - (n+1)^{-s} = s\int_n^{n+1} \frac{\mathrm{d}t}{t^{s+1}}$ so

$$\left| \frac{1}{n^s} - \frac{1}{(n+1)^s} \right| \leq |s| \int_n^{n+1} \frac{\mathrm{d}t}{|t|^{s+1}} = |s| \int_n^{n+1} \frac{\mathrm{d}t}{|t|^{\sigma+1}} \leq \frac{|s|}{n^\sigma}$$

Putting all this together with the result we got from partial summation we get

$$\left| \sum_{n=m}^{M} \frac{a_n}{n^s} \right| \leq B \left(M^{r-\sigma} + m^{r-\sigma} + |s| \sum_{n=m}^{M-1} n^{r-\sigma-1} \right)$$

Since $\sigma > r$, $r-\sigma-1 < -1$, so the sum $\sum_{n=m}^{M-1} n^{r-\sigma-1}$ converges as $M \to \infty$. Also since $\sigma > r$, $M^{r-\sigma}, m^{r-\sigma} \to 0$ as $m, M \to \infty$. Letting $M \to \infty$, then $m \to \infty$ afterwards, we see from the bound just proven that the tail $\sum_{n=m}^{\infty} a_n n^{-s}$ tends to 0. To show uniform convergence, observe that the integral test gives us the bound $\sum_{n=m}^{M-1} n^{r-\sigma-1} \leq \int_{m-1}^{\infty} t^{r-\sigma-1}\,\mathrm{d}t = \frac{(m-1)^{r-\sigma}}{\sigma-r}$ for all M. Furthermore, if we are on a compact subset of the halfplane $\sigma > r$, then necessarily we have to be bounded away from $\sigma = r$ - that is there is an $\varepsilon > 0$ such that $\sigma - r \geq \varepsilon$. Furthermore on a compact subset $|s| \leq B'$ for some constant B'. Putting it all together we have for compact subsets of the halfplane $\sigma > r$ that

$$\left| \sum_{n=m}^{\infty} \frac{a_n}{n^s} \right| \leq B \left(m^{-\varepsilon} + B' \frac{(m-1)^{-\varepsilon}}{\varepsilon} \right)$$

which is independent of s, hence uniform. □

The Riemann zeta function is defined as

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

which by Lemma 4.1.2 defines a holomorphic function for $\sigma > 1$. Observe that we have absolute convergence for $\sigma > 1$ as well. If K is any number field, let j_n denote the number of ideals with ideal norm equal to n. Then we define the Dedekind zeta function of K as

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

Observe that $j_n = 1$ when $K = \mathbb{Q}$ so $\zeta_{\mathbb{Q}} = \zeta$. From Theorem 3.3.4 we have $\sum_{n \leq t} j_n = O(t)$ so ζ_K defines a holomorphic function for $\sigma > 1$. We will however analytically continue this very shortly to $\sigma > 1 - \frac{1}{[K:\mathbb{Q}]}$. To do this we must first analytically continue the Riemann zeta function to $\sigma > 0$. To this end, define the Dirichlet eta function as

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

By the alternating nature of this sum, Lemma 4.1.2 gives that η is a holomorphic function for $\sigma > 0$ and we have absolute convergence for $\sigma > 1$. Now for $\sigma > 1$ we have $\eta(s) = \zeta(s) - 2\frac{1}{2^s}\zeta(s) = \left(1 - 2^{1-s}\right)\zeta(s)$. Hence outside some possible poles we have

$$\zeta(s) = \frac{1}{1 - 2^{1-s}}\eta(s)$$

for $\sigma > 1$. This defines an analytic continuation to $\sigma > 0$, with possible poles where $2^{1-s} = 1$. We claim the only pole of this analytic continuation of ζ is $s = 1$. To that end, let us write $p_k = 1 + \frac{2\pi i k}{\log(2)}$ for $k = \pm1, \pm2, \dots$. Let us define

$$h(s) = 1 + \frac{1}{2^s} - \frac{2}{3^s} + \frac{1}{4^s} + \frac{1}{5^s} - \frac{2}{6^s} + \dots$$

The alternating nature of h gives, by Lemma 4.1.2, that h defines a holomorphic function for $\sigma > 0$, and the convergence is uniform for $\sigma > 1$. Now $h(s) = \zeta(s) - 3\frac{1}{3^s}\zeta(s) = (1 - 3^{1-s})\zeta(s)$, and hence this also provides an analytic continuation for $\zeta(s)$

$$\zeta(s) = \frac{1}{1 - 3^{1-s}}h(s)$$

with possible poles at $\widetilde{p_k} = 1 + \frac{2\pi i k}{\log 3}$ for $k = 0, \pm1, \pm2, \dots$, but we claim that the only pole is when $k = 0$. By the identity theorem in complex analysis,

$$\frac{1}{1 - 3^{1-s}}h(s) = \frac{1}{1 - 2^{1-s}}\eta(s)$$

for $\sigma > 0$ outside the possible poles $p_k, \widetilde{p_k}$, but observe that p_k are distinct from the $\widetilde{p_k}$ whenever $k \neq 0$, so when $s \to p_k$, $k \neq 0$, the right hand side is finite, and vice versa when $s \to \widetilde{p_k}$. Hence we have an analytic continuation of $\zeta(s)$ to $\sigma > 0$ with only one pole in $s = 1$. This pole is simple because $(1 - 2^{1-s})'|_{s=1} = \ln(2)$. We are now ready to prove the class number formula and at the same time analytially continue the Dedekind zeta function.

