Algebraic Number Fields

Ehud de Shalit

Hebrew University, Jerusalem
E-mail address: deshalit@math.huji.ac.il

CHAPTER 1

Number fields and their rings of integers

1. Notation

- $\mathbb{R}$ and $\mathbb{C}$ are the reals and the complex numbers, $\mathbb{Q}$ are the rationals and $\mathbb{Z}$ are the integers. Since we shall encounter "integers" in fields other than $\mathbb{Q}$ in this course, we shall sometimes refer to $\mathbb{Z}$ as the *rational integers.*
- The *degree* of a finite field extension $K \subset L$ is denoted $[L : K]$. Recall that it is the dimension of L as a vector space over K. We also write "L/K" for the extension.
- If L/K is a finite field extension, any $a \in L$ defines a K-linear transformation of L denoted M_a (multiplication by a) by the rule

$$M_a(x) = ax. \tag{1.1}$$

  The *norm* $N_{L/K}(a)$ and the *trace* $Tr_{L/K}(a)$ are by definition $\det(M_a)$ and $tr(M_a)$ respectively.
- A *ring* is usually commutative with 1, unless it is clear from the context that it is noncommutative. Ideals will be often denoted by small gothic letters $\mathfrak{a}, \mathfrak{b}, \mathfrak{c}$ etc. If $\{a_i\}$ is a (usually finite) collection of elements of a ring A, the ideal (a_i) is the ideal generated by the a_i, namely $\sum Aa_i$. We use this notation even if this is not a proper ideal, but the whole ring A.
- If A is a ring and I a (proper) ideal, the quotient ring is denoted by A/I. If I and J are ideals IJ is the ideal consisting of all the sums of products of an element from I by and element from J.
- If A is a commutative ring its multiplicative group of *units* (invertible elements) is denoted by $A^\times$.

2. Introduction

Definition 2.1. *A* number field *(sometimes called an* algebraic number field*) K is a finite field extension of $\mathbb{Q}$.*

Number fields may be viewed abstractly, or as subfields of $\mathbb{C}$. In the latter case, we should be more appropriately speaking of a couple (K, ι), where ι is the embedding into $\mathbb{C}$, but traditionally ι is dropped from the notation.

Classical problems in number theory often "live" in such fields, and are best understood in their context, even though the problem itself may be phrased entirely within $\mathbb{Q}$. We give two examples. Note that in both, in addition to the number field K, certain subrings of it, similar to the subring $\mathbb{Z} \subset \mathbb{Q}$, play a prominent role.

2.1. Sum of two squares.

Theorem 2.1. *(Fermat) A positive integer n is a sum of two squares if and only if for every prime $q \equiv 3 mod 4$, $ord_q(n)$ is even.*

PROOF. The reader should supply the details in the following guided exercise.

- The ring $\mathbb{Z}[i]=\{x+iy; x,y\in\mathbb{Z}\}$ is a principal ideal domain (PID). In fact, it is a Euclidean domain. (If you haven't seen this example, look it up in Amitsur's Algebra.)
- $\mathbb{Z}[i]^\times=\{\pm1,\pm i\}$.
- The rational integers which are sums of two squares are those of the form $z\bar{z}$ for $z\in\mathbb{Z}[i]$.
- Let p be a rational prime. The quotient ring

$$\mathbb{Z}[i]/(p)\simeq\mathbb{Z}[X]/(X^2+1,p)\simeq\mathbb{F}_p[X]/(X^2+1) \tag{2.1}$$

  is a field if and only if $p\equiv 3mod4$. This is the main step, and to prove it you should use the fact that the multiplicative group of $\mathbb{F}_p$ is cyclic. Conclude that p remains prime in $\mathbb{Z}[i]$ if and only if $p\equiv 3mod4$.
- $2=(1+i)(1-i)$ and $1+i$ is prime in $\mathbb{Z}[i]$.
- If $p\equiv 1mod4$, then $p=\pi\bar{\pi}$ for a prime π of $\mathbb{Z}[i]$. *Hint:* if $p=\pi\lambda$ in $\mathbb{Z}[i]$ where π is prime, then $p^2=(\pi\bar{\pi})\left(\lambda\bar{\lambda}\right)$ in $\mathbb{Z}$.
- Any $z\in\mathbb{Z}[i]$ can be written as

$$z=\varepsilon(1+i)^k\prod\pi_i^{m_i}\prod q_j^{n_j} \tag{2.2}$$

  where $\pi_i\bar{\pi}_i=p_i\equiv 1mod4$, $q_j\equiv 3mod4$ and $\varepsilon\in\{\pm1,\pm i\}$.
- For such a z, $z\bar{z}=2^k\prod p_i^{m_i}\prod q_j^{2n_j}$.

For example, a prime is a sum of two squares if and only if it is 2 or $1mod4$, $45=3^2\cdot5$ is a sum of two squares, but $35=7\cdot5$ isn't. □

2.2. Kummer's work on Fermat's Last Theorem. Fermat's Last Theorem (FLT) asserts that for $n\geq3$, the equation $x^n+y^n=z^n$ has no solutions in positive integers x,y,z.

Fermat himself proved it for $n=4$, so the general question is reduced to $n=p$ an odd prime. In the middle of the 19th century Liouville noticed that the equation may be rephrased as

$$x^p=\prod_{k=0}^{p-1}(z-\zeta^k y) \tag{2.3}$$

if $\zeta=e^{2\pi i/p}$ is a primitive pth root of unity. The advantage of this reformulation is that the problem becomes one of decomposing x^p in the ring $\mathbb{Z}[\zeta]$. Kummer made an important contribution when he proved that FLT holds for the exponent p if $\mathbb{Z}[\zeta]$ is a PID. In fact, he proved a stronger theorem. Call p *regular* if the following condition holds:

Whenever I is an ideal of $\mathbb{Z}[\zeta]$ and I^p is principal, then I is already principal.

Kummer's theorem is that if p is regular, FLT holds for p. Clearly, if $\mathbb{Z}[\zeta]$ is a PID, then p is regular, but there are only finitely many p's for which $\mathbb{Z}[\zeta]$ is a PID (the largest one is 19), while there seem to be infinitely many regular primes. Oddly enough, although irregular primes are rare, it is not known that there are infinitely many regular primes, but it *is* known that there are infinitely many irregular ones (the first being 37). In any case, Kummer's result proves FLT for all exponents less than 37, and for many more.

A bit of history: Lamé made the mistake of assuming that unique factorization holds in $\mathbb{Z}[\zeta]$ (which is equivalent to $\mathbb{Z}[\zeta]$ being a PID). He addressed the French

Academy of Science in 1847, claiming to have proven FLT. Liouville noted the gap in the proof immediately. Kummer tried to fix it by inventing "ideal numbers", for which unique factorization will hold. This is how the concept of an ideal in a ring was born, and where it got its name. One of the main results we shall prove later on, is that although unique factorization may fail in $\mathbb{Z}[\zeta]$, ideals in this ring admit unique factorization as products of prime ideals. Thus Kummer's introduction of ideals solved in some sense the unique factorization problem. But ideals are not as nice as numbers, and although one can add and multiply them, they do not form a ring. This is why Kummer could only exploit Liouville's idea for regular primes. This line of attack on FLT was the source of many more developments in number theory and algebra, mostly in the second half of the 19th century. Eventually Taylor and Wiles proved FLT in 1995, by completely different, and much more sophisticated, means.

The lesson from the two examples discussed above is (a) that problems phrased entirely in $\mathbb{Z}$ require studying subrings of number fields such as $\mathbb{Q}(i)$ and $\mathbb{Q}(\zeta)$, and (b) that unique factorization is an important issue in these rings.

3. The geometric embedding of a number field

3.1. The geometric embedding. Let K be a number field of degree $n = [K : \mathbb{Q}]$. We denote by $Emb(K, \mathbb{C})$ the set of embeddings of K into the complex numbers. There are n such embeddings. If $K = \mathbb{Q}(\alpha)$ and the minimal polynomial of α is $f \in \mathbb{Q}[X]$ then f is separable, and if $\alpha = \alpha_1, \ldots, \alpha_n$ is the list of all its roots in $\mathbb{C}$, the distinct embeddings are obtained by sending α to the various α_i. If K is a *normal* extension of $\mathbb{Q}$ then the embeddings are automorphisms of K, and under composition they make up the Galois group $Gal(K/\mathbb{Q})$, but in what follows we nowhere make this assumption.

An embedding σ is called *real* if $\sigma(K) \subset \mathbb{R}$. It is called *complex* if it is not real. The complex conjugate $\bar{\sigma}$ of σ is the embedding defined by $\bar{\sigma}(x) = \overline{\sigma(x)}$. Clearly σ is real if and only if $\bar{\sigma} = \sigma$, so the complex embeddings come in pairs of complex conjugate ones. It is customary to denote by $r_1(K)$ the number of real embeddings of K, and by $2r_2(K)$ the number of complex embeddings, so that $r_1 + 2r_2 = n$.

EXERCISE 3.1. *Determine r_1 and r_2 for quadratic extensions of $\mathbb{Q}$, and for $K = \mathbb{Q}(^3\sqrt{2})$.*

We shall identify $\mathbb{C}$ with $\mathbb{R}^2$ as usual by sending z to (Rez, Imz). Let $\sigma_1, \ldots, \sigma_{r_1}$ be the real embeddings of K, and $\sigma_{r_1+1}, \ldots, \sigma_{r_1+r_2}, \sigma_{r_1+r_2+1} = \bar{\sigma}_{r_1+1}, \ldots, \sigma_{r_1+2r_2} = \bar{\sigma}_{r_1+r_2}$ its complex embeddings. Note that we made an arbitrary choice in ordering the embeddings, and that from each pair of complex conjugate ones, we singled out one which we denote by σ rather than $\bar{\sigma}$. These choices will have little effect on what follows, but it is important to keep them in mind. The map

$$\varphi : K \hookrightarrow \mathbb{R}^{r_1} \times \mathbb{C}^{r_2} = \mathbb{R}^n \tag{3.1}$$

given by $\varphi(x) = (\sigma_1(x), \ldots, \sigma_{r_1+r_2}(x))$, is called *the geometric embedding of K*. It is a $\mathbb{Q}$-linear map whose image, as we shall soon see, is dense in $\mathbb{R}^n$.

3.2. The discriminant of a basis.

LEMMA 3.1. *$\omega_1, \ldots, \omega_n$ is a basis of K over $\mathbb{Q}$ if and only if $\varphi(\omega_1), \ldots, \varphi(\omega_n)$ is a basis of $\mathbb{R}^n$ over $\mathbb{R}$.*

Proof. It is clear that a linear dependence between the ω_j over $\mathbb{Q}$ gets translated into a linear dependence between the $\varphi(\omega_j)$. Conversely, if ω_i are independent over $\mathbb{Q}$ we know that $\det(\sigma_i(\omega_j)) \neq 0$ (Appendix, corollary to Artin's theorem on independence of characters). Denoting $\omega_{ij} = \sigma_i(\omega_j)$ we see that the matrix $B = (\omega_{ij})$ and the matrix $A \in M_n(\mathbb{R})$ whose columns are the vectors $\varphi(\omega_j) \in \mathbb{R}^n$ are very close to each other. If $r_2 = 0$ they are the same. Otherwise, the only difference is that certain pairs of lines $l, \bar{l}$ in the first get replaced by Rel, Iml in the second. It follows that $\det(B) = \pm(2i)^{r_2}\det(A)$, hence $\det(A)$ is nonzero, as we had to prove. □

Definition 3.1. *The discriminant of the basis $\omega_1, \ldots, \omega_n$ is*

$$\Delta(\omega_1, \ldots, \omega_n) = \det(\sigma_i(\omega_j))^2. \tag{3.2}$$

Exercise 3.2. *(1) Δ is well defined, independently of the ordering of the basis or of the embeddings.*

(2) Δ is equal to

$$\left\{(2i)^{r_2} vol\left(\mathbb{R}^n/\varphi(\mathbb{Z}\omega_1 + \cdots + \mathbb{Z}\omega_n)\right)\right\}^2. \tag{3.3}$$

(3) The trace form *$B : K \times K \to \mathbb{Q}$ is the bilinear form defined as*

$$B(x, y) = Tr_{K/\mathbb{Q}}(xy). \tag{3.4}$$

Prove that $\Delta = \det(B(\omega_i, \omega_j))$. In particular the discriminant is a rational number, whose sign is $(-1)^{r_2}$.

Exercise 3.3. *If $K = \mathbb{Q}(\alpha)$ where α is a solution of an irreducible equation $x^2 + bx + c = 0$, then $\Delta(1, \alpha) = b^2 - 4c$.*

Exercise 3.4. *If $K = \mathbb{Q}(\alpha)$ and $\alpha = \alpha_1, \ldots, \alpha_n$ are the conjugates of α, then*

$$\Delta(1, \alpha, \ldots, \alpha^{n-1}) = \det(\alpha_i^{j-1})^2 \tag{3.5}$$

is the square of a famous determinant (the Van der Monde determinant). Evaluate it.

3.3. Lattices.

Definition 3.2. *A* lattice *in K is the additive subgroup spanned by a basis of K over $\mathbb{Q}$:*

$$\Lambda = \mathbb{Z}\omega_1 + \cdots + \mathbb{Z}\omega_n. \tag{3.6}$$

A (geometric) lattice *in $\mathbb{R}^n$ is the additive subgroup spanned by a basis of $\mathbb{R}^n$ over $\mathbb{R}$.*

The lemma can be rephrased by saying that Λ is a lattice if and only if $\varphi(\Lambda)$ is a geometric lattice.

Exercise 3.5. *A subgroup Λ of $\mathbb{R}^n$ is a lattice if and only if it is discrete and cocompact.*

Exercise 3.6. *Is every subgroup of $\mathbb{R}^n$ which is isomorphic to $\mathbb{Z}^n$ a lattice?*

Lemma 3.2. *A finitely generated additive subgroup of K is a lattice if and only if it contains a basis.*

Proof. Suppose Λ is a finitely generated additive subgroup. Since it has no torsion, it is isomorphic to some $\mathbb{Z}^m$. If $m > n$, it has m elements which are linearly independent over $\mathbb{Z}$, hence also over $\mathbb{Q}$ (by clearing denominators any $\mathbb{Q}$-dependence yields a $\mathbb{Z}$-dependence). This is impossible, so $m \leq n$. If Λ contains a basis, then $m = n$. Let $\omega_1, \ldots, \omega_n$ be the basis of Λ over $\mathbb{Z}$. Then it is also a basis of K over $\mathbb{Q}$, so Λ is a lattice. □

Corollary 3.3. *The product of two lattices M and N in K defined as*

$$MN = \left\{ \sum m_i n_i; \ m_i \in M, \ n_i \in N \right\} \tag{3.7}$$

is a lattice.

Proof. If ω_i be a basis of M and ξ_j a basis of N, then $\omega_i \xi_j$ is a (clearly redundant) set of generators for MN, so MN is finitely generated. It clearly contains a basis, for example $\omega_1 \xi_j$, $1 \leq j \leq n$. □

Definition 3.3. *We have seen in part (2) of Exercise 3.2 that $\Delta(\omega_1, \ldots, \omega_n)$ depends only on the lattice spanned by the ω_i. We therefore call it also the* discriminant of the lattice *and denote it by $\Delta(\Lambda)$.*

Lemma 3.4. *If $\Lambda_1 \subset \Lambda_2$ are two lattices then*

$$[\Lambda_2 : \Lambda_1]^2 = \Delta(\Lambda_1)/\Delta(\Lambda_2). \tag{3.8}$$

Proof. This follows from Exercise 3.2, part (2). Alternatively, if $\omega_1, \ldots, \omega_n$ is a basis for Λ_2, and $\omega'_j = \sum_k c_{jk} \omega_k$ ($c_{jk} \in \mathbb{Z}$) is a basis for Λ_1, then clearly $\det(\sigma_i \omega'_j) = \det(c_{jk}) \det(\sigma_i \omega_k)$ so

$$\Delta(\omega'_1, \ldots, \omega'_n) = \det(c_{ij})^2 \Delta(\omega_1, \ldots, \omega_n). \tag{3.9}$$

However, $|\det(c_{ij})| = [\Lambda_2 : \Lambda_1]$ (see appendix). Note that the fact that $\Delta(\Lambda)$ depends only on Λ and not on the basis is a special case of the *proof* of the lemma, when $\Lambda_1 = \Lambda_2$ is the same lattice, but the bases may be different. □

Exercise 3.7. *For any two lattices Λ_1 and Λ_2 in K there exists an integer m such that*

$$m\Lambda_1 \subset \Lambda_2 \subset m^{-1}\Lambda_1. \tag{3.10}$$

3.4. Orders.

Definition 3.4. *A lattice in K is an* order *if it is also a subring (closed under multiplication, and contains 1).*

For example, for any rational number a, $\mathbb{Z}a$ is a lattice in $\mathbb{Q}$, but the only order is $\mathbb{Z}$. More generally, a lattice as above is an order if and only if it contains 1, and the rational numbers c^l_{ij} in the expansion

$$\omega_i \omega_j = \sum c^l_{ij} \omega_l \tag{3.11}$$

are integers.

Orders exist.

Lemma 3.5. *If Λ is an order, $\Delta(\Lambda) \in \mathbb{Z}$.*

PROOF. We use the description of the discriminant as $\det(B(\omega_i, \omega_j))$. Now

$$\omega_i\omega_j\omega_k = \sum_l \sum_m c_{ij}^l c_{lk}^m \omega_m \tag{3.12}$$

so $B(\omega_i, \omega_j) = Tr_{\mathbb{Q}/K}(\omega_i\omega_j) = \sum_k \sum_l c_{ij}^l c_{lk}^k$. If Λ is an order, $B(\omega_i, \omega_j) \in \mathbb{Z}$. □

COROLLARY 3.6. *Any increasing sequence $\Lambda_1 \subset \Lambda_2 \subset \cdots$ of orders must stabilize: there exists a k such that $\Lambda_k = \Lambda_{k+1} = \cdots$.*

PROOF. We have shown that $\Delta(\Lambda_{i+1})$ divides $\Delta(\Lambda_i)$ for each i. □

3.5. Maximal orders.

THEOREM 3.7. *There exists a unique maximal order in K.*

PROOF. We have seen that orders exist, and the last corollary shows that a maximal order exists. It remains to show that it is unique. Suppose that Λ_1 and Λ_2 are orders, spanned by ω_i and η_j respectively. Let Λ be the $\mathbb{Z}$-span of $\omega_i\eta_j$. Since 1 is a linear combination with integer coefficients of the ω_i, $\Lambda_2 \subset \Lambda$ and similarly $\Lambda_1 \subset \Lambda$. Thus Λ is a finitely generated subgroup of K which contains a basis, so it is a lattice. It is easy to see that $\omega_i\eta_j\omega_{i'}\eta_{j'}$ is an integral linear combination of the $\omega_l\eta_k$. It follows that Λ is an order. We have shown that any two orders are contained in another order, and this implies that the maximal order is unique. □

We denote the maximal order of K by $\mathcal{O}_K$. Non-maximal orders are often denoted by the letter $\mathcal{O}$.

4. Algebraic integers

We have arrived at the notion of an order, and a maximal order, from the geometric embedding of K. There is another, purely algebraic, way to arrive at the same notion. Because of its importance, we present it in somewhat greater generality.

4.1. Integral extensions.

DEFINITION 4.1. *Let A be a subring of a field Ω. An element $\alpha \in \Omega$ is said to be* integral *over A if it satisfies a monic polynomial with coefficients from A : if there exists a relation of the form*

$$\alpha^m + a_1\alpha^{m-1} + \cdots + a_m = 0 \tag{4.1}$$

where $a_i \in A$.

A ring $A \subset B \subset \Omega$ is an integral extension *of A if all its elements are integral over A.*

The ring A is said to be integrally closed in Ω *if every element of Ω which is integral over A is already in A.*

A domain A is said to be integrally closed *if it is integrally closed in its field of fractions.*

If A is a field this is the familiar notion of being "algebraic" over A.

EXERCISE 4.1. *Let A be a unique factorization domain. Then A is integrally closed.*

Hint: Let $\alpha = u/v$ be an element of the fraction field of A and assume that it satisfies a monic equation as above with $a_i \in A$. We may assume that u, v are in A and are relatively prime. Multiply by v^m to get that v divides u^m in A, hence must be a unit, so $\alpha \in A$.

It follows that the ring $\mathbb{Z}$ is integrally closed. Recall that a module M over a ring A is *finitely generated* if there are finitely many $\alpha_1, \dots, \alpha_m$ in M such that $M = \sum A\alpha_i$.

THEOREM 4.1. *Let A be a subring of a field Ω, and $\alpha \in \Omega$. The following are equivalent.*

(a) α is integral over A.

(b) The ring generated by α over A, denoted $A[\alpha]$, is finitely generated as an A-module.

(c) There exists a finitely generated A-module $M \subset \Omega$ such that $\alpha M \subset M$.

