Yichao TIAN

LECTURES ON ALGEBRAIC NUMBER THEORY

Yichao TIAN
Morningside Center of Mathematics, 55 Zhong Guan Cun East Road,
Beijing, 100190, China.
E-mail : yichaot@math.ac.cn

LECTURES ON ALGEBRAIC NUMBER THEORY

Yichao TIAN

CONTENTS

CHAPTER 1

NUMBER FIELDS AND ALGEBRAIC INTEGERS

1.1. Algebraic integers

All the rings in this section are supposed to be commutative.

Definition 1.1.1. — Let $A \subset B$ be an extension of rings. We say an element $x \in B$ is integral over A if there exists a monic polynomial $f(T) = T^n + a_1 T^{n-1} + \cdots + a_n \in A[T]$ such that $f(x) = 0$. We say B is integral over A, if every $x \in B$ is integral over A.

Example 1.1.2. — (1) $\mathbb{Z}[i]$ is integral over $\mathbb{Z}$.

(2) Let L/K be an extension of fields. Then L is integral over K if and only if L/K is an algebraic extension.

Proposition 1.1.3. — *Let $A \subset B$ be an extension of rings, $x \in B$. Then the following statements are equivalent:*

1. *x is integral over A.*
2. *the subring $A[x] \subset B$ is a finite generated A-module.*
3. *x belongs to a subring $B' \subset B$ such that B' is finitely generated as an A-module.*

Proof. — (1) $\Longrightarrow$ (2) $\Longrightarrow$ (3) is trivial. We prove now (3) $\Longrightarrow$ (1). Choose generators $\alpha_1, \cdots, \alpha_n$ of the A-module B'. Since $xB' \subset B'$, there exists a $U \in \mathrm{M}_{n\times n}(A)$ such that

$$x(\alpha_1, \cdots, \alpha_n) = (\alpha_1, \cdots, \alpha_n)U \Longleftrightarrow (\alpha_1, \cdots, \alpha_n)(xI_n - U) = 0.$$

Let V be the cofactor matrix of $xI_n - U$. Then one has

$$(\alpha_1, \cdots, \alpha_n)(xI_n - U)V = (\alpha_1, \cdots, \alpha_n)\det(xI_n - U) = 0.$$

As $1 \in B'$ is a linear combination of α_i's, we get $\det(xI_n - U) = x^n + a_1 x^{n-1} + \cdots + a_n = 0$. □

Corollary 1.1.4. — *Let $A \subset B$ be extensions of rings. Then the elements of B which are integral over A form a subring of B.*

Proof. — Given $x, y \in B$ integral over A, we need to show that $x + y$ and xy are also integral over A. Actually, one sees easily that $A[x, y]$ is a finitely generated A-module, and concludes using Proposition 1.1.3(3). □

Corollary 1.1.5. — *Let $A \subset B \subset C$ be extensions of rings. Then C is integral over A if and only if C is integral over B and B is integral over A.*

Proof. — The "only if" part is easy. Prove now that the inverse implication holds. Let $x \in C$. Since C is assumed integral over B, we have $f(x) = x^n + b_1 x^{n-1} + \cdots + b_n = 0$ for some $b_1, \cdots, b_n \in B$. Note that $b_1, \cdots, b_n$ are all integral over A by assumption. One proves easily by induction that $A[b_1, \cdots, b_i]$ is a finitely generated A-module for all $1 \leq i \leq n$. Then $A[b_1, \cdots, b_n, x]$ is a quotient of $A[b_1, \cdots, b_n][T]/(f(T))$, hence it is also finitely generated as A-module. One concludes with Proposition 1.1.3(3). □

Definition 1.1.6. — (1) Let $A \subset R$ be an extension of rings. Define the *integral closure* of A in R to be the subring of R consisting of all integral elements over A.

(2) If the integral closure of A in R is A, we say A is *integrally closed in R.*

(3) Assume A is an integral domain. We say A is *integrally closed* if A is integrally closed in its fraction field.

Example 1.1.7. — (1) $\mathbb{Z}$ is integrally closed. Indeed, let $x = \frac{a}{b} \in \mathbb{Q}$ with $\gcd(a, b) = 1$ and $b > 0$. If x is integral over $\mathbb{Z}$, then there exist some $c_1, \cdots, c_n \in \mathbb{Z}$ such that

$$x^n + c_1 x^{n-1} + \cdots + c_n = 0 \Longleftrightarrow a^n + c_1 a^{n-1} b + \cdots + c_n b^n = 0.$$

If $b \neq 1$, let p denote a prime dividing b. Then the equality above implies that $p | a^n$, hence $p | a$. This contradicts with $\gcd(a, b) = 1$.

(2) Let $\omega = e^{\frac{2\pi i}{3}}$. Then $\mathbb{Z}[\omega]$ is integral over $\mathbb{Z}$ and integrally closed by similarly arguments as above. Actually, every principal ideal domain is integrally closed.

Definition 1.1.8. — (1) An element $x \in \mathbb{C}$ is an *algebraic number* (resp. an *algebraic integer*) if x is integral over $\mathbb{Q}$ (resp. over $\mathbb{Z}$).

(2) A number field is a finite extension of $\mathbb{Q}$. For a number field $K/\mathbb{Q}$, we define $\mathcal{O}_K$ to be the integral closure of $\mathbb{Z}$ in K, and call it the *ring of integers* of K.

Example 1.1.9. — The ring of Gauss integers $\mathbb{Z}[i]$ is the ring of integers of $\mathbb{Q}(i)$, and $\mathbb{Z}[\omega]$ with $\omega = -\frac{1}{2} + \frac{\sqrt{-3}}{2}$ is the ring of integers of $\mathbb{Q}(\sqrt{-3})$.

Proposition 1.1.10. — *Let $x \in \mathbb{C}$ be an algebraic number, and $f(T) = T^n + a_1 T^{n-1} + \cdots + a_n \in \mathbb{Q}[T]$ be its minimal polynomial. Then x in an algebraic integer if and only if $f(T) \in \mathbb{Z}[T]$.*

Proof. — One side implication is clear. Assume now that x is an algebraic integer. Let $\{x = x_1, \cdots, x_n\}$ be the set of complex roots of $f(T)$. We claim that each x_i is also an algebraic integer. Indeed, for each x_i, there exists an embedding of fields $\iota : K = \mathbb{Q}(x) \hookrightarrow \mathbb{C}$ such that $\iota(x) = x_i$. Therefore, if x satisfies $g(x) = 0$ for some monic polynomial $g(T) \in \mathbb{Z}[T]$, then $g(x_i) = \iota(g(x)) = 0$. This proves the claim. But each a_i is a symmetric function of x_j's. It follows that $a_i \in \mathcal{O}_K \cap \mathbb{Q} = \mathbb{Z}$. □

Example 1.1.11. — Let $K = \mathbb{Q}(\sqrt{D})$ with an integer $D \neq 1$ square free. Then we have $\mathcal{O}_K = \mathbb{Z} + \mathbb{Z}\omega_D$, where

$$\omega_D = \begin{cases} \frac{1+\sqrt{D}}{2} & \text{if } D \equiv 1 \mod 4 \\ \sqrt{D} & \text{if } D \equiv 2, 3 \mod 4. \end{cases}$$

It is easy to check that ω_D is an integer. It remains to show that if $x = a + b\sqrt{D}$ with $a, b \neq 0 \in \mathbb{Q}$ is an integer, then $x \in \mathbb{Z} + \mathbb{Z}\omega_D$. Indeed, the minimal polynomial of x over $\mathbb{Q}$ is $T^2 - 2aT + (a^2 - b^2D) = 0$. By the previous Proposition, for x to be an integer, one must have $2a, a^2 - b^2D \in \mathbb{Z}$. If $a \in \mathbb{Z}$, then b must be in $\mathbb{Z}$. Otherwise, if a is a half-integer, then b has to be an half-integer as well and $D \equiv 1 \mod 4$.

1.2. Traces and norms

Definition 1.2.1. — Let L/K be a finite extension of fields, and $x \in L$. We view L as a finite dimensional K-vector space, and denote by

$$\phi_x : L \to L$$

the K-linear endomorphism on L defined by the multiplication by x. We have $\phi_x \in \mathrm{End}_K(L)$. We put $\mathrm{Tr}_{L/K}(x) = \mathrm{Tr}(\phi_x)$, and call it the *trace of x (relative to L/K)*; put $\mathrm{N}_{L/K}(x) = \det(\phi_x)$, and call it the *norm of x (relative to L/K)*.

Lemma 1.2.2. — *Let L/K be a finite extension of fields, and $x \in L$.*

1. *One has*
$$\mathrm{Tr}_{L/K}(x) = [L : K(x)]\mathrm{Tr}_{K(x)/K}(x) \quad \textit{and} \quad \mathrm{N}_{L/K}(x) = \mathrm{N}_{K(x)/K}(x)^{[L:K(x)]}.$$
2. *If $f(T) = T^n + a_1T^{n-1} + \cdots a_n \in K[T]$ is the minimal polynomial of x over K, then $\mathrm{Tr}_{K(x)/K}(x) = -a_1$ and $\mathrm{N}_{K(x)/K}(x) = (-1)^n a_n$.*

Proof. — Exercise. □

Proposition 1.2.3. — *Let L/K be a finite separable extension of fields, and $n = [L : K]$. Fix an algebraically closed field Ω and an embedding $\tau : K \hookrightarrow \Omega$. Then*

1. *there exists exactly n distinct embeddings $\sigma_1, \cdots, \sigma_n : L \hookrightarrow \Omega$ such that $\sigma_i|_K = \tau$ for $1 \leq i \leq n$;*
2. *the n embeddings $\sigma_1, \cdots, \sigma_n$ are linearly independent over Ω.*

Proof. — (1) By induction on n, one reduces to the case where $L = K(x)$ for some $x \in L$. In this case, let $f(T) = T^n + a_1T^{n-1} + \cdots a_n$ be the minimal polynomial of x over K so that $L \cong K[T]/(f(T))$. Put $f^\tau(T) = T^n + \tau(a_1)T^{n-1} + \cdots + \tau(a_n) \in \Omega[T]$, and let $\alpha_1, \cdots, \alpha_n \in \Omega$ be the roots of $f^\tau(T)$. Then the α_i's must be distinct (because $f(T)$ is separable). For each α_i, there exists a unique embedding $\sigma_i : L \hookrightarrow \Omega$ extending τ such that $\sigma_i(x) = \alpha_i$. Conversely, if $\sigma : L \hookrightarrow \Omega$ is an extension of τ, then it must send x to some α_i, hence it must coincide with one of the σ_i's.

(2) The statement is trivial if $n = 1$. Suppose now $n \geq 2$ and in contrary that $\sigma_1, \cdots, \sigma_n$ are linearly independent over Ω. Up to renumbering, we may assume that $\sum_{i=1}^d c_i\sigma_i = 0$

is a linearly relation with $c_i \in \Omega^\times$ such that $d \geq 2$ is minimal. Thus for any $x \in L$, we have $\sum_{i=1}^d c_i\sigma_i(x) = 0$. By dividing c_1, we may assume that $c_1 = 1$. Choose $y \in L$ such that $\sigma_2(y) \neq \sigma_1(y)$. This is possible since σ_1 and σ_2 are distinct. Then $\sum_{i=1}^d c_i\sigma_i(xy) = \sum_{i=1}^d c_i\sigma_i(y)\sigma_i(x) = 0$ for all $x \in L$. Therefore, one obtains

$$\sum_{i=2}^{d} c_i(\sigma_i(y) - \sigma_1(y))\sigma_i(x) = 0, \quad \forall x \in L.$$

This is a non-zero linear relation among σ_i's of length at most $d-1$, whose existence contradicts with the minimality of d. □

Theorem 1.2.4. — *Let L/K be a finite separable extension of fields, τ and σ_i for $1 \leq i \leq n$ be the embeddings as in Proposition 1.2.3. We identify K with its image in Ω via τ. Then one has*

$$\mathrm{Tr}_{L/K}(x) = \sum_{i=1}^{n} \sigma_i(x) \quad \textit{and} \quad \mathrm{N}_{L/K}(x) = \prod_{i=1}^{n} \sigma_i(x), \quad \textit{for all } x \in L.$$

Moreover, the K-bilinear form $L \times L \to K$ given by

$$(x, y) \mapsto \mathrm{Tr}_{L/K}(xy) \quad \textit{for all } x, y \in L$$

is non-degenerate, i.e. if $x \in L$ such that $\mathrm{Tr}_{L/K}(xy) = 0$ for all $y \in L$, then $x = 0$.

Proof. — By the construction of σ_i's in Proposition 1.2.3, $\mathrm{Tr}_{L/K}(x)$ and $\mathrm{N}_{L/K}(x)$ are easily verified if $L = K(x)$. The general case follows from Lemma 1.2.2(1). The non-degeneracy of the K-bilinear form $\mathrm{Tr}_{L/K}(xy)$ follows immediately from Proposition 1.2.3(2). □

Remark 1.2.5. — If L/K is inseparable, then the pairing $\mathrm{Tr}_{L/K}$ is no longer non-degenerate. For instance, if $K = \mathbb{F}_p(x)$ and $L = \mathbb{F}_p(x^{1/p})$, then $\mathrm{Tr}_{L/K}(x) = 0$ for all $x \in L$.

Corollary 1.2.6. — *Let L/K be a separable extension of degree n, and $\alpha_1, \cdots, \alpha_n \in L$. Then $(\alpha_1, \cdots, \alpha_n)$ is a K-basis of L if and only if* $\det(\mathrm{Tr}_{L/K}(\alpha_i\alpha_j)) \neq 0$.

Proof. — Consider the morphisms:

$$K^n \xrightarrow{\phi} L \xrightarrow{\psi} K^n$$

given respectively by $\phi : (x_i)_{1\leq i\leq n} = \sum_i x_i\alpha_i$ and $\psi : x \mapsto (\mathrm{Tr}_{L/K}(x\alpha_i))_{1\leq i\leq n}$. Then the matrix of $\psi\circ\phi$ under the natural basis of K^n is $(\mathrm{Tr}_{L/K}(\alpha_i\alpha_j))_{1\leq i,j\leq n}$. If $(\alpha_i)_{1\leq i\leq n}$ is a basis of L, then ϕ is an isomorphism by definition, and ψ is injective (hence bijective) by the non-degeneracy of $\mathrm{Tr}_{L/K}(xy)$. It follows that $\psi\circ\phi$ is an isomorphism, thus $\det(\mathrm{Tr}_{L/K}(\alpha_i\alpha_j)) \neq 0$. Conversely, if $\det(\mathrm{Tr}_{L/K}(\alpha_i\alpha_j)) \neq 0$, then $\psi \circ \phi$ is an isomorphism. It follows that ϕ is injective, hence bijective since L has the same K-dimension as K^n. □

Given a basis $(\alpha_i)_{1\leq i\leq n}$ of L over K. Let $C=(c_{ij})_{1\leq i,j\leq n}$ denote the inverse matrix of $(\mathrm{Tr}_{L/K}(\alpha_i\alpha_j))_{1\leq i,j\leq n}$, and put $\alpha_i^\vee=\sum_{k=1}^n \alpha_k c_{ki}$ for $1\leq i\leq n$. Then one checks easily that

$$\mathrm{Tr}_{L/K}(\alpha_i\alpha_j^\vee)=\begin{cases}1 & \text{if } i=j\\ 0 & \text{otherwise.}\end{cases}$$

We call $(\alpha_i^\vee)_{1\leq i\leq n}$ *the dual basis of* $(\alpha_i)_{1\leq i\leq n}$ *with respect to* $\mathrm{Tr}_{L/K}$. For any $x\in L$, if we write $x=\sum_i x_i\alpha_i$, then $x_i=\mathrm{Tr}_{L/K}(x\alpha_i)$; similarly if we write $x=\sum_i y_i\alpha_i$, then $y_i=\mathrm{Tr}_{L/K}(x\alpha_i)$.

1.3. Discriminants and integral basis

We apply the theory of previous section to the case of number fields. In this section, let K denote a number field of degree $n=[K:\mathbb{Q}]$, and $\mathcal{O}_K$ be its ring of integers. For $\alpha_1,\cdots,\alpha_n\in K$, we put

$$\mathrm{Disc}(\alpha_1,\cdots,\alpha_n)=\det(\mathrm{Tr}_{K/\mathbb{Q}}(\alpha_i\alpha_j)),$$

and call it the *discriminant of* $\alpha_1,\cdots,\alpha_n$.

Lemma 1.3.1. — (1) *The elements $\alpha_1,\cdots,\alpha_n$ form a basis of K over $\mathbb{Q}$ if and only if* $\mathrm{Disc}(\alpha_1,\cdots,\alpha_n)\neq 0$.