Theorem 4.1.3. (Class Number Formula)
Let K be a number field. Let h be the class number of K, w the number of roots of unity in K, r_1 the number of real embeddings of K into $\mathbb{C}$ and finally r_2 the number of pairs of non-real embeddings of K into $\mathbb{C}$. Then $\zeta_K(s)$ has a simple pole at $s = 1$ with residue

$$\lim_{s\to1}(s-1)\zeta_K(s) = \frac{h2^{r_1+r_2}\pi^{r_2}\mathrm{reg}(\mathcal{O}_K)}{w\sqrt{|\mathrm{disc}(\mathcal{O}_K)|}}$$

Proof. Let $\kappa = \frac{2^{r_1+r_2}\pi^{r_2}\operatorname{reg}(\mathcal{O}_K)}{w\sqrt{|\operatorname{disc}(\mathcal{O}_K)|}}$. The limit is two-sided so we first have to show that this makes sense, i.e. give an analytic continuation of ζ_K to the left for $\sigma = 1$. To this end, observe that we have for $\sigma > 0$ that

$$\zeta_K(s) = h\kappa\zeta(s) + \sum_{n=1}^{\infty} \frac{j_n - h\kappa}{n^s}$$

By Theorem 3.3.4 we have $\sum_{n\leq t}(j_n - h\kappa) = O\left(t^{1-\frac{1}{[K:\mathbb{Q}]}}\right)$, so Lemma 4.1.2 implies that the sum $\sum_{n=1}^{\infty}(j_n - h\kappa)\, n^{-s}$ defines a holomorphic function for $\sigma > 1 - \frac{1}{[K:\mathbb{Q}]}$. In particular this implies

$$\lim_{s\to 1}(s-1)\sum_{n=1}^{\infty}\frac{j_n - h\kappa}{n^s} = 0$$

so it enough to show that $\zeta(s)$ has residue 1 in $s = 1$. Now

$$\lim_{s\to 1}(s-1)\zeta(s) = \lim_{s\to 1}\frac{s-1}{1-2^{1-s}}\eta(s) = \frac{\eta(1)}{\ln(2)} = 1$$

where $\eta(1) = \ln(2)$ follows from considering the Taylor series of ln around 0. Hence we can conclude that $\lim_{s\to 1}(s-1)\zeta_K(s) = h\kappa$ as desired. □

We have given an analytical continuation of $\zeta_K(s)$ to $\sigma > 1 - \frac{1}{[K:\mathbb{Q}]}$ but remark that much more can be done. Erich Hecke was the first to show that $\zeta_K(s)$ can be analytically continued to the whole complex plane with only one pole at $s = 1$ and furthermore it satisifies a functional equation similar to the functional equation of the Riemann zeta function. The proof is similar to the proof for the Riemann zeta function using Poisson summation formula, but this time it is the higher analogue of the Poisson summation formula in multidimensional Fourier analysis. The details can be found in Lang's book [7, Chapter XIII]. Another approach was taken by John Tate in his PhD thesis, where he did Fourier analysis on the locally compact ring of adeles. This recast of Hecke's argument turns out to be very fruitful and there are many interesting results resulting out from it. Tate's thesis can be found in [1, Chapter XV].

Our next and last goal in this thesis is to establish Dirichlet's theorem. To do this we first show that the Dedekind zeta function has an Euler product. The Dedekind zeta function converges absolutely for $\sigma > 1$, and hence we can write

$$\zeta_K(s) = \sum_I \frac{1}{\|I\|^s}$$

where the sum goes over all non-zero ideals of $\mathcal{O}_K$.

Lemma 4.1.4. *Let K be a number field. Then for $\sigma > 1$ we have*

$$\zeta_K(s) = \sum_I \frac{1}{\|I\|^s} = \prod_P \frac{1}{1-\|P\|^{-s}}$$

where the product is over all (non-zero) prime ideals of $\mathcal{O}_K$.

Proof. Let us order the prime ideals by their norms: $P_1, P_2, \dots$ so that $\|P_{i+1}\| \geq \|P_i\|$. We observe that the factors in the product are geometric series:

$$\frac{1}{1-\frac{1}{\|P\|^s}} = \sum_{n=0}^{\infty} \left(\frac{1}{\|P\|^s}\right)^n$$

Fix t. Then by multiplicativeness of the ideal norm and the fact that ideals factorize uniquely into prime ideals in $\mathcal{O}_K$ we get

$$\prod_{n=1}^{t} \frac{1}{1-\frac{1}{\|P_n\|^s}} = \left(\sum_{n=0}^{\infty}\left(\frac{1}{\|P_1\|^s}\right)^n\right)\cdots\left(\sum_{n=0}^{\infty}\left(\frac{1}{\|P_t\|^s}\right)^n\right) = \sum_{\|I\|\leq\|P_t\|} \frac{1}{\|I\|^s} + R_t(s)$$