PROOF. Suppose (a) holds. Then $1, \alpha, \dots, \alpha^{m-1}$ span $A[\alpha]$ as an A-module because α^m is already a linear combination of them with coefficients from A. To show that (b) implies (c) simply take $M = A[\alpha]$. To show that (c) implies (a) pick a system of generators $\omega_1, \dots, \omega_m$ of M as an A module (it need not be a minimal system of generators). Write $\alpha\omega_i = \sum a_{ij}\omega_j$ where $a_{ij} \in A$ (which need not be unique). The matrix $\alpha I - (a_{ij})$ annihilates the vector ${}^t(\omega_1, \dots, \omega_m)$, so it's determinant is 0. This gives the monic polynomial in $A[X]$ satisfied by α. □

COROLLARY 4.2. *The sum, difference and product of two elements of Ω which are integral over A is integral over A.*

PROOF. Let α and β be integral over A. Then $M = A[\alpha]$ and $N = A[\beta]$ are finitely generated A modules, hence so is MN, because it is generated by all the products of a generator of M with a generaor of N. But this A-module is invariant under multiplication by $\alpha + \beta$ and $\alpha\beta$, so by criterion (c) of the theorem, we are done. □

COROLLARY 4.3. *The set of all elements of Ω which are integral over A is a ring. It is called the* integral closure *of A in Ω.*

PROPOSITION 4.4. *If $B = A[\alpha_1, \dots, \alpha_n]$ is finitely generated as a ring over A, and the α_i are integral over A, then B is a finite A-module.*

PROOF. By induction on n. The case $n = 1$ is criterion (b) of the theorem. Suppose we have proved that $A' = A[\alpha_1, \dots, \alpha_{n-1}]$ is a finite A-module, so that $A' = \sum_{i=1}^r A\omega_i$. Since α_n is clearly integral over A', $B = A'[\alpha_n] = \sum_{j=1}^s A'\eta_j$. It is now clear that $B = \sum_{i=1}^r \sum_{j=1}^s A\omega_i\eta_j$. □

PROPOSITION 4.5. *Suppose B is integral over A and α is integral over B. Then α is integral over A.*

PROOF. Let $b_1, \dots, b_r$ be the coefficients of a monic polynomial of α over B. Since they are in B, they are integral over A, so by the previous proposition, $A' = A[b_1, \dots, b_r]$ is a finite A-module. Since α is integral over A', $A'[\alpha]$ is a finite A' module, hence also a finite A module. By criterion (c) of the theorem, α is integral over A. □

COROLLARY 4.6. *The integral closure of A in Ω is integrally closed.*

Corollary 4.7. *If A is integrally closed, and α is integral over A, then already the monic minimal polynomial of α has coefficients from A.*

Proof. Let $K \subset \Omega$ be the field of fractions of A, in which A is integrally closed. All the conjugates of α over A are also integral over A, hence so are the coefficients of the minimal polynomial, which are sums of products of these conjugates. But these coefficents lie in K, so by our assumption, they belong to A. □

Proposition 4.8. *Let K be the field of fractions of A. If α is algebraic over K, then for some $a \in A$, $a\alpha$ is integral over A.*

Proof. Clearing denominators, α satisfies a (non-monic, perhaps) polynomial over A, say

$$a_0\alpha^m + \cdots + a_m = 0. \tag{4.2}$$

Multiplying by a_0^m we see that $a_0\alpha$ satisfies a monic polynomial with coefficients from A. □

4.2. Algebraic integers. We now specialize the algebraic results of the previous section to subrings of number fields. Let A be an order in a number field K. Then A is a finitely generated $\mathbb{Z}$-module, and since it is closed under multiplication, all its elements are integral over $\mathbb{Z}$. Thus the elements of $\mathcal{O}_K$ are integral over $\mathbb{Z}$. Conversely, let α be integral over $\mathbb{Z}$. Let $\omega_1, \ldots, \omega_n$ be a basis of K over $\mathbb{Q}$. Multiplying by a rational integer we may assume that the ω_i are integral over $\mathbb{Z}$. Consider $A = \mathbb{Z}[\alpha, \omega_1, \ldots, \omega_n]$. Since α and the ω_i are integral over $\mathbb{Z}$, by a previous claim, A is a finite $\mathbb{Z}$-module. Since it contains a basis of K, it is a lattice, and since it is a ring, it is an order. Thus $\alpha \in \mathcal{O}_K$.

We have shown that the maximal order of K is simply the ring of all elements of K which are integral over $\mathbb{Z}$. It is called the ring of *algebraic integers* of K. If $K \subset L$ are two number fields then $\mathcal{O}_L$ is simply the integral closure of $\mathcal{O}_K$ in L.

Exercise 4.2. *Let $K = \mathbb{Q}(\sqrt{D})$ be a quadratic extension of $\mathbb{Q}$ where D is a (positive or negative) square-free integer different from 0 or 1. Let*

$$\begin{aligned} \omega_D &= \sqrt{D} \ \textit{if}\ D \equiv 2, 3 mod 4 \\ \omega_D &= \frac{1+\sqrt{D}}{2} \ \textit{if}\ D \equiv 1 mod 4 \end{aligned} \tag{4.3}$$

Prove that $\mathcal{O}_K = \mathbb{Z} + \mathbb{Z}\omega_D$ and compute $\Delta(1, \omega_D)$.

4.3. Dedekind domains. Unlike $\mathbb{Z}$, the rings $\mathcal{O}_K$ need not be PID's or UFD's. We next study their basic properties.

Proposition 4.9. *Let L/K be a finite extension of number fields. Then $N_{L/K}$ and $Tr_{L/K}$ map $\mathcal{O}_L$ into $\mathcal{O}_K$. The norm maps an ideal $\mathfrak{a}$ into $\mathfrak{a} \cap \mathcal{O}_K$.*

Proof. Embed L in a number field M which is Galois over K. There are $[L : K]$ different embeddings of L into M over K, which we denote $Emb_K(L, M)$. If $\alpha \in \mathcal{O}_L$ then the conjugates $\sigma(\alpha)$ of α, for $\sigma \in Emb_K(L, M)$, are in $\mathcal{O}_M$ (because they are in M and they are algebraic integers). Their sum is $Tr_{L/K}(\alpha)$ and their product is $N_{L/K}(\alpha)$ (see Appendix), and these are therefore in $\mathcal{O}_M \cap K = \mathcal{O}_K$.

Suppose now that $\alpha \in \mathfrak{a}$. The product of the conjugates $\sigma(\alpha)$ for $\sigma \neq id.$ is equal to $N_{L/K}(\alpha)\alpha^{-1}$, and is therefore in $\mathcal{O}_M \cap L = \mathcal{O}_L$. It follows that $N_{L/K}(\alpha) \in \mathfrak{a} \cap \mathcal{O}_K$. □

Exercise 4.3. *Similarly, prove that all the coefficients of the characteristic polynomial of $\alpha \in \mathcal{O}_L$ are in $\mathcal{O}_K$.*

Theorem 4.10. *The ring $\mathcal{O}_K$ satisfies: (i) it is a Noetherian domain (ii) it is integrally closed (iii) every nonzero prime ideal of $\mathcal{O}_K$ is maximal.*

Proof. Of the three listed property, the first two have already been proved. Since an ideal of $\mathcal{O}_K$ is a lattice, and any lattice has a finite index in $\mathcal{O}_K$, which is a positive integer, any ascending chain of ideals becomes stationary, proving (i), and (ii) follows from the fact that $\mathcal{O}_K$ is the integral closure of $\mathbb{Z}$ in K. To prove (iii) let $\mathfrak{p}$ be a nonzero prime ideal of $\mathcal{O}_K$. If a is a nonzero element of $\mathfrak{p}$, then $\alpha = N_{K/\mathbb{Q}}(a)$ is a nonzero element of $\mathfrak{p} \cap \mathbb{Z}$. Since $\mathfrak{p} \cap \mathbb{Z}$ is a prime ideal of $\mathbb{Z}$, it must contain a rational prime p and so $\mathcal{O}_K/\mathfrak{p}$ is a vector space over $\mathbb{Z}/p\mathbb{Z}$. It is also finitely generated as a vector space, because $\mathcal{O}_K$ is a finite $\mathbb{Z}$-module. It is therefore a *finite* domain. However, every finite domain is a field, hence $\mathfrak{p}$ is maximal. □

Remark 4.1. *A close examination of the last step in the proof shows that we did not have to use the fact that $\mathcal{O}_K/\mathfrak{p}$ is finite. Instead we could use the following lemma.*

Lemma 4.11. *Let k be a field and R a finite k-algebra which is a domain. Then R is a field.*

Proof. The only problem is to show that every nonzero element of R is invertible. Let a be such an element, and consider as before the linear transformation $M_a : x \mapsto ax$ of R over k. By the Cayley Hamilton theorem a satisfies some polynomial from $k[X]$, namely $char.(M_a)$. Let $f \in k[X]$ be a polynomial of minimal degree satisfied by a. Since R has no zero divisors, the constant term of f is nonzero, so we may assume that it is -1. Moving the constant term to the other side of the equation $f(a) = 0$ we find an expression $ag(a) = 1$, hence a is invertible. □

Definition 4.2. *A* Dedekind domain *is a domain A staisfying (i) A is Noetherian (ii) A is integrally closed (iii) every nonzero prime of A is maximal.*

Exercise 4.4. *Every PID is a Dedekind domain.*

The converse is not true, but as we shall see soon, Dedekind domains are precisely those rings which are "locally" PID's. For that we have to develop somewhat the local-to-global dictionary, which is fundamental in algebraic number theory and, more generally, in commutative algebra and algebraic geometry.

Exercise 4.5. *In each of the examples below determine if it is a Dedekind domain.*

(i) $\mathbb{C}[X, Y]$ *(ii) The ring of all algebraic integers in* $\bar{\mathbb{Q}}$ *(iii)* $\mathbb{C}[X, Y]/(Y^2 - X^3)$ *(iv)* $\mathbb{Z}[\sqrt{-5}]$ *(v)* $\mathbb{Z}[\sqrt{5}]$.

Hint: *for example (iii), show that the ring is a domain and let x and y be the classes of X and Y. Prove that y/x is an element of the field of fractions which is integral over the ring, but does not lie in it.*

Exercise 4.6. $\mathbb{Z}[\sqrt{-5}]$ *is a Dedekind domain, but not a PID.* Hint: $6 = 2 \cdot 3 = (1 + \sqrt{-5})(1 - \sqrt{-5})$.

4.4. Minkowski's lemma. This is a simple but useful lemma in geometry. A subset of $\mathbb{R}^n$ is called *symmetric* if together with x it contains $-x$. It is called *midpoint convex* if with x and y it contains also $(x+y)/2$. Convex sets are midpoint convex.

LEMMA 4.12. *If $U \subset \mathbb{R}^n$ is measurable, symmetric and midpoint convex, and if Λ is a lattice in $\mathbb{R}^n$, such that*

$$2^n vol(\mathbb{R}^n/\Lambda) < vol(U), \tag{4.4}$$

then U contains a nonzero point of Λ.

PROOF. Assume that $U \cap \Lambda = \{0\}$. Let $\omega_1, \dots, \omega_n$ be a basis for Λ, and $\Pi = \{\sum t_i\omega_i;\ 0 \le t_i < 1\}$ the fundamental parallelopiped defined by $\{\omega_i\}$. The subsets $W_\lambda = (\frac{1}{2}U - \lambda) \cap \Pi$, $(\lambda \in \Lambda)$ are disjoint, because if $u/2 - \lambda = v/2 - \mu$ $(u, v \in U$ and $\lambda, \mu \in \Lambda)$ then $\lambda - \mu \in U$ by our assumptions on U, hence $\lambda = \mu$. Since they are also measurable

$$vol(\frac{1}{2}U) = \sum_{\lambda\in\Lambda} vol(W_\lambda) \le vol(\Pi) = vol(\mathbb{R}^n/\Lambda). \tag{4.5}$$

The first equality follows from the fact that $W_\lambda + \lambda = \frac{1}{2}U \cap (\Pi + \lambda)$ make up a disjoint covering of $\frac{1}{2}U$. This contradiction proves that U must contain a nonzero point from Λ. □

4.5. The discriminant and Hermite's theorem.

DEFINITION 4.3. *The* discriminant d_K *of a number field K is the discriminant $\Delta(\omega_1, \dots, \omega_n)$ of a basis of $\mathcal{O}_K$ over $\mathbb{Z}$. It is a well-defined integer, whose sign is $(-1)^{r_2(K)}$.*

THEOREM 4.13. *There are only finitely many number fields K with a bounded degree and discriminant.*

We shall see later that when $[K : \mathbb{Q}] \to \infty$, $|d_K| \to \infty$ too. This will prove

THEOREM 4.14. *(Hermite) There are only finitely many number fields with a bounded discriminant.*

PROOF. Let M be a positive real number . Consider the rectangle

$$U = \left\{x \in \mathbb{R}^n;\ |x_i| \le \frac{1}{2} \text{ for } i \le n-1 \text{ and } |x_n| \le 2^{n-1}M\right\} \tag{4.6}$$

which is convex, symmetric, and whose volume is 2^nM. Let K be any number field of degree n with $\sqrt{|d_K|} < M$, and choose a geometric embedding φ of K in $\mathbb{R}^n$ as above. Then

$$vol(\mathbb{R}^n/\varphi(\mathcal{O}_K)) = 2^{-r_2(K)}\sqrt{|d_K|} < M, \tag{4.7}$$

so by Minkowski's lemma there will be a non-zero $\alpha \in \mathcal{O}_K$ with $\varphi(\alpha) \in U$. For any embedding σ of K in $\mathbb{C}$ other than $\tau = \sigma_{r_1+r_2}$ or $\bar{\tau}$ we have $|\sigma(\alpha)| \le \sqrt{2}/2$ (even $\le 1/2$ if σ is real). Since the norm of α is a non-zero rational integer, we must have $|\tau(\alpha)| > 1$. Moreover if τ is a complex embedding, $|Re(\tau(\alpha))| \le 1/2$, so $\tau(\alpha)$ is not real. We claim that $[\mathbb{Q}(\alpha) : \mathbb{Q}] = n$, hence $K = \mathbb{Q}(\alpha)$. In fact, if α generates a proper subfield of K, then for some $\sigma \ne \tau$ we must have $\sigma(\alpha) = \tau(\alpha)$ because $\tau(\alpha)$, like α, would have less than n conjugates. But we have just seen that this is impossible for $\sigma \ne \tau, \bar{\tau}$ and also for $\sigma = \bar{\tau}$.

We conclude that any K of degree n and $\sqrt{|d_K|} < M$ is generated by an algebraic integer α such that $\varphi(\alpha) \in U$. The (monic) minimal polynomial f_α of α will have $\mathbb{Z}$ coefficients which are symmetric functions of the n conjugates of α. Since these conjugates are bounded (in terms of n and M), so will be the coefficents of f_α. We conclude that there are only finitely many possibilities for f_α, hence for α, hence for K. □

5. Ideal theory in Dedekind domains

5.1. Unique factorization of ideals. As indicated in the introduction the concept of an ideal in a ring grew out of Kummer's attempt to replace unique factorization in rings of algebraic integers by unique factorization between "ideal numbers". The modern set-theoretic definition of an ideal as an additive subgroup closed under multiplication by ring elements is due to Dedekind, who also showed that in the rings named after him unique factorization of ideals indeed holds.

THEOREM 5.1. *Let R be a Dedekind domain. Then every ideal I can be written as a product of prime ideals*

$$I = P_1 \dots P_r \tag{5.1}$$

and this decomposition is unique up to a permutation of the factors.

EXAMPLE 5.1. *Let R be a PID. The prime ideals are precisely those of the form $P = (\pi)$ where π is an irreducible element. Two ideals (a) and (b) are equal if and only if $a = \varepsilon b$ for some unit ε, and of course $(a)(b) = (ab)$. Taken together we recover unique factorization in PID's.*

EXERCISE 5.1. *This example shows how Dedekind's theorem remedies the lack of unique factorization in the ring $\mathbb{Z}[\sqrt{-5}]$, which is a Dedekind domain, but not a PID. We have seen before that*

$$6 = 2 \cdot 3 = (1 + \sqrt{-5})(1 - \sqrt{-5}) \tag{5.2}$$

in this ring, and that 2, 3 or $1 \pm \sqrt{-5}$ are all irreducible elements. The point is that they do not generate prime ideals. Let $P = (2, 1 + \sqrt{-5}) = (2, 1 - \sqrt{-5})$. Let $Q = (3, 1 + \sqrt{-5})$ and $\bar{Q} = (3, 1 - \sqrt{-5})$. Show that these are all primes and that

$$(2) = P^2,\ (3) = Q\bar{Q},\ (1 + \sqrt{-5}) = PQ \text{ and } (1 - \sqrt{-5}) = P\bar{Q}. \tag{5.3}$$

Hint*: If $I = (a_i)$ and $J = (b_j)$ then $IJ = (a_i b_j)$.*

The proof of Dedekind's theorem relies on several lemmas.

LEMMA 5.2. *Let R be a noetherian domain. Then every ideal I of R* contains *a product of prime ideals.*

PROOF. If not, let I be an ideal which is maximal among all the ideals which *do not contain* products of primes. Such an ideal can be found by the noetherianity assumption. Since I is certainly not prime itself, there are a and b not in I, whose product $ab \in I$. But then $I + (a)$ and $I + (b)$ contain products of primes, hence so does their product, which is contained in I. This is a contradiction. □

Let K be the field of fractions of R. For any subset I of K we write

$$I' = \{x \in K \,|\, xI \subset R\}. \tag{5.4}$$

Note that it is an R-submodule of K. In general, it need not be contained in R. Of course, the smaller I is, the larger I'.

LEMMA 5.3. *Let R be Dedekind. If P is a prime ideal, then P' is not contained in R.*

PROOF. Let $a \in P$. By the previous lemma $P \supset (a) \supset P_1 \dots P_r$ a product of primes, and we may assume that their number r is minimal. Since P is prime, it must contain one of the P_i and we may assume $P \supset P_1$. Since P_1 is maximal, $P = P_1$. Now $(a) \not\supset P_2 \dots P_r$, by the minimality of r, so there exists $b \in P_2 \dots P_r$ not in (a). But $(a) \supset Pb$, so ba^{-1} is in P' but not in R. □

LEMMA 5.4. *Let R be Dedekind. If P is a prime ideal, $P'P = R$.*

PROOF. $P'P$ is easily seen to be an ideal of R containing P (since P' contains 1). By the maximality of P, if it is not R itself, it is equal to P. But then, for every $x \in P'$, $xP \subset P$, and since P is a finitely generated R module, this means that x is integral over R. Since R is integrally closed, $x \in R$. This means that P' is contained in R, contradicting the last lemma. □

PROOF. (of theorem). *Existence of decomposition*: Suppose, by way of contradiction, that I is maximal amongst all the ideals not equal to a product of primes. Let P be a maximal ideal containing I. Multiplying by P' we have

$$I \subset P'I \subset R. \tag{5.5}$$

Now $P'I$ is an ideal of R and it is strictly larger than I, otherwise all the elements of P' multiply I into itself, and are therefore integral over R, hence in R. By maximality of I, $P'I$ is already a product of primes. Multiplying by P we see that I is a product of primes as well.

Uniqueness: If $P_1 \dots P_r = Q_1 \dots Q_s$ and all the P_i and the Q_j are prime, then P_1 must contain one of the Q_j, say Q_1, hence they must be equal $P_1 = Q_1 = P$ (say). Multiplying by P' we arrive at a shorter expression and we continue inductively. □

5.2. Fractional ideals. Let R be a Dedekind domain and K its field of fractions.

DEFINITION 5.1. *A* fractional ideal *of R is a finitely generated R submodule of K.*

A fractional ideal contained in R is simply an ideal, and one does not have to say then that it is finitely generated, as this is guaranteed by noetherianity. However, K itself is not a fractional ideal in general.

Let us examine some easy properties of fractional ideals. Fractional ideals can be multiplied, where

$$IJ = \left\{ \sum a_i b_i | \, a_i \in I, \; b_i \in J \right\}. \tag{5.6}$$

This product is associative and commutative. If $a \in K$, $Ra = (a)$ is a fractional ideal, called *principal.* If I is a fractional ideal, then for some $a \in R$, aI is an ordinary ideal in R. Simply take a as the product of the denominators of a finite set of generators of I. An R-submodule of a fractional ideal is also a fractional ideal, since R is noetherian. If I is a nonzero fractional ideal so is I'. Indeed, it is an R-submodule of K, and if $a \in I$, $I' \subset (a)' = (a^{-1})$, so by the last remark, I' is a fractional ideal.

THEOREM 5.5. *Let R be a Dedekind domain, and K its field of fractions. The non-zero fractional ideals of R form a* group $\mathcal{I}$ *under multiplication. The unit is R. The inverse of I is I'. The group $\mathcal{I}$ is a free abelian group on the set of primes as generators. The principal ideals form a subgroup $\mathcal{P}$.*

PROOF. Everything is clear except that the inverse of I is I'. By definition $II' \subset R$. Suppose I is an ideal of R, and let $I = P_1 \dots P_r$ be its decomposition into a product of prime ideals. We see that $I' \supset P'_1 \dots P'_r$ so $II' \supset R$ by the lemma which said that $PP' = R$. If I is an arbitrary fractional ideal, write $I = aJ$ with $a \in K$ and J an ideal of R. Then it is easy to see that $I' = a^{-1}J'$, and the result for I follows from the result for J. □

COROLLARY 5.6. *Every fractional ideal has a unique decomposition as*

$$P_1^{m_1} \dots P_r^{m_r} \tag{5.7}$$

where the m_i are integers and the P_i are distinct prime ideals, and where $P^{-m} = (P')^m$ if $m > 0$. Such a fractional ideal is an ideal of R if and only if all the m_i are non-negative.

EXERCISE 5.2. *Find the expression for $I + J$ and $I \cap J$ in terms of the prime decomposition of I and J. Develop a theory of g.c.d and l.c.m. for fractional ideals.*

5.3. The ideal class group.

DEFINITION 5.2. *The* ideal class group *of R (or of K if the reference to R is clear; note that the same field can be the field of fractions of many Dedekind domains) is the group*

$$Cl(R) = \mathcal{I}/\mathcal{P}. \tag{5.8}$$

The class group measures how far R is from being a PID. Note that $\mathcal{I}$ is a free abelian group of infinite rank, in general. The group $\mathcal{P}$ also has a pretty easy description. Since $(a) = (b)$ if and only if ab^{-1} is a unit of R (this is proved for fractional ideals precisely as it is proved for ideals), the map $K^\times \to \mathcal{P}$ sending a to (a) induces an isomorphism

$$\mathcal{P} \simeq K^\times / R^\times. \tag{5.9}$$

This is however of little help in computing $\mathcal{P}$, and even more mysterious is the way $\mathcal{P}$ sits inside $\mathcal{I}$. The class group $Cl(R)$ is the most important invariant of R. For rings of integers in number fields it will turn out to be finite, but for general Dedekind domains it may be infinite.

6. The local-global principle

There is another description of Dedekind domains that has a geometric flavor and simplifies many of the proofs. To understand it we must recall the notion of localization.

6.1. Localizations (more algebra). Let R be a domain, and K its field of fractions. Let S be a multiplicative set of R (a set closed under multiplication) not containing 0. Common examples of such an S are $\{f^n;\, n \in \mathbb{N}\}$ - the powers of a given element f, or $R - P$, the complement of a prime ideal.

The *localization* of R at S, denoted $S^{-1}R$ is the subring

$$\{r/s;\, r \in R,\, s \in S\} \tag{6.1}$$

of K. When $S = R - P$, P a prime ideal, it is common to denote $S^{-1}R$ by R_P. For example, $\mathbb{Z}_{(p)}$ is the ring of all rational numbers whose denominator is prime to p.

The following facts are easy, and the reader should check them for himself (they are usually taught in Musgei Yesod in Algebra).

- If I is an ideal in R, then $S^{-1}I$ is an ideal in $S^{-1}R$, if I does not intersect R, or is the whole ring $S^{-1}R$ otherwise.
- If J is an ideal of $S^{-1}R$, then $I = R \cap J$ is an ideal of R and $J = S^{-1}I$.
- The two constructions described above establish a *bijection* between the *prime* ideals of R not intersecting S, and the *prime* ideals of $S^{-1}R$. In the case of R_P, the prime ideals of R_P are in one-to-one correspondence with the prime ideals of R contained in P.
- R_P has a unique maximal ideal: PR_P.

A ring R is called *local* if it has a unique maximal ideal M. Equivalently, it is local if there exists an ideal M such that $R^\times = R - M$.

Any domain R can be reconstructed from its localizations at prime ideals (inside K). It is even enough to consider localizations at maximal ideals only.

LEMMA 6.1. *Let R be a domain. Then $R = \bigcap_{M\ maximal} R_M$.*

PROOF. Let x be an element of K lying in the intersection of all the prime ideal localizations of R. Let

$$D(x) = \{d \in R;\, dx \in R\}\,. \tag{6.2}$$

This $D(x)$ is clearly an ideal of R (called the ideal of denominators of x, for obvious reasons). If it is not the whole ring R, it is contained in a maximal ideal M. But by assumption $x = r/s$ for $r \in R$ and $s \notin M$, so $s \in D(x)$. It follows that $D(x) = R$, and $1 \in D(x)$, so $x \in R$. □

This lemma is the basis for many arguments.