(2) *If $\sigma_1,\cdots\sigma_n$ denote the n distinct complex embeddings of K given by Proposition 1.2.3, then*

$$\mathrm{Disc}(\alpha_1,\cdots,\alpha_n)=\det(\sigma_i(\alpha_j))^2.$$

(3) *If $C\in \mathrm{M}_{n\times n}(\mathbb{Q})$ and $(\beta_1,\cdots,\beta_n)=(\alpha_1,\cdots,\alpha_n)C$, then*

$$\mathrm{Disc}(\beta_1,\cdots,\beta_n)=\mathrm{Disc}(\alpha_1,\cdots,\alpha_n)\det(C)^2.$$

Proof. — Statement (1) is Corollary 1.2.6. For (2), one deduces from Theorem 1.2.4 that

$$\mathrm{Tr}_{K/\mathbb{Q}}(\alpha_i\alpha_j)=\sum_{k=1}^n\sigma_k(\alpha_i\alpha_j)=\sum_k\sigma_k(\alpha_i)\sigma_k(\alpha_j).$$

Hence, if A denotes the matrix $(\mathrm{Tr}_{K/\mathbb{Q}}(\alpha_i\alpha_j))_{1\leq i,j\leq n}$ and $U=(\sigma_i(\alpha_j))_{1\leq i,j\leq n}$, then $A=U^T\cdot U$. Statement (2) follows immediately. Let B denote $(\mathrm{Tr}_{K/\mathbb{Q}}(\beta_i\beta_j))_{1\leq i,j\leq n}$. Then one has $B=C^T\cdot A\cdot C$, and hence (3) follows. □

Proposition 1.3.2. — *Let α be an arbitrary element of K, and $f(T)\in\mathbb{Q}[T]$ be its minimal polynomial. Then one has*

$$\mathrm{Disc}(1,\alpha,\cdots,\alpha^{n-1})=\begin{cases}0 & \text{if } \deg(f)<n,\\ (-1)^{\frac{n(n-1)}{2}}\mathrm{N}_{K/\mathbb{Q}}(f'(\alpha)) & \text{if } \deg(f)=n.\end{cases}$$

Proof. — Since $(1, \alpha, \cdots, \alpha^{n-1})$ form a basis of K if and only if $\deg(f) = n$, the first case follows from Lemma 1.3.1(1). Assume now $\deg(f) = n$. Denote by $\sigma_1, \cdots, \sigma_n$ the complex embeddings of K. By Lemma 1.3.1(2), one has

$$\mathrm{Disc}(1, \alpha, \cdots, \alpha^{n-1}) = \det(\sigma_i(\alpha^{j-1})_{1\leq i,j\leq n})^2 = \prod_{i<j}(\sigma_i(\alpha) - \sigma_j(\alpha))^2,$$

where the last equality uses Vandermonde's determinant formula. The Proposition then follows from

$$\mathrm{N}_{K/\mathbb{Q}}(f'(\alpha)) = \prod_{i=1}^{n} \sigma(f'(\alpha)) = \prod_{i=1}^{n}\prod_{j\neq i}(\sigma_i(\alpha) - \sigma_j(\alpha)).$$

□

Theorem 1.3.3. — *The ring of integers $\mathcal{O}_K$ is a free abelian group of rank n.*

Proof. — Choose a basis $(\alpha_i)_{1\leq i\leq n}$ of K over $\mathbb{Q}$. Up to multiplying by an integer, we may assume that $\alpha_i \in \mathcal{O}_K$. Consider the abelian subgroup $M \subset \mathcal{O}_K$ generated by the α_i's. Let $(\alpha_i^\vee)_{1\leq i\leq n}$ be the dual basis of $(\alpha_i)_{1\leq i\leq n}$ with respect to $\mathrm{Tr}_{K/\mathbb{Q}}$, and put $M^\vee = \sum_{i=1}^n \mathbb{Z}\alpha_i^\vee$ as a abelian subgroup of K. It is easy to see that

$$M^\vee = \{x \in K | \mathrm{Tr}_{K/\mathbb{Q}}(xy) \in \mathbb{Z}, \quad \forall y \in M\}.$$

Thus $M \subset \mathcal{O}_K \subset M^\vee$, and one checks easily that $M^\vee/M$ is finite with cardinality $|\mathrm{Disc}(\alpha_1, \cdots, \alpha_n)|$. Since $M^\vee$ is a $\mathbb{Z}$-module free of rank n, the Theorem follows immediately. □

Definition 1.3.4. — A basis $(\alpha_1, \cdots, \alpha_n)$ of K over $\mathbb{Q}$ is called an *integral basis* if it is a basis of $\mathcal{O}_K$ over $\mathbb{Z}$.

Proposition 1.3.5. — *Let $(\alpha_1, \cdots, \alpha_n)$ be an integral basis of K, and $(\beta_1, \cdots, \beta_n)$ be an arbitrary n-tuple of elements in $\mathcal{O}_K$ which form a basis of $K/\mathbb{Q}$. Then* $\mathrm{Disc}(\beta_1, \cdots, \beta_n)$ *equals to* $\mathrm{Disc}(\alpha_1, \cdots, \alpha_n)$ *times a square integer. In particular, if $(\beta_1, \cdots, \beta_n)$ is also an integral basis, if and only if* $\mathrm{Disc}(\beta_1, \cdots, \beta_n) = \mathrm{Disc}(\alpha_1, \cdots, \alpha_n)$.

Proof. — Write each β_i as a $\mathbb{Z}$-linear combination of α_j's, then there exists a matrix $C \in \mathrm{M}_{n\times n}(\mathbb{Z})$ with $\det(C) \neq 0$ and $(\beta_1, \cdots, \beta_n) = (\alpha_1, \cdots, \alpha_n)\cdot C$. Note that $\det(C) \in \mathbb{Z}$, then the Proposition follows from Lemma 1.3.1. □

Definition 1.3.6. — The discriminant of K, denoted by $\Delta_K \in \mathbb{Z}$, is the discriminant of an integral basis of K.

By the previous Lemma, this definition does not depend on the choice of the integral basis. For instance, if $K = \mathbb{Q}(\sqrt{D})$ where $D \neq 0, 1$ is a square free integer, then $(1, \omega_D)$ considered in Example 1.1.11 is an integral basis of K. Thus $\Delta_K = \mathrm{Disc}(1, \omega_D)$ which equals to $4D$ if $D \equiv 2, 3 \mod 4$, and to D if $D \equiv 1 \mod 4$.

We now give a practical criterion for n-elements of $\mathcal{O}_K$ to be an integral basis.

Lemma 1.3.7. — *Let $\beta_1, \cdots, \beta_n$ be n elements of $\mathcal{O}_K$ which form a basis of K. Then $(\beta_1, \cdots, \beta_n)$ is not an integral basis if and only if there exists a rational prime p with $p^2|\mathrm{Disc}(\beta_1, \cdots, \beta_n)$ and some $x_i \in \{0, 1, \cdots, p-1\}$ for $1 \leq i \leq n$ such that not all of x_i are zero and $\sum_{i=1}^n x_i\beta_i \in p\mathcal{O}_K$.*

Proof. — Choose an integral basis $(\alpha_1, \cdots, \alpha_n)$ and write $(\beta_1, \cdots, \beta_n) = (\alpha_1, \cdots, \alpha_n) \cdot C$ for some matrix $C \in \mathrm{M}_{n\times n}(\mathbb{Z})$ with $\det(C) \neq 0$. Then $(\beta_1, \cdots, \beta_n)$ is an integral basis if and only if $\det(C) = \pm 1$. Assume $(\beta_1, \cdots, \beta_n)$ is not an integral basis. Let p be a prime dividing $\det(C)$. Then $p^2|\mathrm{Disc}(\beta_1, \cdots, \beta_n) = \det(C)^2\Delta_K$. Denote by $\bar{C}$ the reduction of C modulo p. Let $(\bar{x}_1, \cdots, \bar{x}_n)^T \in \mathbb{F}_p^n$ be a non-zero column vector such that $\bar{C}(\bar{x}_1, \cdots, \bar{x}_n)^T = 0$. If x_i denotes the unique lift of $\bar{x}_i$ in $\{0, 1, \cdots, p-1\}$, then we see that $\sum_i x_i\beta_i \in p\mathcal{O}_K$. Conversely, if such a nonzero $\sum_i x_i\beta_i \in p\mathcal{O}_K$ exists, then $0 \neq (\bar{x}_1, \cdots, \bar{x}_n) \in \mathrm{Ker}(\bar{C})$. Hence, $\det(C)$ is divisible by p, and $(\beta_1, \cdots, \beta_n)$ is not an integral basis.

□

Proposition 1.3.8. — *Let $\alpha \in \mathcal{O}_K$ such that $K = \mathbb{Q}(\alpha)$, and $f(T) \in \mathbb{Z}[T]$ be its minimal polynomial. Assume that for each prime p with $p^2|\mathrm{Disc}(1, \alpha, \cdots, \alpha^{n-1})$, there exists an integer i (which may depend on p) such that $f(T+i)$ is an Eisenstein polynomial for p. Then $\mathcal{O}_K = \mathbb{Z}[\alpha]$.*

Here, recall that a polynomial $f(T) = T^n + a_1T^{n-1} + \cdots + a_n$ is called an Eisenstein polynomial for p if $p|a_i$ for all $1 \leq i \leq n$ and $p^2 \nmid a_n$.

Proof. — Note that $\mathbb{Z}[\alpha] = \mathbb{Z}[\alpha - i]$ for all integer $i \in \mathbb{Z}$. Up to replacing α by $\alpha - i$ and using Lemma 1.3.7, it suffices to show that if $f(T) = T^n + a_1T^{n-1} + \cdots + a_n$ is an Eisentein polynomial for some prime p, then $x = \frac{1}{p}\sum_{i=0}^{n-1} x_i\alpha^i \notin \mathcal{O}_K$ for $x_i \in \{0, 1, \cdots, p-1\}$ not all zero. Put $j = \min\{i | x_i \neq 0\}$. Then

$$\mathrm{N}_{K/\mathbb{Q}}(x) = \frac{\mathrm{N}_{K/\mathbb{Q}}(\alpha)^j}{p^n}\mathrm{N}_{K/\mathbb{Q}}(\sum_{i=j}^{n-1} x_i\alpha^{i-j}).$$

We claim that $\mathrm{N}_{K/\mathbb{Q}}(\sum_{i=j}^{n-1} x_i\alpha^{i-j}) \equiv x_j^n \mod p$. But the denominator of $\frac{a_n^j}{p^n}$ is divisible by p, since $p||\mathrm{N}_{K/\mathbb{Q}}(\alpha) = (-1)^n a_n$. Therefore, it follows that $\mathrm{N}_{K/\mathbb{Q}}(x) \notin \mathbb{Z}$, and hence $x \notin \mathcal{O}_K$. To prove the claim, let $\sigma_1, \cdots, \sigma_n$ denote the complex embeddings of K. Then

$$\mathrm{N}_{K/\mathbb{Q}}(\sum_{i=j}^{n-1} x_i\alpha^{i-j}) = \prod_{k=1}^{n}(x_j + x_{j+1}\sigma_k(\alpha)^{i-j} + \cdots x_{n-1}\sigma_k(\alpha)^{n-1-j}).$$

Expanding the product, we see easily that all terms, except for x_j^n, are divisible by p, since they can be expressed as linearly combinations of $(-1)^k a_k$ for $k \geq 1$, which is k-elementary symmetric functions of $\alpha_1, \cdots, \alpha_n$'s.

□

Example 1.3.9. — Let $K = \mathbb{Q}(\alpha)$ with $\alpha^3 = 2$. We see easily that $\mathrm{Disc}(1, \alpha, \alpha^2) = -3^3 2^2$. But $f(T) = T^3 - 2$ is Eisenstein for $p = 2$ and $f(T-1) = T^3 - 3T^2 + 3T - 3$ is Eisenstein for $p = 3$. Hence, we get $\mathcal{O}_K = \mathbb{Z}[\alpha]$ by the previous Proposition.

We now give another property on the sign of the discriminant Δ_K. Let $\sigma : K \hookrightarrow \mathbb{C}$ be a complex embedding. We say that σ is a real embedding if $\sigma(K) \subseteq \mathbb{R}$; otherwise, we say σ is complex. Genuine complex embeddings of K always come in pairs. Actually, the composition of σ with the complex conjugation, denoted by $\bar{\sigma}$, is another complex embedding of K. We denote usually by r_1 the number of real embeddings of K, and by r_2 the number of pairs of genuine complex embeddings so that $n = r_1 + 2r_2$.

Proposition 1.3.10. — *The sign of* Δ_K *is* $(-1)^{r_2}$.

Proof. — We label the n embeddings of K into $\mathbb{C}$ as $\sigma_1, \cdots, \sigma_n$ such that $\sigma_1, \cdots, \sigma_{r_1}$ are real, and $\sigma_{r_1+2i} = \bar{\sigma}_{r_1+2i-1}$ for $1 \leq i \leq r_2$. Let $(\alpha_1, \cdots, \alpha_n)$ denote an integral basis of K. Then

$$\overline{\det(\sigma_i(\alpha_j))} = \det(\overline{\sigma_i}(\alpha_j)) = (-1)^{r_2} \det(\sigma_i(\alpha_j)),$$

because the matrix $(\bar{\sigma}_i(\alpha_j))_{1 \leq i,j \leq n}$ is obtained from $(\sigma_i(\alpha_j))_{1 \leq i,j \leq n}$ by swiping the $r_1 + 2i - 1$-th and $r_1 + 2i$-th rows for all $1 \leq i \leq r_2$. Therefore, if r_2 is even then $\det(\sigma_i(\alpha_j))$ is real, hence $\Delta_K = (\det(\sigma_i(\alpha_j)))^2$ is positive; and if r_2 is odd then $\det(\sigma_i(\alpha_j))$ is purely imaginary, thus Δ_K is negative. □

Remark 1.3.11. — By Lemma 1.3.1, the discriminant of any $\mathbb{Q}$-basis of K has the sign as Δ_K.

1.4. Cyclotomic fields

Let $N \geq 3$ be an integer, and $\zeta_N \in \mathbb{C}$ be a primitive n-th root of unity. Consider the number field $\mathbb{Q}(\zeta_N)$. Then for any $\sigma \in \mathrm{Aut}_{\mathbb{Q}}(\mathbb{C})$, $\sigma(\zeta_N)$ must be also a primitive N-th root of unity, hence of the form ζ_N^a for some a coprime to N. Therefore, $\mathbb{Q}(\zeta_N)$ is a Galois extension over $\mathbb{Q}$, and we have an injective map of groups $\varphi : \mathrm{Gal}(\mathbb{Q}(\zeta_N)/\mathbb{Q}) \hookrightarrow (\mathbb{Z}/N\mathbb{Z})^\times$.

Proposition 1.4.1. — *The homomorphism* φ *is an isomorphism* $\mathrm{Gal}(\mathbb{Q}(\zeta_N)/\mathbb{Q}) \cong (\mathbb{Z}/N\mathbb{Z})^\times$.

Proof. — To prove the subjectivity of φ, it suffices to show that the image of every prime p with $p \nmid N$ in $(\mathbb{Z}/N\mathbb{Z})^\times$ lies in the image of ϕ (since such elements generate the group $(\mathbb{Z}/N\mathbb{Z})^\times$). It is equivalent to showing that ζ_N^p is a conjugate of ζ_N. Let $f(T) \in \mathbb{Z}[T]$ denote the minimal polynomial of ζ_N, and write $T^N - 1 = f(T)g(T)$ with $g(T) \in \mathbb{Z}[T]$. Suppose in contrary that ζ_N^p is not conjugate to ζ_N. Then one has $g(\zeta_N^p) = 0$, that is ζ_n is a root of $g(T^p)$. Since $f(T)$ is the minimal polynomial of ζ_n, one has $f(T) | g(T^p)$. Let $\bar{f}, \bar{g} \in \mathbb{F}_p[T]$ denote the reduction modulo p of f and g respectively. Note that $\bar{g}(T)^p = \bar{g}(T^p)$, so we get $\bar{f}(T) | \bar{g}(T)^p$. If α is any root of $\bar{f}(T)$ in an algebraic closure $\bar{\mathbb{F}}_p$ of $\mathbb{F}_p$, then $\bar{g}(\alpha) = 0$. Then means that α is a multiple root of $\bar{F}(T) = \bar{f}(T)\bar{g}(T)$. But $\bar{F}'(\alpha) = N\alpha^{N-1} \neq 0$ in $\bar{\mathbb{F}}_p$, hence $\bar{F}(T)$ has no multiple root. This is a contradiction. □

Corollary 1.4.2. — *If $N, M \geq 2$ are integers with $\gcd(N, M) = 1$, then we have $\mathbb{Q}(\zeta_N) \cap \mathbb{Q}(\zeta_M) = \mathbb{Q}$.*

Proof. — Note that $\mathbb{Q}(\zeta_{NM}) = \mathbb{Q}(\zeta_M)\mathbb{Q}(\zeta_N)$ as subfields of $\mathbb{C}$. By field theory, one has

$$[\mathbb{Q}(\zeta_{MN}) : \mathbb{Q}(\zeta_N)] = [\mathbb{Q}(\zeta_M) : \mathbb{Q}(\zeta_M) \cap \mathbb{Q}(\zeta_N)].$$

Therefore, $\mathbb{Q}(\zeta_M) \cap \mathbb{Q}(\zeta_N) = \mathbb{Q}$ if and only if $[\mathbb{Q}(\zeta_{MN}) : \mathbb{Q}(\zeta_N)] = \phi(M)$, where $\phi(M) := \#(\mathbb{Z}/M\mathbb{Z})^\times$ is the Euler function. But this follows from

$$[\mathbb{Q}(\zeta_{MN}) : \mathbb{Q}(\zeta_N)] = [\mathbb{Q}(\zeta_{MN}) : \mathbb{Q}]/[\mathbb{Q}(\zeta_N : \mathbb{Q})] = \phi(MN)/\phi(N) = \phi(M).$$

□

We manage to compute the discriminant of $\mathbb{Q}(\zeta_N)$. We put

$$\Phi_N(T) = \prod_{a \in (\mathbb{Z}/N\mathbb{Z})^\times} (T - \zeta_N^a) \in \mathbb{Z}[T],$$

and call it the N-cyclotomic polynomial.