Here we must have $R_t(s) \leq \sum_{I>\|P_t\|} \|I\|^{-s}$ and this tends to 0 when $t \to \infty$ as ζ_K converges. Hence letting $t \to \infty$ we get the desired conclusion. □

4.2 Dirichlet's Theorem

To prove Dirichlet's theorem we first need to study the factorization of ideals a bit more. More specifically we want to take a prime $p \in \mathbb{Z}$ and study how it *splits* in $\mathcal{O}_K$. By unique factorization we can write

$$p\mathcal{O}_K = P_1^{e_1} \cdots P_t^{e_t}$$

We say that the prime ideals P_i lie over p. The exponent e_i is called the ramification index, often denoted $e(P_i/p)$. We say that p ramifies in K if $e_i > 1$ for some i. If not then p is said to be unramified in K. Recall that $\mathcal{O}_K$ has Krull dimension 1, so P_i is a maximal ideal, hence $\mathcal{O}_K/P_i$ is a field. Of course $\mathbb{Z}/p\mathbb{Z}$ is also a field. We have a natural inclusion homomorphism $\mathbb{Z} \hookrightarrow \mathcal{O}_K$ inducing a homomorphism $\mathbb{Z} \to \mathcal{O}_K/P_i$. The kernel of this map is clearly $\mathbb{Z} \cap P_i$. We claim this intersection is indeed $p\mathbb{Z}$. Since $P_i \mid p\mathcal{O}_K$ it is clear that $p\mathbb{Z} \subseteq p\mathcal{O}_K \subseteq P_i$. Hence $p\mathbb{Z} = p\mathbb{Z} \cap \mathbb{Z} \subseteq P_i \cap \mathbb{Z}$. Now $\mathbb{Z} \cap P_i$ is an ideal of $\mathbb{Z}$, and as $p\mathbb{Z}$ is a maximal ideal we must have $\mathbb{Z} \cap P_i = p\mathbb{Z}$ or $\mathbb{Z} \cap P_i = \mathbb{Z}$. The latter cannot be true because then $1 \in P_i$. Hence $\mathbb{Z} \cap P_i = p\mathbb{Z}$. By the first isomorphism theorem we get an embedding of $\mathbb{Z}/p\mathbb{Z} \hookrightarrow \mathcal{O}_K/P_i$. Since the ideal norm is finite, this is an embedding of a finite field into another finite field. We denote the degree of this field extension by f_i or $f(P_i/p)$ and we call it the inertia degree of P_i over p. We have

$$p\mathcal{O}_K = \prod_{i=1}^{t} P_i^{e_i}$$

so taking ideal norms and remembering that the ideal norm of a principal ideal is the norm of the generator itself, we get

$$p^n = \|P_1\|^{e_1} \cdots \|P_t\|^{e_t}$$

where $n = [K : \mathbb{Q}]$. Finally, as $\|P_1\| = p^{k_i}$ for some k_i and $\mathcal{O}_K/P_i$ is a field extension of $\mathbb{Z}/p\mathbb{Z}$ of degree f_i it follows that $k_i = f_i$. Comparing both exponents we see that we get $n = f_1e_1 + \cdots + f_ne_n$, and so we see that $\|P_i\| = p^{f_i}$. The theory of prime splitting is rich and beautiful, and we will only merely touch it. The interested reader can find a nice account of this in [8].

It turns out that there is a close relation between the primes that do ramify and the discriminant. We will not prove this, but remark that one can prove this by building the theory of splitting of primes further.

Theorem 4.2.1. *A prime number $p \in \mathbb{Z}$ ramifies in K if and only if $p \mid \operatorname{disc}(\mathcal{O}_K)$.*

For a proof, see [8, Theorem 24 & Theorem 34].

We now assume that $K = \mathbb{Q}(\zeta_n)$ for a primitive n-th root of unity ζ_n. We proved earlier that

$$\zeta_K(s) = \prod_P \frac{1}{1 - \|P\|^{-s}}$$

The plan is to split this up into those primes $p \in \mathbb{Z}$ that do ramify and these that are unramified. We can do this by calculating the discriminant, but this is a somewhat complicated task. Luckily for us Theorem 4.2.1 implies that at most finitely many primes do ramify and this is enough for us. We need to show that the ramification degree and inertia degree is the same for all P lying over p. To this end, we study the extension $K/\mathbb{Q}$ which is Galois with Galois group canonically isomorphic to $(\mathbb{Z}/n\mathbb{Z})^*$ (see [4, Theorem 18.1.4]).