EXERCISE 6.1. *Immitate the proof of the lemma to show that for any ideal I, $I = \bigcap_{M\ maximal} IR_M$.*

6.2. Valued fields and DVR's (review).

DEFINITION 6.1. *Let K be a field. A (non-archemidean) valuation v on K is a map $v : K^\times \to \mathbb{R}$ satisfying*
(i) $v(xy) = v(x) + v(y)$
(ii) $v(x + y) \geq \min\{v(x), v(y)\}$.

One sometimes writes $v(0) = \infty$. The *value group* of the valuation is $v(K^\times)$. It is a subgroup of $\mathbb{R}$. If the value group is discrete, v is called a *discrete valuation.* In such a case we may scale v so that $v(K^\times) = \mathbb{Z}$. We say that v is *normalized.* If v is a normalized discrete valuation, an element π such that $v(\pi) = 1$ is called a *uniformizer* or a prime.

Suppose v is a discrete valuation. The set $R_v = \{x;\, v(x) \geq 0\}$ is a subring of K called the *valuation ring of v*. The set $P_v = \{x;\, v(x) > 0\}$ is called the *valuation ideal.* It is a principal ideal generated by any uniformizer. Since $R_v - P_v$ are precisely the units of R_v, one concludes easily that R_v is a local PID. Its only nonzero ideals are the positive powers of P_v, and the only nonzero prime ideal is P_v. The field $\kappa_v = R_v/P_v$ is called the *residue field of v.*

A local PID is called a *discrete valuation ring* (DVR). It is always of the form described above (exercise: let R be a DVR and define a valuation on its field of fractions so that R becomes R_v.)

Let (K, v) be a valued field (a field with a valuation). Let $e > 1$ be any constant, and define a *norm* and a *metric* on K by

$$|x| = e^{-v(x)},\ d(x,y) = |x-y|\,. \tag{6.3}$$

This defines a *topology* on K. Two valuations are called equivalent if they induce the same topology.

EXAMPLE 6.1. *(i) Let p be a prime of $\mathbb{Q}$ and define $v(x) = ord_p(x)$.*

(ii) On $\mathbb{C}(t)$ let $v(f) = ord_\omega(f)$, where $\omega \in \mathbb{C}$, or $v(f) = -\deg(f)$ ("the valuation placed at ∞").

6.3. Dedekind rings and DVR's. Let R be a Dedekind domain and K its field of fractions. For each nonzero prime P of K let $v_P(x)$ denote the power of P in the prime ideal factorization of the fractional ideal (x). One easily checks that this is a valuation on K, and that its valuation ring is R_P, the localization of R at P, which is therefore a DVR.

THEOREM 6.2. *All the localizations R_P of a Dedekind domain at nonzero primes are DVR's. Conversely, if R is a noetherian domain all of whose localizations are DVR's, then R is a Dedekind domain.*

PROOF. We have just seen the first direction. For the converse, recall that R is the intersection of all the R_P, where P is maximal. If P is a nonmaximal prime of R, let M be a maximal prime containing it, and look in R_M. Then PR_M is a nonmaximal prime in the DVR R_M, so must be 0. We conclude that $P = 0$. This proves that every nonzero prime of R is maximal.

To get that R is integrally closed, recall that $R = \bigcap R_P$. Each R_P is integrally closed (being a PID) and the intersection of any number of integrally closed rings is again integrally closed. □

6.4. Dedekind rings with finitely many primes. A Dedekind ring with only one prime is a DVR, hence a PID. We shall now see that a Dedekind ring with *finitely many primes* (a *semilocal* Dedekind ring) is also a PID.

THEOREM 6.3. *Let R be a Dedekind domain, $P_1, \dots, P_r$ prime ideals and e_i non-negative integers. Then there exists an $\alpha \in R$ with $v_{P_i}(\alpha) = e_i$ for each i.*

PROOF. For each i choose $\alpha_i \in P_i^{e_i} - P_i^{e_i+1}$. By the Chinese remainder theorem, since the $P_i^{e_i+1}$ are co-prime, there is an $\alpha \in R$ such that $\alpha \equiv \alpha_i mod P_i^{e_i+1}$. Clearly $v_{P_i}(\alpha) = e_i$. □

COROLLARY 6.4. *A Dedekind domain with only finitely many primes is a PID.*

PROOF. Let $P_1, \dots, P_r$ be the nonzero primes of R. Every ideal is of the form $I = P_1^{e_1} \dots P_r^{e_r}$. Choose an α as in the theorem. Comparing valuations, we see that α generates I. □

EXERCISE 6.2. *Prove that every ideal in a Dedekind domain is generated by two elements.*

Semilocal Dedekind domains are obtained in the following way. Consider a rational prime p and $S = \mathbb{Z} - p\mathbb{Z}$. Let $\mathcal{O}_{K,(p)} = S^{-1}\mathcal{O}_K$ be the localization of $\mathcal{O}_K$ in S. Then $\mathcal{O}_{K,(p)}$ is semilocal. Indeed, by the general theory of localizations we have to show that there are only finitely many primes of $\mathcal{O}_K$ containing p, because any prime of $\mathcal{O}_K$ contains a rational prime, and any rational prime other than p is in S so becomes invertible. But for every $a \in \mathcal{O}_K$ there are only finitely many primes containing it, namely those in the prime ideal decomposition of (a).

6.5. The absolute norm of an ideal. Let $\mathfrak{a}$ be a fractional ideal of the ring of integers $\mathcal{O}_K$ in a number field K. Let $\omega_1, \dots, \omega_n$ be a $\mathbb{Z}$-basis of $\mathcal{O}_K$ and $\sum_j a_{ij}\omega_j$ (where the $a_{ij} \in \mathbb{Q}$, $1 \le i \le n$) a $\mathbb{Z}$-basis of $\mathfrak{a}$. We define the *absolute norm* $N\mathfrak{a}$ of $\mathfrak{a}$ to be

$$N\mathfrak{a} = |\det(a_{ij})|. \tag{6.4}$$

In other words, the norm is the determinant of the matrix transforming a basis of $\mathcal{O}_K$ to a basis of $\mathfrak{a}$. The norm is well defined, because the matrix (a_{ij}) is well-defined up to a matrix from $GL_n(\mathbb{Z})$. Here are its main properties.

- If $\mathfrak{a}$ is an integral ideal, $N\mathfrak{a} \in \mathbb{Z}$ and $N\mathfrak{a} = [\mathcal{O}_K : \mathfrak{a}]$.
  This is clear since the matrix (a_{ij}) is integral. The relation to the index is well-known (see Proposition x.xx in the appendix).
- If $\mathfrak{p}$ is a prime ideal, $N\mathfrak{p}$ is the number of elements in the residue field $\kappa(\mathfrak{p}) = \mathcal{O}_K/\mathfrak{p}$, so it is a power of the characteristic p.
- If $\beta \in K^\times$ then $N(\beta\mathfrak{a}) = |N_{K/\mathbb{Q}}\beta| N\mathfrak{a}$.

If A is the transition matrix from a basis of $\mathcal{O}_K$ to a basis of $\mathfrak{a}$, and B the matrix describing multiplication by β in the given basis of $\mathfrak{a}$, then BA transforms a basis of $\mathcal{O}_K$ to a basis of $\beta\mathfrak{a}$. By definition $N_{K/\mathbb{Q}}\beta = \det B$, so

$$N(\beta\mathfrak{a}) = \det(BA) = \det B \det A = |N_{K/\mathbb{Q}}\beta| N\mathfrak{a} \tag{6.5}$$

follows. The next property needs some attention.

PROPOSITION 6.5. *The norm is multiplicative: $N(\mathfrak{a}\mathfrak{b}) = N\mathfrak{a}N\mathfrak{b}$.*

PROOF. If one of the ideals is principle, then this is essentially the last property. The problem is how to overcome the difficulties introduced by the fact that $\mathcal{O}_K$ need not be a PID. The solution is to localize and use the local-global principle. Multiplying by an integer we may assume that both $\mathfrak{a}$ and $\mathfrak{b}$ are integral ideals. We have to show

$$[\mathcal{O}_K : \mathfrak{a}\mathfrak{b}] = [\mathcal{O}_K : \mathfrak{a}][\mathcal{O}_K : \mathfrak{b}], \tag{6.6}$$

or, equivalently

$$[\mathcal{O}_K : \mathfrak{b}] = [\mathfrak{a} : \mathfrak{a}\mathfrak{b}]. \tag{6.7}$$

If $\Lambda \supset \Lambda'$ are any two lattices and p is a prime we denote by $[\Lambda : \Lambda']_p$ the p-part of the index $[\Lambda : \Lambda']$. It is clearly enough to show that for each p separately, $[\mathcal{O}_K : \mathfrak{b}]_p = [\mathfrak{a} : \mathfrak{a}\mathfrak{b}]_p$. Denote by $\Lambda_{(p)}$ the localization of Λ at p, namely the free $\mathbb{Z}_{(p)}$ module generated by Λ. If $\omega_1, \dots, \omega_n$ is a basis of Λ over $\mathbb{Z}$ then it is also a basis of $\Lambda_{(p)}$ over the DVR $\mathbb{Z}_{(p)}$. We now forget the fact that $\Lambda_{(p)}$ came from Λ. A $\mathbb{Z}_{(p)}$ lattice M in K is any $\mathbb{Z}_{(p)}$ module spanned by a basis of K. If M and M' are two such $\mathbb{Z}_{(p)}$-lattices then we may consider a matrix A in $GL_n(\mathbb{Q})$ that transforms a basis if M to a basis of M'. It is well-defined up to multiplication (on both sides) by matrices from $GL_n(\mathbb{Z}_{(p)})$. In particular, its determinant is well-defined not up to ±1 this time, but up to a p-adic unit in $\mathbb{Q}$. It follows that the *p-index*

$$[M : M']_p = p^{val_p(\det A)} \tag{6.8}$$

is well defined. The notation is not ambiguous because if $M = \Lambda_{(p)}$ and $M' = \Lambda'_{(p)}$ then their p-index is the p-part of $[\Lambda : \Lambda']$. One last obvious remark is that the process of localization at p respects products: $(\mathfrak{a}\mathfrak{b})_{(p)} = \mathfrak{a}_{(p)}\mathfrak{b}_{(p)}$. Putting together all these remarks, we have to show that

$$[\mathcal{O}_{K,(p)} : \mathfrak{b}_{(p)}]_p = [\mathfrak{a}_{(p)} : \mathfrak{a}_{(p)}\mathfrak{b}_{(p)}]_p. \tag{6.9}$$

The advantage now is that $\mathcal{O}_{K,(p)}$ is semilocal, hence a PID, so $\mathfrak{a}_{(p)} = \alpha\mathcal{O}_{K,(p)}$ and the identity becomes easy: If we use the bases ω_i and ω'_i for $\mathcal{O}_{K,(p)}$ and $\mathfrak{b}_{(p)}$ respectively, then we may use the bases $\alpha\omega_i$ and $\alpha\omega'_i$ for $\mathfrak{a}_{(p)}$ and $\mathfrak{a}_{(p)}\mathfrak{b}_{(p)}$ and the matrix transforming the bases is the same on both sides of the equation. □

6.6. Zariski topology on Spec$\mathcal{O}_K$ (optional).

7. Nonmaximal orders (optional)

CHAPTER 2

Ideal class group and Units

1. Minkowski's bound

1.1. Minkowski's theorem. Let K be a number field of degree n with r_1 real embeddings and r_2 complex ones. Let

$$N : \mathbb{R}^n = \mathbb{R}^{r_1} \times \mathbb{C}^{r_2} \to \mathbb{R} \tag{1.1}$$

be the map

$$N(x_1, \dots, x_{r_1}, z_1, \dots, z_{r_2}) = \prod x_i \prod z_j \bar{z}_j. \tag{1.2}$$

If $\varphi : K \to \mathbb{R}^n$ is a geometric embedding, then $N(\varphi\alpha) = N_{K/\mathbb{Q}}(\alpha)$. We simply write it as $N\alpha$, the norm of α. Note that if α is in $\mathcal{O}_K$ and nonzero then $|N\alpha| \geq 1$.

Theorem 1.1. *Let K be a number field and Λ a lattice in K. Then there exists an $\alpha \in \Lambda$ such that*

$$|N\alpha| \leq \left(\frac{4}{\pi}\right)^{r_2} \frac{n!}{n^n} \sqrt{|\Delta(\Lambda)}. \tag{1.3}$$

Proof. Let U be a compact, convex, symmetric set in $\mathbb{R}^n$. Suppose $|N(x)| \leq 1$ for every $x \in U$. Let $t > 0$ be such that

$$t^n vol(U) = vol(tU) > 2^n vol(\mathbb{R}^n/\varphi(\Lambda)) = 2^n 2^{-r_2} \sqrt{|\Delta(\Lambda)}. \tag{1.4}$$

By Minkowski's lemma there exists an $0 \neq \alpha \in \Lambda$ such that $\varphi(\alpha) \in tU$, and hence $|N\alpha| \leq t^n$. Since U is compact, the intersection of tU with Λ is finite, and we may assume that t satisfies the equality

$$t^n vol(U) = 2^n 2^{-r_2} \sqrt{|\Delta(\Lambda)|}\,. \tag{1.5}$$

We conclude that there exists an $\alpha \in \Lambda$ with

$$|N\alpha| \leq 2^{r_1 + r_2} \sqrt{|\Delta(\Lambda)|} vol(U)^{-1}. \tag{1.6}$$

It is clear now that the larger we make U (subject to the restriction $|N(x)| \leq 1$ for $x \in U$), the better our estimate will be. To obtain Minkowski's bound, one needs a clever choice of U. Let

$$U = \left\{ (x_i, z_j) \in \mathbb{R}^{r_1} \times \mathbb{C}^{r_2};\ \frac{1}{n} \left(\sum |x_i| + 2 \sum |z_j| \right) \leq 1 \right\}. \tag{1.7}$$

Clearly U is compact symmetric and convex. The assertion that $|N(x)| \leq 1$ for $x \in U$ follows directly from the inequality between the geometric and the arithmetic mean. We leave it as an exercise in calculus to check that

$$vol(U) = \frac{2^{r_1 - r_2} \pi^{r_2} n^n}{n!}. \tag{1.8}$$

(*Hint:* start with the case $r_2 = 0$, then use induction on r_2.) Inserting this into the formula obtained above we get the theorem. □

1.2. Applications to the ring of integers.

COROLLARY 1.2. *(Minkowski's discriminant theorem). The discriminant of K satisfies*

$$|d_K| \geq \left(\frac{\pi}{4}\right)^{2r_2} \frac{n^{2n}}{(n!)^2}. \tag{1.9}$$

PROOF. Take $\Lambda = \mathcal{O}_K$ in the theorem and observe that we must have $|N\alpha| \geq 1$. □

COROLLARY 1.3. *The only number field with $|d_K| = 1$ is $\mathbb{Q}$. As $n \to \infty$ we must have $d_K \to \infty$.*

PROOF. The function

$$f(n) = \left(\frac{\pi}{4}\right)^n \frac{n^n}{n!} \tag{1.10}$$

is monotone increasing to infinity as

$$\frac{f(n+1)}{f(n)} = \left(\frac{\pi}{4}\right)\left(1 + \frac{1}{n}\right)^n \geq \frac{2\pi}{4} > 1. \tag{1.11}$$

□

As we observed before Hermite's theorem follows from this and from the fact that there is only a finite number of number fields with a bounded discriminant *and* degree.

2. Finiteness of the class number

2.1. Orders associated to lattices. If Λ is any lattice in K we define

$$\mathcal{O}(\Lambda) = \{x \in K;\, x\Lambda \subset \Lambda\}\,. \tag{2.1}$$

This is an *order* in K. In fact, it is a subring containing 1. For every $x \in K$, some multiple $Nx \in \mathcal{O}(\Lambda)$ so $\mathcal{O}(\Lambda)$ contains a basis. By one of the characterizations of integrality, $\mathcal{O}(\Lambda) \subset \mathcal{O}_K$. These facts are enough to show that $\mathcal{O}(\Lambda)$ is an order. We call it the *order associated to* Λ, or say that Λ *belongs* to $\mathcal{O}(\Lambda)$. Lattices belonging to the maximal order $\mathcal{O}_K$ are just fractional ideals of K. However, for a non-maximal order $\mathcal{O}$, an ideal in $\mathcal{O}$ need not necessarily belong to $\mathcal{O}$. It may belong to a larger order.

Let $\mathcal{O}$ be an order in K. Two lattices belonging to $\mathcal{O}$, Λ and Λ', ar called equivalent, if there exists an $\alpha \in K$ such that $\Lambda' = \alpha\Lambda$. If $\mathcal{O}$ is the maximal order, then the equivalence classes of lattices belonging to it form the class group.

2.2. Finiteness of the class number.

THEOREM 2.1. *Let $\mathcal{O}$ be an order. In any class of lattices belonging to $\mathcal{O}$ there exists a lattice $\Lambda \supset \mathcal{O}$ with*

$$[\Lambda : \mathcal{O}] \leq \left(\frac{4}{\pi}\right)^{r_2} \frac{n!}{n^n}\sqrt{|\Delta(\mathcal{O})|}. \tag{2.2}$$

PROOF. Let M be a lattice belonging to $\mathcal{O}$ and contained in it. Find an element $\alpha \in M$ with

$$|N\alpha| \leq \left(\frac{4}{\pi}\right)^{r_2} \frac{n!}{n^n} \sqrt{|\Delta(M)|} = [\mathcal{O} : M] \left(\frac{4}{\pi}\right)^{r_2} \frac{n!}{n^n} \sqrt{|\Delta(\mathcal{O})|}. \tag{2.3}$$

If $x \in \mathcal{O}$ then $x\alpha \in M$ so $x = \alpha^{-1} x\alpha \in \alpha^{-1}M$. It follows that $\mathcal{O} \subset \alpha^{-1}M$. Moreover

$$[\alpha^{-1}M : \mathcal{O}][\mathcal{O} : M] = [\alpha^{-1}M : M] = |N\alpha| \tag{2.4}$$

becasue the determinant of the integral matrix representing multiplication by α with respect to a basis of M is $N\alpha$. Setting $\Lambda = \alpha^{-1}M$ we get the theorem. □

Recall that the norm of an integral ideal $\mathfrak{c}$, denoted $N\mathfrak{c}$, is the index $[\mathcal{O}_K : \mathfrak{c}]$. For a prime ideal $\mathfrak{p}$ this is the number of elements in the field $\kappa(\mathfrak{p}) = \mathcal{O}_K/\mathfrak{p}$ so it is a power of the characteristic p. In general

COROLLARY 2.2. *In any class $C \in Cl_K$ there exists an integral ideal $\mathfrak{c}$ whose norm satisfies*

$$N\mathfrak{c} \leq \left(\frac{4}{\pi}\right)^{r_2} \frac{n!}{n^n} \sqrt{|d_K|}. \tag{2.5}$$

PROOF. Let $\mathfrak{a} \in C^{-1}$ be an ideal in the inverse class of C, containing R, and satisfying

$$[\mathfrak{a} : \mathcal{O}_K] \leq \left(\frac{4}{\pi}\right)^{r_2} \frac{n!}{n^n} \sqrt{|d_K|}. \tag{2.6}$$

Then $\mathfrak{c} = \mathfrak{a}^{-1}$ is an integral ideal in C and $[\mathcal{O}_K : \mathfrak{c}] = [\mathfrak{a} : \mathcal{O}_K]$. □

THEOREM 2.3. *The class group of a number field K is finite.*

PROOF. Since $N\mathfrak{c} = [\mathcal{O}_K : \mathfrak{c}]$, we only have to show that there are only finitely many ideals in $\mathcal{O}_K$ with a bounded index. This is clear because an ideal is in particular an additive subgroup, and $\mathcal{O}_K \simeq \mathbb{Z}^n$ as an additive subgroup, so it has only finitely many subgroups of bounded index. □

2.3. Examples. Minkowski's bound is a pretty effective tool in determining the class number, as well as representatives of the various classes and the full structure of the class group. Let us give two examples.

Example 1. $K = \mathbb{Q}(\sqrt{-5})$. Here $d_K = -20$, $n = 2$, $r_2 = 1$. Minkowski's bound is

$$\left(\frac{4}{\pi}\right)^{r_2} \frac{n!}{n^n} \sqrt{|d_K|} = \frac{4}{2\pi}\sqrt{20} = 2.8471 \tag{2.7}$$

Every class in Cl_K therefore contains an ideal of norm ≤ 2. If $\mathfrak{a} = \prod \mathfrak{p}_i$ then $N\mathfrak{a} = \prod N\mathfrak{p}_i$ and it follows that the only possible ideals of norm ≤ 2 are $\mathcal{O}_K$ and $\mathfrak{p}_2 = (2, 1 + \sqrt{-5})$, the unique prime factor of $(2) = \mathfrak{p}_2^2$. Since we have shown that $\mathcal{O}_K$ is not a PID, $\mathfrak{p}_2$ is not principal, and the class group is $\mathbb{Z}/2\mathbb{Z}$.

Example 2. $K = \mathbb{Q}(\sqrt{-23})$. Here $d_K = -23$, $n = 2$, $r_2 = 1$, so Minkowski's bound is $\frac{2}{\pi}\sqrt{23} = 3.0531$ and we have to consider all the ideals of norm ≤ 3. These are the ideals dividing 2 and 3. Let

$$\mathfrak{p}_2 = \left(2, \frac{1+\sqrt{-23}}{2}\right), \quad \mathfrak{p}_3 = \left(3, \frac{1+\sqrt{-23}}{2}\right). \tag{2.8}$$

One finds that $(2) = \mathfrak{p}_2\mathfrak{p}_2'$ and $(3) = \mathfrak{p}_3\mathfrak{p}_3'$ where the $'$ denotes complex conjugation. Computing norms we find that $N\mathfrak{p}_2 = 2$ and $N\mathfrak{p}_3 = 3$ hence they are prime, and so are $\mathfrak{p}_2'$ and $\mathfrak{p}_3'$. Since the equation $8 = x^2 + 23y^2$ has no solutions in integers, $\mathfrak{p}_2$ is non-principal. Since likewise $12 = x^2 + 23y^2$ has no solutions, $\mathfrak{p}_3$ is non-principal. Since $\mathfrak{p}_2\mathfrak{p}_3 = \left(\frac{1+\sqrt{-23}}{2}\right)$, $[\mathfrak{p}_2'] = [\mathfrak{p}_2]^{-1} = [\mathfrak{p}_3]$ and likewise $[\mathfrak{p}_3'] = [\mathfrak{p}_2]$. It follows that every ideal class is represented by one of the three ideals (1), $\mathfrak{p}_2$ or $\mathfrak{p}_3$. Since we already saw that $\mathfrak{p}_2$ and $\mathfrak{p}_3$ are not principal we only have to check whether they represent the same class or not. Now $\mathfrak{p}_2^2 = \left(1 + \sqrt{-23}, \frac{-3+\sqrt{-23}}{2}\right)$. If it were principal, it would be of the form $\mathfrak{p}_2^2 = \left(\frac{x+\sqrt{-23}y}{2}\right)$ and computing norms, $16 = x^2 + 23y^2$, so $x = 4$ and $y = 0$, but clearly $\mathfrak{p}_2^2 \neq (2)$. This shows that the class group must be $\mathbb{Z}/3\mathbb{Z}$, and that we must have $[\mathfrak{p}_3] = [\mathfrak{p}_2^2]$.