Lemma 1.4.3. — *The discriminant of $\mathbb{Q}(\zeta_N)$ divides $N^{\phi(N)}$.*

Proof. — Since $\Delta_{\mathbb{Q}(\zeta_N)} | \mathrm{Disc}(1, \zeta_N, \cdots, \zeta_N^{\phi(N)-1})$, it suffices to prove the latter divides $N^{\phi(N)}$. Write $T^N - 1 = \Phi_N(T)F(T)$ for some $F(T) \in \mathbb{Z}[T]$. Then we get

$$NT^{N-1} = \Phi_N'(T)F(T) + \Phi_N(T)F'(T).$$

Thus $\mathrm{N}_{\mathbb{Q}(\zeta_N)/\mathbb{Q}}(\Phi'(\zeta_N)) | \mathrm{N}_{\mathbb{Q}(\zeta_N)/\mathbb{Q}}(N\zeta_N^{N-1}) = N^{\phi(N)}$. We conclude by Proposition 1.3.2. □

Corollary 1.4.4. — *If p is a prime, then the ring of integers of $\mathbb{Q}(\zeta_{p^n})$ is $\mathbb{Z}[\zeta_{p^n}]$.*

Proof. — Indeed, $\Phi_{p^n}(X+1)$ is an Eisenstein polynomial for p, and the statement follows from Proposition 1.3.8. □

In order to generalize the previous Corollary to arbitrary $\mathbb{Q}(\zeta_N)$, we need some preparation. Let K and L be two number fields, and KL be the composite field (inside $\mathbb{C}$). Consider the subring

$$\mathcal{O}_K\mathcal{O}_L = \{x_1y_1 + \cdots + x_ry_r | x_i \in \mathcal{O}_K, y_j \in \mathcal{O}_L\}.$$

We have always $\mathcal{O}_K\mathcal{O}_L \subset \mathcal{O}_{KL}$, but they are not equal in general. However, we have the following

Proposition 1.4.5. — *Assume that $K \cap L = \mathbb{Q}$, and put $d = \gcd(\Delta_K, \Delta_L)$. Then we have $\mathcal{O}_{KL} \subset \frac{1}{d}\mathcal{O}_K\mathcal{O}_L$.*

Proof. — Let $(\alpha_1, \cdots, \alpha_n)$ and $(\beta_1, \cdots, \beta_m)$ be integral basis of K and L respectively. Any $x \in \mathcal{O}_{KL}$ writes as

$$x = \sum_{i,j} \frac{x_{i,j}}{r} \alpha_i\beta_j, \quad \text{with } x_{i,j}, r \in \mathbb{Z} \text{ and } \gcd(x_{i,j}, r) = 1.$$

We have to show that $r|d$. By symmetry, it suffices to prove that $r|\Delta_L$. Let $(\alpha_i^\vee)_{1\leq i\leq n}$ be the dual basis of $(\alpha_i)_{1\leq i\leq n}$ with respect to $\mathrm{Tr}_{K/\mathbb{Q}}$. Then we have

$$\mathrm{Tr}_{KL/L}(x\alpha_i^\vee) = \sum_{k,l} \frac{x_{k,l}}{r}\mathrm{Tr}_{KL/L}(\alpha_k\beta_l\alpha_i^\vee) = \sum_l \frac{x_{i,l}}{r}\beta_l.$$

On the other hand, we have $\alpha_i^\vee \in \frac{1}{\Delta_K}\mathcal{O}_K$ by definition of $\alpha_i^\vee$ and Cramer's rule. So $x\alpha_i^\vee \in \frac{1}{\Delta_K}\mathcal{O}_{KL}$, and hence $\mathrm{Tr}_{KL/L}(x\alpha_i^\vee) \in \frac{1}{\Delta_K}\mathrm{Tr}_{KL/L}(\mathcal{O}_{KL}) \subset \frac{1}{\Delta_K}\mathcal{O}_L$, i.e. $\Delta_K\mathrm{Tr}_{KL/L}(x\alpha_i^\vee) \in \mathcal{O}_L$. But $(\beta_j)_{1\leq j\leq m}$ is a basis of $\mathcal{O}_L$ over $\mathbb{Z}$, thus $\Delta_K \cdot \frac{x_{i,j}}{r} \in \mathbb{Z}$ for all i, j, and so $r|\Delta_K$. □

Corollary 1.4.6. — *Assume that $K \cap L = \mathbb{Q}$ and $\gcd(\Delta_K, \Delta_L) = 1$. Then we have $\mathcal{O}_{KL} = \mathcal{O}_K\mathcal{O}_L$.*

We can now prove

Theorem 1.4.7. — *The ring of integers of $\mathbb{Q}(\zeta_N)$ is $\mathbb{Z}[\zeta_N]$.*

Proof. — We prove the statement by induction on the number of prime factors of N. When N is a power of some prime p, then this is proved in Corollary 1.4.4. If N contains several prime factors, then write $N = nm$ with $n, m > 1$ and $\gcd(n, m) = 1$. By Corollary 1.4.2 and Lemma 1.4.3, the assumptions of Corollary 1.4.6 are satisfied. We conclude by induction hypothesis that $\mathcal{O}_{\mathbb{Q}(\zeta_N)} = \mathcal{O}_{\mathbb{Q}(\zeta_n)}\mathcal{O}_{\mathbb{Q}(\zeta_m)} = \mathbb{Z}[\zeta_n, \zeta_m] = \mathbb{Z}[\zeta_N]$. □

We can also compute the exact value of $\Delta_{\mathbb{Q}(\zeta_N)}$ when $N = p^n$, with p a prime.

Proposition 1.4.8. — *We have*

$$\Delta_{\mathbb{Q}(\zeta_{p^n})} = \mathrm{Disc}(1, \zeta_{p^n}, \cdots, \zeta_{p^n}^{p^{n-1}(p-1)-1}) = \pm p^{p^{n-1}(pn-n-1)},$$

where we have $-$ if $p \equiv 3 \mod 4$ or $p^n = 4$, and we have $+$ otherwise.

Proof. — The statement for the sign follows easily from Remark 1.3.11. Compute now $|\mathrm{Disc}(1, \zeta_{p^n}, \cdots, \zeta_{p^n}^{p^{n-1}(p-1)-1})|$, which is equal to $|\mathrm{N}_{\mathbb{Q}(\zeta_{p^n})/\mathbb{Q}}(\Phi'_{p^n}(\zeta_{p^n}))|$ by Proposition 1.3.2, where

$$\Phi_{p^n}(T) = \frac{T^{p^n}-1}{T^{p^{n-1}}-1} = \sum_{i=0}^{p-1} T^{p^{n-1}i}.$$

If $p = 2$, then $\Phi'_{2^n}(\zeta_{2^n}) = 2^{n-1}\zeta_{2^n}^{2^{n-1}-1}$ and $|\mathrm{N}_{\mathbb{Q}(\zeta_{2^n})/\mathbb{Q}}(\Phi'_{2^n}(\zeta_{2^n}))| = 2^{2^{n-1}(n-1)}$. If $p \geq 3$, then

$$\begin{aligned}\Phi'(\zeta_{p^n}) &= p^{n-1}\zeta_{p^n}^{p^{n-1}-1}\sum_{i=1}^{p-1} i\zeta_{p^n}^{p^{n-1}(i-1)} = p^{n-1}\zeta_{p^n}^{p^{n-1}-1}\Phi'_p(\zeta_p)\\ &= p^{n-1}\zeta_{p^n}^{p^{n-1}+(p-3)}\prod_{i=1}^{p-2}(1-\zeta_p^i).\end{aligned}$$

Therefore, $|\mathrm{N}_{\mathbb{Q}(\zeta_{p^n})/\mathbb{Q}}| = p^{p^{n-1}(p-1)(n-1)} \prod_{i=1}^{p-2} |\mathrm{N}_{\mathbb{Q}(\zeta_p)/\mathbb{Q}}(\zeta_p^i - 1)|^{p^{n-1}}$. But the minimal polynomial of $\zeta_p^i - 1$ over $\mathbb{Q}$ is

$$\Phi_p(X+1) = X^{p-1} + pX^{p-2} + \cdots + \binom{p}{2} X + p.$$

Thus we have $|\mathrm{N}_{\mathbb{Q}(\zeta_p)/\mathbb{Q}}(\zeta_p^i - 1)| = p$, and the Lemma follows immediately. □

CHAPTER 2

DEDEKIND DOMAINS

2.1. Preliminaries on Noetherian rings

All rings in this section are commutative.

Proposition 2.1.1. — *Let R be a ring, and M be an R-module. The following statements are equivalent:*

1. *Every submodule of M (including M itself) is finitely generated.*
2. *For any increasing chain of submodules $N_1 \subseteq N_2 \subseteq \cdots \subseteq N_n \subseteq N_{n+1} \subseteq \cdots$ in M, there exists an integer m such that $N_n = N_{n+1}$ for all $n \geq m$.*
3. *Every non-empty subset $\mathcal{S}$ of submodules of M contains a maximal element N under inclusion, i.e. if $N' \in \mathcal{S}$ contains N, then $N = N'$.*

Proof. — We prove first (1) $\Longrightarrow$ (2). Given an increasing chain of submodules $N_1 \subseteq N_2 \subseteq \cdots N_n \subseteq \cdots$, put $N_\infty = \cup_{n\geq 1} N_n$. Write $N_\infty = (x_1, \cdots, x_r)$. If $m \geq 1$ is large enough so that all $x_i \in N_m$, then $N_n = N_\infty$ for all $n \geq m$.

For (2) $\Longrightarrow$ (3), we assume that $\mathcal{S}$ does not contain any maximal element. Take an arbitrary $N_1 \in \mathcal{S}$. Since N_1 is not maximal, there exists $N_2 \in \mathcal{S}$ such that $N_1 \subsetneq N_2$. Continuing this process, we produce an increasing chain of ideals $N_1 \subsetneq N_2 \subsetneq \cdots N_n \subsetneq N_{n+1} \subsetneq \cdots$, whose existence contradicts with (2).

Finally, we prove (3) $\Longrightarrow$ (1). It is enough to prove that M is finitely generated, since the same arguments apply with M replaced by any submodule $N \subseteq M$. Consider the set $\mathcal{S}$ consisting of all finitely generated submodules of M. Then $\mathcal{S}$ is non-empty, because $(0) \in \mathcal{S}$. Let $N \in \mathcal{S}$ be a maximal element. For any $x \in M$, $N' = N + R \cdot x$ is also finitely generated and $N \subseteq N'$. Then one has $N = N'$ by the maximality of N. This implies that $x \in N$, i.e. $N = M$. □

Definition 2.1.2. — (1) We say an R-module M is *Noetherian* if it satisfies the equivalent conditions in the previous Proposition.

(2) We say a ring R is *Noetherian*, if R itself is Noetherian as an R-module.

Proposition 2.1.3. — *Let $0 \to M_1 \to M \to M_2 \to 0$ be a short exact sequence of R-modules. Then M is Noetherian if and only if both M_1 and M_2 are Noetherian.*

Proof. — The "only if" part is easy and left as an Exercise. Assume now both M_1 and M_2 are Noetherian. Let N be a submodule of M. Put $N_1 = M_1 \cap N$ and $N_2 \subseteq M_2$ to be the image of N. By assumption, both N_1 and N_2 are finitely generated. Let $N_1 = (x_1, \cdots, x_r)$, and $x_{r+1}, \cdots, x_{r+s} \in N$ be such that their image in N_2 generate N_2. Then we claim that N is generated by $x_1, \cdots, x_{r+s}$. Indeed, for any $x \in N$, there exist $a_{r+1}, \cdots, a_{r+s} \in R$ such that the image of $x - \sum_{i=1}^{s} a_{r+i}x_{r+i}$ in N_2 is zero, i.e. $x - \sum_{i=1}^{s} a_{r+i}x_{r+i} \in N_1$. Thus there exist $a_1, \cdots, a_r \in R$ such that $x - \sum_{i=1}^{s} a_{r+i}x_{r+i} = \sum_{i=1}^{r} a_i x_i$, that is $x \in (x_1, \cdots, x_{r+s})$. □

Corollary 2.1.4. — *Any finitely generated R-module is Noetherian.*

Proof. — Indeed, any finitely generated R-module is a quotient of $R^{\oplus n}$ for some n. □

Corollary 2.1.5. — *(1) If R is a Noetherian ring, then any quotient of R is also Noetherian.*

(2) If R_1 and R_2 are Noetherian rings, then so is $R_1 \oplus R_2$.

Proof. — (1) If $\bar{R}$ is quotient of R, then any ideal of $\bar{R}$ is finitely generated as R-module, hence as $\bar{R}$-module.

(2) It suffices to note that any ideal of $R_1 \oplus R_2$ is of the form $I_1 \oplus I_2$ where I_j is an ideal of R_j. □

Example 2.1.6. — (1) Principal ideal domains such as $\mathbb{Z}$, $\mathbb{Q}[X]$ are Noetherian.

(2) The ring $\mathbb{Q}[X_1, X_2, \cdots, X_n, \cdots]$ is non-Noetherian. The ring of all algebraic numbers is also non-Noetherian.

Finally, we have famous theorem of Hilbert.

***Theorem 2.1.7* (Hilbert basis theorem)**. — *If R is Noetherian, then $R[X]$ is also Noetherian.*

Proof. — Let $J \subset R[X]$ be an ideal. Let $I \subseteq R$ denote the subset consisting of $\in R$ such that there exists some $f \in J$ whose top coefficient is a. Then we see easily that I is an ideal of R. Choose $f_i = a_{i,d_i}X^{d_i} + \cdots + a_{i,0} \in J$ for $1 \leq i \leq r$ such that $I = (a_{1,d_1}, \cdots, a_{r,d_r})$. Let $d = \max_i\{d_i\}$. The polynomials of degree $< d$ contained in J form a finitely generated R-module; let $\{g_1, \cdots, g_s\}$ denote a set of generators over R. Let $f \in J$ of degree $n = \deg(f)$. We claim that f is generated by $\{f_1, \cdots, f_r, g_1, \cdots, g_s\}$. If $n < d$, then f is generated by $\{g_1, \cdots, g_s\}$ (even over R). If $n \geq d$, then there exist $b_1, \cdots, b_r \in R$ such that $f' = f - \sum_{i=1}^{r} b_i X^{n-d_i} f_i$ has degree strict less than n. Repeating the process with f replaced by f', one may finally reduce to the case of degree $< d$. □

Combining with Corollary 2.1.5, we have the following

Corollary 2.1.8. — *If R is a finitely generated algebra over $\mathbb{Z}$ or over a field, then R is Noetherian.*

2.2. Dedekind domains

Definition 2.2.1. — An integral domain A is called *a Dedekind domain* if it is Noetherian and integrally closed, and every non-zero prime ideal is maximal.

Example 2.2.2. — (1) Every principal ideal domain is a Dedekind domain, e.g. $\mathbb{Z}$, $\mathbb{F}_p[X]$, $\mathbb{C}[X]$.

(2) For any number field K, $\mathcal{O}_K$ is a Dedekind domain.

(3) Let k be a field, $F(x,y) \in k[x,y]$ such that $F(x,y), F'_x(x,y)$ and $F'_y(x,y)$ has no common zeros. Then $k[x,y]/(F(x,y))$ is a Dedekind domain.

Proposition 2.2.3. — *Let A be a Dedekind domain with fraction field K. Let L/K be a finite separable extension, and B be the integral closure of A in L. Then B is also a Dedekind domain, and is finitely generated as an A-module.*

Proof. — It suffices to show that

(a) B is a finitely generated A-module (hence neotherian as an B-module) and,
(b) every non-zero prime ideal of B is maximal.

The proof of (a) is similar to that of Theorem 1.3.3. Choose $\alpha_1, \cdots, \alpha_n \in B$ which form a basis of L/K, and let $(\alpha_1^\vee, \cdots, \alpha_n^\vee)$ be the dual basis with respect to $\mathrm{Tr}_{L/K}$. Put $M^\vee = \sum_i \alpha_i^\vee A$. Then one has $B \subseteq M^\vee$. As A is Noetherian and M is finitely generated, it follows that B is a finitely generated A-module. For (b), let $\mathfrak{P}$ be a non-zero prime ideal of B, and let $\mathfrak{p} = \mathfrak{P} \cap A$. Then $\mathfrak{p}$ is a non-zero prime ideal of A, and $B/\mathfrak{P}$ is integral over $A/\mathfrak{p}$. Since A is Dedekind by assumption, $A/\mathfrak{p}$ is a field. Now (b) follows from the following Lemma.

□

Lemma 2.2.4. — *Let $A \subseteq B$ be an extension of domains, and assume that B is integral over A. Then B is a field if and only if A is a field.*

Proof. — Assume first that A is a field. Let $x \in B$. As x is integral over A, it satisfies a monic polynomial equation

$$x^n + a_1 x^{n-1} + \cdots + a_n = 0, \quad \text{with } a_i \in A$$

Up to canceling some powers of x, we may assume that $a_n \neq 0$. Then a_n is invertible, so is x. Assume now B is a field. Let $y \in A$. Then y^{-1} exists as an element in B. Then it must satisfy some equation

$$y^{-m} + b_1 y^{-m+1} + \cdots + b_m = 0, \quad \text{with } b_i \in A.$$

Multiplying both sides by y^{m-1}, we see that

$$y^{-1} = -b_1 + \cdots - b_m y^{m-1} \in A.$$

□

Definition 2.2.5. — Let A be a domain with fractional field K. Then a *fractional ideal* I of A is a sub-A-module of K such that there exists $d \in A$ with $dI \subset A$.

If I and J are both fractional ideals of A, then

$$I+J=\{x\in K|x=a+b,\ a\in I,b\in J\},\quad I\cdot J=\{x=\sum_i a_ib_i|a_i\in I,b_i\in J\}$$

are both fractional ideals.

The main result of this section is the following

Theorem 2.2.6. — *Let A be a Dedekind domain. Every ideal I of A has a factorization $I=\mathfrak{p}_1^{a_1}\cdots\mathfrak{p}_r^{a_r}$ where $\mathfrak{p}_i$ are distinct prime ideals and $a_i\in\mathbb{Z}_{\geq 0}$; moreover, the factorization of I is unique up to order, i.e. if I has two such factorizations $\mathfrak{p}_1^{a_1}\cdots\mathfrak{p}_r^{a_r}=\mathfrak{q}_1^{b_1}\cdots\mathfrak{q}_s^{b_s}$, then $r=s$ and for each $1\leq i\leq r$, there exists a unique j such that $\mathfrak{p}_i=\mathfrak{q}_j$ and $a_i=b_j$.*

To prove this theorem, we need some preparation.

Lemma 2.2.7. — *Let A be a Noetherian ring. Then every ideal $I\neq 0$ of A contains a product of non-zero prime ideals.*

Proof. — Let $\mathcal{S}$ be the set of ideals that do not contain any product of non-zero prime ideals. Suppose that $\mathcal{S}$ is non-empty. Since A is Noetherian, $\mathcal{S}$ admits a maximal element, say I. Then I must not be a prime ideal. Thus there exist $a,b\in R$ such that $a,b\notin I$ but $ab\in I$. Then consider $I_1=I+(a)$ and $I_2=I+(b)$. Then $I\subsetneq I_i$ for $i=1,2$. By the maximality of I, both I_1 and I_2 will contain a product of prime ideals. But it follows from

$$I_1I_2\subseteq(ab)+aI+bI+I^2\subseteq I$$

that I should also contain a product of prime ideals. This is a contradiction.

□

Lemma 2.2.8. — *Let A be a Dedekind domain, and $\mathfrak{p}\subseteq A$ be a non-zero prime ideal. Then*

$$\mathfrak{p}^{-1}:=\{x\in K|x\cdot\mathfrak{p}\subset A\}$$

is a fractional ideal of A, and $\mathfrak{p}^{-1}\mathfrak{p}=A$.