Lemma 4.2.2. *Let L be a Galois extension of $\mathbb{Q}$ and p a prime in $\mathbb{Z}$. Then $\operatorname{Gal}(L/\mathbb{Q})$ permutes the primes in $\mathcal{O}_L$ lying over p transitively. As a consequence, if Q and P are primes in $\mathcal{O}_L$ lying over p then $e(Q|p) = e(P|p)$ and $f(Q|p) = f(P|p)$.*

Proof. First we show that $\sigma \in \operatorname{Gal}(L/\mathbb{Q})$ restricts to an automorphism of $\mathcal{O}_L$. Since σ is the identity on $\mathbb{Z}$ it must map algebraic integers to algebraic integers so $\sigma(\mathcal{O}_L) \subseteq \mathcal{O}_L$. If $\alpha \in \mathcal{O}_L$ then $a_0 + a_1\alpha + \cdots + a_{t-1}\alpha^{t-1} + \alpha^t = 0$ for some $a_0, \ldots, a_{t-1} \in \mathbb{Z}$. Since σ is an automorphism of L there is some $\beta \in L$ such that $\alpha = \sigma(\beta)$. In other words $a_0 + a_1\sigma(\beta) + \cdots + \sigma(\beta)^t = 0$ and after applying σ^{-1} to this equation we get $a_0 + a_1\beta + \cdots + \beta^n = 0$ so $\beta \in \mathcal{O}_L$, which shows that σ restricts to an automorphism of $\mathcal{O}_L$.

Now let P be a prime of $\mathcal{O}_L$ lying over p and σ as before. Then $\sigma(p\mathbb{Z}) = p\mathbb{Z}$ and so $\sigma(P)$ is still a prime ideal lying over p. If Q is another prime lying over p we claim that there exists a $\sigma \in \operatorname{Gal}(L/\mathbb{Q})$ such that $\sigma(P) = Q$. Assume $\sigma(P) \neq Q$ for all $\sigma \in \operatorname{Gal}(L/\mathbb{Q})$. By the Chinese Remainder Theorem we then get the existence of $x \in \mathcal{O}_L$ such that

$$x \equiv 0 \pmod{Q} \qquad\qquad x \equiv 1 \pmod{\sigma(P)}\ \forall \sigma \in \operatorname{Gal}(L/\mathbb{Q})$$

Since Q is an ideal and $x \in Q$ we have $N(x) \in Q$. Furthermore $N(x)$ is an integer by Theorem 1.1.9. Hence $N(x) \in Q \cap \mathbb{Z} = p\mathbb{Z}$. However by the congruences above we have $x \notin \sigma(P)$ for all $\sigma \in \operatorname{Gal}(L/\mathbb{Q})$. Hence $\sigma^{-1}(x) \notin P$ for all σ. We can write $N(x)$ as a product over all $\sigma^{-1}(x)$ and none of these are in P. Since P is a prime ideal this implies that $N(x) \notin P$, but this is a contradicition since $N(x) \in p\mathbb{Z} \subseteq P$.

Finally let us show the claims on ramification and inertia. Write $p\mathcal{O}_L = P_1^{e_1}\cdots P_t^{e_t}$. By the above we can find $\sigma_\ell \in \mathrm{Gal}(L/\mathbb{Q})$ so that $\sigma_\ell(P_1) = P_\ell$ for any $1 \le \ell \le t$. Remembering that $\sigma_\ell(p\mathcal{O}_L) = p\mathcal{O}_L$ and then applying σ_ℓ gives $p\mathcal{O}_L = P_\ell^{e_1}\cdots\sigma_\ell(P_t)^{e_t}$ so $e(P_\ell|p) = e(P_1|p)$. This shows the claim about the ramification index. To show the claim on inertia index, let σ_ℓ be as before and define $\psi : \mathcal{O}_L \to \mathcal{O}_L/\sigma_\ell(P_1)$ by $\alpha \mapsto \sigma_\ell(\alpha) + \sigma_\ell(P_1)$. This is a surjective group homomorphism with kernel P_1, hence we get an isomorphism $\mathcal{O}_L/P_1 \cong \mathcal{O}_L/\sigma_\ell(P_1) = \mathcal{O}_L/P_\ell$ and thus $f(P_1|p) = f(P_\ell|p)$, which finishes the proof. □

Recall that we saw earlier that if there are r primes lying over p then $n = f_1e_1 + \cdots + f_ne_n$ where f_i, e_i are the inertia index and ramification index of the primes lying over p. As we have now shown, in Galois extensions the ramification and inertia is only dependent of the one prime lying under and hence we get that $n = r_pe_pf_p$ in this case. We now state without proof what the discriminant of $\mathbb{Q}(\zeta_n)$ is.

Lemma 4.2.3. *Let $n > 2$ and ζ_n be a primitive nth root of unity. Then*

$$\mathrm{disc}(\mathbb{Q}(\zeta_n)) = (-1)^{\varphi(n)/2}\frac{n^{\varphi(n)}}{\prod_{p|n} p^{\varphi(n)/(p-1)}}$$

For a proof see [11, Proposition 2.7]. When $n = 2$, the cyclotomic extension is just $\mathbb{Q}$ itself, which has discriminant 1. From this we see that the only primes that ramify are those dividing n. Take any P lying over an unramified p, in other words $p \nmid n$. Then it can be proven that $f_p = \mathrm{ord}_n(p)$. For a proof see [11, Theorem 2.13].