3. Dirichlet's unit theorem

3.1. The logarithmic embedding. To study the group of units $\mathcal{O}_K^\times$ one uses the map

$$\lambda : \mathcal{O}_K^\times \to \mathbb{R}^{r_1+r_2} \tag{3.1}$$

defined by

$$\lambda(u) = (e_i \log|\sigma_i(u)|)_{1\leq i\leq r_1+r_2} \tag{3.2}$$

where $e_i = 1$ for a real embedding, and $e_i = 2$ for a complex embedding.

Proposition 3.1. *(i) The map λ is a homomorphism.*

(ii) $\ker(\lambda) = \mu_K$ is the group of roots of unity in K. It is finite and cyclic.

(iii) $Im(\lambda)$ is discrete and contained in the hyperplane H defined by $\sum_{i=1}^{r_1+r_2} x_i = 0$.

Lemma 3.2. *An element $u \in \mathcal{O}_K$ is a unit if and only if $N_{K/\mathbb{Q}}(u) = \pm 1$.*

Proof. If $uv = 1$ for $v \in \mathcal{O}_K$ then $N(u)N(v) = 1$ but $N(u)$ and $N(v)$ are in $\mathbb{Z}$. Conversely, if $N(u) = \pm 1$ then u^{-1} is an algebraic integer, because up to a sign it is a product of conjugates of u, so it belongs to $\mathcal{O}_K$. □

Proof. (of the proposition). The map λ is clearly a group homomorphism. Its kernel μ_K is therefore a subgroup of $\mathcal{O}_K^\times$. If φ is the geometric embedding of $\mathcal{O}_K$ in $\mathbb{R}^n$ then $\varphi(\mu_K)$ is bounded because all the coordinates are bounded by 1. It follows from the discreteness of $\varphi(\mathcal{O}_K)$ that μ_K is finite. But every finite subgroup of a field is a cyclic group of roots of unity. Conversely, if ζ is a root of unity, then $\lambda(\zeta) = 0$ because $\mathbb{R}^{r_1+r_2}$ has no torsion. The fact that $Im(\lambda)$ is contained in the hyperplane as indicated follows at once from the lemma. It remains to show that $Im(\lambda)$ is discrete. Let $M > 0$ and consider the box B defined by $|x_i| \leq M$. If $\lambda(u) \in B$ then

$$|\sigma_i(u)| \leq e^{M/e_i} \tag{3.3}$$

so $\varphi(u)$ also lies in a bounded domain, and there are only finitely many such u. □

A discrete subgroup of a real vector space H of dimension $r = r_1 + r_2 - 1$ is isomorphic to $\mathbb{Z}^t$ for some $t \leq r$. It is of rank r if and only if it is also cocompact, and it is then a lattice in H.

3.2. Dirichlet's unit theorem.

THEOREM 3.3. *The image of λ is a lattice in H. Consequently*

$$\mathcal{O}_K^\times \simeq \mu_K \times \mathbb{Z}^r \tag{3.4}$$

where $r = r_1 + r_2 - 1$.

PROOF. We shall find $u_1, \dots, u_{r_1+r_2}$ such that if $\lambda(u_i) = (x_{i,1}, \dots, x_{i,r_1+r_2})$ then $x_{i,j} < 0$ for $j \neq i$ and (necessarily) $x_{i,i} > 0$. Note that this means that u_i is small in all the $r_1 + r_2$ embeddings except σ_i and large in σ_i. We claim that r of the vectors $\lambda(u_i)$ must then be linearly independent. Indeed, suppose the rank of the matrix $(x_{i,j})$ was less than r. Then a linear dependence between the first r columns

$$\sum_{j=1}^{r} c_j x_{i,j} = 0 \tag{3.5}$$

$(1 \leq i \leq r_1 + r_2)$ must exist. Assume that $|c_j| \leq c_{j_0}$ for all j. Looking at the $i = j_0$ row, $x_{j_0,j_0} = -\sum_{j \neq j_0} x_{j_0,j} \geq -\sum_{j \neq j_0, j \leq r} x_{j_0,j}$ so

$$\begin{aligned} c_{j_0} x_{j_0,j_0} &\geq -c_{j_0} \sum_{j \neq j_0, j \leq r} x_{j_0,j} \\ &\geq -\sum_{j \neq j_0, j \leq r} c_j x_{j_0,j} \end{aligned} \tag{3.6}$$

with a strict inequality unless all the c_j are equal to each other (and $x_{j_0,r+1} = 0$). But this contradicts the equation $\sum_{j=1}^{r} c_j x_{i,j} = 0$ for $i = r_1 + r_2$ because all the entries have the same sign.

It is therefore enough to find a unit which is small in all the embeddings except one, say the first one (hence necessarily large in the first one). Let $\varepsilon > 0$ be small, M a fixed positive number, and consider the rectangle

$$[-M\varepsilon^{1-n}, M\varepsilon^{1-n}] \times [-\varepsilon, \varepsilon]^{n-1} \subset \mathbb{R}^n \tag{3.7}$$

whose volume is $2^n M$. If M is large enough (depending on n and d_K) this rectangle will contain non-zero points from $\mathcal{O}_K$, by Minkowski's lemma. These points will have norm which is a rational integer bounded in terms of M. Shrinking ε we get infinitely many such points $\alpha_i \in \mathcal{O}_K$, with first coordinate of $\varphi(\alpha_i)$ tending to ∞ and all the others tending to 0. Since their norms are all integers bounded by a fixed number, passing to a subsequence we may assume that $N\alpha_i$ are all equal. Consider now the ideal factorization of (α_i). There are only finitely many integral ideals of $\mathcal{O}_K$ of a given norm, so there are only finitely many possible ideal factorizations. Passing again to a subsequence we may assume that the ideals (α_i) are all equal. This means that $\alpha_j \alpha_i^{-1}$ is a *unit* for all $i < j$. But if j is large enough we can make the first coordinate of $\alpha_j \alpha_i^{-1}$ arbitrarily large and all the others arbitrarily small, as desired. □

3.3. The regulator. Just like $\sqrt{|d_K|}$ was (up to a factor of 2^{r_2}) the volume of the fundamental parallelopiped of the lattice $\varphi(\mathcal{O}_K)$ in $\mathbb{R}^n$, we get another invariant, called the *regulator* R_K of K, by considering the volume of a fundamental parallelopiped for the lattice $\lambda(\mathcal{O}_K^\times)$ in $H \subset \mathbb{R}^{r_1+r_2}$.

Definition 3.1. *Let $u_1, \ldots, u_r$ be r units in $\mathcal{O}_K$ such that $\lambda(u_i)$ are linearly independent (equivalently, such that the group generated by them is of finite index in $\mathcal{O}_K^\times$). Such a set is called a* fundamental set of units. *Define their regulator,* $\mathrm{Reg}(u_1, \ldots, u_r)$ *by*

$$\mathrm{Reg}(u_1, \ldots, u_r) = \left|\det(e_j \log|\sigma_j(u_i)|)_{1\le i,j\le r}\right|. \tag{3.8}$$

The regulator of $\mathcal{O}_K$, denoted R_K, is the regulator of a basis over $\mathbb{Z}$ for $\mathcal{O}_K^\times$ modulo μ_K.

We chose to ignore the last column in the $r \times (r+1)$ matrix whose rows are the $\lambda(u_i)$, but since the sum of the columns is 0 (the logarithmic embedding falls inside the hyperplane H) we could have omitted any other column without affecting the result.

Note that if $u_1, \ldots, u_r$ is a fundamental set of units then

$$Reg(u_1, \ldots, u_r)/R_K = [\mathcal{O}_K^\times : \mu_K \langle u_1, \ldots, u_r\rangle].$$

Exercise 3.1. *Let $\mathcal{O}$ be any order in K. Prove that its units, $\mathcal{O}^\times$, make up a group of finte index in $\mathcal{O}_K^\times$.*

4. Class groups and units in quadratic fields

4.1. Quadratic fields. The simplest examples to consider, besides $\mathbb{Q}$, are the quadratic number fields K, those for which $n = [K : \mathbb{Q}] = 2$. Such a K is either *real* ($r_1 = 2$, $r_2 = 0$) or *imaginary* ($r_1 = 0$, $r_2 = 1$). They are important not only as examples of the general theory but also for the study of binary quadratic forms over $\mathbb{Q}$. We have seen that if we write

$$K = \mathbb{Q}(\sqrt{D}) \tag{4.1}$$

where D is a square free integer, then

$$\mathcal{O}_K = \mathbb{Z} + \mathbb{Z}\omega_D \tag{4.2}$$

where $\omega_D = \sqrt{D}$ if $D \equiv 2, 3 mod 4$ and $\omega_D = (1 + \sqrt{D})/2$ if $D \equiv 1 mod 4$.

Exercise 4.1. *Show that every order in K is of the form*

$$\mathcal{O} = \mathbb{Z} + \mathbb{Z}f\omega_D \tag{4.3}$$

for a unique $f = 1, 2, 3, \ldots$

The class groups of quadratic imaginary fields are the subject of much research, begun by Gauss in the context of definite binary quadratic forms. When $D < 0$ and $|D| \to \infty$ then also $h_K \to \infty$. Gauss found the quadratic imaginary fields with $h_K = 1$ (i.e. such that $\mathcal{O}_K$ is a PID). There are nine such fields, the last one with $d_K = -163$. It was an open problem for over a century whether these are *all*, and it was finally settled by Heegner that there is no tenth quadratic imaginary field with $h_K = 1$. Heegner got the credit for his proof only late, not before Stark re-proved it, apparently without seeing Heegner's older work, which was written in German.

Class groups of real quadratic fields are connected with the classification of indefinite binary quadratic forms, and it is still a major open problem if there are infinitely many real quadratic fields of class number one.

The story of units is much easier. If K is imaginary, then there are only finitely many units, namely the roots of unity, and it is easy to see that they are ±1, except if $K = \mathbb{Q}(i)$ where there are 4, or $K = \mathbb{Q}(\rho)$, $\rho^2 + \rho + 1 = 1$, where there are 6.

EXERCISE 4.2. *Prove this! Hint: consider the equation $N_{K/\mathbb{Q}}(x + y\omega_D) = 1$ and find all its solutions in $x, y \in \mathbb{Z}$.*

If K is real quadratic then Dirichlet's unit theorem says that

$$\mathcal{O}_K^\times = \{\pm 1\} \langle \varepsilon \rangle , \tag{4.4}$$

where ε is uniquely determined if we insist that it is > 1. This ε is called the *fundamental unit* of K. If $D \equiv 2, 3 mod 4$ then ε is the smallest number $x + \sqrt{D}y > 1$ solving

$$x^2 - Dy^2 = \pm 1. \tag{4.5}$$

The equation $x^2 - Dy^2 = 1$ is known as Pell's equation. Note that if there is an ε of norm -1, a fundamental solution to Pell's equation (a smallest solution > 1) is the *square* of a fundamental unit, but if all units have norm $+1$, then it is a fundamental unit. In any case, the theory tells us that all solutions to Pell's equation are gotten from a fundamental solution by raising to some integral power, and adding a sign.

If $D \equiv 1 mod 4$ (square free and positive) then still the solutions to $x^2 - Dy^2 = \pm 1$ give us all the units in the order $\mathbb{Z}[\sqrt{D}]$, but there may be more units in $\mathcal{O}_K$, namely the solutions of $N_{K/\mathbb{Q}}(x + y\omega_D) = \pm 1$, which is the equation

$$x^2 + xy + \left(\frac{1-D}{4}\right) y^2 = \pm 1. \tag{4.6}$$

EXERCISE 4.3. *Prove that if $D \equiv 3 mod 8$ then all the units in K have norm 1.*

Here are some examples. The fundamental unit of $\mathbb{Q}(\sqrt{2})$ is $1 + \sqrt{2}$. The fundamental unit of $\mathbb{Q}(\sqrt{19})$ is $170 + 39\sqrt{19}$, whose norm is 1 (see the previous exercise) and the fundamental unit of $\mathbb{Z}[\sqrt{61}]$ is

$$1766319049 + 226153980\sqrt{61}. \tag{4.7}$$

This solution to Pell's equation with $D = 61$ was obtained by Fermat, who never revealed his method. Today we know, thanks to Lagrange, a relatively easy algorithm to obtain the fundamental solution using continued fractions.

4.2. Pell's equation and continued fractions.

5. Binary quadratic forms (optional)

CHAPTER 3

Extensions of number fields and Hilbert's theory of ramification

1. The decomposition of primes

1.1. The degree formula. Let L/K be a finite extension of number fields and $\mathfrak{p}$ a prime of $\mathcal{O}_K$. The ideal $\mathfrak{p}\mathcal{O}_L$ need not be prime, but it decomposes as

$$\mathfrak{p}\mathcal{O}_L = \mathfrak{P}_1^{e_1} \dots \mathfrak{P}_g^{e_g}. \tag{1.1}$$

The primes $\mathfrak{P}_i$ are the primes *dividing* (or sitting above, or containing) $\mathfrak{p}$. Each prime $\mathfrak{P}$ of $\mathcal{O}_L$ divides a unique prime of $\mathcal{O}_K$, namely $\mathfrak{P} \cap \mathcal{O}_K = \mathfrak{p}$, so by going over all the primes of K and decomposing them in L, we list all the primes of L. This is particular useful when $K = \mathbb{Q}$.

The residue field $\kappa(\mathfrak{P}) = \mathcal{O}_L/\mathfrak{P}$ is a finite extension of $\kappa(\mathfrak{p}) = \mathcal{O}_K/\mathfrak{p}$ (because $\mathcal{O}_L$ is a finite $\mathcal{O}_K$ module) and the relative degree

$$f(\mathfrak{P}/\mathfrak{p}) = [\kappa(\mathfrak{P}) : \kappa(\mathfrak{p})] \tag{1.2}$$

is called the *inertial degree* of $\mathfrak{P}$ over $\mathfrak{p}$. The exponent e in which $\mathfrak{P}$ appears in the decomposition of $\mathfrak{p}$ is called the *ramification degree* of $\mathfrak{P}$ over $\mathfrak{p}$, and it is denoted by $e(\mathfrak{P}/\mathfrak{p})$. When $K = \mathbb{Q}$, we denote these by $f(\mathfrak{P})$ and $e(\mathfrak{p})$ and call them the *absolute* residual degree or ramification degree.

THEOREM 1.1. *Let $\mathfrak{p}$ decompose as above. With the obvious notation,*

$$\sum_{i=1}^{g} e_i f_i = [L : K]. \tag{1.3}$$

PROOF. We give the proof when $K = \mathbb{Q}$ and $\mathfrak{p} = (p)$, and then show how to modify it in the relative case. From the Chinese Remainder Theorem we know that

$$\mathcal{O}_L/p\mathcal{O}_L \simeq \prod \mathcal{O}_L/\mathfrak{P}_i^{e_i}. \tag{1.4}$$

The idea is to compare dimensions over $\mathbb{F}_p$. Since $\mathcal{O}_L$ is a free $\mathbb{Z}$-module of rank n, the dimension on the left is n. The dimension of $\mathcal{O}_L/\mathfrak{P}_i$ is f_i by definition. The proof will be complete if we show that $\mathfrak{P}^m/\mathfrak{P}^{m+1}$ is a 1-dimensional vector space over $\mathcal{O}_L/\mathfrak{P}$. Indeed, if so, then $\dim_{\mathbb{F}_p} \mathfrak{P}_i^m/\mathfrak{P}_i^{m+1} = f_i$ and looking at the filtration

$$\mathcal{O}_L/\mathfrak{P}_i^{e_i} \supset \mathfrak{P}_i/\mathfrak{P}_i^{e_i} \supset \dots \supset \mathfrak{P}_i^{e_i-1}/\mathfrak{P}_i^{e_i} \supset 0 \tag{1.5}$$

by subspaces, we see that

$$\dim_{\mathbb{F}_p} \mathcal{O}_L/\mathfrak{P}_i^{e_i} = \sum_{m=0}^{e_i-1} \mathfrak{P}_i^m/\mathfrak{P}_i^{m+1} = e_i f_i. \tag{1.6}$$

Let $a \in \mathfrak{P}^m - \mathfrak{P}^{m+1}$. Then

$$(a) + \mathfrak{P}^{m+1} = \gcd((a), \mathfrak{P}^{m+1}) = \mathfrak{P}^m. \tag{1.7}$$

it follows that $\mathfrak{P}^m/\mathfrak{P}^{m+1}$ is generated as an $\mathcal{O}_L$ module by a single element a, as was to be shown. □

The same proof works in the relative case, except that we cannot use the fact that $\mathcal{O}_L$ is free over $\mathcal{O}_K$, because $\mathcal{O}_K$ need not be a PID. However, we are allowed to localize at $\mathfrak{p}$ (namely, at the multiplicative set $S = \mathcal{O}_K - \mathfrak{p}$) because the decomposition of $\mathfrak{p}$ in $\mathcal{O}_L$ is the same as the decomposition of $\mathfrak{p}\mathcal{O}_{K,\mathfrak{p}}$ in $\mathcal{O}_{L,\mathfrak{p}}$. Now $\mathcal{O}_{K,\mathfrak{p}}$ is a DVR, and any torsion-free finitely generated module over it is free, so we can argue as before.

THEOREM 1.2. *Suppose $K = \mathbb{Q}(\alpha)$ for an algebraic integer α and f is the monic irreducible polynomial. If p does not divide $d_K^{-1}\Delta(1, \alpha, \ldots, \alpha^{n-1})$, $\bar{f} = fmodp \in \mathbb{F}_p[X]$ and*

$$\bar{f} = \prod h_i^{e_i} \tag{1.8}$$

is its prime decomposition, $\deg(h_i) = f_i$, *then*

$$(p) = \prod \mathfrak{p}_i^{e_i} \tag{1.9}$$

and the inertial degree of $\mathfrak{p}_i$ is f_i.

PROOF. Assume first that $\mathcal{O}_K = \mathbb{Z}[\alpha]$, so $d_K = \Delta(1, \alpha, \ldots, \alpha^{n-1})$. On the one hand

$$\mathcal{O}_K/p\mathcal{O}_K \simeq \prod \mathcal{O}_K/\mathfrak{p}_i^{e_i'}. \tag{1.10}$$

On the other hand

$$\mathcal{O}_K/p\mathcal{O}_K \simeq \mathbb{Z}[X]/(f,p) \simeq \mathbb{F}_p[X]/(\bar{f}) \simeq \prod \mathbb{F}_p[X]/(h_i)^{e_i}. \tag{1.11}$$

Both decompositions are as direct sums of rings that can not be further decomposed as direct sums (because they are local - a local ring is not a product of two subrings). Such decompositions are unique up to ordering, because they correspond 1:1 to decompositions

$$1 = \sum \varepsilon_i \tag{1.12}$$

of 1 as a sum of mutually orthogonal minimal idempotents ($\varepsilon_i^2 = \varepsilon_i$, $\varepsilon_i\varepsilon_j = 0$ if $i \neq j$, and ε_i is not the sum of two mutually orthogonal idempotents). [Given a decomposition $R = \prod R_i$ let ε_i be the unit of R_i. Given a decomposition $1 = \sum \varepsilon_i$ let $R_i = R\varepsilon_i$.] It is an easy exercise to show that such a decomposition of 1 is unique. It follows that we must have the same number of factors and we may assume

$$\mathcal{O}_K/\mathfrak{p}_i^{e_i'} \simeq \mathbb{F}_p[X]/(h_i)^{e_i}. \tag{1.13}$$

This is a local Artinian ring and $e_i = e_i'$ because it is intrinsically characterized as the first power of the maximal ideal that vanishes. Finally $[\kappa(\mathfrak{p}_i) : \mathbb{F}_p] = \deg(h_i)$ because comparing the residue fields of the two isomorphic local rings

$$\kappa(\mathfrak{p}_i) \simeq \mathbb{F}_p[X]/(h_i). \tag{1.14}$$

The general case follows by localization at (p) because if p does not divide the index $[\mathcal{O}_K : \mathbb{Z}[\alpha]]$ and we localize at (p) the two $\mathbb{Z}_{(p)}$-modules become equal. □

1.2. The meaning of ramification. We have seen that $\mathfrak{p}$ gives a discrete valuation $v_\mathfrak{p}$ on K. Likewise, if $\mathfrak{P}|\mathfrak{p}$ it gives a discrete valuation $v_\mathfrak{P}$ on L. From the very definition of these valuations it is clear that for $a \in K$

$$v_\mathfrak{P}(a) = e(\mathfrak{P}/\mathfrak{p})v_\mathfrak{p}(a). \tag{1.15}$$

Thus $v = e(\mathfrak{P}/\mathfrak{p})^{-1}v_\mathfrak{P}$ extends the valuation $v_\mathfrak{p}$ from K to L. Since the valuations $v_\mathfrak{p}$ and $v_\mathfrak{P}$ are normalized, this means that the ramification degree $e(\mathfrak{P}/\mathfrak{p})$ is the index of the value group $[v(L^\times) : v(K^\times)]$.

The prime $\mathfrak{P}$ is *ramified* in the extension L/K if $e(\mathfrak{P}/\mathfrak{p}) > 1$, and *unramified* if it is not ramified. Being unramified is equivalent to the assertion that a uniformizer at $v_\mathfrak{p}$ remains a uniformizer at $v_\mathfrak{P}$.

We say that $\mathfrak{p}$ is *ramified* in L if at least one of the primes $\mathfrak{P}$ above it is ramified.