Proof. — It is easy to see that $\mathfrak{p}^{-1}$ is a fractional ideal and $A\subseteq\mathfrak{p}^{-1}$. Then we have $\mathfrak{p}\subset\mathfrak{p}\mathfrak{p}^{-1}\subset A$. Since the ideal $\mathfrak{p}$ is maximal, we have either $\mathfrak{p}=\mathfrak{p}^{-1}\mathfrak{p}$ or $\mathfrak{p}^{-1}\mathfrak{p}=A$. We have to exclude the first case. Suppose in contrary that $\mathfrak{p}=\mathfrak{p}^{-1}\mathfrak{p}$. Let $\{\alpha_1,\cdots,\alpha_r\}\subset\mathfrak{p}$ be a subset of generators. Then for any $x\in\mathfrak{p}^{-1}$, we have

$$x\alpha_i=\sum_j c_{i,j}\alpha_j,\quad\text{for } c_{i,j}\in A.$$

If C denotes the matrix $(c_{i,j})_{1\leq i,j\leq n}$, we have $\det(xI_r-C)=0$. Thus x is integral over A. But A is integrally closed, we get $x\in A$. This shows that $\mathfrak{p}^{-1}=A$.

Now to get a contradiction, it suffices to construct an element $x\in\mathfrak{p}^{-1}$ but $x\notin A$. Choose $0\neq b\in\mathfrak{p}$. Let r be the minimal integer such that $\mathfrak{p}\supset(b)\supset\mathfrak{p}_1\cdots\mathfrak{p}_r$, where all $\mathfrak{p}_i$ are non-zero prime ideals. Such a r exists by Lemma 2.2.7. Then there exists a $\mathfrak{p}_i$, say $i=1$, such that $\mathfrak{p}\supset\mathfrak{p}_i=\mathfrak{p}_1$, so $\mathfrak{p}=\mathfrak{p}_1$ since every non-zero prime ideal in A is maximal. Then $\mathfrak{p}_2\cdots\mathfrak{p}_r\subsetneq(b)$ so that there exists $a\in\mathfrak{p}_2\cdots\mathfrak{p}_r$ but $a\notin(b)$, i.e. $a/b\notin A$. But we have $a/b\mathfrak{p}\subseteq\frac{1}{b}\mathfrak{p}_1\cdots\mathfrak{p}_r\subseteq A$, i.e. $\frac{a}{b}\in\mathfrak{p}^{-1}$. □

Proof of Theorem 2.2.6. — We show the existence of the factorization. Let $\mathcal{S}$ denote the set of ideals of A that are not product of primes. Suppose that $\mathcal{S}$ is non-empty. Denote by I a maximal element in $\mathcal{S}$ by the Noetherianness of A. Then I can not be a prime. Thus there exists a prime ideal $I \subsetneq \mathfrak{p}$. By Lemma 2.2.8, we have $A = \mathfrak{p}^{-1}\mathfrak{p} \supsetneq \mathfrak{p}^{-1}I \supsetneq I$. By the maximality of I, we see that $\mathfrak{p}^{-1}I$ is a product of primes, that is $\mathfrak{p}^{-1}I = \prod_{i=2}^{r} \mathfrak{p}_i$. It then follows immediately that $I = \mathfrak{p}\prod_i \mathfrak{p}_i$.

Now we prove the uniqueness of the factorization. Suppose that $\prod_{i=1}^{r} \mathfrak{p}_i = \prod_j \mathfrak{q}_j$. If $r \geq 1$, then $\mathfrak{p}_1 \supset \prod_{j=1}^{s} \mathfrak{q}_j$. It follows that $\mathfrak{p}_1 \supset \mathfrak{q}_j$ for some j. Since every non-zero prime of A is maximal, we see that $\mathfrak{p}_1 = \mathfrak{q}_j$. We may assume that $j = 1$. By Lemma 2.2.8, we get $\prod_{i=2}^{r} \mathfrak{p}_i = \prod_{j=2}^{s} \mathfrak{q}_j$. By induction, we see that every $\mathfrak{p}_i$ has to coincide with some $\mathfrak{q}_j$ and vice-versa. □

Corollary 2.2.9. — *A Dedekind domain is a unique factorization ring if and only if it is a principal ideal domain.*

Proof. — Let A be a Dedekind domain. We have already seen that if A is a principal ideal domain, then A is necessarily a unique factorization domain (without assuming A is Dedekind). We suppose conversely that A is a unique factorization domain. By Theorem 2.2.6, it suffices to prove that every prime ideal $\mathfrak{p} \subset A$ is principal. Choose $0 \neq x \in \mathfrak{p}$, and let $x = p_1 \cdots p_r$ be a prime factorization of x. Then each of the principal ideal (p_i) is prime, and we have $\mathfrak{p}|(x) = \prod_{i=1}^{r}(p_i)$. There exists thus a p_i such that $\mathfrak{p}|(p_i)$. But (p_i) is maximal, so we get $\mathfrak{p} = (p_i)$. □

Let A be a Dedekind domain with fraction field K. If I is a fractional ideal of A, we put

$$I^{-1} := \{x \in K | xI \subseteq A\}.$$

Then I^{-1} is also a fractional ideal, and if $I = \prod_{i=1}^{r} \mathfrak{p}_i^{a_i}$ with $a_i \in \mathbb{Z}$, then $I^{-1} = \prod_{i=1}^{r} \mathfrak{p}_i^{-a_i}$. For any integer $a \in \mathbb{Z}_{>0}$, we put $I^{-a} = (I^a)^{-1} = (I^{-1})^a$.

Corollary 2.2.10. — *Let I be a fractional ideal of a Dedekind domain. Then I has a unique factorization $I = \prod_{i=1}^{r} \mathfrak{p}_i^{a_i}$, where each $\mathfrak{p}_i$ is a prime of A and distinct with each other and $a_i \in \mathbb{Z}$. Moreover, I is an ideal if and only if $a_i \geq 0$ for all i.*

Proof. — The existence and uniqueness of the factorization follows from Theorem 2.2.6. If all $a_i \geq 0$, I is clearly an ideal. Suppose now that I is an ideal. If some of a_i's are negative, say $a_1, \cdots, a_s < 0$ and $a_{s+1}, \cdots, a_r \geq 0$ for some $1 \leq s \leq r$, then $\prod_{i=s+1}^{r} \mathfrak{p}_i^{a_i} \subseteq \prod_{i=1}^{s} \mathfrak{p}_i^{-a_i} \subseteq \mathfrak{p}_1$. But this implies that $\mathfrak{p}_1 \supset \mathfrak{p}_i$ for some $s+1 \leq i \leq r$. This contradicts with the fact that $\mathfrak{p}_i$'s are distinct.

□

It is convenient to introduce the following notation: For two fractional ideals I, J, we say that I divides J and write $I|J$, if $J \subseteq I$. For a fractional ideal I and a prime $\mathfrak{p}$, we denote by $v_\mathfrak{p}(I)$ the exponent index of $\mathfrak{p}$ appearing in the prime decomposition of I. For $x \in K$, we put $v_\mathfrak{p}(x) = v_\mathfrak{p}((x))$ if $x \neq 0$, and $v_\mathfrak{p}(x) = \infty$ if $x = 0$.

Lemma 2.2.11. — *Let I, J be fractional ideals of a Dedekind domain A. Then $I|J$ if and only if $v_\mathfrak{p}(I) \leq v_\mathfrak{p}(J)$ for all prime $\mathfrak{p}$.*

Proof. — Note that $I|J$ is equivalent to saying that $J' = I^{-1}J$ is an ideal of A. Since $v_\mathfrak{p}(JI^{-1}) = v_\mathfrak{p}(J) - v_\mathfrak{p}(I)$, we are reduced to proving that J' is an ideal of A if and only if $v_\mathfrak{p}(J') \geq 0$ for all prime $\mathfrak{p}$. But this follows from Corollary 2.2.10. □

The prime ideals in A behave like the usual prime numbers in $\mathbb{Z}$, that is if $\mathfrak{p}$ is a prime ideal of A and $\mathfrak{p}|IJ$ then either $\mathfrak{p}|I$ or $\mathfrak{p}|J$. Note also that $\mathfrak{p}$ appears in the prime factorization of an ideal I, if and only if $\mathfrak{p}|I$.

Corollary 2.2.12. — *Let I, J be fractional ideals of a Dedekind domain A. Then*

1. $I = \{x \in K | v_\mathfrak{p}(x) \geq v_\mathfrak{p}(I)$ *for all prime* $\mathfrak{p}\}$;
2. $v_\mathfrak{p}(I + J) = \min(v_\mathfrak{p}(I), v_\mathfrak{p}(J))$ *for all prime* $\mathfrak{p}$;
3. $v_\mathfrak{p}(x + y) \geq \min(v_\mathfrak{p}(x), v_\mathfrak{p}(y))$ *for all* $x, y \in K$ *and prime* $\mathfrak{p}$;
4. $v_\mathfrak{p}(I \cap J) = \max(v_\mathfrak{p}(I), v_\mathfrak{p}(J))$.

Proof. — (1) $x \in I \Longrightarrow I|(x) \Longrightarrow v_\mathfrak{p}(x) \geq v_\mathfrak{p}(I)$.

(2) If K is a fractional ideal containing both I and J, then $v_\mathfrak{p}(I) \leq v_\mathfrak{p}(K)$ and $v_\mathfrak{p}(J) \leq v_\mathfrak{p}(K)$ by the previous Lemma. Since $I + J$ is the minimal fractional ideal with this property, the statement follows.

(3) $v_\mathfrak{p}(x + y) \geq v_\mathfrak{p}((x) + (y)) = \min(v_\mathfrak{p}(x), v_\mathfrak{p}(y))$.

(4) Similar to (2), $I \cap J$ is the maximal fractional ideal contained in I and J. □

Example 2.2.13. — (1) If $A = \mathbb{C}[x]$, then A can be viewed as the algebraic functions on $\mathbb{C}$ and its fraction field $k(x)$ is the set of meromorphic functions on $\mathbb{C}$. The set of primes of A is naturally identified with $\mathbb{C}$. If $\mathfrak{p} = (x - a)$ and $f \in \mathbb{C}(x)^\times$, then $v_\mathfrak{p}(f)$ is the vanishing order of f at $x = a$.

(2) Consider $A = \mathbb{Z}[\sqrt{-5}]$. We have factorizations

$$(2) = (2, 1 + \sqrt{-5})^2, \quad (3) = (3, 1 + \sqrt{-5})(3, 1 - \sqrt{-5}).$$

Definition 2.2.14. — Let A be a Dedekind domain with fraction field K.

(1) The set of fractional ideals of A form an abelian group (with addition given by multiplication), which we denote by $\mathcal{I}$.

(2) A fractional ideal is called principal if it is of the form xA with $x \in K^\times$. Principal fractional ideals of A form clearly a subgroup of $\mathcal{I}$, and we denote it by $\mathcal{P}$.

(3) We define $\mathcal{C}l_K$ to be the quotient group $\mathcal{I}/\mathcal{P}$, and call it the ideal class group of A or of K.

By Theorem 2.2.6, the group $\mathcal{I}$ is isomorphic to the free abelian group with basis given by the set of primes of A, which is usually of infinite rank. However, in the case of number fields, we have

Theorem 2.2.15. — *Let K be a number field, then $\mathcal{C}l_K$ is a finite abelian group.*

The proof of this fundamental Theorem will be given in Section 4.1. For a number field, we usually denote by h_K the cardinality of $\mathcal{C}l_K$, and call it the class number of K.

2.3. Localization

In this section, let A be a domain and K be its fraction field.

Definition 2.3.1. — (1) A subset $S \subseteq A$ is called multiplicative if $s_1, s_2 \in S$ implies $s_1 s_2 \in S$. For a multiplicative subset S, we define

$$S^{-1}A := \{\frac{a}{s} | a \in A, s \in S\} \subseteq K,$$

which a subring of K.

(2) If $\wp$ is a prime ideal of A, then $S_\wp = A \backslash \wp$ is multiplicative, and we put $A_\wp = S_\wp^{-1} A$. We call $A_\wp$ the local ring of A at $\wp$, or the localization of A at $\wp$.

Example 2.3.2. — (1) If $f \in A$ is non-zero, then $S_f = \{f^n | n \in \mathbb{Z}_{\geq 0}\}$ is multiplicative. We have $S_f^{-1} A = A[\frac{1}{f}]$. For instance, the prime ideals of $\mathbb{Z}[\frac{1}{6}]$ are (0) and $p\mathbb{Z}[\frac{1}{6}]$ for any $p \nmid 6$.

(2) For a prime p, then $\mathbb{Z}_{(p)} = \{\frac{m}{n} \in \mathbb{Q} | p \nmid n\}$. Note that $\mathbb{Z}_{(p)}$ has two prime ideals, namely $p\mathbb{Z}_{(p)}$ and (0).

Let $S \subseteq A$ be a multiplicative subset, $A' = S^{-1}A$. Let I be an ideal of A, then

$$IA' = S^{-1}I = \{\frac{a}{s} | a \in I, s \in S\}$$

is an ideal of A'. It is clear that if $S \cap I \neq \emptyset$, then $IA' = A'$. Conversely, if $I' \subsetneq A'$, then $I = A \cap I'$ is also an ideal of A disjoint with S. It is always true that $(I' \cap A)A' = I'$, but, in general, it is not true that $IA' \cap A = I$. For instance, if $A = \mathbb{Z}$, $S = \{4^n | n \in \mathbb{Z}_{\geq 0}\}$ and $I = (10)$, $I' \cap A = (5)$.

Lemma 2.3.3. — *Let $S \subset A$ be a multiplicative subset, and $A' = S^{-1}A$. Then $\mathfrak{p} \mapsto \mathfrak{p}A'$ establishes an order preserving bijection with set of prime ideals of A disjoint with S and the set of prime ideals of A', and its inverse is given by $\mathfrak{p}' \mapsto \mathfrak{p}' \cap A$. In particular, if $A' = A_\wp$ for some prime ideal $\wp \subseteq A$, then $\wp A_\wp$ is the unique maximal ideal of $A_\wp$, and $A_\wp / \wp A_\wp$ is the fraction field of $A/\wp$.*

Proof. — Let $\mathfrak{p}$ be a prime of A. It is clear that $\mathfrak{p} \subseteq \mathfrak{p}A' \cap A$. If $x = \frac{a}{s} \in \mathfrak{p}A' \cap A$, then $a = xs \in \mathfrak{p}$. But $\mathfrak{p} \cap S = \emptyset$, it follows that $x \in \mathfrak{p}$. This proves that $\mathfrak{p}A' \cap A = \mathfrak{p}$, i.e. $\mathfrak{p}A' \cap A = \mathfrak{p}$. If $A' = A_\wp$, then any ideal of $A_\wp$ has the form $IA_\wp$ for some ideal $I \subset A$. But $IA' \neq A'$ if and only if $I \cap (A \backslash \wp) = \emptyset$, i.e. $I \subseteq \wp$ or equivalently $IA_\wp \subseteq \wp A_\wp$.

□

Proposition 2.3.4. — *Under the notation of Lemma 2.3.3, the following holds:*

1. *If A is a Noetherian ring, then so is $A' = S^{-1}A$.*
2. *If B is the integral closure of A is a finite extension L/K, then $S^{-1}B$ is the integral closure of $S^{-1}A$ in L.*
3. *If A is integrally closed, then so A'.*
4. *If A is Dedekind, then so is A'.*

Proof. — Statement (1) follows from the fact that every ideal of A' has the form IA' with $I \subset A$ an ideal.

For (2), let $x = b/s \in S^{-1}B$. If $f(T) = T^n + a_1T^{n-1} + \cdot + a_n \in A[T]$ is a monic polynomial such that $f(b) = 0$, then $g(x) = 0$ with $g(T) = T^n + a_1s^{-1}T^{n-1} + \cdots + a_ns^{-n} \in S^{-1}A$. Conversely, if $x \in L$ is integral over $S^{-1}A$ and $g(x) = 0$ for some monic polynomial $g(T) = T^n + c_1T^{n-1} + \cdots + c_n$ with $c_i \in S^{-1}A$, then there exists $s \in S$ such that $sc_i \in A$. Then sx is the root of $f(T) = T^n + sc_1T^{n-1} + \cdots + s^nc_n \in A[T]$. Therefore, $sx \in B$ and $x \in S^{-1}B$. Statement (3) is a special case of (2).

For (4), we note that if $\mathfrak{m} \subseteq A$ is a maximal ideal disjoint with S then $S^{-1}\mathfrak{m} \subseteq A'$ is also maximal, because the image of S in $A/\mathfrak{m}$ is already invertible. Combined (1) and (3), we see that A' is also a Dedekind domain.

□

Proposition 2.3.5. — *Let A be a Dedekind domain, and $A' = S^{-1}A$ for some multiplicative subset S.*

1. *Let $\mathfrak{p} \subseteq A$ be a non-zero prime, and $\mathfrak{p}' = \mathfrak{p}A'$. Then $\mathfrak{p}' = A'$ if $S \cap \mathfrak{p} \neq \emptyset$, and $\mathfrak{p}' \subset A'$ is a maximal ideal of A' with $A/\mathfrak{p} \cong A'/\mathfrak{p}'$ if $\mathfrak{p}$ is disjoint with S.*
2. *If I is a fractional ideal of A with prime decomposition $I = \prod_{i=1}^r \mathfrak{p}_i^{a_i}$, then $I' = IA'$ is a fractional ideal of A' with prime decomposition $I' = \prod_{i=1}^r \mathfrak{p}_i'^{a_i}$, where $\mathfrak{p}_i' = \mathfrak{p}_iA'$.*
3. *Let $I = \prod_{i=1}^r \mathfrak{p}_i^{e_i}$ be an ideal of A such that each $e_i > 0$ and $\mathfrak{p}_i$ is disjoint from S. Then the natural map $A/I \xrightarrow{\sim} A'/IA'$ is an isomorphism.*
4. *Assume that $S = A\backslash\wp$ for some prime $\wp \subset A$. Then all the nonzero ideals of $A' = A_\wp$ is of the form $\wp'^n$ with $\wp' = \wp A_\wp$ for $n \geq 1$. Moreover, if $\pi \in \wp\backslash\wp^2$, then we have $\wp'^n = \pi^n A_\wp$; in particular, $A_\wp$ is a principal ideal domain.*

Proof. — Statement (1) is clear.

(2) It is clear that I' is a fractional ideal of A'. Note that $J^{-1}A' = (JA')^{-1}$ for any fractional ideal J of A, and $(J_1J_2)A' = J_1A'J_2A'$ for any fractional ideals J_1, J_2 of A. The statement (2) follows immediately.