There are only finitely many primes that ramify and hence there can be at most finitely many prime ideals P lying over primes that do ramify. Since every prime ideal P lies over one unique prime p we split up the product of the Dedekind zeta function as

$$\zeta_K(s) = \prod_{P \text{ lying over ramified } p} \frac{1}{1-\|P\|^{-s}} \prod_{P \text{ lying over unramified } p} \frac{1}{1-\|P\|^{-s}}$$

We are not going to try to get good control over the first part of the product, and since it is finite we will see that we can really ignore it in the application to Dirichlet's theorem that we are aiming for. For all unramified primes p, we still let r_p denote the number of prime ideals P lying over p, and f_p their inertia degree (only dependent on p as we have showed). Then

$$\prod_{P \text{ lying over unramified } p} \frac{1}{1-\|P\|^{-s}} = \prod_{p \text{ unramified}} \left(\frac{1}{1-p^{-f_ps}}\right)^{r_p}$$

We now define Dirichlet characters and their corresponding Dirichlet L-functions.

Definition 4.2.4. A Dirichlet character modulo m is a multiplicative homomorphism $\chi : (\mathbb{Z}/m\mathbb{Z})^* \to \mathbb{C}^*$. We can extend this to a function $\chi : \mathbb{Z} \to \mathbb{C}$ by extending χ on residue classes modulo m and defining $\chi(k) = 0$ whenever $\gcd(k, m) > 1$. The Dirichlet L-function $L(s, \chi)$ associated to χ is defined as

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

Since $(\mathbb{Z}/m\mathbb{Z})^*$ has order $\varphi(m)$ we observe that for $\ell \in (\mathbb{Z}/m\mathbb{Z})^*$ that $1 = \chi(\ell^{\varphi(m)}) = \chi(\ell)^{\varphi(m)}$ so the image of χ consists of $\varphi(m)$th roots of unity. We will now define characters on any finite abelian group - especially we want to study the characters of $\mathrm{Gal}(\mathbb{Q}(\zeta_m)/\mathbb{Q})$, but as we will explain this is the same as studying the Dirichlet characters modulo m. First we remark that since the Dirichlet character χ modulo m is multiplicative we get an Euler product for $L(s,\chi)$:

$$L(s,\chi) = \prod_{p \text{ prime}} \frac{1}{1-\chi(p)p^{-s}} = \prod_{p \text{ prime}, p\nmid m} \frac{1}{1-\chi(p)p^{-s}}$$

This can be proven by observing that the factors in the Euler product are geometric series and multiplying all these together yield the Dirichlet L-series by the fundamental theorem of arithmetic.

Definition 4.2.5. Let G be a finite abelian group. Then a character ψ on G is a homomorphism $G \to \mathbb{C}^*$. The set of all characters of G is denoted by $\widehat{G}$.

Since G is a finite group the image of ψ is actually contained in the unit circle as for Dirichlet characters. By pointwise multiplication $\widehat{G}$ is a group, in fact $G \cong \widehat{G}$.

Theorem 4.2.6. *Let G be a finite abelian group. Then $\widehat{G} \cong G$.*

Proof. We first prove the statement in the case G is a cyclic group, then let the fundamental theorem of finite abelian groups do the rest of the work, which says G is isomorphic to a product of cyclic groups, that is $G \cong \mathbb{Z}/n_1\mathbb{Z} \times \cdots \times \mathbb{Z}/n_k\mathbb{Z}$. This is enough because $\mathrm{Hom}_{\mathbb{Z}}(\prod_{i=1}^k \mathbb{Z}/n_i\mathbb{Z}, \mathbb{C}^*) \cong \prod_{i=1}^k \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}/n_i\mathbb{Z}, \mathbb{C}^*)$. So let G be a cyclic group of order n. Let a generate G. A character on G is uniquely defined by where it sends a. Let $\zeta_\ell = \exp\left(\frac{2\pi i \ell}{n}\right)$, and define a map on G by $\chi_\ell(a) = \zeta_\ell$. To check that this is a homomorphism we see that

$$\chi_\ell(a^k a^m) = \zeta_\ell^{k+m} = \zeta_\ell^k \zeta_\ell^m = \chi_\ell(a^k)\chi_\ell(a^m)$$

All the homomorphism χ_ℓ are different, so $|\widehat{G}| \geq n$. Conversely let $\chi \in \widehat{G}$. Then $\chi(a) = \exp(\frac{2\pi i \ell}{n})$ for some ℓ and so actually $\chi = \chi_\ell$. Hence $|\widehat{G}| = n$. Finally observe that for any ℓ, we actually have $\chi_\ell = \chi_1^\ell$, so χ_1 generate $\widehat{G}$. Since there is only one cyclic group of each finite order up to isomorphism, it follows that $\widehat{G} \cong G$. □

Hence there are $\varphi(m)$ Dirichlet characters modulo m. Since $(\mathbb{Z}/m\mathbb{Z})^*$ is a group, multiplication by an element k with gcd 1 to m induces a bijection of $(\mathbb{Z}/m\mathbb{Z})^*$. Hence

$$\sum_{n=1}^{m} \chi(n) = \sum_{n \in (\mathbb{Z}/m\mathbb{Z})^*} \chi(n) = \sum_{n \in (\mathbb{Z}/m\mathbb{Z})^*} \chi(kn) = \chi(k) \sum_{n \in (\mathbb{Z}/m\mathbb{Z})^*} \chi(n)$$