1.3. Ramification and the discriminant.

THEOREM 1.3. *Let K be a number field. A rational prime p ramifies in K if and only if it divides d_K. In particular, only finitely many primes ramify.*

PROOF. We use the ring structure on $R = \mathcal{O}_K/p\mathcal{O}_K$. Let $Tr(r)$, for $r \in R$, be the trace of the linear transformation (over $\mathbb{F}_p$) "multiplication by r". Thus $Tr(r) \in \mathbb{F}_p$ and if r is nilpotent, $Tr(r) = 0$. Let $\bar{\omega}_1, \dots, \bar{\omega}_n$ be a basis of R over $\mathbb{F}_p$ and define its discriminant

$$\Delta(\bar{\omega}_1, \dots, \bar{\omega}_n) = \det\left(Tr(\bar{\omega}_i\bar{\omega}_j)\right). \tag{1.16}$$

The discriminants of two different bases of R over $\mathbb{F}_p$ differ by the square of a non-zero element in $\mathbb{F}_p$. In particular, if the discriminant vanishes for one basis, it always vanishes. If $\{\bar{\omega}_i\}$ is obtained by reduction modulo p of a basis $\{\omega_i\}$ of $\mathcal{O}_K$ over $\mathbb{Z}$, then clearly

$$\Delta(\bar{\omega}_1, \dots, \bar{\omega}_n) = \Delta(\omega_1, \dots, \omega_n) modp. \tag{1.17}$$

Observe that p ramifies in K if and only if R has nilpotent elements. Suppose first that p ramifies. Let $\bar{\omega}_1, \dots, \bar{\omega}_n$ be a basis of R over $\mathbb{F}_p$ such that $\bar{\omega}_1 \in R$ is nilpotent. Then $\bar{\omega}_1\bar{\omega}_j$ are nilpotent and $Tr(\bar{\omega}_1\bar{\omega}_j) = 0$. A whole row in the matrix whose determinant is $\Delta(\bar{\omega}_1, \dots, \bar{\omega}_n)$ then vanishes, so $\Delta(\bar{\omega}_1, \dots, \bar{\omega}_n) = 0$. Now start with any basis $\{\omega_i\}$ of $\mathcal{O}_K$ over $\mathbb{Z}$. Reducing $\Delta(\omega_1, \dots, \omega_n)$ modulo p we must get 0, so p divides d_K.

Conversely, if R has no nilpotents it is a product of fields $\kappa_l = \mathcal{O}_K/\mathfrak{p}_l$ and we may choose a basis $\bar{\omega}_i$ of R which is the union of bases of the κ_l over $\mathbb{F}_p$. An easy computation shows then that $Tr(\bar{\omega}_i\bar{\omega}_j) = 0$ unless $\bar{\omega}_i$ and $\bar{\omega}_j$ both lie in the same field κ_l, and then this trace is $Tr_{\kappa_l/\mathbb{F}_p}(\bar{\omega}_i\bar{\omega}_j)$. From the non-degeneracy of the trace pairing in the separable extensions $\kappa_l/\mathbb{F}_p$ (see Appendix) we deduce that $\Delta(\bar{\omega}_1, \dots, \bar{\omega}_n) \neq 0$. Now this result remains valid for *any basis* $\bar{\omega}_i$ of R over $\mathbb{F}_p$, not necessarily of the type chosen above. In particular we may start with a basis of $\mathcal{O}_K$ over $\mathbb{Z}$ and reduce it modulo p. But then p does not divide d_K. □

COROLLARY 1.4. *Only finitely many primes ramify in K. In every number field at least one prime ramifies.*

The last assertion follows from $d_K > 1$. In contrast, a relative extension L/K can well be unramified: all the primes of K are unramified in L. We shall see momentarily examples where K is a quadratic field.

1.4. Multiplicativity in towers. Let $K \subset L \subset M$ be number fields. Let $\mathfrak{p} \subset P \subset \mathfrak{P}$ be primes in the respective rings of integers.

Proposition 1.5. *$e(\mathfrak{P}/\mathfrak{p}) = e(\mathfrak{P}/P)e(P/\mathfrak{p})$ and $f(\mathfrak{P}/\mathfrak{p}) = f(\mathfrak{P}/P)f(P/\mathfrak{p})$.*

Proof. For f this follows from the multiplicativity in towers of field extension degrees. For e this is clear if we simply extend the ideal $\mathfrak{p}$ to M in two steps: first to L and then from L to M. □

We can now give an example of an unramified field extension. Let $F = \mathbb{Q}(i)$ and $K = \mathbb{Q}(\sqrt{5})$. Then $d_F = -4$ and $d_K = 5$ so only 2 ramifies in F and only 5 ramifies in K. Consider $M = KF$ and $L = \mathbb{Q}(\sqrt{-5})$ which is the third quadratic field contained in M. Note that $d_L = -20$, so both 2 and 5 ramify in L. We claim that M/L is everywhere unramified.

Lemma 1.6. *The discriminant of M divides $4^2 5^2$.*

Proof. Let ω_1, ω_2 be a basis over $\mathbb{Z}$ of $\mathcal{O}_F$ and η_1, η_2 a basis over $\mathbb{Z}$ of $\mathcal{O}_K$. Then $\omega_i \eta_j$ are four linearly independent elements of $\mathcal{O}_M$. Since traces are sums of Galois conjugates,

$$Tr_{M/\mathbb{Q}}(\omega_i \eta_j \omega_{i'} \eta_{j'}) = Tr_{F/\mathbb{Q}}(\omega_i \omega_{i'}) Tr_{K/\mathbb{Q}}(\eta_j \eta_{j'}) \tag{1.18}$$

and computing determinants we get that

$$\Delta(\omega_i \eta_j) = d_F^{[K:\mathbb{Q}]} d_K^{[F:\mathbb{Q}]} = 4^2 5^2. \tag{1.19}$$

However, d_M divides this quantity. □

We conclude that any prime other than 2 or 5 is unramified in M, so a fortiori, primes above it are unramified in M/L. Primes above 2 are unramified in K, so their index of ramification in M is at most 2, but is also at least 2, since they ramify in F, so it must be exactly 2. Similarly for primes above 5. Since these primes ramify (with index 2) in L, the primes above them are unramified in M/L.

1.5. Different and the discriminant. Consider the set

$$\mathcal{O}'_K = \left\{ a \in K;\, Tr_{K/\mathbb{Q}}(ab) \in \mathbb{Z} \text{ for all } b \in \mathcal{O}_K \right\}. \tag{1.20}$$

This is clearly a fractional ideal of K : it is a sub-$\mathcal{O}_K$-module containing $\mathcal{O}_K$, and since $B(x, y) = Tr(xy)$ is a non-degenerate pairing on K (as a vector space over $\mathbb{Q}$), it is a lattice. Its *inverse*

$$\mathcal{D}_K = (\mathcal{O}'_K)^{-1} \tag{1.21}$$

is an integral ideal called the *different* of K.

If $\omega_1, \dots, \omega_n$ is a basis of $\mathcal{O}_K$ and $\omega'_1, \dots, \omega'_n$ is the dual basis w.r.t. B, then it spans $\mathcal{O}'_K$. If we write

$$\omega_i = \sum m_{ij} \omega'_j \tag{1.22}$$

then $B(\omega_i, \omega_j) = m_{ij}$ and therefore $d_K = \det(m_{ij}) = [\mathcal{O}'_K : \mathcal{O}_K] = [\mathcal{O}_K : \mathcal{D}_K] = N\mathcal{D}_K$.

We have proven the following lemma.

Lemma 1.7. *The different is an ideal of K whose norm is the discriminant.*

We can now refine the theorem saying that p ramifies if and only if it divides the discriminant to deal with primes of K. We shall state it without a proof.

PROPOSITION 1.8. *A prime $\mathfrak{p}$ is ramified in K if and only if it divides the different $\mathcal{D}_K$.*

2. Relative norm, different and discriminant

2.1. Relative norm. The relative norm of an ideal $\mathfrak{A}$ of L is defined as follows. First, if $\mathcal{O}_K$ is a PID, choose bases ω_i and η_i of $\mathcal{O}_L$ and of $\mathfrak{A}$ as free $\mathcal{O}_K$ modules of rank n and write

$$\eta_i = \sum a_{ij}\omega_j \tag{2.1}$$

with $a_{ij} \in K$. The matrix (a_{ij}) is well-defined up to multiplication on both sides by matrices from $GL_n(\mathcal{O}_K)$, corresponding to changing the bases. Its determinant therefore gives a well-defined ideal in $\mathcal{O}_K$ which we call the norm of $\mathfrak{A}$

$$N_{L/K}\mathfrak{A} = (\det(a_{ij})). \tag{2.2}$$

In general, localize at a prime $\mathfrak{p}$ of K. Then $\mathcal{O}_{K,\mathfrak{p}}$ is a PID and we apply the same procedure to define $N_{L/K}\mathfrak{A}_\mathfrak{p}$ (note that we localize $\mathcal{O}_L$ too at the multiplicative set $\mathcal{O}_K - \mathfrak{p}$, and then it becomes a free $\mathcal{O}_{K,\mathfrak{p}}$ module of rank n). We then define $N_{L/K}\mathfrak{A}$ as the unique ideal of K whose localization at $\mathfrak{p}$ is $N_{L/K}\mathfrak{A}_\mathfrak{p}$, i.e.

$$(N_{L/K}\mathfrak{A})_\mathfrak{p} = N_{L/K}(\mathfrak{A}_\mathfrak{p}) \tag{2.3}$$

for all $\mathfrak{p}$. (Recall that such an ideal is simply the intersection of all its localizations inside L). The properties of the norm, which were proven for the absolute norm in detail, hold for the relative norm as well, with the obvious modifications in the proof, always using localization to be able to work over a PID. We claim that

$$N_{L/K}\mathfrak{P} = \mathfrak{p}^f \tag{2.4}$$

where $f = f(\mathfrak{P}/\mathfrak{p})$. Note that this agrees with the absolute norm (given by the number of elements in $\kappa(\mathfrak{P})$) in case $K = \mathbb{Q}$, $\mathfrak{p} = (p)$. Indeed, localizing we may assume that $\mathcal{O}_K$ is a PID, so $\mathfrak{p} = (\pi)$. Since

$$\mathcal{O}_L/\mathfrak{P} \simeq (\mathcal{O}_K/\mathfrak{p})^f \tag{2.5}$$

as an $\mathcal{O}_K$ module, we can choose the bases ω_i and η_i so that

$$\eta_i = \pi\omega_i \tag{2.6}$$

for $1 \le i \le f$ and $\eta_i = \omega_i$ for $f < i \le n$. Here we have used the theorem on elementary divisors (structure of finitely generated modules over a PID). The claim then becomes obvious.

PROPOSITION 2.1. *$N_{L/K}\mathfrak{A}$ is the fractional ideal of K generated by $N_{L/K}(a)$ for all $a \in \mathfrak{A}$.*

PROOF. We may localize on the base, replacing $\mathcal{O}_K$ by $\mathcal{O}_{K,\mathfrak{p}}$. The Dedekind ring $\mathcal{O}_{L,\mathfrak{p}}$ is then semi-local, hence a PID. But then $\mathfrak{A}_\mathfrak{p} = (a)$ and the norm of a principal ideal is, be definition, the ideal generated by the norm of its generator, since we can take for bases $\eta_i = a\omega_i$. □

2.2. Relative discriminant and different. The discriminant of a K-basis ω_i of L is defined as before, as

$$\Delta(\omega_1, \dots, \omega_n) = \det(Tr_{L/K}(\omega_i \omega_j)). \tag{2.7}$$

If $\mathcal{O}_K$ is a PID, let ω_i be a basis of $\mathcal{O}_L$ over $\mathcal{O}_K$, and define the *discriminant ideal* $d_{L/K}$ as the ideal of K generated by the discriminant of a that basis. The definition is independent of the basis. In fact, we can do better: we can define the discriminant as an element of $\mathcal{O}_K$ modulo $(\mathcal{O}_K^\times)^2$. In general use localization. The proof of the following theorem is the same as in the absolute case.

THEOREM 2.2. *A prime ideal* $\mathfrak{p}$ *of* K *is ramified in* L *if and only if it divides* $d_{L/K}$.

We define the *relative different* $\mathcal{D}_{L/K}$ as the inverse of the fractional ideal in L

$$\left\{a \in L;\, Tr_{L/K}(ab) \in \mathcal{O}_K \text{ for al } b \in \mathcal{O}_L\right\}. \tag{2.8}$$

As before, this is an integral ideal of L whose norm to K is $d_{L/K}$. An ideal of L is ramified in L/K if and only if it divides the different.

The different is easy to compute if $\mathcal{O}_L = \mathcal{O}_K[\alpha]$ (namely, it is free, and admits a basis over $\mathcal{O}_K$ consisting of powers of an algebraic integer α). More generally, given any algebraic integer α of degree n we may compute the different except above the primes "supporting" the module $\mathcal{O}_L/\mathcal{O}_K[\alpha]$.

THEOREM 2.3. *Suppose* $\mathfrak{p}$ *is a prime of* K *and* $\mathcal{O}_{L,\mathfrak{p}} = \mathcal{O}_{K,\mathfrak{p}}[\alpha]$. *Let* f *be the monic irreducible polynomial of* α. *Then*

$$\mathcal{D}_{L/K,\mathfrak{p}} = (f'(\alpha))\mathcal{O}_{L,\mathfrak{p}}. \tag{2.9}$$

PROOF. Let

$$\frac{f(X)}{X-\alpha} = \sum_{i=0}^{n-1} b_i X^i. \tag{2.10}$$

We show that the dual basis of $1, \alpha, \dots, \alpha^{n-1}$ w.r.t. the trace pairing is $\frac{b_i}{f'(\alpha)}$. We have to show

$$Tr_{L/K}\left(\frac{b_i \alpha^j}{f'(\alpha)}\right) = \delta_{ij}. \tag{2.11}$$

Multiplying by X^i and summing we have to show

$$Tr_{L/K}\left(\frac{f(X)\alpha^j}{(X-\alpha)f'(\alpha)}\right) = X^j \tag{2.12}$$

$(0 \le j \le n-1)$. However, both sides of the equation are polynomials of degree $\le n-1$ having n distinct roots (all the conjugates of α), so they must coincide.

It follows that the inverse of $\mathcal{D}_{L/K,\mathfrak{p}}$ (the dual of $\mathcal{O}_{L,\mathfrak{p}}$) is spanned over $\mathcal{O}_{K,\mathfrak{p}}$ by $\frac{b_i}{f'(\alpha)}$. Solving successively for the b_i we get that the $\mathcal{O}_{K,\mathfrak{p}}$-span of the b_i is the same as the $\mathcal{O}_{K,\mathfrak{p}}$-span of the α^i, and is therefore $\mathcal{O}_{L,\mathfrak{p}}$. The theorem follows. □

3. Galois extensions

3.1. The decomposition group. Assume from now on that L/K is Galois with $Gal(L/K) = G$. Let $\sigma \in G$. Since σ induces an automorphism of $\mathcal{O}_L$, it carries any prime $\mathfrak{P}$ to a prime $\sigma(\mathfrak{P})$, and $\sigma(\mathfrak{P}) \cap \mathcal{O}_K = \sigma(\mathfrak{P} \cap \mathcal{O}_K) = \mathfrak{P} \cap \mathcal{O}_K$, so σ induces a permutation of the primes dividing any given $\mathfrak{p}$ of K.

LEMMA 3.1. *G acts transitively on the primes dividing a given prime $\mathfrak{p}$.*

PROOF. Let $\mathfrak{P}$ be a prime dividing $\mathfrak{p}$ and assume that $\mathfrak{Q}$ is another prime dividing $\mathfrak{p}$ which is not in the G-orbit of $\mathfrak{P}$. By the Chinese Remainder Theorem we may choose $a \in \mathcal{O}_L$ such that $a \equiv 1 mod \sigma(\mathfrak{P})$ for every $\sigma \in G$, but $a \equiv 0 mod \mathfrak{Q}$. The norm of a is then in $\mathfrak{Q} \cap \mathcal{O}_K = \mathfrak{p}$, but we also have

$$\sigma^{-1}(a) \equiv 1 mod \mathfrak{P} \tag{3.1}$$

for all σ, so $N_{L/K}(a) \equiv 1 mod \mathfrak{P}$. This is a contradiction. □

DEFINITION 3.1. *The* decomposition group *$G_{\mathfrak{P}}$ of $\mathfrak{P}$ is its stabilizer in G.*

PROPOSITION 3.2. *(i) $G_{\tau(\mathfrak{P})} = \tau G_{\mathfrak{P}} \tau^{-1}$ (ii) all the primes above $\mathfrak{p}$ have the same e and f, and their number is $g = \frac{n}{ef}$ (iii) the order of $G_{\mathfrak{P}}$ is ef.*

PROOF. G acts on the prime factors of $\mathfrak{p}$ transitively. Looking at its action on the decomposition of $\mathfrak{p}\mathcal{O}_L$ we get part (ii), noting also that conjugate primes should have the same f since σ induces and isomorphism of $\kappa(\mathfrak{P})$ onto $\kappa(\sigma(\mathfrak{P}))$. Part (i) is clear, and part (iii) follows from the fact that the order of the orbit should be the index of the stabilizer. □

We caution that different $\mathfrak{P}$ above the same $\mathfrak{p}$ will have conjugate decomposition groups, but they need not be equal. However, if $G_{\mathfrak{P}}$ is normal for any reason, for example, if G is abelian, then all the decomposition groups are equal, so they depend only on $\mathfrak{p}$.

DEFINITION 3.2. *The* decomposition field *of $\mathfrak{P}$ is the fixed field of $G_{\mathfrak{P}}$. We denote it by $Z_{\mathfrak{P}}$.*

PROPOSITION 3.3. *(i) $Z_{\mathfrak{P}}$ is an extension of degree g of K, and $[L : Z_{\mathfrak{P}}] = ef$.*
(ii) $Gal(L/Z_{\mathfrak{P}}) = G_{\mathfrak{P}}$.
(iii) Let $\mathfrak{P}_Z = Z_{\mathfrak{P}} \cap \mathfrak{P}$ be the prime of $Z_{\mathfrak{P}}$ below $\mathfrak{P}$. Then $\mathfrak{P}$ is the only *prime above $\mathfrak{P}_Z$ in L, and*

$$e(\mathfrak{P}_Z/\mathfrak{p}) = f(\mathfrak{P}_Z/\mathfrak{p}) = 1. \tag{3.2}$$

PROOF. Parts (i) and (ii) are clear from Galois theory. For part (iii) note that for every $\sigma \in G_{\mathfrak{P}}$, $\sigma(\mathfrak{P}) = \mathfrak{P}$ by definition, and $G_{\mathfrak{P}}$ acts transitively on the primes above $\mathfrak{P}_Z$. It follows that

$$e(\mathfrak{P}/\mathfrak{P}_Z) f(\mathfrak{P}/\mathfrak{P}_Z) = [L : Z_{\mathfrak{P}}] = e(\mathfrak{P}/\mathfrak{p}) f(\mathfrak{P}/\mathfrak{p}). \tag{3.3}$$

By the multiplicativity of e and f in towers, we must have equalities

$$e(\mathfrak{P}/\mathfrak{P}_Z) = e(\mathfrak{P}/\mathfrak{p}) \tag{3.4}$$

and similarly for f. The same multiplicativity in towers now proves the last statement. □

3.2. The inertia group and Frobenius. Let us focus now on $G_{\mathfrak{P}}$. Every $\sigma \in G_{\mathfrak{P}}$ induces an automorphism $\bar{\sigma} \in Gal(\kappa(\mathfrak{P})/\kappa(\mathfrak{p}))$. The map $\sigma \mapsto \bar{\sigma}$ is a group homomorphism. Its kernel, the elements of the decomposition group inducing the identity on $\mathcal{O}_L mod \mathfrak{P}$, is called the inertia group $I_{\mathfrak{P}}$.

Recal that $Gal(\kappa(\mathfrak{P})/\kappa(\mathfrak{p})) = Gal(\mathbb{F}_{q^f}/\mathbb{F}_q)$ is cyclic of degree f, with a canonical generator

$$Fr_q : x \mapsto x^q. \tag{3.5}$$

Here $q = N\mathfrak{p}$.

Lemma 3.4. *The homomorphism $\sigma \mapsto \bar{\sigma}$ is* surjective.

Proposition 3.5. *We may assume that $K = Z_{\mathfrak{P}}$ and $G = G_{\mathfrak{P}}$. Let $\alpha \in \mathcal{O}_L$ be such that $\bar{\alpha}$, its image in $\kappa(\mathfrak{P})$, generates the residue field extension. Let f be the monic minimal polynomial of α, and $\bar{f}$ its reduction modulo $\mathfrak{p}$. The minimal polynomial of $\bar{\alpha}$ is then a factor of $\bar{f}$, so its roots are of the form $\bar{\beta}$ for some conjugates β of α. But G acts transitively on the roots, so for any $\tau \in Gal(\kappa(\mathfrak{P})/\kappa(\mathfrak{p}))$ there is a $\sigma \in G$ with*

$$\bar{\sigma}(\bar{\alpha}) = \tau(\bar{\alpha}). \tag{3.6}$$

But then $\bar{\sigma} = \tau$, as $\bar{\alpha}$ generates $\kappa(\mathfrak{P})$.

Corollary 3.6. *(i) The inertia group is trivial if and only if $\mathfrak{P}$ is unramified, and in general $|I_{\mathfrak{P}}| = e$.*

(ii) $G_{\mathfrak{P}}/I_{\mathfrak{P}} = Gal(\kappa(\mathfrak{P})/\kappa(\mathfrak{p}))$ is cyclic of order f, and has a canonical generator mapping to Fr_q.

We call this generator the *Frobenius* of $\mathfrak{P}$, and denote it $(L/K, \mathfrak{P})$ or

$$\left(\frac{L/K}{\mathfrak{P}}\right). \tag{3.7}$$

It is a *class* modulo $I_{\mathfrak{P}}$, but if $\mathfrak{P}$ is unramified it is an element of G, which generates $G_{\mathfrak{P}}$.

If $\sigma \in G$ is any element, it induces an isomorphism of $\kappa(\mathfrak{P})$ onto $\kappa(\sigma\mathfrak{P})$, hence by conjugation an isomorphism of $G_{\mathfrak{P}}$ onto $G_{\sigma\mathfrak{P}}$. It is not difficult to check that conjugation by σ takes $I_{\mathfrak{P}}$ to $I_{\sigma\mathfrak{P}}$ and that

$$\left(\frac{L/K}{\sigma\mathfrak{P}}\right) = \sigma \left(\frac{L/K}{\mathfrak{P}}\right) \sigma^{-1}. \tag{3.8}$$

In the special case where G is *abelian* the Frobenius class (or automorphism, if $\mathfrak{p}$ is unramified) does not depend on $\mathfrak{P}$ but only on $\mathfrak{p}$, and we denote it (only then!) $(L/K, \mathfrak{p})$.

Definition 3.3. *Let $T_{\mathfrak{P}}$ be the fixed field of $I_{\mathfrak{P}}$. It is called the inertia subfield of $\mathfrak{P}$.*

Proposition 3.7. *(i) $T_{\mathfrak{P}}$ is a normal extension of degree f of $Z_{\mathfrak{P}}$.*

(ii) $Gal(L/T_{\mathfrak{P}}) = I_{\mathfrak{P}}$ and $Gal(T_{\mathfrak{P}}/Z_{\mathfrak{P}}) = G_{\mathfrak{P}}/I_{\mathfrak{P}}$.

(iii) Let $\mathfrak{P}_T = \mathfrak{P} \cap T_{\mathfrak{P}}$. Then $\mathfrak{P}_Z$ is unramified in $T_{\mathfrak{P}}$ and $\mathfrak{P}_T$ is totally ramified in L.