To prove (3), we proceed by induction on $\ell(I) := \sum_i e_i \geq 1$. When $\ell(I) = 1$, then $I = \mathfrak{p}$ is a prime disjoint with S. The statement is verified in (1). Now assume that $\ell(I) = n > 1$ and the statement is true for any ideal $J \subset A$ with $\ell(J) = n - 1$. Put $J = \mathfrak{p}_1^{e_1-1} \prod_{i=2}^r \mathfrak{p}_i^{e_i}$, $I' = IA'$ and $J' = JA'$. Then we have a commutative diagram of exact sequences:

$$\begin{array}{ccccccccc} 0 & \longrightarrow & J/I & \longrightarrow & A/I & \longrightarrow & A/J & \longrightarrow & 0 \\ & & \downarrow\cong & & \downarrow & & \downarrow\cong & & \\ 0 & \longrightarrow & J'/I' & \longrightarrow & A'/I' & \longrightarrow & A'/J' & \longrightarrow & 0 \end{array}$$

By inductive assumption, the last vertical arrow is an isomorphism. Note that J/I and J'/I' are both vector spaces over $k(\mathfrak{p}_1) := A/\mathfrak{p}_1$. If $x \in J\backslash I$, then $J = xA + I$ and $J' = xA' + I'$. Thus the $k(\mathfrak{p}_1)$-dimensions of J/I and J'/I' are both one. Choose $x_1 \in J\backslash I$. Then the image of x_1 in J'/I' is non-zero, since $v_{\mathfrak{p}_1'}(x_1) = v_\mathfrak{p}(x)$ by (2). Therefore, the first vertical arrow is also an isomorphism. It follows from an easy diagram chasing that so is the middle vertical one.

(4) If I is an ideal of A, we have $IA_\wp = \wp'^{v_\wp(I)} = \wp^{v_\wp(I)}A_\wp$ by statement (2). Thus all the ideals of A' are of the form $\wp'^n$. Let $x \in \wp'^n$. Then $v_\wp(x/\pi^n) = v_\wp(x) - n \geq 0$. Note that $\wp'$ is the unique prime of the Dedekind domain A'. The statement follows from Corollary 2.2.12(1).

□

CHAPTER 3

DECOMPOSITION OF PRIMES IN NUMBER FIELDS

3.1. Norms of ideals

Let K be a number field, and $\mathcal{O}_K$ be its ring of integers.

Definition 3.1.1. — Let $0 \neq I \subseteq \mathcal{O}_K$ be an ideal. Define the norm of I to be

$$\mathrm{N}(I) = \#(\mathcal{O}_K/I) = [\mathcal{O}_K : I].$$

Proposition 3.1.2. — 1. *If $I = (x)$ for some $x \in \mathcal{O}_K$, then $\mathrm{N}(I) = |\mathrm{N}_{K/\mathbb{Q}}(x)|$.*
2. *We have $\mathrm{N}(IJ) = \mathrm{N}(I)\mathrm{N}(J)$ for any ideals $I, J \subseteq \mathcal{O}_K$.*
3. *For $n \in \mathbb{Z}_{\geq 0}$, there exist only finitely many ideals $I \subseteq \mathcal{O}_K$ such that $\mathrm{N}(I) = n$.*

Proof. — (1) Let $(\alpha_1, \cdots, \alpha_n)$ be a $\mathbb{Z}$-basis of $\mathcal{O}_K$. Then there exists a matrix $C \in \mathrm{M}_{n\times n}(\mathbb{Z})$ such that

$$(x\alpha_1, \cdots, x\alpha_n) = (\alpha_1, \cdots, \alpha_n)C.$$

It follows that

$$\mathrm{N}(I) = [\mathcal{O}_K : I] = [\sum_i \mathbb{Z} \cdot \alpha_i : \sum_i \mathbb{Z} \cdot x\alpha_i] = |\det(C)|.$$

But by definition, $\mathrm{N}_{K/\mathbb{Q}}(x) = \det(C)$.

(2) By Theorem 2.2.6, it suffices to show that

$$\mathrm{N}(\prod_{i=1}^{r} \mathfrak{p}_i) = \mathrm{N}(\mathfrak{p}_1)\mathrm{N}(\prod_{i=2}^{r} \mathfrak{p}_i)$$

for any prime ideals $\mathfrak{p}_1, \cdots, \mathfrak{p}_r$. First, note that $k(\mathfrak{p}_1) := \mathcal{O}_K/\mathfrak{p}_1$ is a finite field, since $\mathfrak{p}_1 \subseteq \mathcal{O}_K$ is maximal. We claim that $\prod_{i=2}^{r} \mathfrak{p}_i / \prod_{i=1}^{r} \mathfrak{p}_i$ is a $k(\mathfrak{p}_1)$-vector space of dimension 1. Assuming this claim, we see that

$$\frac{[\mathcal{O}_K : \prod_{i=1}^{r} \mathfrak{p}_i]}{[\mathcal{O}_K : \prod_{i=2}^{r} \mathfrak{p}_i]} = [\prod_{i=2}^{r} \mathfrak{p}_i : \prod_{i=1}^{r} \mathfrak{p}_i] = \#k(\mathfrak{p}_1) = \mathrm{N}(\mathfrak{p}_1),$$

which is clearly equivalent to the assertion needed. It remains to prove the claim. Since $\prod_{i=1}^{r} \mathfrak{p}_i \neq \prod_{i=2}^{r} \mathfrak{p}_i$ by Theorem 2.2.6, there exists $x \in \prod_{i=2}^{r} \mathfrak{p}_i$ but $x \notin \prod_{i=1}^{r} \mathfrak{p}_i$. Then we

have

$$\prod_{i=1}^{r}\mathfrak{p}_i \subsetneq (x)+\prod_{i=1}^{r}\mathfrak{p}_i \subseteq \prod_{i=2}^{r}\mathfrak{p}_i \Longrightarrow \mathfrak{p}_1 \subsetneq (x)\prod_{i=2}^{r}\mathfrak{p}_i^{-1}+\mathfrak{p}_1 \subseteq A.$$

It follows immediately that $(x)\prod_{i=2}^{r}\mathfrak{p}_i^{-1}+\mathfrak{p}_1 = A$, that is $(x)+\prod_{i=1}^{r}\mathfrak{p}_i = \prod_{i=2}^{r}\mathfrak{p}_i$.

(3) If $I \subseteq \mathcal{O}_K$ if an ideal of norm n, then $(n) \subseteq I \subseteq \mathcal{O}_K$. Note that $\mathcal{O}_K/(n)$ is finite of cardinality $n^{[K:\mathbb{Q}]}$. Therefore, there are only finitely many possibilities for I. □

If $I = \mathfrak{a}\mathfrak{b}^{-1}$ is a fractional ideal with $\mathfrak{a}, \mathfrak{b} \subset A$ ideals, then we define the norm of I as

$$\mathrm{N}(I) := \frac{\mathrm{N}(\mathfrak{a})}{\mathrm{N}(\mathfrak{b})} \in \mathbb{Q}^{\times}.$$

Using the previous Proposition, we check easily that $\mathrm{N}(I)$ is independent of the writing $I = \mathfrak{a}\mathfrak{b}^{-1}$.

Definition 3.1.3. — We put

$$\delta_K^{-1} = \{x \in K | \mathrm{Tr}_{K/\mathbb{Q}}(xy) \in \mathbb{Z},\ \forall y \in \mathcal{O}_K\}.$$

This is a fractional ideal containing $\mathcal{O}_K$. We define the (absolute) *different* of K to be $\delta_K = (\delta_K^{-1})^{-1}$.

Proposition 3.1.4. — *We have* $\mathrm{N}(\delta_K) = |\Delta_K|$.

Proof. — Let $(\alpha_1, \cdots, \alpha_n)$ be a $\mathbb{Z}$-basis of $\mathcal{O}_K$, and $(\alpha_1^\vee, \cdots, \alpha_n^\vee)$ be its dual basis with respect to $\mathrm{Tr}_{K/\mathbb{Q}}$. Then $\delta_K^{-1} = \bigoplus_i \mathbb{Z}\cdot\alpha^\vee$, and $\alpha_i = \sum_j \mathrm{Tr}_{K/\mathbb{Q}}(\alpha_i\alpha_j)\alpha_j^\vee$. Therefore,

$$\begin{aligned}|\Delta_K| &= |\det(\mathrm{Tr}_{K/\mathbb{Q}}(\alpha_i\alpha_j))| = [\bigoplus_i \mathbb{Z}\alpha_i^\vee : \bigoplus_i \mathbb{Z}\alpha_i] \\ &= [\delta_K^{-1} : \mathcal{O}_K] = [\mathcal{O}_K : \delta_K] = \mathrm{N}(\delta_K).\end{aligned}$$

□

3.2. Decomposition of primes in extension of number fields

Let L/K be a finite extension of number fields, and $\mathfrak{p} \neq 0$ be a prime of $\mathcal{O}_K$. We have a prime decomposition

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

Definition 3.2.1. — (1) We put

$$e(\mathfrak{P}_i/\mathfrak{p}) = e_i = v_{\mathfrak{P}_i}(\mathfrak{p}\mathcal{O}_L),$$

and call it the *ramification index* of $\mathfrak{P}_i$ above $\mathfrak{p}$.

(2) Note that $k(\mathfrak{P}_i) = \mathcal{O}_L/\mathfrak{P}_i$ is a finite extension of $k(\mathfrak{p}) = \mathcal{O}_K/\mathfrak{p}$. We put

$$f(\mathfrak{P}_i|\mathfrak{p}) = [k(\mathfrak{P}_i) : k(\mathfrak{p})],$$

and call it the *residue degree* of $\mathfrak{P}_i$ above $\mathfrak{p}$.

(3) We say that $\mathfrak{p}$ is

– *unramified in* L/K, if $e(\mathfrak{P}_i|\mathfrak{p}) = 1$ for all i,

– *split* in L/K, if $e(\mathfrak{P}_i|\mathfrak{p}) = f(\mathfrak{P}_i|\mathfrak{p}) = 1$ for all i;
– *inert* in L/K, if $g = 1$ and $e(\mathfrak{P}_1|\mathfrak{p}) = 1$;
– *ramified* in L/K, if $e(\mathfrak{P}_i|\mathfrak{p}) > 1$ for some i;
– *totally ramified* in L/K, if $g = 1$ and $f(\mathfrak{P}_1|\mathfrak{p}) = 1$.

Proposition 3.2.2. — *Under the above assumptions, then the following statements hold:*

1. *A prime $\mathfrak{P}$ of $\mathcal{O}_L$ appears in $\mathfrak{p}\mathcal{O}_L$ if and only if $\mathfrak{P} \cap \mathcal{O}_K = \mathfrak{p}$.*
2. *We have $\sum_{i=1}^{g} e(\mathfrak{P}_i|\mathfrak{p})f(\mathfrak{P}_i|\mathfrak{p}) = [L : K]$.*

Proof. — (1) Note that $\mathfrak{P} \cap \mathcal{O}_K$ is always a non-zero prime of $\mathcal{O}_K$. Statement (1) follows immediately from the prime decomposition of $\mathfrak{p}\mathcal{O}_L$.

(2) Let q denote the cardinality of the residue field $k(\mathfrak{p}) = \mathcal{O}_K/\mathfrak{p}$. Then we have

$$[\mathcal{O}_L : \mathfrak{p}\mathcal{O}_L] = \mathrm{N}(\mathfrak{p}\mathcal{O}_L) = \prod_{i=1}^{g} \mathrm{N}(\mathfrak{P}_i)^{e_i} = \prod_{i=1}^{g} q^{e_i f_i} = q^{\sum_{i=1}^{g} e_i f_i}.$$

Note that $\mathcal{O}_L/\mathfrak{p}\mathcal{O}_L$ is a finite dimensional vector space over $k(\mathfrak{p})$. Thus the above computation shows that

$$\dim_{k(\mathfrak{p})} \mathcal{O}_L/\mathfrak{p}\mathcal{O}_L = \sum_{i=1}^{g} e_i f_i.$$

To conclude the proof, we have to show that $\dim_{k(\mathfrak{p})}(\mathcal{O}_L/\mathfrak{p}\mathcal{O}_L) = [L : K]$.

– Consider first the special case that $\mathcal{O}_L$ is a free module over $\mathcal{O}_K$ (e.g. $K = \mathbb{Q}$). Then the rank of $\mathcal{O}_L$ over $\mathcal{O}_K$ must be $[L : K]$ (because $L = \mathcal{O}_L \otimes_{\mathcal{O}_K} K$), and $\mathcal{O}_L/\mathfrak{p}\mathcal{O}_L = \mathcal{O}_L \otimes_{\mathcal{O}_K} \mathcal{O}_K/\mathfrak{p}\mathcal{O}_K$ is also free of rank $[L : K]$ over $k(\mathfrak{p}) = \mathcal{O}_K/\mathfrak{p}\mathcal{O}_K$. Thus our proof is finished in this case.
– In the general case, we consider the localizations of $\mathcal{O}_K$ and $\mathcal{O}_L$ with respect to the multiplicative subset $S = \mathcal{O}_K \backslash \mathfrak{p}$; we denote the localized rings respectively by $\mathcal{O}_{K,\mathfrak{p}}$ and $\mathcal{O}_{L,\mathfrak{p}}$. By Proposition 2.3.4, both $\mathcal{O}_{K,\mathfrak{p}}$ and $\mathcal{O}_{L,\mathfrak{p}}$ are Dedekind domains; moreover, $\mathcal{O}_{K,\mathfrak{p}}$ is a principal ideal domain by 2.3.5(4). Since $\mathcal{O}_{L,\mathfrak{p}}$ is a finitely generated torsion free $\mathcal{O}_K$-module ($\mathcal{O}_L$ is finitely generated over $\mathbb{Z}$ hence over $\mathcal{O}_K$), thus $\mathcal{O}_{L,\mathfrak{p}}$ must be free over $\mathcal{O}_{K,\mathfrak{p}}$ of rank $[L : K]$. It follows that $\mathcal{O}_{L,\mathfrak{p}}/\mathfrak{p}\mathcal{O}_{L,\mathfrak{p}}$ is of dimension $[L : K]$ over $k(\mathfrak{p})$. But $\mathcal{O}_L/\mathfrak{p}\mathcal{O}_L \cong \mathcal{O}_{L,\mathfrak{p}}/\mathfrak{p}\mathcal{O}_{L,\mathfrak{p}}$ by Proposition 2.3.5, this implies that $\dim_{k(\mathfrak{p})}(\mathcal{O}_L/\mathfrak{p}\mathcal{O}_L) = [L : K]$.

□

Theorem 3.2.3 **(Kummer)**. — *Let $\alpha \in \mathcal{O}_L$ be such that $\mathcal{O}_L/\mathfrak{p}\mathcal{O}_L = (\mathcal{O}_K/\mathfrak{p})[\bar{\alpha}]$, where $\bar{\alpha}$ denote the image of α. Let $f(X) \in \mathcal{O}_K[X]$ be the minimal polynomial of α. Assume that*

$$f(X) \equiv \prod_{i=1}^{g} g_i(X)^{e_i} \mod \mathfrak{p}\mathcal{O}_K[X],$$

where $e_i \geq 1$, and $g_i(X)$ is a monic polynomial whose image in $k(\mathfrak{p})[X]$ is irreducible and distinct with each other. Then $\mathfrak{P}_i = \mathfrak{p}\mathcal{O}_L + g_i(\alpha)\mathcal{O}_L$ is a maximal ideal of $\mathcal{O}_L$ for each i, and we have the prime decomposition

$$\mathfrak{p}\mathcal{O}_L = \mathfrak{P}_1^{e_1} \cdots \mathfrak{P}_g^{e_g}, \tag{3.2.3.1}$$

with residue degrees $f(\mathfrak{P}_i|\mathfrak{p}) = \deg(g_i)$.

Proof. — Put $k(\mathfrak{p}) = \mathcal{O}_K/\mathfrak{p}$. We have

$$\mathcal{O}_L/\mathfrak{P}_i = (\mathcal{O}_L/\mathfrak{p}\mathcal{O}_L)/(\bar{g}_i(\bar{\alpha})) = (\mathcal{O}_K/\mathfrak{p})[\bar{\alpha}]/(\bar{g}_i(\bar{\alpha})) \cong k(\mathfrak{p})[X]/(\bar{g}_i(X)).$$

Since $\bar{g}_i(X)$ is irreducible in $k(\mathfrak{p})[X]$, the quotient $k(\mathfrak{p})[X]/(\bar{g}_i(X))$ is a field. This shows that $\mathfrak{P}_i$ is a maximal ideal of $\mathcal{O}_L$. Moreover, we have

$$f(\mathfrak{P}_i|\mathfrak{p}) = [\mathcal{O}_L/\mathfrak{P}_i : \mathcal{O}_K/\mathfrak{p}] = \deg(\bar{g}_i) = \deg(g_i).$$

To prove the decomposition 3.2.3.1, we note that

$$\mathcal{O}_L/\mathfrak{p}\mathcal{O}_L = k(\mathfrak{p})[\bar{\alpha}] \cong k(\mathfrak{p})[X]/(\bar{f}(X)) \cong \prod_{i=1}^{g} k(\mathfrak{p})[X]/(\bar{g}_i^{e_i}(X)).$$

Here, the last step used Chinese reminder theorem. On the other hand, note that

$$k(\mathfrak{p})[X]/(\bar{g}_i^{e_i}(X)) \xrightarrow{\sim} (\mathcal{O}_L/\mathfrak{p}\mathcal{O}_L)/(\bar{g}_i(\bar{\alpha})) \cong \mathcal{O}_L/(\mathfrak{p}\mathcal{O}_L + g_i^{e_i}(\alpha)).$$

Hence, to finish the proof, it suffices to show that $\mathfrak{P}_i^{e_i} = (\mathfrak{p}, g_i^{e_i}(\alpha))$ for $1 \leq j \leq e_i$. We have $\mathfrak{P}_i^{e_i} = (\mathfrak{p}, g_i(\alpha))^{e_i} \subseteq (\mathfrak{p}, g_i^{e_i}(\alpha))$. We deduce $\mathfrak{P}_i^{e_i} = (\mathfrak{p}, g_i^{e_i}(\alpha))$ from the equality

$$\begin{aligned}\dim_{k(\mathfrak{p})} \mathcal{O}_L/(\mathfrak{p}, g_i^{e_i}(\alpha)) &= \dim_{k(\mathfrak{p})} k(\mathfrak{p})[X]/(\bar{g}_i^{e_i}(X)) = e_i \dim_{k(\mathfrak{p})} k(\mathfrak{p})[X]/(\bar{g}_i(X)) \\ &= e_i \dim_{k(\mathfrak{p})} \mathcal{O}_L/\mathfrak{P}_i = \dim_{k(\mathfrak{p})} \mathcal{O}_L/\mathfrak{P}_i^{e_i}.\end{aligned}$$

□

Remark 3.2.4. — We have two important special cases where the assumption $\mathcal{O}_L/\mathfrak{p}\mathcal{O}_L \cong (\mathcal{O}_K/\mathfrak{p})[\bar{\alpha}]$ is satisfied:

1. If $\mathcal{O}_L = \mathcal{O}_K[\alpha]$, then Theorem 3.2.3 can be applied to any prime $\mathfrak{p}$ of $\mathcal{O}_K$.
2. If $\alpha \in \mathcal{O}_L$ such that $\mathfrak{p} \nmid \mathrm{N}_{L/K}(f'(\alpha))$, then $\mathcal{O}_L/\mathfrak{p}\mathcal{O}_L = (\mathcal{O}_K/\mathfrak{p})[\bar{\alpha}]$.