As long as χ is not the principal character (the trivial character defined by $\chi(n) = 1$ for all n with $\gcd(n,m) = 1$), there exists a k with $\gcd(k,m) = 1$ so that $\chi(k) \neq 1$, and hence in this case we can see by comparing the second sum and the last sum that the sum is 0. If χ is the principal character it is easy to see that the sum is $\varphi(m)$. We have proved:

Lemma 4.2.7. *Let χ be a Dirichlet character modulo m. Then*

$$\sum_{n=1}^{m}\chi(n)=\begin{cases}\varphi(m) & \text{if } \chi \text{ is principal}\\ 0 & \text{otherwise}\end{cases}$$

From this lemma it follows by the residue class nature of a Dirichlet character modulo m that as long as χ is non-principal then

$$\left|\sum_{n\le t}\chi(n)\right|\le\varphi(m)=O(1)$$

and hence by Lemma 4.1.2 we conclude that $L(s,\chi)$ defines an analytic function for $\mathrm{Re}(s)>0$. When χ is principal, let us call it χ_0, we see that

$$L(s,\chi_0)=\prod_{p\text{ prime},p\nmid m}\frac{1}{1-\frac{1}{p^s}}=\zeta(s)\prod_{p|m}\left(1-\frac{1}{p^s}\right)$$

so taking the analytic continuation of ζ from the last chapter we get that $L(s,\chi_0)$ is a meromorphic function for $\mathrm{Re}(s)>0$ with a simple pole at $s=1$. We are now done with the prerequisites we need for showing non-vanishing of $L(s,\chi)$ for non-principal χ.

Theorem 4.2.8. *Let χ be a non-principal Dirichlet character modulo m. Then $L(1,\chi)=C_\chi\neq 0$ where C_χ is some constant depending on χ.*

Proof. We consider the function $F(s)=\prod_\chi L(s,\chi)$ where the product runs over all Dirichlet characters modulo m. As we have seen, as long as χ is non-principal $L(s,\chi)$ defines an analytic function for $\mathrm{Re}(s)>0$. Hence there is one simple pole at $s=1$ in the product coming from the principal character. If $L(1,\chi)=0$ for some other χ then this will cancel out the pole and the product defines an analytic function for $\mathrm{Re}(s)>0$. When $K=\mathbb{Q}(\zeta_m)$ we will see that ζ_K is more or less $F(s)$, but by the class number formula ζ_K has a simple pole at $s=1$, and hence so must $F(s)$ also have, so we cannot have $L(1,\chi)=0$ for any non-principal character.

Our next claim is the following: if G is a finite abelian group and we fix $g\in G$ and let χ run through $\widehat{G}$ then $\chi(g)$ will run through all the $\mathrm{ord}_G(g)$th roots of unity and moreover it will take each value equally many times. Here $\mathrm{ord}_G(g)$ denotes the order of g in G. To prove this, define $\psi_g:\widehat{G}\to\mathbb{C}^*$ by $\psi_g(\chi)=\chi(g)$. Then ψ_g is a group homomorphism. We have

$$\ker\psi_g=\{\chi\in\widehat{G}:\chi(g)=1\}\cong\widehat{G/(g)}$$

By the first isomorphism theorem we then get

$$|\mathrm{Im}(\psi_g)|=\frac{|\widehat{G}|}{|\widehat{G/(g)}|}=\frac{|G|}{\frac{|G|}{\mathrm{ord}_G(g)}}=\mathrm{ord}_G(g)$$

Since the image has order $\mathrm{ord}_G(g)$ it consists of all of the $\mathrm{ord}_G(g)$th roots of unity. Above we showed there are $\frac{|G|}{\mathrm{ord}_G(g)}$ characters that take the value 1 in g. Two characters χ_1,χ_2 will take the same value on g if and only if $(\chi_1^{-1}\chi_2)(g) = 1$, so two characters will take the same value exactly $\frac{|G|}{\mathrm{ord}_G(g)}$ times. Since the image is of order $\mathrm{ord}_G(g)$ this implies that when χ runs through $\widehat{G}$ it takes on all the $\mathrm{ord}_G(g)$th roots of unity equally many times.

We now apply this on $G = \mathrm{Gal}(\mathbb{Q}(\zeta_m)/\mathbb{Q})$, which is canonically isomorphic to $(\mathbb{Z}/m\mathbb{Z})^*$ as mentioned earlier. Hence the characters of G are exactly the Dirichlet characters modulo m. From now, let f denote the order of p modulo m. Letting $\omega_1, \dots, \omega_f$ denote the fth roots of unity, we then get

$$\begin{aligned}
\prod_{\chi \in \widehat{G}} \left(1 - \frac{\chi(p)}{p^s}\right) &= \left(1 - \frac{\omega_1}{p^s}\right)^{\varphi(m)/f} \cdots \left(1 - \frac{\omega_f}{p^s}\right)^{\varphi(m)/f} \\
&= \left(\left(1 - \frac{\omega_1}{p^s}\right) \cdots \left(1 - \frac{\omega_f}{p^s}\right)\right)^{\varphi(m)/f} \\
&= \frac{1}{p^{\varphi(m)s}} \left((p^s - \omega_1) \cdots (p^s - \omega_f)\right)^{\varphi(m)/f} \\
&= \frac{1}{p^{\varphi(m)s}} \left(p^{fs} - 1\right)^{\varphi(m)/f} = \left(1 - \frac{1}{p^{fs}}\right)^{\varphi(m)/f}
\end{aligned}$$