Proof. Parts (i) and (ii) follow from Galois theory, since $I_{\mathfrak{P}}$ is a normal subgroup of order e. Consider $I_{\mathfrak{P}}$. It acts trivially on $\kappa(\mathfrak{P})$ but it should also map *onto* $Gal(\kappa(\mathfrak{P})/\kappa(\mathfrak{P}_T))$, so it follows that $\mathfrak{P}_T$ is totally ramified in $L/T_{\mathfrak{P}}$. □

3.3. Decomposition in non-Galois extensions. Let L/K be an arbitrary finite extension of number fields, and embed it in a Galois extension M/K. Let

$$G = Gal(M/K), \ \ H = Gal(M/L). \tag{3.9}$$

Let $\mathfrak{P}$ be a prime of M, $P = \mathfrak{P} \cap L$ and $\mathfrak{p} = \mathfrak{P} \cap K$.

PROPOSITION 3.8. *The association $\sigma \mapsto P_\sigma = \sigma\mathfrak{P} \cap L$ is a bijection between the primes of L dividing $\mathfrak{p}$ and the set of double cosets $H\backslash G/G_{\mathfrak{P}}$.*

PROOF. If P' is a prime lying above $\mathfrak{p}$ in L choose a prime $\mathfrak{P}'$ dividing it in M, and let $\sigma \in G$ carry $\mathfrak{P}$ to $\sigma\mathfrak{P} = \mathfrak{P}'$. Then $P' = P_\sigma$, so all primes above $\mathfrak{p}$ are of this form. If $P_\sigma = P_\tau$ then by the transitivity of H in its action on the primes of M above this prime, there exists a $\gamma \in H$ such that $\gamma\sigma\mathfrak{P} = \tau\mathfrak{P}$. This means that $\tau = \gamma\sigma\delta$ for some $\delta \in G_{\mathfrak{P}}$ or that the double cosets of σ and τ coincide. □

The decomposition (resp. inertia) subgroup of $\mathfrak{P}$ in M/L is the intersection of H with the corresponding groups in M/K. This is immediate from the definition. If H is normal, then the decomposition (inertia) subgroup of $\mathfrak{P}$ in L/K is the homomorphic image of the same group in M/K.

Recall that P is *unramified* in L/K if $e(P/\mathfrak{p}) = 1$ and is *split* if $e(P/\mathfrak{p}) = f(P/\mathfrak{p}) = 1$. The prime $\mathfrak{p}$ is called unramified (resp. totally split) if all primes above it are unramified (resp. split). In *Galois extensions* being unramified is equivalent to the fact that the corresponding inertia group is trivial, and being split is equivalent to the (stronger) fact that the decomposition group is trivial. If these hold for one $\mathfrak{P}$ above a given $\mathfrak{p}$, then by the transitivity of the Galois action, they hold for all of them.

PROPOSITION 3.9. *The prime P is unramified (resp. split) in L/K if and only if $I_{\mathfrak{P}}$ (resp. $G_{\mathfrak{P}}$) is contained in H.*

PROOF. If $I_{\mathfrak{P}}$ is contained in H then the inertia subgroup for M/K or M/L is the same, hence the inertia degree in these two extensions are the same, so the inertia degree in L/K is 1. Conversely, if the inertia degrees in M/L and M/K are the same, then we must have $I_{\mathfrak{P}} \subset H$. The proof for "split" is the same, except that we argue on $G_{\mathfrak{P}}$ instead of $I_{\mathfrak{P}}$. □

COROLLARY 3.10. *Let L_1 and L_2 be two extensions of K, and $L = L_1L_2$ their compositum (inside $\mathbb{C}$). Then a prime $\mathfrak{p}$ of K is unramified (resp. totally split) in L if and only if it is unramified (resp. totally split) in both L_i.*

COROLLARY 3.11. *Let L/K be an arbitrary finite extension and F an arbitrary extension of K. Then if a prime $\mathfrak{P}$ is unramified (resp. split) in L/K, every prime above it in FL is unramified (resp. split) in FL/F.*

PROOF. The proof of the two corollaries is similar. We prove the second and leave the first as an exercise to the reader. Embed L and F in a Galois extension M/K and let $G = Gal(M/K)$. Let $\mathfrak{P}_{LF}$ be a prime of LF above $\mathfrak{P}$ and $\mathfrak{P}_M$ a prime of M above $\mathfrak{P}_{LF}$. Since $\mathfrak{P}$ is unramified, $I(\mathfrak{P}_M) \subset Gal(M/L)$. The inertia group of $\mathfrak{P}_M$ in the Galois extension M/F is $I(\mathfrak{P}_M) \cap Gal(M/F)$ which is therefore contained in $Gal(M/LF)$. It follows that $\mathfrak{P}_{LF}$ is unramified. The proof for the "split" condition is the same, replacing the inertia group by the decomposition group. □

The group $I_{\mathfrak{P}}$ can be further analyzed and has a filtration measuring how bad the ramification is. This is the subject of Hilbert's "higher ramification theory", into which we shall not enter. It is best studied in the framework of local fields.

CHAPTER 4

Cyclotomic fields

1. The ring of integers

1.1. The cyclotomic polynomial. Let $\zeta = \zeta_m = \exp(2\pi i/m)$. The field $K = \mathbb{Q}(\zeta_m)$ is called the m^{th} cyclotomic field. It is a Galois extension of $\mathbb{Q}$, since it is the splitting field of $X^m - 1$.

Let $G = Gal(\mathbb{Q}(\zeta_m)/\mathbb{Q})$. Define the *cyclotomic character* $\chi : G \to (\mathbb{Z}/m\mathbb{Z})^\times$ by

$$\sigma(\zeta) = \zeta^{\chi(\sigma)}. \tag{1.1}$$

It is clear that this equation determines $\chi(\sigma)$ as an integer modulo m, since $\sigma(\zeta)$ is also an m^{th} root of unity. It is multiplicative $\chi(\sigma\tau) = \chi(\sigma)\chi(\tau)$, so it is a group homomorphism into the multiplicative group of the invertible residue classes modulo m.

The cyclotomic polynomial

$$\Phi_m(X) = \prod_{(k,m)=1} (X - \zeta^k) \tag{1.2}$$

is invariant under G (which permutes its roots, the primitive roots of unity of order m). Its coefficients are therefore rational. Since they are also algebraic integers, they are in $\mathbb{Z}$. Clearly $\deg(\Phi_m) = \varphi(m)$.

THEOREM 1.1. *(i) The cyclotomic polynomial is irreducible.*
(ii) $[K : \mathbb{Q}] = \varphi(m)$ *and* $\chi : G \to (\mathbb{Z}/m\mathbb{Z})^\times$ *is an isomorphism.*

PROOF. Part (ii) is a direct consequence of (i) and standard Galois theory. We shall prove (i) only for $m = p$ (prime). The proof for $m = p^a$ (a prime power) is very similar, and is left to the reader. The proof for an m which is divisible by more than one prime is by induction on the number of prime divisors. If $m = m'p^a$ where p does not divide m', then we shall see that p does not ramify in $\mathbb{Q}(\zeta_{m'})$ but is ramified completely in $\mathbb{Q}(\zeta_{p^a})$, so its index of ramification in K is $[\mathbb{Q}(\zeta_{p^a}) : \mathbb{Q}]$, hence $[K : \mathbb{Q}(\zeta_{m'})]$ is at least $[\mathbb{Q}(\zeta_{p^a}) : \mathbb{Q}]$. Since we may assume, by induction, that $[\mathbb{Q}(\zeta_{m'}) : \mathbb{Q}] = \varphi(m')$, and since $\varphi(m) = \varphi(m')\varphi(p^a)$, we get $[\mathbb{Q}(\zeta_m) : \mathbb{Q}] \geq \varphi(m)$, hence an equality.

For $m = p$ the cyclotomic polynomial is $(X^p - 1)/(X - 1)$. Substituting $X = 1 + Y$ we get $\Phi_p(1 + Y) = Y^{p-1} + \cdots + p \equiv Y^{p-1} mod p$ so $\Phi_p(1 + Y)$ is irreducible by Eisenstein's criterion, hence $\Phi_p(X)$ is also irreducible. □

From now on we assume that $m = p$*is an odd prime.*

1.2. Cyclotomic units.

LEMMA 1.2. *If $1 \leq i\ , j \leq p-1$ then $u = (1-\zeta^i)/(1-\zeta^j)$ is a unit in $\mathbb{Z}[\zeta]$.*

PROOF. It is enough to show that it belongs to $\mathbb{Z}[\zeta]$, because then by symmetry u^{-1} also lies there, hence u is a unit in that ring. Write $i \equiv jkmodp$. Then

$$u = \frac{1-\zeta^{jk}}{1-\zeta^j} = 1 + \zeta^j + \cdots + (\zeta^j)^{k-1} \in \mathbb{Z}[\zeta]. \tag{1.3}$$

□

COROLLARY 1.3. *We have $(p) = (1-\zeta)^{p-1}$ in $\mathcal{O}_K$ and $\mathfrak{p} = (1-\zeta)$ is prime.*

PROOF. We have

$$p = \Phi_p(1) = \prod_{i=1}^{p-1}(1-\zeta^i), \tag{1.4}$$

but by the lemma there is an equality of ideals $(1-\zeta) = (1-\zeta^i)$ for all i not divisible by p. Comparing the decomposition $(p) = (P_1 P_2 \dots P_g)^e$, $p-1 = [K:\mathbb{Q}] = efg$ with the one just obtained we see that there is only one prime above p, that it is principal and generated by $1-\zeta$, and that p is totally ramified. □

1.3. The discriminant. The elements $1, \zeta, \zeta^2, \dots, \zeta^{p-2}$ make up a basis of K over $\mathbb{Q}$. Its discriminant is given by the van-der-Monde

$$\begin{aligned} \Delta(1,\zeta,\dots,\zeta^{p-2}) &= \left(\det(\zeta^{ij})_{0\leq i\leq p-2, 1\leq j\leq p-1}\right)^2 \\ &= \prod_{1\leq i<j\leq p-1} (\zeta^i - \zeta^j)^2 \\ &= (-1)^{(p-1)(p-2)/2} \prod_{1\leq i\neq j\leq p-1} (\zeta^i - \zeta^j) \\ &= (-1)^{(p-1)/2} \prod_{i\neq 0} \prod_{l\neq 0,-i} \zeta^i(1-\zeta^l) \\ &= (-1)^{(p-1)/2} \prod_{i\neq 0} \left(\zeta^{i(p-2)}p/(1-\zeta^{-i})\right) \\ &= (-1)^{(p-1)/2} p^{p-1} / \prod_{i\neq 0}(1-\zeta^{-i}) \\ &= (-1)^{(p-1)/2} p^{p-2}. \end{aligned} \tag{1.5}$$

where we have used the formula above for $\Phi_p(1)$.

COROLLARY 1.4. *The index $[\mathcal{O}_K : \mathbb{Z}[\zeta]]$ is not divisible by any prime different than p.*

PROOF. The square of this index is the ratio $\Delta(1,\zeta,\dots,\zeta^{p-2})/\Delta(\mathcal{O}_K)$, which is a power of p by the above. □

1.4. The integers.

PROPOSITION 1.5. *We have $\mathcal{O}_K = \mathbb{Z}[\zeta]$.*

The discriminant is $d_K = (-1)^{(p-1)/2}p^{p-2}$. The prime p is the only ramified prime in K, and it is totally ramified.

PROOF. Since no prime other than p can divide the index $[\mathcal{O}_K : \mathbb{Z}[\zeta]]$, it is enough to prove that this index is not divisible by p. For this we may localize at (p) (w.r.t. $\mathbb{Z} - p\mathbb{Z}$), and show $\mathcal{O}_{K,(p)} = \mathbb{Z}_{(p)}[\zeta]$. Let $\pi = 1 - \zeta$, the generator of the ideal $\mathfrak{p}$. For any l the $\mathbb{Z}$-span of $1, \zeta, \ldots, \zeta^l$ is the same as the $\mathbb{Z}$-span of $1, \pi, \ldots, \pi^l$. In particular $\mathbb{Z}_{(p)}[\zeta] = \mathbb{Z}_{(p)}[\pi]$, and $1, \pi, \ldots, \pi^{p-2}$ is a basis of K. Let $\alpha \in \mathcal{O}_K$ and write it as

$$\alpha = a_0 + a_1\pi + \cdots + a_{p-2}\pi^{p-2}. \tag{1.6}$$

Let $v = v_{\mathfrak{p}}$ be the normalized discrete valuation of K associated with $\mathfrak{p}$. Since $(p) = \mathfrak{p}^{p-1}$, if $a \in \mathbb{Q}$ then

$$v(a) = (p-1)ord_p(a) \equiv 0mod(p-1). \tag{1.7}$$

It follows that if we denote $ord_p(a_i) = k_i \in \mathbb{Z}$ then

$$v(a_i\pi^i) = i + (p-1)k_i. \tag{1.8}$$

These numbers are all distinct, since they are even distinct modulo $p - 1$. For a non-archemidean valuation, if $v(x) \neq v(y)$ then $v(x + y) = \min(v(x), v(y))$. By induction, the same holds for a sum of more than two elements, and we find out that

$$0 \leq v(\alpha) = \min(i + (p-1)k_i). \tag{1.9}$$

Every $i + (p-1)k_i \geq 0$, and since $i \leq p - 2$, we must have $k_i \geq 0$, that is the $a_i \in \mathbb{Z}_{(p)}$ as we had to show.

The assertion about the discriminant now follows from the computation of $\Delta(1, \zeta, \ldots, \zeta^{p-2})$, and since every ramified prime must divide the discriminant, primes different from p are unramified. □

2. The decomposition of primes

2.1. The decomposition group of a prime l. Let l be a rational prime different from p. Since l is unramified in $K = \mathbb{Q}(\zeta_p)$, we know that

$$(l) = \mathfrak{l}_1 \ldots \mathfrak{l}_g \tag{2.1}$$

in $\mathcal{O}_K$ and G acts transitively on the $\mathfrak{l}_i$. Let $f = f(\mathfrak{l}_i/l)$, so that $fg = p - 1$, and denote by G_l the decomposition group of any of the $\mathfrak{l}_i$. Since G is abelian, G_l depends only on l and not on $\mathfrak{l}_i$. The same may be said about the Frobenius at l,

$$\sigma_l = (K/\mathbb{Q}, \mathfrak{l}_i) \tag{2.2}$$

which is characterized by

$$\sigma_l(x) \equiv x^l mod\mathfrak{l}_i \tag{2.3}$$

for any of the $\mathfrak{l}_i$ (any $x \in \mathcal{O}_K$). This in particular holds for ζ, so we get

$$\zeta^{\chi(\sigma_l)} \equiv \zeta^l mod\mathfrak{l}_i. \tag{2.4}$$

However, if $i \neq j$ then $(\zeta^i - \zeta^j) = \mathfrak{p}$ is relatively prime to l, and we get that $\chi(\sigma_l) = l$. Since χ is an isomorphism from G to $(\mathbb{Z}/p\mathbb{Z})^\times$, the order of σ_l is the (multiplicative) order of $lmodp$. We have obtained the following theorem.

THEOREM 2.1. *Let l be a prime different than p. Then under the identification of G with $(\mathbb{Z}/p\mathbb{Z})^\times$, σ_l maps to l and the order of G_l is the minimal f such that $l^f \equiv 1 mod p$.*

EXAMPLE 2.1. *Let $p = 7$, so that G is cyclic of order 6. It has four subgroups, $\{1\}$, $H_2 = \{\pm 1\}$, $H_3 = \{1, 2, 4\}$ and G. Their fixed fields are, respectively, K, $\mathbb{Q}(\cos(2\pi/7))$, $\mathbb{Q}(\sqrt{-7})$ and $\mathbb{Q}$ (proving this is a good exercise in Galois theory - that the unique quadratic field inside K must be $\mathbb{Q}(\sqrt{-7})$ follows from the fact that the only prime ramifying in it is 7). If $l \equiv 1 mod 7$, then l splits completely ($f = 1$, $g = 7$). If $l \equiv 2, 4 mod 7$ then its decomposition group is H_3, its decomposition field is $\mathbb{Q}(\sqrt{-7})$ and $l = \mathfrak{l}_1\mathfrak{l}_2$ in K, and so on. Note that whether there exists a prime l whose Frobenius σ_l is a given σ becomes a question whether there are primes in arbitrary arithmetic progressions $7n+a$, $(a, 7) = 1$. This is in fact a famous theorem of Dirichlet.*

COROLLARY 2.2. *The way l decomposes (namely f and g) is determined by the congruence class of l modulo p.*

All of the above remains true in any cyclotomic extension, with some rather obvious modifications. The *Kronecker-Weber* theorem says that any abelian extension of $\mathbb{Q}$ is contained in some cyclotomic field, and appropriately formulated, the results remain valid in any abelian extension $K/\mathbb{Q}$. *Class field theory* extends these results to abelian extension L/K where the ground field K may be arbitrary, but the proofs are much more difficult.

2.2. Quadratic reciprocity. As an application of the discussion above, we shall prove Gauss' famous law of quadratic reciprocity. We first need to find out which quadratic field lies inside K.

LEMMA 2.3. *The unique quadratic subfield of K is $F = \mathbb{Q}(\sqrt{p^*})$ where $p^* = (-1)^{(p-1)/2}p$.*

PROOF. Since K is a cyclic extension of degree $p-1$, it has a unique quadratic subextension, the fixed field of the unique subgroup of index 2 in G. The only prime ramifying in it is p. It follows that the discriminant of F must be divisible only by p, and since a discriminant of a quadratic field is either 0 or 1 modulo 4, the sign is forced upon us as in the lemma. □

Alternatively, let $\left(\frac{a}{p}\right)$ be *Legendre's symbol*, which is 1 if $a mod p$ is a square in $(\mathbb{Z}/p\mathbb{Z})^\times$, -1 if $a mod p$ is a non-square in the same group, and 0 if $p|a$. Consider the *Gauss sum*

$$\tau = \sum_{a=1}^{p-1} \left(\frac{a}{p}\right) \zeta^a \in K. \tag{2.5}$$

It is not difficult to show that

$$\tau^2 = p^*, \tag{2.6}$$

hence $\mathbb{Q}(\sqrt{p^*}) \subset K$.

Let l be another odd prime. On the one hand, we know that l splits in F (into the product of two primes of inertial degree 1) if and only if its decomposition group $G_l \subset G$ fixes F, and this happens if and only if its generator σ_l fixes F. But the $\sigma's$

fixing F are precisely the squares in G, so l splits in F if and only if l is a square modulo p, namely $(l/p) = 1$.

On the other hand, $\mathcal{O}_{F,(l)} = \mathbb{Z}_{(l)}[\sqrt{p^*}]$ so by a theorem proved some time ago, l splits in F if and only if $X^2 - p^*$ splits mod l. However, this is the case if and only if $(p^*/l) = 1$. Now it is easy to see that $(-1/l) = (-1)^{(l-1)/2}$ (-1 is a quadratic residue mod l precisely when $l \equiv 1 mod 4$) and that Legendre's symbol is multiplicative in the numerator.

Theorem 2.4. *(Gauss' law of quadratic reciprocity)*

$$\left(\frac{l}{p}\right) = (-1)^{(p-1)(l-1)/2}\left(\frac{p}{l}\right). \tag{2.7}$$

Proof. We have seen that $(l/p) = (p^*/l)$ where $p^* = (-1)^{(p-1)/2}p$. By the multiplicativity

$$\left(\frac{p^*}{l}\right) = \left(\frac{(-1)}{l}\right)^{(p-1)/2}\left(\frac{p}{l}\right) = (-1)^{(p-1)(l-1)/2}\left(\frac{p}{l}\right). \tag{2.8}$$

□

CHAPTER 5

Zeta and L functions

There are many results in algebraic number theory for which the only known proof is analytical. Zeta and L functions play an important role, and the introduction of complex analysis is indispensible. We shall examine the simplest results of this kind. Our ultimate goal is to prove Dirichlet's theorem on primes in arithmetic progressions. It was the motivation for Dirichlet himself to introduce the series and L functions that bear his name in the first half of the 19th century.

1. The Riemann zeta function

1.1. The Euler product. The Riemann Zeta function is defined for $s \in \mathbb{C}$, $Re(s) > 1$, by the convergent series

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}. \tag{1.1}$$

By unique factorization in $\mathbb{Z}$ we have the *Euler product formula*

$$\zeta(s) = \prod_{p \in P} \left(1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \dots\right) = \prod_{p \in P} \left(1 - \frac{1}{p^s}\right)^{-1}. \tag{1.2}$$

Here P is the set of primes. One can prove that P is infinite from here, becuase otherwise the product would be finite, and $\zeta(s)$ would have a finite limit as $s \to 1$ from the right, contradicting the divergence of the harmonic series. The factor $\left(1 - \frac{1}{p^s}\right)^{-1}$ is called the Euler factor at p (of $\zeta(s)$). Incidentally, the Euler product formula shows that $\zeta(s) \neq 0$ for $Re(s) > 1$.

1.2. The Gamma function. The Gamma function is defined for $Re(s) > 0$ by the convergent integral

$$\Gamma(s) = \int_0^{\infty} e^{-t} t^s \frac{dt}{t}. \tag{1.3}$$

Note that dt/t is invariant under the change of variables $t \mapsto ct$, $c > 0$. Integration by parts yields

$$\Gamma(s+1) = s\Gamma(s). \tag{1.4}$$

This wonderful formula has two immediate consequences. First, for a positive integer n, $\Gamma(n+1) = n!$, because it is easy to see that $\Gamma(1) = 1$. Second, we may use the formula to meromorphically continue $\Gamma(s)$ to the whole complex plane. It will have then simple poles at $s = 0, -1, -2, \ldots$ and will be holomorphic elsewhere. Indeed, if it has already been defined on $Re(s) > m$, extend it to $Re(s) > m - 1$ by setting $\Gamma(s) = \Gamma(s+1)/s$. This does not lead to a contradiction in $Re(s) > m$, and by the very construction extends the relation $\Gamma(s+1) = s\Gamma(s)$ to the larger domain $Re(s) > m-1$.

The Gamma function has many remarkable properties. The function

$$\pi^{-s/2}\Gamma(s/2) \tag{1.5}$$

is sometimes called "the Euler factor at infinity", because when we multiply $\zeta(s)$ by it we get an even nicer function. Let

$$Z(s) = \pi^{-s/2}\Gamma(\frac{s}{2})\zeta(s). \tag{1.6}$$

THEOREM 1.1. *$Z(s)$ has a meromorphic continuation to the whole complex plane, with simple poles at $s = 0, 1$, and no other poles. It satisfies the functional equation*

$$Z(s) = Z(1-s). \tag{1.7}$$

1.3. Poisson summation formula. The proof which we shall give to the theorem relies on some Fourier analysis. Let $\mathcal{S}$ be the class of Schwartz functions on $\mathbb{R}$. These are the smooth (complex valued) functions f such that $|x|^n f^{(m)}(x) \to 0$ as $|x| \to \infty$ for any $m, n \geq 0$. The *Fourier transform* is a bijective map from $\mathcal{S}$ to itself defined by

$$\widehat{f}(y) = \int_{-\infty}^{\infty} f(x)e^{2\pi ixy}dx. \tag{1.8}$$

We recall its well-known properties. For the moment denote $\widehat{f}$ also by $\mathcal{F}f$

- If $f \in \mathcal{S}$ then $\mathcal{F}f \in \mathcal{S}$ too and $\mathcal{F}\mathcal{F}f(x) = f(-x)$.
- Let $\mathcal{M}f(x) = 2\pi ixf(x)$ and $\mathcal{D}f(x) = f'(x)$. Then

$$\mathcal{F}\mathcal{M}f = \mathcal{D}\mathcal{F}f, \;\; \mathcal{F}\mathcal{D}f = -\mathcal{M}\mathcal{F}f. \tag{1.9}$$

- If $g(x) = f(x+a)$ then $\widehat{g}(y) = e^{-2\pi iay}\widehat{f}(y)$.