Theorem 3.2.5. — *Let* $K = \mathbb{Q}(\sqrt{D})$ *with* D *a square-free integer. Let* p *be a rational prime. Then*

1. p *is ramified in* K *if and only if* $p|\Delta_K$*; in particular,* 2 *is ramified in* K *if and only if* $D \equiv 2, 3 \mod 4$*;*
2. *if* p *is odd and unramified in* K*, then* p *splits in* K *if and only if* $\left(\frac{D}{p}\right) = 1$*;*
3. *when* $D \equiv 1 \mod 4$*, then* 2 *splits in* K *if and only if* $D \equiv 1 \mod 8$.

Proof. — We write $\mathcal{O}_K = \mathbb{Z}[\alpha]$, for $\alpha = \sqrt{D}$ if $D \equiv 2, 3 \mod 4$ and $\alpha = \frac{-1+\sqrt{D}}{2}$ if $D \equiv 1 \mod 4$. Then the minimal polynomial of α is

$$f(x) = \begin{cases} x^2 + x + \frac{1-D}{4} & \text{if } D \equiv 1 \mod 4, \\ x^2 - D & \text{if } D \equiv 2, 3 \mod 4, \end{cases}$$

and Δ_K coincides with the discriminant of $f(x)$.

(1) By Theorem 3.2.3, p is ramified in K if and only if $\bar{f}(x) = (x-a)^2$ for some $a \in \mathbb{F}_p$, where $\bar{f}(x) \in \mathbb{F}_p[x]$ denotes the image of $f(x)$. The latter condition is equivalent to saying that $p|\Delta_K$.

(2) Assume p odd and unramified in D. We have $p \nmid \Delta_K$ by (1). By Theorem 3.2.3, we have the following equivalence:

$$p \text{ splits in } K \quad \Leftrightarrow \quad \bar{f}(x) \text{ has distinct roots in} \mathbb{F}_p.$$

So if $\bar{f}(x) = (x-a)(x-b)$ with $a, b \in \mathbb{F}_p$ and $a \neq b$, then $\Delta_K \equiv (a-b)^2 \mod p$, or equivalently $\left(\frac{D}{p}\right) = 1$. Conversely, if $\left(\frac{D}{p}\right) = 1$, assume that $D \equiv c^2 \mod p$ with $p \nmid c$. Then $\frac{1 \pm c}{2}$ (resp. $\pm c$) are two distinct roots of $\bar{f}(x)$ in $\mathbb{F}_p$ if $D \equiv 1 \mod 4$ (resp. if $D \equiv 2, 3 \mod 4$).

(3) If $D \equiv 1 \mod 8$, then $\bar{f}(X) = X^2 + X$ has two distinct roots in $\mathbb{F}_2$. If $D \equiv 5 \mod 8$, then $\bar{f}(X) = X^2 + X + 1$ is the unique irreducible polynomial of degree 2 in $\mathbb{F}_2[X]$. □

We have the following transitivity of ramification and residue indexes:

Proposition 3.2.6. — *Let L/K be as above, and M/L be another finite extension. Let $\mathfrak{P}_M$ be a prime ideal of M, $\mathfrak{P}_L = \mathfrak{P}_M \cap \mathcal{O}_L$ and $\mathfrak{p} = \mathcal{O}_K \cap \mathfrak{P}_M$. Then we have*

$$f(\mathfrak{P}_M|\mathfrak{p}) = f(\mathfrak{P}_M|\mathfrak{P}_L) f(\mathfrak{P}_L|\mathfrak{p}), \quad e(\mathfrak{P}_M|\mathfrak{p}) = e(\mathfrak{P}_M|\mathfrak{P}_L) e(\mathfrak{P}_L|\mathfrak{p}).$$

Proof. — The equality for $f(\mathfrak{P}_M|\mathfrak{p})$ follows from

$$[\mathcal{O}_M/\mathfrak{P}_M : \mathcal{O}_K/\mathfrak{p}] = [\mathcal{O}_M/\mathfrak{P}_M : \mathcal{O}_L/\mathfrak{P}_L][\mathcal{O}_L/\mathfrak{P}_L : \mathcal{O}_K/\mathfrak{p}].$$

For the equalities on ramification indexes, we have

$$\mathfrak{p}\mathcal{O}_M = \mathfrak{p}\mathcal{O}_L \cdot \mathcal{O}_M = \prod_{\mathfrak{P}_L|\mathfrak{p}} \mathfrak{P}_L^{e(\mathfrak{P}_L|\mathfrak{p})} \mathcal{O}_M = \prod_{\mathfrak{P}_L|\mathfrak{p}\mathcal{O}_L} \Big(\prod_{\mathfrak{P}_M|\mathfrak{P}_L} \mathfrak{P}_M^{e(\mathfrak{P}_M|\mathfrak{P}_L)} \Big)^{e(\mathfrak{P}_L|\mathfrak{p})}.$$

Note that for a fixed $\mathfrak{P}_M$, there exists a unique prime $\mathfrak{P}_L$ of $\mathcal{O}_L$ such that $\mathfrak{P}_M|\mathfrak{P}_L$, namely $\mathfrak{P}_L = \mathfrak{P}_M \cap \mathcal{O}_L$. Therefore, we get

$$\mathfrak{p}\mathcal{O}_M = \prod_{\mathfrak{P}_M|\mathfrak{p}} \mathfrak{P}_M^{e(\mathfrak{P}_M|\mathfrak{P}_L) e(\mathfrak{P}_L|\mathfrak{p})},$$

that is, $e(\mathfrak{P}_M|\mathfrak{p}) = e(\mathfrak{P}_M|\mathfrak{P}_L) e(\mathfrak{P}_L|\mathfrak{p})$. □

Finally, we give a criterion for a prime p to be ramified in a number field.

Theorem 3.2.7. — *Let K be a number fields, p be a rational prime. Then the following statements are equivalent:*

1. *p is unramifies in K.*
2. *The ring $\mathcal{O}_K/p\mathcal{O}_K$ is reduced (i.e. it has no nilpotent elements).*
3. *The $\mathbb{F}_p$-bilinear map $\overline{\mathrm{Tr}}_{K/\mathbb{Q}} : \mathcal{O}_K/(p) \times \mathcal{O}_K/(p) \to \mathbb{F}_p$ sending (x, y) to $\mathrm{Tr}_{K/\mathbb{Q}}(xy)$ mod p is non-degenerate.*
4. *$p \nmid \Delta_K$, where Δ_K denotes the discriminant of K.*

Proof. — (1) ⇔ (2): By Chinese remainder Theorem, we have

$$\mathcal{O}_K/p\mathcal{O}_K \cong \prod_{\mathfrak{p}|p} \mathcal{O}_K/\mathfrak{p}^{e(\mathfrak{p}/p)}. \tag{3.2.7.1}$$

Note that each $\mathcal{O}_K/\mathfrak{p}^e$ is reduced if and only if $e(\mathfrak{p}|p) = 1$, because $\bar{x}^{e(\mathfrak{p}/p)} = 0$ for any $x \in \mathfrak{p}/\mathfrak{p}^2$.

(2) $\Leftrightarrow$ (3): Note that if $x \in \mathcal{O}_K/p\mathcal{O}_K$ is nilpotent, then xy is also nilpotent for any $y \in \mathcal{O}_K/\mathfrak{p}\mathcal{O}_K$, hence $\overline{\mathrm{Tr}}_{K/\mathbb{Q}}(xy) = 0$. Hence, if $\overline{\mathrm{Tr}}_{K/\mathbb{Q}}$ is non-degenerate, then $\mathcal{O}_K/p\mathcal{O}_K$ is reduced. Conversely, if $\mathcal{O}_K/p\mathcal{O}_K$ is reduced, then we have necessarily $\mathcal{O}_K/p\mathcal{O}_K = \bigoplus_{\mathfrak{p}|p} k(\mathfrak{p})$ by (3.2.7.1), where $k(\mathfrak{p}) = \mathcal{O}_K/\mathfrak{p}$ is a finite extension of $\mathbb{F}_p$. Since $\mathbb{F}_p$ is perfect, $\mathrm{Tr}_{k(\mathfrak{p})/\mathbb{F}_p}$ is non-degenerate by Theorem 1.2.4. It follows that $\overline{\mathrm{Tr}}_{K/\mathbb{Q}} = \bigoplus_{\mathfrak{p}|p} \mathrm{Tr}_{k(\mathfrak{p})/\mathbb{F}_p}$ is non-degenerate.

(3) $\Leftrightarrow$ (4): Let $(\alpha_i)_{1\leq i\leq n}$ denote a basis of $\mathcal{O}_K$ over $\mathbb{Z}$, and $\bar{\alpha}_i \in \mathcal{O}_K/(p)$ denote the image of α_i. The pairing $\overline{\mathrm{Tr}}_{K/\mathbb{Q}}$ on $\mathcal{O}_K/(p)$ induces an $\mathbb{F}_p$-linear map: $\phi : \mathcal{O}_K/(p) \to (\mathcal{O}_K/(p))^\vee$, where $(\mathcal{O}_K/(p))^\vee$ denotes the $\mathbb{F}_p$-dual of $\mathcal{O}_K/(p)$. If $(\bar{\alpha}_i^\vee)_{1\leq i\leq n}$ denotes the basis of $(\mathcal{O}_K/(p))^\vee$ dual to $(\bar{\alpha}_i)_{1\leq i\leq n}$, then the matrix of ϕ under the basis $(\bar{\alpha}_i)_i$ and $(\bar{\alpha}_i^\vee)_i$ is $\overline{\mathrm{Tr}}_{K/\mathbb{Q}}(\bar{\alpha}_i\bar{\alpha}_j)$. Hence, the pairing $\overline{\mathrm{Tr}}_{K/\mathbb{Q}}$ is non-degenerate if and only if $\det(\overline{\mathrm{Tr}}_{K/\mathbb{Q}}(\bar{\alpha}_i\bar{\alpha}_j)) \neq 0$ in $\mathbb{F}_p$, i.e. $p \nmid \det(\mathrm{Tr}_{L/K}(\alpha_i\alpha_j)) = \Delta_K$. This finishes the proof.

□

3.3. Relative different and discriminant

Let L/K be a finite extension of number fields.

Definition 3.3.1. — For a non-zero prime ideal $\mathfrak{P}$ of $\mathcal{O}_L$, we put

$$\mathrm{N}_{L/K}(\mathfrak{P}) = \mathfrak{p}^{f(\mathfrak{P}/\mathfrak{p})},$$

where $\mathfrak{p} = \mathfrak{P} \cap \mathcal{O}_K$, and $f(\mathfrak{P}|\mathfrak{p}) = [\mathcal{O}_L/\mathfrak{P} : \mathcal{O}_K/\mathfrak{p}]$ is the residue degree of $\mathfrak{P}/\mathfrak{p}$. For an arbitrary fractional ideal $I = \prod_{i=1}^r \mathfrak{P}_i^{a_i}$, we put

$$\mathrm{N}_{L/K}(I) := \prod_{i=1}^r \mathrm{N}(\mathfrak{P}_i)^{a_i}.$$

Then $\mathrm{N}_{L/K}(I)$ is a fractional ideal of K, and we call it *the norm of I relative to L/K*.

Lemma 3.3.2. —
1. *We have* $\mathrm{N}_{L/K}(IJ) = \mathrm{N}_{L/K}(I)\mathrm{N}_{L/K}(J)$ *for any fractional ideals* I, J *of* L.
2. *When* $K = \mathbb{Q}$, *then we have* $\mathrm{N}_{L/\mathbb{Q}}(I) = (\mathrm{N}(I))$ *for any fractional ideal* I *of* L, *where* $\mathrm{N}(I) \in \mathbb{Q}^\times$ *is the absolute norm of* I *defined in Section 3.1.*
3. *If* $I = J\mathcal{O}_L$ *for some ideal* $J \subseteq \mathcal{O}_K$, *then* $\mathrm{N}_{L/K}(I) = J^{[L:K]}$.
4. *If* M/L *is another finite extension, then one has*

$$\mathrm{N}_{M/K}(I) = \mathrm{N}_{L/K}(\mathrm{N}_{M/L}(I))$$

for any fractional ideal I *of* M.

Proof. — Statement (1) is immediate from the definition. Statement (2) follows from the fact that, if $\mathfrak{P}$ is a prime of $\mathcal{O}_L$ above p, then $p^{f(\mathfrak{P}|p)} = \#(\mathcal{O}_L/\mathfrak{P})$. To prove (3), we may

assume that $J = \mathfrak{p}$ is a prime of $\mathcal{O}_K$. If $\mathfrak{p}\mathcal{O}_L = \prod_{i=1}^{g} \mathfrak{P}_i^{e_i}$ is the prime decomposition of $\mathfrak{p}$ in $\mathcal{O}_L$, then

$$\mathrm{N}_{L/K}(\mathfrak{p}\mathcal{O}_L) = \prod_{i=1}^{g} \mathrm{N}_{L/K}(\mathfrak{P}_i)^{e_i} = \mathfrak{p}^{\sum_{i=1}^{g} e_i f_i} = \mathfrak{p}^{[L:K]}.$$

Finally, (4) is an easy consequence of Proposition 3.2.6. □

Definition 3.3.3. — We put

$$\delta_{L/K}^{-1} := \{x \in L | \mathrm{Tr}_{L/K}(xy) \in \mathcal{O}_K, \quad \forall y \in \mathcal{O}_L\}.$$

This is fractional ideal of L containing $\mathcal{O}_L$. We put $\delta_{L/K} := (\delta_{L/K}^{-1})^{-1}$, and call it the *relative different of* L/K. We define *the relative discriminant of* L/K to be

$$\mathrm{Disc}_{L/K} = \mathrm{N}_{L/K}(\delta_{L/K}).$$

It is clear that $\delta_{L/\mathbb{Q}} = \delta_L$ defined in Definition 3.1.3, and $\mathrm{Disc}_{L/\mathbb{Q}} = (\Delta_L)$ by Proposition 3.1.4.

Lemma 3.3.4. — *For any fractional ideal I of $\mathcal{O}_K$, we have*

$$I\mathcal{O}_L \cdot \delta_{L/K}^{-1} = \{x \in L | \mathrm{Tr}_{L/K}(x) \in I\}.$$

Proof. — The statement follows from the following equivalences:

$$\mathrm{Tr}_{L/K}(x) \in I \Leftrightarrow \mathrm{Tr}_{L/K}(xI^{-1}) \in \mathcal{O}_K \Leftrightarrow xI^{-1}\mathcal{O}_L \subseteq \delta_{L/K}^{-1} \Leftrightarrow x \in I\mathcal{O}_L \cdot \delta_{L/K}^{-1}.$$

□

Proposition 3.3.5. — *If M/L is a further finite extension, then $\delta_{M/K} = \delta_{L/K}\mathcal{O}_M \cdot \delta_{M/L}$.*

Proof. — This follows from:

$$\begin{aligned} x \in \delta_{M/K} &\Leftrightarrow \mathrm{Tr}_{M/K}(xy) \in \mathcal{O}_K, \quad \forall y \in \mathcal{O}_M \\ &\Leftrightarrow \mathrm{Tr}_{L/K} \circ \mathrm{Tr}_{M/L}(xy) \in \mathcal{O}_K, \quad \forall y \in \mathcal{O}_M \\ &\Leftrightarrow \mathrm{Tr}_{M/L}(xy) \in \delta_{L/K}^{-1}, \quad \forall y \in \mathcal{O}_M \\ &\Leftrightarrow x \in \delta_{M/L}^{-1} \cdot \delta_{L/K}^{-1}\mathcal{O}_L \quad \text{(by the previous Lemma)}. \end{aligned}$$

□

Corollary 3.3.6. — *Under the situation of the Proposition, one has*

$$\mathrm{Disc}_{M/K} = \mathrm{N}_{L/K}(\mathrm{Disc}_{M/L})\mathrm{Disc}_{L/K}^{[M:L]}.$$

Proof. — This follows easily from the Proposition by applying $\mathrm{N}_{M/K}$. □

Proposition 3.3.7. — *Let L_1 and L_2 be number fields such that $L_1 \cap L_2 = \mathbb{Q}$, and $M = L_1L_2$. Then*

1. $\delta_{L_2}\mathcal{O}_M \subseteq \delta_{M/L_1}$;
2. Δ_M *divides* $\Delta_{L_1}^{[L_2:\mathbb{Q}]} \cdot \Delta_{L_2}^{[L_1:\mathbb{Q}]}$;

3. *if* $\gcd(\Delta_{L_1}, \Delta_{L_2}) = 1$ *and* $[M : \mathbb{Q}] = [L_1 : \mathbb{Q}][L_2 : \mathbb{Q}]$, *then* $|\Delta_M| = |\Delta_{L_1}|^{[L_2:\mathbb{Q}]} \cdot |\Delta_{L_2}|^{[L_1:\mathbb{Q}]}$.