Now recall that we proved at the start of this section that $f_p = f$ for all prime ideals lying over primes p that does not ramify. Furthermore for any prime P lying over p that does not ramify, write $p\mathcal{O}_K = P_1 \cdots P_t$. Then since the P_i are comaximal, taking the quotient and applying the Chinese Remainder Theorem gives $\mathcal{O}_K/p\mathcal{O}_K \cong \mathcal{O}_K/P_1 \times \cdots \mathcal{O}_K/P_t$. Taking norms we get $p^{\varphi(m)} = \|(p)\| = p^{f_p} \cdots p^{f_p} = p^{r_p f_p}$ (recall r_p is the number of prime ideals P lying over p), and hence $\varphi(m) = r_p f_p = r_p f$, or more importantly $r_p = \varphi(m)/f$. Hence

$$\prod_{\chi \in \widehat{G}} L(s, \chi) = \prod_{p \nmid m} \left(\frac{1}{1 - \frac{1}{p^{fs}}}\right)^{\varphi(m)/f}$$

Recall that we split ζ_K up into two products; one for ramified primes, the other for unramified primes. The product of P lying over primes p that may be ramified primes, let us call it $\Gamma_{\mathbf{ram}}$, is finite and each factor is analytic and non-vanishing in a neighborhood of $s = 1$, hence $\Gamma_{\mathbf{ram}}$ is as well. Using the splitting of ζ_K into ramified and unramified primes we finally arrive from this that

$$\zeta_K(s) = \Gamma_{\mathbf{ram}}(s) \prod_{\chi \in \widehat{G}} L(s, \chi)$$

which finishes the proof by the first paragraph of this proof. □

Now that we know $L(1, \chi) \neq 0$ for χ non-principal, Dirichlet's theorem is not that far anyway. However, we first remark that the factorization of ζ_K into Dirichlet L-functions (modulo some finite product coming from ramified primes) can be generalized to any abelian extension. This comes basically down to the fact that any abelian extension can be embedded into a cyclotomic extension (this is the Kronecker-Weber theorem). More is true: if $K/\mathbb{Q}$ is Galois, then ζ_K factors into Artin L-functions $L(\rho, s)$ where ρ is a representation of the Galois group $\mathrm{Gal}(K/\mathbb{Q})$.

Lemma 4.2.9. *Let $\chi_1, \ldots, \chi_{\varphi(m)}$ be the Dirichlet characters modulo m and $k, \ell \in (\mathbb{Z}/m\mathbb{Z})^*$. Then*

$$\sum_{n=1}^{\varphi(m)} \chi_n(k)\overline{\chi_n(\ell)} = \begin{cases} \varphi(m) & \text{if } k \equiv \ell \pmod m \\ 0 & \text{if } k \not\equiv \ell \pmod m \end{cases}$$

Proof. Let $a \in (\mathbb{Z}/m\mathbb{Z})^*$. Then $1 = |\chi(a)|^2 = \chi(a)\overline{\chi(a)}$. Hence $\overline{\chi} = \chi^{-1}$. Hence we can rewrite the sum as

$$\sum_{n=1}^{\varphi(m)} \chi_n(k)\overline{\chi_n(\ell)} = \sum_{n=1}^{\varphi(m)} \chi_n(k)\chi_n(\ell)^{-1} = \sum_{n=1}^{\varphi(m)} \chi_n(k\ell^{-1})$$

If $k \equiv \ell \pmod m$, this sum is clearly $\varphi(m)$. If not there exists a Dirichlet character χ_t modulo m so that $\chi_t(k\ell^{-1}) \neq 1$, and since the group of Dirichlet characters modulo m is finite, multiplication by this induces a bijection on this group and we get

$$\sum_{n=1}^{\varphi(m)} \chi_n(k\ell^{-1}) = \sum_{n=1}^{\varphi(m)} (\chi_t\chi_n)(k\ell^{-1}) = \chi_t(k\ell^{-1}) \sum_{n=1}^{\varphi(m)} \chi_n(k\ell^{-1})$$