Let $f \in \mathcal{S}$. The function $F(x) = \sum_{m\in\mathbb{Z}} f(x+m)$ is smooth an 1-periodic. It has therefore a convergent Fourier expansion

$$F(x) = \sum_{n\in\mathbb{Z}} a_n e^{2\pi inx}. \tag{1.10}$$

Compute

$$a_n = \int_0^1 F(x)e^{-2\pi inx}dx = \int_{-\infty}^{\infty} f(x)e^{-2\pi inx}dx = \widehat{f}(-n). \tag{1.11}$$

Now compare the two experessions that we obtained for $F(0)$. We get:

THEOREM 1.2. *Let $f \in \mathcal{S}$. Then*

$$\sum_{n\in\mathbb{Z}} f(n) = \sum_{n\in\mathbb{Z}} \widehat{f}(n). \tag{1.12}$$

1.4. An important example. Let $f(x) = e^{-\pi t x^2}$ where $t > 0$ is a fixed parameter. This function belongs to $\mathcal{S}$, and it's Fourier transform is

$$\widehat{f}(y) = \frac{1}{\sqrt{t}} e^{-\pi y^2/t}. \tag{1.13}$$

EXERCISE 1.1. *Prove this! You will have to use Cauchy's theorem to shift the line of integration from $iy/t + \mathbb{R}$ back to $\mathbb{R}$, and you will also have to use the fact that $\int_{-\infty}^{\infty} e^{-\pi x^2} dx = 1$. Alternatively, if you only want to use real analysis, separate real and imaginary parts, and use integration by parts twice.*

Let

$$\theta(t) = \sum_{n \in \mathbb{Z}} e^{-\pi t n^2}. \tag{1.14}$$

This is *Riemann's theta function.* The Poisson summation formula now gives

$$\theta(t) = \frac{1}{\sqrt{t}} \theta(1/t). \tag{1.15}$$

Since $\theta(t) - 1 \to 0$ exponentially as $t \to \infty$, this shows that $\theta(t) - \frac{1}{\sqrt{t}} \to 0$ as $t \to 0$.

1.5. Analytic continuation and functional equation. We are now ready to prove the theorem. Consider, for $Re(s) > 1$

$$\begin{aligned} Z(s) &= \pi^{-s/2} \Gamma(\frac{s}{2}) \sum_{n=1}^{\infty} \frac{1}{n^s} \\ &= \sum_{n \geq 1} \int_0^{\infty} e^{-t} \left(\frac{t}{\pi n^2} \right)^{s/2} \frac{dt}{t} \\ &= \int_0^{\infty} \sum_{n \geq 1} e^{-\pi t n^2} t^{s/2} \frac{dt}{t} \\ &= \int_0^{\infty} \left(\frac{\theta(t) - 1}{2} \right) t^{s/2} \frac{dt}{t}. \qquad (1.16) \end{aligned}$$

We now break the domain of integration to $[1, \infty)$, where we do not change anything, and to $(0, 1]$ where we use the functional equation of θ and then a change of variables $t \mapsto 1/t$:

$$\begin{aligned} Z(s) &= \int_0^1 \left(\frac{\theta(1/t) - \sqrt{t}}{2\sqrt{t}} \right) t^{s/2} \frac{dt}{t} + \int_1^{\infty} \left(\frac{\theta(t) - 1}{2} \right) t^{s/2} \frac{dt}{t} \\ &= \int_1^{\infty} \left(\frac{\theta(t) - t^{-1/2}}{2} \right) t^{(1-s)/2} \frac{dt}{t} + \int_1^{\infty} \left(\frac{\theta(t) - 1}{2} \right) t^{s/2} \frac{dt}{t} \\ &= \int_1^{\infty} \left(\frac{1 - t^{-1/2}}{2} \right) t^{(1-s)/2} \frac{dt}{t} + \int_1^{\infty} \left(\frac{\theta(t) - 1}{2} \right) \left(t^{s/2} + t^{(1-s)/2} \right) \frac{dt}{t}. \qquad (1.17) \end{aligned}$$

In the last expression, valid for $Re(s) > 1$, the second integral makes sense for every s since $\frac{\theta(t)-1}{2}$ decays exponentially to 0 as $t \to \infty$. It is moreover holomorphic in s and invariant under $s \mapsto 1 - s$. The first integral is evaluated directly; it is equal to $-\frac{1}{1-s} - \frac{1}{s}$, and therefore contributes simple poles with residues 1 (at 1) and -1 (at 0), and is again invariant under $s \mapsto 1 - s$. This proves the theorem.

2. Dirichlet L functions

2.1. Dirichlet characters. Let m be an integer ≥ 1, and $\chi : (\mathbb{Z}/m\mathbb{Z})^\times \to \mathbb{C}^\times$ a character of order d. In other words, χ is a homomorphism of multiplicative groups, and d is the minimal number such that $\chi^d = 1$. The image of χ is the cyclic group of d^{th} roots of unity in $\mathbb{C}$, and d must divide $\varphi(m)$. We extend χ to a function on $\mathbb{Z}$ by setting $\chi(n) = 0$ if $(n, m) \neq 1$. Recalling the canonical identification

$$\omega : Gal(\mathbb{Q}(\zeta_m)/\mathbb{Q}) = G \simeq (\mathbb{Z}/m\mathbb{Z})^\times \tag{2.1}$$

via the cyclotomic character ($\sigma(\zeta_m) = \zeta_m^{\omega(\sigma)}$), we can associate to χ a character $\chi_{Gal} = \chi \circ \omega$ on G, and vice versa. If we define $H_\chi = \ker(\chi)$ then $G/H_\chi \simeq Im(\chi)$ is cyclic of order d. By Galois theory the fixed field

$$K_\chi \subset \mathbb{Q}(\zeta_m) \tag{2.2}$$

of H_χ is a cyclic extension of $\mathbb{Q}$ of degree d, and χ_{Gal} induces and isomorphism between $Gal(K_\chi/\mathbb{Q}) = G/H_\chi$ and the group of d^{th} roots of unity in $\mathbb{C}$. For example, if χ is *quadratic,* meaning $\chi^2 = 1$ (this is equivalent to χ being real, $\chi = \bar{\chi}$), then K_χ is a quadratic extension of $\mathbb{Q}$.

2.2. Dirichlet L-functions $L(\chi, s)$. The Dirichlet L function of χ is the function

$$L(\chi, s) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}. \tag{2.3}$$

Like the Riemann Zeta function, it converges absolutely and uniformly on compact subsets in $Re(s) > 1$. However, if χ is not trivial, Abel summation shows that $L(\chi, s)$ converges, uniformly on compact subsets, in $Re(s) > 0$. Indeed

$$\sum_{n=1}^{N+1} \frac{\chi(n)}{n^s} = \sum_{n=1}^{N} \left(\sum_{m=1}^{n} \chi(m) \right) \left(\frac{1}{n^s} - \frac{1}{(n+1)^s} \right) + \left(\sum_{m=1}^{N+1} \chi(m) \right) \frac{1}{(N+1)^s} \tag{2.4}$$

Now the partial sums $\sum_1^M \chi(n)$ are bounded and

$$\frac{1}{n^s} - \frac{1}{(n+1)^s} = \frac{s}{(n+\theta)^{s+1}} \tag{2.5}$$

for some $0 < \theta < 1$, so the first sum converges absolutely, and the remainder term tends to 0.

Dirichlet L series have an *Euler product* similar to the zeta function:

$$L(\chi, s) = \prod_{(p,m)=1} \left(1 - \frac{\chi(p)}{p^s} \right)^{-1}. \tag{2.6}$$

The proof is the same as for the zeta function, taking into account the multiplicativity of χ.

Theorem 2.1. *For non-trivial* χ, $L(\chi, 1) \neq 0$.

We shall prove this important theorem soon, but we first give its most famous application.

2.3. Primitive and non-primitive characters. We say that χ is *primitive* of level m if there does not exists an m' dividing m and a character $\chi' : (\mathbb{Z}/m'\mathbb{Z})^\times \to \mathbb{C}^\times$ such that χ is the pull-back of χ' via the canonical projection of $\mathbb{Z}/m\mathbb{Z}$ to $\mathbb{Z}/m'\mathbb{Z}$. Equivalently, χ does not have a period smaller than m when restricted to integers relatively prime to m. If χ comes from $\chi' mod m'$ and also from $\chi'' mod m''$ then it comes from a character modulo $\gcd(m', m'')$. This proves that there exists a minimal $m'|m$ and $\chi' mod m'$ such that χ comes from χ'. We call χ' the *primitive character* associated to χ, and m' its *conductor*, and write

$$\chi' = \chi^{prim}. \tag{2.7}$$

This seems trivial, but one should post a small warning sign: when we extend χ and χ' to functions on $\mathbb{Z}$, they will agree on integers prime to m, but will differ at any n which is relatively prime to m' but not to m. If there are new primes dividing m which did not divide m', there will be such integers.

Comparing the Euler products for χ and χ' we see that

$$L(\chi, s) = \prod_{p|m,\, p\nmid m'} \left(1 - \frac{\chi'(p)}{p^s}\right) L(\chi', s). \tag{2.8}$$

Thus the primitive and the non-primitive L functions differ only by finitely many Euler factors (and do not differ at all if m and m' have the same prime factors). In particular the question of *vanishing at* $s = 1$ is the same for both, because none of these Euler factors vanishes at $s = 1$.

2.4. Dirichlet's theorem on primes in arithmetic progressions.

THEOREM 2.2. *Let $(a, m) = 1$. Then there are infinitely many primes of the form $a + km$.*

PROOF. Consider

$$F(s) = \sum_\chi \bar{\chi}(a) \log L(\chi, s) \tag{2.9}$$

where the sum is over the $\varphi(m)$ Dirichlet characters modulo m, and $s > 1$. The sum is real, because together with χ there appears $\bar{\chi}$ and the corresponding terms are complex conjugates. Also note that there are several branches of the logarithm that one can consider, but the Euler product for $L(\chi, s)$ allows one to make a natural choice as below.

As $s \to 1$ from the right, the trivial character gives $\log L(1, s)$ which, up to the logarithm of the Euler factors dividing m is just $\log \zeta(s)$, hence tends to $+\infty$. The other terms tend to the finite limit $\bar{\chi}(a) \log L(\chi, 1)$. It follows that $\lim_{s\to 1^+} F(s) = \infty$.

On the other hand, use the Euler product to find that

$$\log L(\chi, s) = -\sum_p \log\left(1 - \frac{\chi(p)}{p^s}\right) = \sum_p \sum_{m=1}^\infty \frac{\chi(p^m)}{mp^{ms}} = \sum_p \frac{\chi(p)}{p^s} + O(1) \tag{2.10}$$

where $O(1)$ denotes an expression which is bounded as $s \to 1$ from the right. To see this we estimate

$$\frac{1}{2mp^{2m}} + \frac{1}{(2m+1)p^{2m+1}} \le \frac{1}{mp^{2m}}, \tag{2.11}$$

hence

$$\sum_p \sum_{m=2}^{\infty} \frac{1}{mp^m} \le \sum_p \sum_{m=1}^{\infty} \frac{1}{mp^{2m}} = \log \prod_p (1 - \frac{1}{p^2})^{-1} = \log(\pi^2/6). \tag{2.12}$$

It follows that

$$F(s) = \sum_p \left(\sum_\chi \bar{\chi}(a)\chi(p) \right) \frac{1}{p^s} + O(1). \tag{2.13}$$

Now the inner sum (over χ) is 0 if $p \neq amodm$, by the orthogonality relations for characters. If $p \equiv amodm$ it gives $\varphi(m)$. We therefore get

$$F(s) = \varphi(m) \sum_{p \equiv amodm} \frac{1}{p^s} + O(1). \tag{2.14}$$

The fact that $F(s)$ is not $O(1)$ imples that there are infinitely many $p's$ - not only that, but sufficiently many to make the series in question diverge. □

Remark. Let A be a set of primes. The *Dirichlet density* of A is the limit (if it exists)

$$\delta(A) = \lim_{s \to 1+} \frac{\sum_{p \in A} p^{-s}}{\sum_p p^{-s}}. \tag{2.15}$$

The proof of the theorem shows that the primes congruent to $amodm$ have Dirichlet density $1/\varphi(m)$. In other words, each of the $\varphi(m)$ residue classes has on the average equally many primes.

There is a more naive notion of density, namely

$$\lim_{x \to \infty} \frac{\#\{p \in A | \, p \le x\}}{\#\{p | \, p \le x\}}. \tag{2.16}$$

It can be shown that if the latter exists and is equal δ, then $\delta(A)$ exists too and is equal to the naive density δ.

2.5. Non-vanishing of $L(\chi, 1)$. There are several proofs of this nowadays. We shall follow Dirichlet, who gave an easy argument for χ which is non-quadratic ($\chi \neq \bar{\chi}$), and then a separate more complicated argument for quadratic non-prinicpal χ.

Assume that $\chi \neq \bar{\chi}$, but $L(\chi, 1) = 0$. Then $L(\bar{\chi}, 1) = 0$ too and the product $\prod_\chi L(\chi, s)$ has a zero at $s = 1$, because $\zeta(s)$ contributes a simple pole, but at least two other factors contribute zeros. Letting $s \to 1$ from the right we see that $\sum_\chi \log L(\chi, s)$ must tend to $-\infty$. But we have computed this expression in the proof and found it to be $\varphi(m) \sum_{p \equiv 1modm} p^{-s} + O(1)$. This is a contradiction, so there is at most *one* χ for which $L(\chi, 1) = 0$.

Incidentally, this argument proves that among *all* Dirichlet characters, of all conductors, there is at most one, quadratic or not, for which $L(\chi, 1) = 0$ (why?). This is very unlikely, because who would be the fortunate χ, had it existed??

Let now χ be quadratic, $\chi^2 = 1$. Then as we saw, K_χ is quadratic.

LEMMA 2.3. *Let $(p, m) = 1$. Then p is unramified in K_χ. It splits if and only if $\chi(p) = 1$.*

PROOF. If p is ramified in K_χ then it must be ramified in $\mathbb{Q}(\zeta_m)$, so must divide m (we have not proved this for composite m, but only primes dividing m ramify in $\mathbb{Q}(\zeta_m)$). Now p splits in K_χ if and only if its Frobenius σ_p lies in $Gal(\mathbb{Q}(\zeta_m)/K_\chi) = \ker\chi$, if and only if $\chi_{Gal}(\sigma_p) = 1$. But $\chi_{Gal}(\sigma_p) = \chi(\omega(\sigma_p)) = \chi(p)$. □

Consider

$$\zeta_m(s)L(\chi, s) = \prod_{\chi(p)=1} (1 - p^{-s})^{-2} \prod_{\chi(p)=-1} (1 - p^{-2s})^{-1}. \tag{2.17}$$

The subscript m indicates that we have removed from the Riemann Zeta function the Euler factors of $p|m$. The expression on the right is a Dirichlet series (a series of the form $\sum a_n n^{-s}$) with non-negative coefficients ($a_n \geq 0$) which converges for $Re(s) > 1$, but definitely not in all $Re(s) > 0$, as a simple estimate at $s = 1/2$ shows: the terms $(1-p^{-1/2})^{-2} > (1-p^{-1})^{-1}$, so we can compare the product with the usual Euler product for $\zeta(s)$, which diverges for $s = 1$.

On the other hand, *if we assume* that $L(\chi, 1) = 0$, then $\zeta_m(s)L(\chi, s)$ is analytic in $Re(s) > 0$, because $L(\chi, s)$ is analytic there and the only pole of $\zeta_m(s)$ in that region is at $s = 1$. The following lemma gives us the desired contradiction, and finishes the proof.

LEMMA 2.4. *Suppose $F(s) = \sum a_n n^{-s}$ is a Dirichlet series with non-negative coefficients which converges in some domain $Re(s) > \sigma_0$. Suppose $F(s)$ has an analytic continuation to $Re(s) > \sigma_1$. Then the series converges also for $Re(s) > \sigma_1$.*

PROOF. Clearly we may assume that σ_0 is minimal such that the series converges for $Re(s) > \sigma_0$ and that $\sigma_1 < \sigma_0$, otherwise there is nothing to prove. Let $\alpha > \sigma_0$ and $r > \alpha - \sigma_0$ be such that $F(s)$ is analytic in the open disc $D(\alpha, r)$. Note that this disc contains points to the left of σ_0 on the real axis. The Taylor expansion of $F(s)$ around α is

$$F(s) = \sum_{k=0}^{\infty} \sum_{n=1}^{\infty} \frac{a_n (\log n)^k n^{-\alpha}}{k!} (\alpha - s)^k. \tag{2.18}$$

Picking a point s such that $\alpha - r < s < \sigma_0$ we get that the Taylor expansion converges at s. However, this is a series with non-negative terms, so it must converge after a re-arrangement of its terms. Rearranging we get

$$\begin{aligned} F(s) &= \sum_{n=1}^{\infty} a_n \left(\sum_{k=0}^{\infty} \frac{(\log n)^k n^{-\alpha}}{k!} (\alpha - s)^k \right) \\ &= \sum_{n=1}^{\infty} a_n n^{-s}. \end{aligned} \tag{2.19}$$

This is a contradiction, as the Dirichlet series was asumed to diverge at s. □

2.6. Formula for $L(\chi, 1)$. Assume that χ is a primitive non-trivial character modulo m, so that $m > 1$. By Fourier analysis on the finite group $\mathbb{Z}/m\mathbb{Z}$, any complex function f on it can be written as

$$f(a) = \sum_b \hat{f}(b) e^{2\pi i ab/m} \tag{2.20}$$

where

$$\hat{f}(b) = \frac{1}{m}\sum_{a} f(a)e^{-2\pi iab/m}. \tag{2.21}$$

Apply this with $f = \chi$. If $d = (b, m) > 1$ let $m' = m/d$ and $b' = m/d$. Since χ is *primitive*, it does not come from a character modulo m', so there exists a c such that $(c, m) = 1$ and $c \equiv 1 modm'$ but $\chi(c) \neq 1$. Then

$$\begin{aligned}
& \sum_{a=0}^{m-1} \chi(a)e^{-2\pi iab/m} \\
= & \sum_{a=0}^{m-1} \chi(a)e^{-2\pi iab'/m'} \\
= & \sum_{a=0}^{m-1} \chi(a)e^{-2\pi iacb'/m'} \\
= & \bar{\chi}(c)\sum_{a=0}^{m-1} \chi(ac)e^{-2\pi iacb'/m'} \\
= & \bar{\chi}(c)\sum_{a=0}^{m-1} \chi(a)e^{-2\pi iab'/m'}, \qquad (2.22)
\end{aligned}$$

so the sum vanishes and $\hat{f}(b) = 0$. On the other hand, if $(b, m) = 1$ we easily see that

$$\hat{\chi}(b) = \bar{\chi}(-b)\tau(\chi)/m \tag{2.23}$$

where the *Gauss sum* $\tau(\chi)$ is defined by

$$\tau(\chi) = \sum_{amodm} \chi(a)e^{2\pi ia/m}. \tag{2.24}$$

Substituting into the expression giving $L(\chi, 1)$ we get

$$\begin{aligned}
L(\chi, 1) &= \frac{\tau(\chi)}{m}\sum_{bmodm} \bar{\chi}(-b)\sum_{n=1}^{\infty} \frac{e^{2\pi ibn/m}}{n} \\
&= -\frac{\tau(\chi)}{m}\sum_{bmodm} \bar{\chi}(-b)\log(1-\zeta^{b}) \qquad (2.25)
\end{aligned}$$

where $\zeta = e^{2\pi i/m}$. Note that $\zeta^b \neq 1$ since if $m|b$ then $\bar{\chi}(b) = 0$. The branch of the logarithm is that branch which is analytic in the right half plane and satisfies $\log(1) = 0$.

From here on the discussion diverges according to whether $\chi(-1) = 1$ (χ is even, and as a Galois character factors through a real subfield of $\mathbb{Q}(\zeta)$) or $\chi(-1) = -1$ (χ is odd).

2.7. Odd χ. This corresponds to K_χ imaginary, $\chi(-1) = -1$. Here we group b with $-b$, so

$$\begin{aligned} L(\chi,1) &= \frac{1}{2}\frac{\tau(\chi)}{m}\sum_{b=1}^{m-1}\bar{\chi}(b)\log\left(\frac{1-\zeta^b}{1-\zeta^{-b}}\right) \\ &= \frac{1}{2}\frac{\tau(\chi)}{m}\sum_{b=1}^{m-1}\bar{\chi}(b)\log(-\zeta^b) \\ &= \frac{\pi i\tau(\chi)}{m^2}\sum_{b=1}^{m-1}\bar{\chi}(b)b. \qquad (2.26) \end{aligned}$$

2.8. Even χ. This corresponds to K_χ real. Again we group b with $-b$, this time getting

$$L(\chi,1) = -\frac{\tau(\chi)}{m}\sum_{b mod m}\bar{\chi}(b)\log|1-\zeta^b|. \qquad (2.27)$$

Suppose that χ is non-trivial, quadratic ($\chi(b) = \bar{\chi}(b) = \pm 1$) and even. We then define

$$\varepsilon_\chi = \prod_{b mod m}(1-\zeta^b)^{\chi(b)} = \prod_{b mod m}\sin\left(\frac{\pi b}{m}\right)^{\chi(b)}. \qquad (2.28)$$

Then $\varepsilon_\chi > 0$ and

$$L(\chi,1) = -\frac{\tau(\chi)}{m}\log\varepsilon_\chi. \qquad (2.29)$$

If $\sigma \in H_\chi$ then $\omega(\sigma) = a^2$ for some a so

$$\begin{aligned} \sigma(\varepsilon_\chi) &= \prod_{b mod m}(1-\zeta^{a^2 b})^{\chi(b)} \\ &= \prod_{b mod m}(1-\zeta^b)^{\chi(ba^{-2})} \\ &= \prod_{b mod m}(1-\zeta^b)^{\chi(b)} = \varepsilon_\chi. \qquad (2.30) \end{aligned}$$

It follows that $\varepsilon_\chi \in K_\chi$. In fact, it is a *unit*. This is because we can write

$$\varepsilon_\chi = \prod_{b mod m}\left(\frac{1-\zeta^b}{1-\zeta}\right)^{\chi(b)} \qquad (2.31)$$

and $(1-\zeta^b)/(1-\zeta)$ is a unit.