Proof. — (1) The $\mathbb{Q}$-algebra $L_1 \otimes L_2$ (or simply $L_1 \otimes L_2$) has a decomposition

$$L_1 \otimes L_2 \cong \bigoplus_{i=1}^{r} M_i,$$

such that M is one of the direct factors. Let $p : L_1 \otimes L_2 \to M$ denote the canonical projection. Then $\mathcal{O}_{L_1 \otimes L_2} := \bigoplus_i \mathcal{O}_{M_i}$ is the integral closure of $\mathbb{Z}$ in $L_1 \otimes L_2$, and we have $\mathcal{O}_{L_1} \otimes \mathcal{O}_{L_2} \subseteq \mathcal{O}_{L_1 \otimes L_2}$ and $\mathcal{O}_{L_1}\mathcal{O}_{L_2} = p(\mathcal{O}_{L_1} \otimes \mathcal{O}_{L_2})$. Note that $\mathrm{Tr}_{L_1 \otimes L_2/L_1} = \bigoplus_i \mathrm{Tr}_{M_i/L_1}$, and that $\delta^{-1}_{L_1 \otimes L_2/L_1} := \bigoplus_{i=1}^{r} \delta^{-1}_{M_i/L_1}$ consists of elements $x \in L_1 \otimes L_2$ such that $\mathrm{Tr}_{L_1 \otimes L_2/L_1}(xy) \in \mathcal{O}_{L_1}$ for all $y \in \mathcal{O}_{L_1 \otimes L_2}$. Let $(\beta_j)_{1 \le j \le m}$ be an integral basis for $\mathcal{O}_{L_2}$, and $(\beta_j^\vee)_j$ denote its dual basis with respect to the pairing on L_2 induced by $\mathrm{Tr}_{L_2/\mathbb{Q}}$. Then every $x \in L_1 \otimes L_2$ writes uniquely as $x = \sum_j x_j \otimes \beta_j^\vee$ with $x_j \in L_1$. If $\mathrm{Tr}_{L_1 \otimes L_2/L_1}(xy) \in \mathcal{O}_{L_1}$ for all $y \in \mathcal{O}_{L_1} \otimes \mathcal{O}_{L_2}$, then we have $x_j = \mathrm{Tr}_{M/L_1}(x\beta_j) \in \mathcal{O}_{L_1}$. Since $\delta^{-1}_{L_2} = \sum_j \mathbb{Z}\beta_j^\vee$, we have $x \in \mathcal{O}_{L_1} \otimes \delta^{-1}_{L_2} \subseteq \delta^{-1}_{L_2}\mathcal{O}_{L_1 \otimes L_2}$. Hence, $\delta^{-1}_{L_1 \otimes L_2/L_1} \subseteq \delta^{-1}_{L_2}\mathcal{O}_{L_1 \otimes L_2}$. Applying p, we have $\delta^{-1}_{M/L_1} \subseteq \delta^{-1}_{L_2}\mathcal{O}_M$, or equivalently $\delta_{L_2}\mathcal{O}_M \subseteq \delta_{M/L_1}$.

(2) Taking N_{M/L_1}, we see from (1) that Disc_{M/L_1} divides $\mathrm{N}_{M/L_1}(\delta_{L_2}\mathcal{O}_M)$. Therefore, by Corollary 3.3.6, Δ_M must divide

$$\begin{aligned} \mathrm{N}_{L_1/\mathbb{Q}}(\mathrm{N}_{M/L_1}(\delta_{L_2}\mathcal{O}_M))\Delta_{L_1}^{[M:L_1]} &= \mathrm{N}_{L_2/\mathbb{Q}} \circ \mathrm{N}_{M/L_2}(\delta_{L_2}\mathcal{O}_M)\Delta_{L_1}^{[M:L_1]} \\ &= \Delta_{L_2}^{[M:L_2]}\Delta_{L_1}^{[M:L_1]}. \end{aligned}$$

The statement now follows immediately from $[M : L_1] \le [L_2 : \mathbb{Q}]$ and $[M : L_2] \le [L_1 : \mathbb{Q}]$.

(3) By Corollary 1.4.6, we have $\mathcal{O}_M = \mathcal{O}_{L_1}\mathcal{O}_{L_2}$. It follows easily that $\delta_{M/L_1} = \delta_{L_2}\mathcal{O}_M$, thus $\mathrm{Disc}_{M/L_1} = \Delta_{L_2}\mathcal{O}_M$. Then statement (3) follows immediately from the arguments for (2) above.

□

Corollary 3.3.8. — *Under the situation of the Proposition, a rational prime p is unramified in M if and only if p is unramified in both L_1 and L_2.*

Proof. — If p is unramified in M, then it is clearly unramified in both L_1 and L_2 by the transitivity of ramification index. Conversely, if p is unramified in both L_1 and L_2, then p is coprime to $\Delta_{L_1}\Delta_{L_2}$ by Theorem 3.2.7, so it is coprime to Δ_M by the Proposition. By Theorem 3.2.7 again, p is unramified in M. □

3.4. Decomposition of primes in Galois extensions

Let L/K be a finite Galois extension of number fields with $G = \mathrm{Gal}(L/K)$. Two fractional ideals I_1 and I_2 are called *conjugate under* G, if there exists $\sigma \in G$ such that $\sigma(I_1) = I_2$.

Let $\mathfrak{p}$ be a prime of $\mathcal{O}_K$ with prime decomposition in $\mathcal{O}_L$:

$$\mathfrak{p}\mathcal{O}_L = \prod_{i=1}^{g} \mathfrak{P}_i^{e_i}, \quad \text{with } e_i \geq 1 \text{ and } \mathfrak{P}_i \text{ distinct.}$$

Since $\mathfrak{p}\mathcal{O}_L$ is invariant under G, the group G stabilizes the set $\{\mathfrak{P}_1, \cdots, \mathfrak{P}_g\}$.

Proposition 3.4.1. — *Any two primes $\mathfrak{P}_i$ and $\mathfrak{P}_j$ are conjugate under G, and we have $e := e_1 = \cdots = e_g$, $f := f(\mathfrak{P}_1|\mathfrak{p}) = \cdots = f(\mathfrak{P}_g|\mathfrak{p})$, and $[L : K] = efg$.*

Proof. — Note that for any $\sigma \in G$, we have $\mathfrak{p}\mathcal{O}_L = \sigma(\mathfrak{p})\mathcal{O}_L$, which implies that $\prod_{i=1}^{g} \mathfrak{P}_i^{e_i} = \prod_{i=1}^{g} \sigma(\mathfrak{P}_i)^{e_i}$. Hence, $e_i = e_{\sigma^{-1}(i)}$ by the uniqueness of the prime decomposition. Moreover, if $\sigma(\mathfrak{P}_i) = \mathfrak{P}_j$, then σ induces an isomorphism

$$\sigma : \mathcal{O}_L/\mathfrak{P}_i \xrightarrow{\sim} \mathcal{O}_L/\mathfrak{P}_j,$$

and hence $f(\mathfrak{P}_i/\mathfrak{p}) = f(\mathfrak{P}_j/\mathfrak{p})$. By Proposition 3.2.2, to complete the proof, it suffices to show that any $\mathfrak{P}_i$, there exists $\sigma \in G$ such that $\sigma(\mathfrak{P}_1) = \mathfrak{P}_i$. Suppose in contrary that some $\mathfrak{P}' = \mathfrak{P}_j$ is not conjugate to $\mathfrak{P}_1$, i.e. for any $\sigma \in G$, $\sigma(\mathfrak{P}_1) \neq \mathfrak{P}_1$. By Lemma 3.4.2 below, there exists $x \in \mathfrak{P}'$ such that $x \notin \sigma(\mathfrak{P}_1)$ for all $\sigma \in G$, or equivalently $\sigma(x) \notin \mathfrak{P}_1$ for all $\sigma \in G$. But then $\mathrm{N}_{L/K}(x) = \prod_{\sigma \in G} \sigma(x) \notin \mathfrak{P}_1 \cap \mathcal{O}_K = \mathfrak{p}$, which contradict with the fact that $\mathrm{N}_{L/K}(x) \in \mathfrak{P}' \cap \mathcal{O}_K = \mathfrak{p}$.

□

Lemma 3.4.2. — *Let R be a commutative ring, $\mathfrak{p}_1, \cdots, \mathfrak{p}_r$ be prime ideals of R, and $\mathfrak{b} \subseteq R$ be an ideal such that $\mathfrak{b} \nsubseteq \mathfrak{p}_i$ for any $1 \leq i \leq n$. Then there exists $x \in \mathfrak{b}$ such that $x \notin \mathfrak{p}_i$ for any i.*

Proof. — We may assume that $\mathfrak{p}_i \nsubseteq \mathfrak{p}_j$ for any $i \neq j$. Take $x_{i,j} \in \mathfrak{p}_j \backslash \mathfrak{p}_i$ and $a_i \in \mathfrak{b} \backslash \mathfrak{p}_i$ for $1 \leq i \leq n$ since $\mathfrak{b} \nsubseteq \mathfrak{p}_i$. Then $b_i = a_i \prod_{j \neq i} x_{i,j}$ belongs to $\mathfrak{b} \cap (\bigcap_{j \neq i} \mathfrak{p}_j)$ but not $\mathfrak{p}_i$. Put $x = \sum_{i=1}^{r} b_i$. Then $x \in \mathfrak{b}$ and $x \equiv b_i \mod \mathfrak{p}_i$ for all i. □

Definition 3.4.3. — For a prime ideal $\mathfrak{P}$ of $\mathcal{O}_L$ with $\mathfrak{p} = \mathfrak{P} \cap \mathcal{O}_K$, we put

$$D(\mathfrak{P}|\mathfrak{p}) = \{\sigma \in G | \sigma(\mathfrak{P}) = \mathfrak{P}\},$$

and call it the *decomposition group* at $\mathfrak{P}$ relative to $\mathfrak{p}$. Any $\sigma \in D(\mathfrak{P}|\mathfrak{p})$ induces an automorphism

$$\sigma : k(\mathfrak{P}) = \mathcal{O}_L/\mathfrak{P} \xrightarrow{\sim} \mathcal{O}_L/\sigma(\mathfrak{P}) = k(\mathfrak{P}).$$

which fixes the subfield $\mathcal{O}_K$. We get thus a homomorphism

$$\varphi_{\mathfrak{P}} : D(\mathfrak{P}|\mathfrak{p}) \to \mathrm{Gal}(k(\mathfrak{P})/k(\mathfrak{p})).$$

We define

$$I(\mathfrak{P}|\mathfrak{p}) := \mathrm{Ker}(\varphi_{\mathfrak{P}}) = \{\sigma \in D(\mathfrak{P}|\mathfrak{p}) | \sigma(x) \equiv x \mod \mathfrak{P},\ \forall x \in \mathcal{O}_L\},$$

and call it the *inertia subgroup* of $\mathfrak{P}$ relative to $\mathfrak{p}$.

Proposition 3.4.4. — 1. *The extension $k(\mathfrak{P})/k(\mathfrak{p})$ is Galois, and the map $\varphi_{\mathfrak{P}}$ is surjective, i.e. we have an exact sequence*

$$1 \to I(\mathfrak{P}|\mathfrak{p}) \to D(\mathfrak{P}|\mathfrak{p}) \to \mathrm{Gal}(k(\mathfrak{P})/k(\mathfrak{p})) \to 1.$$

Moreover, one has $e(\mathfrak{P}|\mathfrak{p}) = \#I(\mathfrak{P}|\mathfrak{p})$, and $f(\mathfrak{P}|\mathfrak{p})e(\mathfrak{P}|\mathfrak{p}) = \#D(\mathfrak{P}|\mathfrak{p})$.

2. *For any $\tau \in G$, we have $D(\tau(\mathfrak{P})|\mathfrak{p}) = \tau D(\mathfrak{P}|\mathfrak{p})\tau^{-1}$ and $I(\tau(\mathfrak{P})|\mathfrak{p}) = \tau I(\mathfrak{P}|\mathfrak{p})\tau^{-1}$.*

Proof. — We denote simply $D_{\mathfrak{P}} = D(\mathfrak{P}|\mathfrak{p})$ and $I_{\mathfrak{P}} = I(\mathfrak{P}|\mathfrak{p})$. Statement (2) is immediate by the definition of $D(\mathfrak{P}|\mathfrak{p})$ and $I(\mathfrak{P}|\mathfrak{p})$. It remains to prove (1). By Proposition 3.4.1, G acts transitively on the set $\{\mathfrak{P} = \mathfrak{P}_1, \cdots, \mathfrak{P}_g\}$ of primes above $\mathfrak{p}$, and $D_{\mathfrak{P}}$ is the stabilizer of G on $\mathfrak{P}$. We see that $g = [G : D_{\mathfrak{P}}]$, but $\#G = efg$, with $e = e(\mathfrak{P}|\mathfrak{p})$ and $f = f(\mathfrak{P}|\mathfrak{p})$. It follows that $\#D_{\mathfrak{P}} = ef$.

Let $M = L^{D_{\mathfrak{P}}}$ and $N = L^{I_{\mathfrak{P}}}$, and $\mathfrak{P}_D = \mathfrak{P} \cap \mathcal{O}_M$ and $\mathfrak{P}_I = \mathfrak{P} \cap \mathcal{O}_N$. Then for any $\sigma \in D_{\mathfrak{P}} = \mathrm{Gal}(L/M)$, we have

$$\sigma(\mathfrak{P}_D) = \sigma(\mathfrak{P}) \cap \mathcal{O}_M = \mathfrak{P} \cap \mathcal{O}_M = \mathfrak{P}_D.$$

It follows from Proposition 3.4.1 that $\mathfrak{P}$ is the only prime above $\mathfrak{P}_D$; so is the same for $\mathfrak{P}_I$. Suppose that $\mathfrak{P}_D\mathcal{O}_L = \mathfrak{P}^{e'}$. Then by the transitivity of ramification and residue indexes, we have $e'|e$ and $f' := f(\mathfrak{P}|\mathfrak{P}_D)|f$. However, by Proposition 3.2.2, it follows that

$$ef = \#D_{\mathfrak{P}} = [L : M] = e'f'.$$

Hence, we get $e = e'$ and $f = f'$. Similarly, assume that $\mathfrak{P}_I\mathcal{O}_L = \mathfrak{P}^{e''}$. Then one has $e''|e' = e$. Let $\bar{\alpha} \in k(\mathfrak{P})$ be an arbitrary element. Take $\alpha \in \mathcal{O}_L$ a lift of $\bar{\alpha}$, and let $f(X) \in \mathcal{O}_N[X]$ be the minimal polynomial of α over N. Then we have

$$f(X) = \prod_{\sigma \in I_{\mathfrak{P}}} (X - \sigma(\alpha)).$$

By definition of $I_{\mathfrak{P}}$, we have $\sigma(\alpha) \equiv \alpha \mod \mathfrak{P}$, that is $\sigma(\bar{\alpha}) = \bar{\alpha}$. If $\bar{f}(X) \in k(\mathfrak{P}_I)[X]$ denotes the reduction of $f(X)$, then $\bar{f}(X) = (X - \bar{\alpha})^{\#I_{\mathfrak{P}}}$. Since any Galois conjugate of $\bar{\alpha}$ must be a root of $\bar{f}(X)$, it follows that $\bar{\alpha} \in k(\mathfrak{P}_I)$. Hence, we see that $f(\mathfrak{P}/\mathfrak{P}_I) = 1$. By Proposition 3.2.2, we have $\#I_{\mathfrak{P}} = e'' \leq e$. By definition, $\varphi_{\mathfrak{P}}$ induces an injection $D_{\mathfrak{P}}/I_{\mathfrak{P}} \hookrightarrow \mathrm{Gal}(k(\mathfrak{P})/k(\mathfrak{p}))$. But note that

$$\#(D_{\mathfrak{P}}/I_{\mathfrak{P}}) = \#D_{\mathfrak{P}}/\#I_{\mathfrak{P}} = ef/e'' \geq f$$

and $\#\,\mathrm{Gal}(k(\mathfrak{P})/k(\mathfrak{p})) = f$. Hence, $\varphi_{\mathfrak{P}}$ must be surjective and $\#I_{\mathfrak{P}} = e$.

□

The following Proposition is very useful when considering the problem of sub-extensions.

Proposition 3.4.5. — *Let K'/K be a sub-extension of L/K, and $H \subseteq G$ denote the subgroup such that $K' = L^H$. We fix a prime $\mathfrak{P}$ of $\mathcal{O}_L$, let $\mathfrak{p} = \mathcal{O}_K \cap \mathfrak{P}$ and $\mathfrak{p}' = \mathfrak{P} \cap \mathcal{O}_{K'}$. Then $\mathfrak{P}$ is the only prime above $\mathfrak{p}'$ if and only if $H \subseteq D(\mathfrak{P}|\mathfrak{p})$ or equivalently $L^{D(\mathfrak{P}|\mathfrak{p}')} \subseteq K'$, and $e(\mathfrak{p}'|\mathfrak{p}) = 1$ if and only if $I(\mathfrak{P}|\mathfrak{p}) \subseteq H$ or equivalently $K' \subseteq L^{I(\mathfrak{P}|\mathfrak{p})}$.*

Proof. — By definition, it is clear that $D(\mathfrak{P}|\mathfrak{p}') = D(\mathfrak{P}|\mathfrak{p}) \cap H$ and $I(\mathfrak{P}|\mathfrak{p}') = I(\mathfrak{P}|\mathfrak{p}) \cap H$. Then by Proposition 3.4.4, $\mathfrak{P}$ is the only prime of $\mathcal{O}_L$ above $\mathfrak{p}'$ if and only if $H = D(\mathfrak{P}|\mathfrak{p}')$, or equivalently $H \subseteq D(\mathfrak{P}|\mathfrak{p})$; and $e(\mathfrak{p}'|\mathfrak{p}) = 1$ if and only if $e(\mathfrak{P}|\mathfrak{p}') = e(\mathfrak{P}|\mathfrak{p})$ by the transitivity of ramification index. By Proposition 3.4.4, this is equivalent to $I(\mathfrak{P}|\mathfrak{p}) \cap H = I(\mathfrak{P}|\mathfrak{p})$, that is $I(\mathfrak{P}|\mathfrak{p}) \subseteq H$. □

The following is a generalization of Corollary 3.3.8.