Since $\chi_t(k\ell^{-1}) \neq 1$ the desired conclusion follows. □

Let $s > 1$ be real. We want to take the logarithm of $L(s, \chi)$, but since $L(s, \chi)$ is complex-valued we must watch out. Why we can take the logarithm is justified in [6, p. 256,261]. Starting from the Euler product of $L(s, \chi)$ and taking the logarithm we get

$$\ln L(s, \chi) = -\sum_{p \text{ prime}, \, p \nmid m} \ln\left(1 - \frac{\chi(p)}{p^s}\right) = \sum_{p \text{ prime}, \, p \nmid m} \sum_{n=1}^{\infty} \frac{\chi(p^n)}{np^{ns}}$$

where we in the last term have used the power series of ln. To ease the notation from now on, p under the sum means a summation over the prime numbers. Assume $\gcd(a, m) = 1$ or in other words $a \in (\mathbb{Z}/m\mathbb{Z})^*$. Multiply the expression above with $\overline{\chi(a)}$ and then a summation over all Dirichlet characters $\chi_1, \ldots, \chi_{\varphi(m)}$ modulo m yields

$$\sum_{k=1}^{\varphi(m)} \overline{\chi_k(a)} \ln L(s, \chi_k) = \sum_{k=1}^{\varphi(m)} \sum_{p \nmid m} \sum_{n=1}^{\infty} \frac{\chi_k(p^n)\overline{\chi_k(a)}}{np^{ns}} = \sum_{p \nmid m} \sum_{n=1}^{\infty} \sum_{k=1}^{\varphi(m)} \frac{\chi_k(p^n)\overline{\chi_k(a)}}{np^{ns}}$$

The interchanging of summation order is justified by absolute convergence. Using Lemma 4.2.9 on the inner sum yields

$$\sum_{k=1}^{\varphi(m)} \overline{\chi_k(a)} \ln L(s, \chi_k) = \varphi(m) \sum_{p \equiv a \pmod m} \frac{1}{p^s} + \varphi(m) R_\chi(s)$$

Now we bound $R_\chi(s)$ with $\mathrm{Re}(s) > 1$: observe that $|\chi_k(p^k)\overline{\chi_k(a)}\frac{1}{n}p^{-ns}| \leq p^{-ns}$ and hence

$$|R_\chi(s)| \leq \sum_p \sum_{n=2}^{\infty} \left(\frac{1}{p^s}\right)^n = \sum_p \frac{p^{-2s}}{1-\frac{1}{p^s}} \leq \sum_p \frac{p^{-2s}}{1-\frac{1}{2^s}} \leq 2\zeta(2)$$

and hence $|R_\chi(s)|$ remains bounded as $s \to 1^+$. Say χ_1 is the principal character - we do a final manipulation of our sum formula to make the conclusion clearer:

$$\overline{\chi_1(a)} \ln L(s, \chi_1) = \varphi(m) \sum_{p \equiv a \pmod m} \frac{1}{p^s} + \left[\varphi(m)R_\chi(s) - \sum_{k=2}^{\varphi(m)} \overline{\chi_k(a)} \ln L(s, \chi_k)\right]$$

Since $L(s,\chi) \neq 0$ for non-principal χ and $R_\chi(s)$ is bounded as $s \to 1^+$ it follows that the expression inside $[\cdot]$ is bounded as $s \to 1^+$. But as we have seen $L(s,\chi_1)$ has a simple pole at $s = 1$ so as $s \to 1^+$, we have $\ln L(s,\chi_1) \to \infty$ and hence the same has to be true for the right hand side - the only way this can be true is if the sum has infinitely many terms, hence the set $\{p \text{ prime} : p \equiv a \pmod m\}$ has to be infinite. But this set is exactly the primes in the arithmetic progression $a + m, a + 2m, a + 3m, \ldots$. We have proved:

Theorem 4.2.10. (Dirichlet's theorem)
Let $\gcd(a, m) = 1$. *Then there are infinitely many prime numbers* p *in the sequence* $a + m, a + 2m, a + 3m, \ldots$.

Bibliography

[1] *Algebraic number theory.* Proceedings of an instructional conference organized by the London Mathematical Society (a NATO Advanced Study Institute) with the support of the International Mathematical Union. Edited by J. W. S. Cassels and A. Fröhlich. Academic Press, London; Thompson Book Co., Inc., Washington, D.C., 1967.

[2] Lars V. Ahlfors. *Complex analysis.* McGraw-Hill Book Co., New York, third edition, 1978. An introduction to the theory of analytic functions of one complex variable, International Series in Pure and Applied Mathematics.

[3] M. F. Atiyah and I. G. Macdonald. *Introduction to commutative algebra.* Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1969.

[4] P. B. Bhattacharya, S. K. Jain, and S. R. Nagpaul. *Basic abstract algebra.* Cambridge University Press, Cambridge, second edition, 1994.

[5] Thomas W. Hungerford. *Algebra*, volume 73 of *Graduate Texts in Mathematics.* Springer-Verlag, New York-Berlin, 1980. Reprint of the 1974 original.

[6] Kenneth Ireland and Michael Rosen. *A classical introduction to modern number theory*, volume 84 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, second edition, 1990.

[7] Serge Lang. *Algebraic number theory*, volume 110 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, second edition, 1994.

[8] Daniel A. Marcus. *Number fields.* Universitext. Springer, Cham, 2018. Second edition, With a foreword by Barry Mazur.

[9] Joseph J. Rotman. *An introduction to homological algebra.* Universitext. Springer, New York, second edition, 2009.

[10] Ian Stewart and David Tall. *Algebraic number theory and Fermat's last theorem.* A K Peters, Ltd., Natick, MA, third edition, 2002.

[11] Lawrence C. Washington. *Introduction to cyclotomic fields*, volume 83 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, second edition, 1997.

NTNU
Norwegian University of
Science and Technology