3. Dedekind's Zeta function

Let K be a number field, $[K:\mathbb{Q}] = n$, r_1 and r_2 as before, the number of real and pairs of complex embeddings, $r = r_1+r_2-1$ the unit rank. Let μ_K be the group of roots of unity in K, $w = |\mu_K|$, and let $\varepsilon_1,\ldots,\varepsilon_r$ be a system of representatives for $\mathcal{O}_K^\times/\mu_K$ (a system of fundamental units). Let $h = h_K$ be the class number of K, and $R = R_K$ its regulator. Let d_K be the discriminant of K.

3.1. Definition and basic properties. Recall that for every ideal $\mathfrak{a}$ of $\mathcal{O}_K$ we denoted by $\mathbb{N}\mathfrak{a} = [\mathcal{O}_K : \mathfrak{a}]$ the absolute norm of $\mathfrak{a}$. Define

$$\zeta_K(s) = \sum \frac{1}{\mathbb{N}\mathfrak{a}^s}, \tag{3.1}$$

the summations extending over all the integral ideals of K. By unique factorization of ideals, we have the Euler product

$$\zeta_K(s) = \prod_{\mathfrak{p}} \left(1 - \frac{1}{\mathbb{N}\mathfrak{p}^s}\right)^{-1}. \tag{3.2}$$

This also proves convergence for $Re(s) > 1$, absolutely and uniformly on compact sets, because we can write, for $s > 1$ real

$$\zeta_K(s) = \prod_p \prod_{\mathfrak{p}|p} \left(1 - \frac{1}{\mathbb{N}\mathfrak{p}^s}\right)^{-1} \leq \prod_p \left(1 - \frac{1}{p^s}\right)^{-n} = \zeta_{\mathbb{Q}}(s)^n. \tag{3.3}$$

Define

$$Z_K(s) = |d_K|^{s/2} \left(\pi^{-s/2}\Gamma(s/2)\right)^{r_1} \left((2\pi)^{-s}\Gamma(s)\right)^{r_2} \zeta_K(s). \tag{3.4}$$

THEOREM 3.1. *The completed Zeta function $Z_K(s)$ extends meromorphically to all $s \in \mathbb{C}$, has simple poles at $s = 0$ and $s = 1$ and nowhere else, and satisfies the functional equation*

$$Z_K(s) = Z_K(1 - s). \tag{3.5}$$

The residue at $s = 1$ is

$$\rho_K = \frac{2^{r_1} h_K R_K}{w_K} \tag{3.6}$$

and at $s = 0$ the residue is $-\rho$.

COROLLARY 3.2. *The Zeta function $\zeta_K(s)$ has a simple pole at $s = 1$ with residue*

$$\frac{2^{r_1+r_2}\pi^{r_2} h_K R_K}{\sqrt{|d_K|} w_K}. \tag{3.7}$$

Thus all the fundamental invariants of K show up in this marvelous formula. Note that the theorem generalizes what we have already seen for the Riemann Zeta function. Hecke's proof from the 1930's uses several-variable versions of Poisson summation and the Riemann theta function. Except for the complications arising from the class number and the units, the underlying ideas are very similar to the ones used by Riemann almost a century earlier. You can read the proof in any standard book on alegbaric number theory.

4. The class number formula for quadratic fields

4.1. The relation between ζ_K and Dirichlet's L functions. Let K be a quadratic field.

LEMMA 4.1. *We have*

$$\zeta_K(s) = \zeta(s) L(\chi, s) \tag{4.1}$$

where χ is the quadratic character associated to K.

PROOF. In terms of the Legendre symbol, $K \subset \mathbb{Q}(\zeta_m)$ with $m = |d_K|$ and

$$\chi(a) = \left(\frac{d_K}{a}\right). \tag{4.2}$$

We have seen in the course of the proof of the non-vanishing of $L(\chi, 1)$ that the product of the Euler factors at p of the right hand side is $(1 - p^{-s})^{-2}$ if $\chi(p) = 1$ and $(1 - p^{-2s})^{-1}$ if $\chi(p) = -1$. If $p|d_K$ then it is $(1 - p^{-s})^{-1}$. In each of the three cases this is also the product of the Euler factors in ζ_K at primes of K dividing p. This proves the desired equality. □

REMARK 4.1. *A similar relation exists between L-functions of characters of Gal$(K/\mathbb{Q})$ and ζ_K for any* abelian *number field K. The* Kronecker-Weber theorem *asserts that $K \subset \mathbb{Q}(\zeta_m)$ for some m. The minimal m with this property is called the* conductor *of K. Let*

$$H_K \subset Gal(\mathbb{Q}(\zeta_m)/\mathbb{Q}) \simeq (\mathbb{Z}/m\mathbb{Z})^\times \tag{4.3}$$

be the subgroup fixing K elementwise. The characters of Gal$(K/\mathbb{Q})$ can be identified with Dirichlet characters mod m which are trivial on H_K. To each such character we attach the corresponding primitive *Dirichlet character mod m' (for an appropriate $m'|m$) and we denote the resulting collection of Dirichlet characters simply by $\widehat{G_K}$. We then have*

$$\zeta_K(s) = \prod_{\widehat{G_K}} L(\chi, s). \tag{4.4}$$

4.2. The class number formula for quadratic imaginary fields. Let K be quadratic imaginary and χ, as before, the non-trivial character of $Gal(K/\mathbb{Q})$, viewed as a Dirichlet character mod $m = |d_K|$. We have found

$$L(\chi, 1) = \frac{\pi i \tau(\chi)}{m^2} \sum_{b=1}^{m} \chi(b) b \tag{4.5}$$

(note that $\chi = \bar{\chi}$). On the other hand, from the lemma we get

$$res_{s=1}\zeta_K(s) = L(\chi, 1). \tag{4.6}$$

Comparing, we find a formula for h_K :

$$h_K = \frac{w}{2} \frac{i\tau(\chi)}{\sqrt{|d_K|}} \frac{1}{|d_K|} \sum_{b=1}^{|d_K|} \chi(b) b. \tag{4.7}$$

One can show that $|\tau(\chi)| = \sqrt{|d_K|}$. (Exercise: prove it when d_K is prime!) This implies

THEOREM 4.2. *(analytic class number formula)*

$$h_K = \frac{w_K}{2} \frac{1}{|d_K|} \left| \sum_{b=1}^{|d_K|} \chi(b) b \right|. \tag{4.8}$$

4.3. Analytic class number formula for real quadratic fields. Assume now K is real quadratic. This time

$$L(\chi,1) = -\frac{\tau(\chi)}{m}\log(\varepsilon_\chi) \tag{4.9}$$

where ε_χ is the circular unit defined before. Comparing with $res_{s=1}\zeta_K(s)$ we get

$$\frac{-\tau(\chi)}{|d_K|}\log(\varepsilon_\chi) = \frac{h_K\log(\varepsilon_K)}{\sqrt{|d_K|}} \tag{4.10}$$

where ε_K is the fundamental unit. We have shown the following.

THEOREM 4.3. *(analytic class number formula, real quadratic case).*

$$h_K = \left|\frac{\log\varepsilon_\chi}{\log\varepsilon_K}\right|. \tag{4.11}$$

In other words, $[\mathcal{O}_K^\times : \langle\pm\varepsilon_\chi\rangle] = h_K$. The inability to seperate h_K from $\log(\varepsilon_K) = R_K$ is the main obstacle in studying class numbers of real quadratic fields.

CHAPTER 6

Appendix

1. Some facts on abelian groups

1.1. Indices.

PROPOSITION 1.1. *Let $\Lambda \subset \mathbb{Z}^n$ be the lattice spanned by the column vectors ${}^t(c_{1j}, \dots, c_{nj})$ of the matrix $C = (c_{ij})$. Then*

$$[\mathbb{Z}^n : \Lambda] = |\det(C)|. \tag{1.1}$$

PROOF. This proposition is usually proved as a corollary of the Theorem on Elementary Divisors. The latter asserts that there exists a basis $\varepsilon_1, \dots, \varepsilon_n$ of $\mathbb{Z}^n$ over $\mathbb{Z}$ (not necessarily the standard basis), and uniquely determined natural numbers $d_1, \dots, d_n$ satisfying $d_{i+1}|d_i$ (called the elementary divisors of $\Lambda \subset \mathbb{Z}^n$), such that $d_1\varepsilon_1, \dots, d_n\varepsilon_n$ is a basis of Λ. The index $[\mathbb{Z}^n : \Lambda]$ is then clearly $d_1 \cdots d_n$. (Moreover, the elementary divisors give us the structure of $\mathbb{Z}^n/\Lambda$ as a product of cyclic groups.) On the other hand if P is the matrix changing the basis $\varepsilon_1, \dots, \varepsilon_n$ of $\mathbb{Z}^n$ to the standard basis, and Q is the matrix changing the basis of Λ which is given by the columns of C to $d_1\varepsilon_1, \dots, d_n\varepsilon_n$, then in terms of matrices the Theorem on Elementary Divisors says

$$D = QCP \tag{1.2}$$

where D is the diagonal matrix with the d_i on the diagonal. Both P and Q are unimodular matrices (integral matrices with determinant ±1), so we see that $d_1 \cdot \cdot \cdot d_n = \det(D) = \pm \det(C)$.

Let us give *another* proof that does not use the Elementary Divisors Theorem. Define a function g on $M_n(\mathbb{Q})$ as follows. To define $g(C)$ consider the columns of C, and let $\Lambda(C)$ be the abelian subgroup generated by them: all their $\mathbb{Z}$-linear combinations. Multiplying by some natural number N we may assume that C is integer-valued, so that $N \cdot \Lambda(C) \subset \mathbb{Z}^n$. We then set

$$g(C) = \pm[\mathbb{Z}^n : N \cdot \Lambda(C)]/N^n. \tag{1.3}$$

Here we use the convention that the index is 0 if $N \cdot \Lambda(C)$ is not of finite index. The ± sign is the sign of $\det(C)$. It is clear that $g(C)$ is independent of N, because $[\mathbb{Z}^n : NM \cdot \Lambda] = M^n[\mathbb{Z}^n : N \cdot \Lambda]$.

We now check that (i) g gets the value 1 on the identity matrix.

(ii) g changes sign if we permute two columns of C.

(iii) $g(C') = g(C)$ if C' is obtained from C by adding the j^{th} column to the i^{th} column for $i \neq j$.

(iv) $g(C') = Ng(C)$ if C' is obtained from C by multiplying a certain column by an integer N.

All these assertions are very easy to check straight from the definitions. However, it is known that the only function satisfying them is $\det(C)$ (this is how we compute determinants by Gauss' elementary substitutions.) □

2. Complements from Galois theory

2.1. Artin's theorem on linear independence of characters.

THEOREM 2.1. *Let K and L be any two fields, and $\chi_1, \dots, \chi_m : K^\times \to L^\times$ distinct mutiplicative homomorphisms (characters). Let $c_i \in L$. Then if*

$$c_1\chi_1(x) + \cdots + c_m\chi_m(x) = 0 \tag{2.1}$$

holds for all $x \in K$, $x \neq 0$, all $c_i = 0$.

PROOF. Assume such linear dependencies between characters exist, and pick a shortest one (smallest m). Clearly all the c_i are then nonzero. Let $a \in K^\times$ be such that $\chi_1(a) \neq \chi_2(a)$. Sustituting ax for x and using the multiplicativity of the χ_i we get the relation

$$c_1\chi_1(a)\chi_1 + \cdots + c_m\chi_m(a)\chi_m = 0. \tag{2.2}$$

Multiplying the original relation by $\chi_1(a)$ and subtracting from this second one we get

$$c_2(\chi_2(a) - \chi_1(a))\chi_2 + \cdots + c_m(\chi_m(a) - \chi_1(a))\chi_m = 0 \tag{2.3}$$

which is a shorter non-trivial relation. This contradiction proves the theorem. □

COROLLARY 2.2. *Let K/F be a separable extension of degree n, and let $\sigma_1, \dots, \sigma_n$ be n distinct embeddings of K in another extension L/F, which are the identity on F. Let $\omega_1, \dots, \omega_n$ be a basis of K over F. Then*

$$\det(\sigma_i(\omega_j)) \neq 0. \tag{2.4}$$

PROOF. If not, there is a linear dependence between the rows of the matrix

$$\sum_{i=1}^{n} c_i\sigma_i(\omega_j) = 0. \tag{2.5}$$

Since the σ_i are F-linear, and the ω_j form a basis of K over F, it follows that $\sum c_i\sigma_i = 0$ identically on K. This contradicts Artin's theorem. □

2.2. Norm and Trace. Let L/K be a finite separable field extension and embed it in a Galois extension M/K. Let $\Gamma = Emb_K(L, M)$ be the set of $n = [L : K]$ embeddings of L into M over K.

PROPOSITION 2.3. *If $a \in L$, then $N_{L/K}(a) = \prod_{\sigma\in\Gamma} \sigma(a)$ and $Tr_{L/K}(a) = \sum_{\sigma\in\Gamma} \sigma(a)$.*

PROOF. Let $M \otimes_K L$ be the tensor product of the two fields over K. Recall that if $L = \sum_{i=1}^n K\omega_i$ (a direct sum) then as a vector space

$$M \otimes_K L = \sum_{i=1}^{n} M\omega_i \tag{2.6}$$

(a direct sum) and the product structure is defined by using the product in M between the coefficients and the formulas $\omega_i\omega_j = \sum_{k=1}^n c_{ij}^k \omega_k$ ($c_{ij}^k \in K$). Each embedding $\sigma : L \hookrightarrow M$ extends M-linearly to an M-algebra homomorphism

$$\sigma : M \otimes_K L \to M, \qquad \sigma(\sum_{i=1}^n m_i\omega_i) = \sum_{i=1i}^n m_i\sigma(\omega_i). \tag{2.7}$$

Taken together we get an M-algebra homomorphism

$$\iota : M \otimes_K L \to M^\Gamma \simeq M^n \tag{2.8}$$

where M^Γ is the ring of maps from Γ to M, with pointwise addition and multiplication. Both $M \otimes_K L$ and M^Γ are n-dimensional over M, and the matrix of this map with repsect to $\{\omega_j\}$ as a basis of $M \otimes_K L$ and the standard basis of M^Γ, is $(\sigma_i(\omega_j))$ where we have written $\Gamma = \{\sigma_1, \dots, \sigma_n\}$. Artin's theorem on independence of characters shows that it is nonsingular, hence ι is an isomorphism.

Let now $a \in L$ and consider the map "multiplication by a". Via the isomorphism ι it coresponds to the map which sends $(x_\sigma)_{\sigma\in\Gamma}$ to $(\sigma(a)x_\sigma)_{\sigma\in\Gamma}$, and is therefore represented by the diagonal matrix with $\sigma_i(a)$, $1 \le i \le n$, on the diagonal. Taking the determinant and the trace of this linear transformation, the proposition follows. □

EXERCISE 2.1. *Prove that* $N_{L/K}(ab) = N_{L/K}(a)N_{L/K}(b)$ *and* $Tr_{L/K}(a+b) = Tr_{L/K}(a) + Tr_{L/K}(b)$.

EXERCISE 2.2. *Let* K/F *be a finite extension too. Prove that* $N_{L/F} = N_{K/F} \circ N_{L/K}$ *and* $Tr_{L/F} = Tr_{K/F} \circ Tr_{L/K}$ *(you may assume* L/F *is separable to use the proposition above, but the result holds without this assumption).*

EXERCISE 2.3. *Find the generalization of the above proposition to non-separable extensions.*

2.3. The trace form. Let L/K be a finite extension of fields, and define

$$B(x,y) = Tr_{L/K}(xy). \tag{2.9}$$

This is a symmetric K-bilinear form on L.

EXERCISE 2.4. *Prove that* B *is non-degenerate if and only if* L/K *is separable.*

EXERCISE 2.5. *Let* $O(L/K, B)$ *be the group of all linear transformations of* L *over* K *that preserve* B*. Prove that if* L/K *is Galois, it admits a representation into* $O(L/K, B)$.

CHAPTER 7

Review Problems

1) Let p be a rational prime. Prove that the equation $4p = x^2 + 27y^2$ has a unique solution $(x, y) \in \mathbb{Z}^2$, $x > 0$, $y > 0$ if $p \equiv 1 mod 3$, and no solution if $p \equiv 2 mod 3$. (*Hint*: Prove that $\mathbb{Z}[\omega]$, $\omega = \frac{-1+\sqrt{-3}}{2}$, is a unique factorization domain.)

2) Let p be a prime, $a \in \mathbb{Z}$, and assume that a is not a pth power in $\mathbb{Z}$. Let $\theta^p = a$. Prove that $[\mathbb{Q}(\theta) : \mathbb{Q}] = p$, that primes l not dividing pa are unramified in $\mathbb{Q}(\theta)$, and that primes $l|a$ for which $ord_l(a)$ is not divisible by p are fully ramified in $\mathbb{Q}(\theta)$.

3) (i) Let K be an algebraic number field of degree n, $\omega_i \in \mathcal{O}_K$ $(1 \leq i \leq n)$ and suppose that the discriminant $\Delta(\omega_1, \ldots, \omega_n)$ is square-free. Prove that the ω_i are a basis for $\mathcal{O}_K$ over $\mathbb{Z}$.

(ii) Let α be a solution of $\alpha^3 - \alpha - 1 = 0$. Let $K = \mathbb{Q}(\alpha)$. Show that $\Delta(1, \alpha, \alpha^2) = -23$ and conclude that $\mathcal{O}_K = \mathbb{Z}[\alpha]$.

(iii) Find the prime factorization in K of the primes 2, 5 and 23.

4) Let K be an algebraic number field of degree n, $\alpha \in \mathcal{O}_K$ a generator of K, and suppose that the minimal polynomial of α is Eisenstein with respect to the prime p. Prove

(i) p does not divide $[\mathcal{O}_K : \mathbb{Z}[\alpha]]$.

(ii) p is fully ramified in K.

Hint for (i): Suppose that $p^{-1} \sum_{i=0}^{n-1} c_i \alpha^i$ $(c_i \in \mathbb{Z})$ is integral, but not all c_i are divisible by p. Multiply by a suitable power of α to get an integral element of such a form with $p|c_i$ for $i < n-1$, p not dividing c_{n-1}. Conclude that $p^{-1} c_{n-1} \alpha^{n-1}$ is integral. Take the norm to $\mathbb{Q}$ to get a contradiction.

An alternative hint for (i): Let v be a valuation on K extending the p-adic valuation on $\mathbb{Q}$, normalized by $\nu(p) = 1$. Prove that $\nu(\alpha) = 1/n$, and that $\nu(\sum_{i=0}^{n-1} c_i \alpha^i) = \inf \left\{ \frac{i}{n} + \nu(c_i) \right\}$ $(c_i \in \mathbb{Q})$.

5) Let K and F be two number fields and assume that their discriminants are relatively prime.

(i) Let L be the normal closure of K. Prove that d_L is divisible only by the primes dividing d_K, hence is still relatively prime to d_F.

(ii) Prove that $[LF : \mathbb{Q}] = [L : \mathbb{Q}][F : \mathbb{Q}]$.

(iii) Prove that $[KF : \mathbb{Q}] = [K : \mathbb{Q}][F : \mathbb{Q}]$.

Why was it necessary to pass from K to L and back to K?

6) Show that every ideal in a Dedekind domain is generated by two elements.

7) Let K be an algebraic number field of discriminant d_K.

(i) Prove that the normal closure of K contains $\sqrt{d_K}$.

(ii) Obtain from (i) a formula for the quadratic subfield of $\mathbb{Q}(\zeta_p)$, where $\zeta_p = e^{2\pi i/p}$.

8) Show that the class number of $\mathbb{Q}(\sqrt{-19})$ is 1.

9) Show that the class number of $\mathbb{Q}(\sqrt{10})$ is 2, and find representatives of the class group.

10) Let D be a square free positive integer, $D \equiv 5 mod 12$. Prove that $(3) = \mathfrak{p}\mathfrak{p}'$ splits in $K = \mathbb{Q}(\sqrt{-D})$ and that if $D > 3^n$ then the order of the class of $\mathfrak{p}$ in Cl_K is at least n. Conclude that the class number of quadratic imaginary fields is unbounded.

11) Let $p_1, \ldots, p_n$ be distinct primes congruent to $1 mod 4$ and $m = \prod p_i$. Which primes ramify in $K = \mathbb{Q}(\sqrt{-m})$? Use that to show that the class number h_K is divisible by 2^{n-1}.

12) Let F be a totally real field (all the embeddings of F are real) and $d \in F$ a totally negative element ($\sigma(d) < 0$ for every embedding σ of F in $\mathbb{R}$). Let $K = F(\sqrt{d})$.

(i) Show that the ranks of the unit groups of F and K are equal.

(iii) Let μ_K be the group of roots of unity in K. Prove that $[\mathcal{O}_K^\times : \mu_K \mathcal{O}_F^\times] = 1$ or 2. (*Hint:* consider the map sending a unit u to $u/\bar{u}$ and prove that its image consists of roots of unity.)

13) Let $S = \{P_1, \ldots, P_s\}$ be a finite set of primes in K. An S-unit is $u \in K^\times$ such that the fractional ideal $(u) = P_1^{e_1} \ldots P_s^{e_s}$ for some $e_i \in \mathbb{Z}$. Show (i) the S-units form a multiplicative group $\mathcal{O}_{K,S}^\times$. (ii) The e_i define a homomorphism from $\mathcal{O}_{K,S}^\times$ to $\mathbb{Z}^s$ whose kernel is $\mathcal{O}_K^\times$. (iii) The group $\mathcal{O}_{K,S}^\times$ is a finitely generated abelian group whose rank is $r_1(K) + r_2(K) + s - 1$.

14) Let $\zeta = e^{2\pi i/7}$ and $K = \mathbb{Q}(\zeta)$. Consider triples of integers (a, b, c) such that $a + b + c = 0$ and let

$$\varepsilon_{a,b,c} = (\zeta - 1)^a (\zeta^2 - 1)^b (\zeta^4 - 1)^c. \tag{0.10}$$

(i) Show that these elements form a group C of units in K (it is called the group of *circular* units).

(ii) Show that U is of finite index in $\mathcal{O}_K^\times$. (*Hint*: let $\sigma \in Gal(K/\mathbb{Q})$ be defined by $\sigma(\zeta) = \zeta^2$. What are $\sigma(\varepsilon_{a,b,c})$ and $\sigma^2(\varepsilon_{a,b,c})$? Consider the logarithmic embedding

$$\lambda(\varepsilon) = (\log|\varepsilon|, \log|\sigma(\varepsilon)|, \log|\sigma^2(\varepsilon)|) \tag{0.11}$$

into $\{(x_1, x_2, x_3);\ x_1 + x_2 + x_3 = 0\}$ and prove that U spans a lattice there.)

15) Let K be a Galois extension of $\mathbb{Q}$ whose Galois group is S_3 (the symmetric group on 3 letters). What are all the possible factorizations in K of an unramified prime p? In each case, what are the decomposition fields of the various primes above p?