Corollary 3.4.6. — *Let L_1 and L_2 be finite (not necessarily Galois) extensions of a number field K, and L_1L_2 be their composite inside an algebraic closure of K. Then a prime $\mathfrak{p}$ of $\mathcal{O}_K$ is unramified in L_1L_2 if and only if it is unramified in both L_1 and L_2.*

Proof. — Choose a finite Galois extension M/K containing both L_1L_2. Let H_1 and H_2 denote the subgroups of $\mathrm{Gal}(M/K)$ that fix L_1 and L_2 respectively. Then L_1L_2 is the fixed field of $H_1 \cap H_2$. By the Proposition, $\mathfrak{p}$ is unramified in L_1L_2 if and only if $I(\mathfrak{P}|\mathfrak{p}) \subseteq H_1 \cap H_2$ for every prime $\mathfrak{P}$ of M above $\mathfrak{p}$, or equivalently $I(\mathfrak{P}|\mathfrak{p}) \subset H_1$ and $I(\mathfrak{P}|\mathfrak{p}) \subset H_2$. By the Proposition again, the latter condition is exactly equivalent to that $\mathfrak{p}$ is unramified in L_1 and L_2. □

Now assume that the prime $\mathfrak{p}$ is unramified in $\mathcal{O}_L$, and $\mathfrak{P}$ be a prime of $\mathcal{O}_L$ above $\mathfrak{p}$. Then we have $I(\mathfrak{P}|\mathfrak{p}) = 1$ and $D(\mathfrak{P}|\mathfrak{p}) \xrightarrow{\sim} \mathrm{Gal}(k(\mathfrak{P})/k(\mathfrak{p}))$. Let $q = \mathrm{N}(\mathfrak{p})$, and $q^f = \mathrm{N}(\mathfrak{P})$. Then it is well known that $\mathrm{Gal}(k(\mathfrak{P})/k(\mathfrak{p})) \cong \mathbb{Z}/f\mathbb{Z}$ with a canonical generator given by $\sigma_q : x \to x^q$ for any $x \in k(\mathfrak{P})$. We denote by

$$\sigma_{\mathfrak{P}} = \left(\frac{L/K}{\mathfrak{P}}\right) \in D(\mathfrak{P}|\mathfrak{p})$$

the element corresponding to σ_q, that is the unique element of $D(\mathfrak{P}|\mathfrak{p})$ with

$$\sigma_{\mathfrak{P}}(x) \equiv x^q \mod \mathfrak{P}, \quad \forall x \in \mathcal{O}_L.$$

We call $\sigma_{\mathfrak{P}}$ the *Frobenius element* of $\mathfrak{P}$ over $\mathfrak{p}$. It is clear that $\sigma_{\mathfrak{P}}$ is a generator of $D(\mathfrak{P}|\mathfrak{p}) \cong \mathbb{Z}/f\mathbb{Z}$, and the Frobenius elements verify the following properties:

1. For any $\tau \in \mathrm{Gal}(L/K)$, we have
$$\left(\frac{L/K}{\tau(\mathfrak{P})}\right) = \tau\left(\frac{L/K}{\mathfrak{P}}\right)\tau^{-1}$$
2. If M/K is a Galois sub extension of L/K and $\mathfrak{P}_M = \mathfrak{P} \cap \mathcal{O}_M$, then we have
$$\left(\frac{L/K}{\mathfrak{P}}\right)\Big|_M = \left(\frac{M/K}{\mathfrak{P}_M}\right), \qquad \left(\frac{L/M}{\mathfrak{P}}\right) = \left(\frac{L/K}{\mathfrak{P}}\right)^{f(\mathfrak{P}_M/\mathfrak{p})}$$

It follows from (1) and Proposition 3.4.1 that if $\mathfrak{P}'$ is another prime of $\mathcal{O}_L$ above $\mathfrak{p}$, then the Frobenius element of $\mathfrak{P}'$ is conjugate to that of $\mathfrak{P}$. Therefore, if $\mathrm{Gal}(L/K)$ is abelian, then these two Frobenius elements coincide; in that case, we denote it common by $\sigma_{\mathfrak{p}} = \left(\frac{L/K}{\mathfrak{p}}\right)$.

Example 3.4.7. — We put $L = \mathbb{Q}(\sqrt[3]{2}, \sqrt{-3})$. Then $G = \mathrm{Gal}(L/\mathbb{Q}) \cong \langle \sigma, \tau \rangle / (\sigma^3 = \tau^2 = 1, \sigma\tau = \tau\sigma^2)$, where $\sigma(\sqrt[3]{2}) = \omega\sqrt[3]{2}$ with $\omega = \frac{-1+\sqrt{-3}}{2}$ and $\sigma(\sqrt{-3}) = \sqrt{-3}$, and $\tau(\sqrt[3]{2}) = \sqrt[3]{2}$ and $\tau(\sqrt{-3}) = -\sqrt{-3}$. A rational prime p ramifies in L if and only if $p = 2, 3$.

(1) The prime 2 is inert in $\mathbb{Q}(\sqrt{-3})$ and ramifies in $\mathbb{Q}(\sqrt[3]{2})$. So there exists a unique prime $\mathfrak{p}_2$ in $\mathcal{O}_L$ of degree above 2 such that $2\mathcal{O}_L = \mathfrak{p}_2^3$. We have $D(\mathfrak{p}_2|2) = G$, and $I(\mathfrak{p}_2|2) = \mathrm{Gal}(L/\mathbb{Q}(\sqrt{-3})) = \langle \sigma \rangle$.

(2) The prime 3 is ramified in both $\mathbb{Q}(\sqrt{-3})$ and $\mathbb{Q}(\sqrt[3]{2})$, so its ramification degree in $L/\mathbb{Q}$ is divisible by 6. Thus we see that $3\mathcal{O}_L = \mathfrak{p}_3^6$ for some prime $\mathfrak{p}_3$ of residue degree 1 above 3. We have $D(\mathfrak{p}_3|3) = I(\mathfrak{p}_3|3) = G$.

(3) It is easy to see that $p = 5$ is inert in $K = \mathbb{Q}(\sqrt{-3})$ so that $\mathcal{O}_K/(5) \cong \mathbb{F}_{25}$. Note that $x^3 - 2$ has 3 distinct solutions in $\mathbb{F}_{25}$, and exactly one of them is in $\mathbb{F}_5$, namely $x = 3 \in \mathbb{F}_5$. Therefore, there are 3 distinct primes of $\mathcal{O}_L$ above 5: $\mathfrak{p}_5^{(1)} = (5, \sqrt[3]{2} - 3)$, $\mathfrak{p}_5^{(2)} = (5, \sqrt[3]{2} - 3\omega)$ and $\mathfrak{p}_5^{(3)} = (5, \sqrt[3]{2} - 3\omega^2)$ with $\omega = \frac{-1+\sqrt{-3}}{2}$, and each of them has residue degree 2 over 5. The decomposition group of $\mathfrak{p}_5^{(1)}$, $\mathfrak{p}_5^{(2)}$ and $\mathfrak{p}_5^{(3)}$ are respectively $\mathrm{Gal}(L/\mathbb{Q}(\sqrt[3]{2})) = \langle \tau \rangle$, $\mathrm{Gal}(L/\mathbb{Q}(\sqrt[3]{2}\omega^2)) = \langle \sigma\tau \rangle$ and $\mathrm{Gal}(L/\mathbb{Q}(\sqrt[3]{2}\omega)) = \langle \sigma^2\tau \rangle$. The Frobenius elements of $\mathfrak{p}_5^{(1)}$, $\mathfrak{p}_5^{(2)}$ and $\mathfrak{p}_5^{(3)}$ are respectively $\tau, \sigma\tau, \sigma^2\tau$.

(4) Consider the case $p = 7$. Then 7 is split in $\mathbb{Q}(\sqrt{-3})$ and inert in $\mathbb{Q}(\sqrt[3]{2})$. Thus 7 splits in $\mathcal{O}_K$ into two primes of degree 2, namely $\mathfrak{p}_7^{(1)} = (7, \sqrt{-3} + 2)$ and $\mathfrak{p}_7^{(2)} = (7, \sqrt{-3} - 2)$. The decomposition groups of both $\mathfrak{p}_7^{(1)}$ and $\mathfrak{p}_7^{(2)}$ are both $\mathrm{Gal}(K/\mathbb{Q}(\sqrt{-3})) = \langle \sigma \rangle$. The Frobenius element $\sigma_{\mathfrak{p}_7^{(i)}}$ is the unique element of $\mathrm{Gal}(K/\mathbb{Q}(\sqrt{-3}))$ such that

$$\sigma_{\mathfrak{p}_7^{(i)}}(x) \equiv x^7 \mod \mathfrak{p}_7^{(i)}, \quad \forall x \in \mathcal{O}_L.$$

Since $\omega \equiv 2 \mod \mathfrak{p}_7^{(1)}$ and $\omega \equiv 4 \mod \mathfrak{p}_7^{(2)}$, we have $(\sqrt[3]{2})^7 \equiv \sqrt[3]{2}\omega^2 \mod \mathfrak{p}_7^{(1)}$ and $(\sqrt[3]{2})^7 \equiv \sqrt[3]{2}\omega \mod \mathfrak{p}_7^{(2)}$. Thus it follows that $\sigma_{\mathfrak{p}_7^{(1)}} = \sigma^2$ and $\sigma_{\mathfrak{p}_7^{(2)}} = \sigma$.

Example 3.4.8. — Let $m \neq -1$ be a non-square integer, and $\sqrt[4]{m} > 0$ be its positive real 4-th root. Put $K = \mathbb{Q}(\sqrt[4]{m})$ and $L = \mathbb{Q}(\sqrt[4]{m}, i)$. Then $L/\mathbb{Q}$ is the Galois closure of $K/\mathbb{Q}$ with Galois group $G = \mathrm{Gal}(L/\mathbb{Q}) \cong \langle \sigma, \tau \rangle / (\sigma^4 = \tau^2 = 1, \tau\sigma\tau = \sigma^{-1})$, where σ and τ are defined by

$$\begin{cases} \sigma(\sqrt[4]{m}) = \sqrt[4]{m}\, i, & \sigma(i) = i; \\ \tau(\sqrt[4]{m}) = \sqrt[4]{m}, & \tau(i) = -i. \end{cases}$$

A rational prime p is unramified in $L/\mathbb{Q}$ if and only if $p \nmid 2m$. Let p be such a prime and $\mathfrak{P}$ is a prime of $\mathcal{O}_L$ above p. Assume that $\sigma_{\mathfrak{P}} = \left(\frac{L/\mathbb{Q}}{\mathfrak{P}}\right) = \tau$. Then we have

$$p\mathcal{O}_L = \mathfrak{P}_1\mathfrak{P}_2\mathfrak{P}_3\mathfrak{P}_4,$$

where $\mathfrak{P}_i = \sigma^{i-1}(\mathfrak{P})$ for $i = 1, \cdots, 4$, and each $\mathfrak{P}_i$ has degree 2 over p. Moreover, the decomposition groups of $\mathfrak{P}_i$ is

$$D(\mathfrak{P}_i|p) = \begin{cases} \langle \tau \rangle & \text{for } i = 1, 3; \\ \langle \sigma^2 \tau \rangle & \text{for } i = 2, 4. \end{cases}$$

Put $\mathfrak{p}_i = \mathfrak{P}_i \cap \mathcal{O}_K$. Since $K = L^H$ with $H = \langle \tau \rangle$, both $\mathfrak{p}_1$ and $\mathfrak{p}_3$ have degree 1 over p, and $\mathfrak{P}_1$ (resp. $\mathfrak{P}_3$) are the unique prime of $\mathcal{O}_L$ above $\mathfrak{p}_1$ (resp. $\mathfrak{p}_3$). In particular, $\mathfrak{p}_1 \neq \mathfrak{p}_3$. As $H \cap D(\mathfrak{P}_i|p) = \{1\}$ for $i = 2, 4$, we see that $\mathfrak{p}_2$ and $\mathfrak{p}_4$ are both degree 2. For degree reasons, we have necessarily $\mathfrak{p}_2 = \mathfrak{p}_4$. Actually, the equality $\mathfrak{p}_2 = \mathfrak{p}_4$ can also be proved using the fact that $\mathfrak{P}_2$ and $\mathfrak{P}_4$ are conjugate under the action of H.

3.5. Prime decompositions in cyclotomic fields

Let $K = \mathbb{Q}(\zeta_N)$ for some integer $N \geq 3$.

Proposition 3.5.1. — *A rational prime l is ramified in $\mathbb{Q}(\zeta_N)$ if and only if $l|N$.*

Proof. — When $N = p^n$, then the statement follows from Theorem 3.2.7 and Proposition 1.4.8. Since $\mathbb{Q}(\zeta_N) = \prod_{p|N} \mathbb{Q}(\zeta_{p^{v_p(N)}})$, the general cases follows from Corollary 3.3.8. □

Let l be a prime with $\gcd(l, N) = 1$, and $\mathfrak{l}_1, \cdots, \mathfrak{l}_g$ be the primes of $\mathbb{Q}(\zeta_N)$ above l. Recall that $\mathrm{Gal}(\mathbb{Q}(\zeta_N)/\mathbb{Q}) \cong (\mathbb{Z}/N\mathbb{Z})^\times$. We denote by $\sigma_l \in \mathrm{Gal}(\mathbb{Q}(\zeta_N)/\mathbb{Q})$ the Frobenius element at l. Then σ_l is characterized by the following property:

$$\sigma_l(x) \equiv x^l \mod \mathfrak{l}_i, \quad \forall 1 \leq i \leq g.$$

Lemma 3.5.2. — *If $i \neq j$ in $(\mathbb{Z}/N\mathbb{Z})$, then ζ_N^i is not congruent to $\zeta_N^j \mod \mathfrak{l}_k$ for all k.*

Proof. — Consider the polynomial $f(X) = X^N - 1$. For any N-th root of unit ζ, we have $f'(\zeta) = N\zeta^{N-1} = N\zeta^{-1}$. For any prime $\mathfrak{l}_k$ above l of $\mathbb{Q}(\zeta_N)$, $f'(\zeta)$ is non-vanishing modulo $\mathfrak{l}_k$. It follows that $f(X)$ has no multiple roots in an algebraic closure of $\mathbb{F}_l$. Therefore, if $\zeta_N^i \neq \zeta_N^j$ in $\mathbb{Q}(\zeta_N)$, then ζ_N^i is not congruent to ζ_N^j modulo $\mathfrak{l}_k$. □

Proposition 3.5.3. — *The Frobenius element $\sigma_l \in \mathrm{Gal}(\mathbb{Q}(\zeta_N)/\mathbb{Q})$ is given by $\sigma_l(\zeta_N) = \zeta_N^l$. The decomposition group of each $\mathfrak{l}_i$ is $D_l = \langle l \rangle \subseteq (\mathbb{Z}/N\mathbb{Z})^\times$, and the residue degree $f(\mathfrak{l}_i|l)$ is the order of l in $(\mathbb{Z}/N\mathbb{Z})^\times$, that is, the minimal integer $f > 0$ such that $N|(l^f - 1)$.*

Example 3.5.4. — Consider the number field $\mathbb{Q}(\zeta_{31})$ and $l = 2$. Since 2 has order 5 in $(\mathbb{Z}/31\mathbb{Z})^\times$, it splits into 6 primes in $\mathcal{O}_K$ and each of them has residue degree 5. Let $H = \langle 2 \rangle \subseteq (\mathbb{Z}/31\mathbb{Z})^\times$, and $K = \mathbb{Q}(\zeta_{31})^H$. Then K is the decomposition field of each prime above 2. Thus 2 splits into 6 primes, namely $\mathfrak{p}_1, \cdots, \mathfrak{p}_6$, in $\mathcal{O}_K$, and each $\mathfrak{p}_i$ has degree 1. We claim that there exists no $\alpha \in \mathcal{O}_K$ such that $\mathcal{O}_K = \mathbb{Z}[\alpha]$. Otherwise, let $f(X) \in \mathbb{Z}[X]$ denote the minimal polynomial of α. Then by Kummer's Theorem 3.2.3, $\bar{f}(X)$ has 6 distinct roots in $\mathbb{F}_2$. But this is impossible since $\#\mathbb{F}_2 = 2$.

Proposition 3.5.3 should be viewed as the reciprocity law for cyclotomic fields. One can use it to give a proof of the quadratic reciprocity law.

Lemma 3.5.5. — *Let p be an odd prime. Then $\mathbb{Q}(\zeta_p)$ contains a unique quadratic field K, which is*

$$K = \begin{cases} \mathbb{Q}(\sqrt{p}) & \text{if } p \equiv 1 \mod 4, \\ \mathbb{Q}(\sqrt{-p}) & \text{if } p \equiv 3 \mod 4. \end{cases}$$

Proof. — The Galois group $\mathrm{Gal}(\mathbb{Q}(\zeta_p)/\mathbb{Q}) \cong (\mathbb{Z}/p\mathbb{Z})^\times$ is cyclic of order $p-1$. It contains a unique subgroup H of index 2. Thus $\mathbb{Q}(\zeta_p)$ contains a unique quadratic field K. Explicitly, if $a \in (\mathbb{Z}/p\mathbb{Z})^\times$ denotes a generator, then $H = \langle a^2 \rangle$, that is H consists of the quadratic residues of in $\mathbb{F}_p^\times$. Since p is the only prime ramified in $\mathbb{Q}(\zeta_p)$, so every prime different from p must be unramified in K. Therefore, by Theorem 3.2.5, we see that $K = \mathbb{Q}(\sqrt{p})$ if $p \equiv 1 \mod 4$ and $K = \sqrt{-p}$ if $p \equiv 3 \mod 4$.

□

Theorem 3.5.6 **(Quadratic Reciprocity Law)**. — *Let p, q be odd primes. Then we have*

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{(p-1)(q-1)}{2}},$$

Proof. — The statement is equivalent to saying that $\left(\frac{p^*}{q}\right) = \left(\frac{q}{p}\right)$, where $p^* = p$ if $p \equiv 1 \mod 4$, and $p^* = -p$ if $p \equiv 3 \mod 4$. The statement is deduced from the following equivalences:

$$\begin{aligned} \left(\frac{p^*}{q}\right) = 1 &\Leftrightarrow x^2 - p^* \equiv 0 \mod q \text{ has solutions.} \\ &\Leftrightarrow q \text{ splits in } \mathbb{Q}(\sqrt{p^*}) \text{ by Theorem 3.2.5.} \\ &\Leftrightarrow \left(\frac{\mathbb{Q}(\sqrt{p^*})/\mathbb{Q}}{q}\right) = \left(\frac{\mathbb{Q}(\zeta_p)/\mathbb{Q}}{q}\right)\Bigg|_{\mathbb{Q}(\sqrt{p^*})} = 1 \\ &\Leftrightarrow \sigma_q = \left(\frac{\mathbb{Q}(\zeta_p)/\mathbb{Q}}{q}\right) \in H \\ &\Leftrightarrow q \text{ is a quadratic residue in } \mathbb{F}_p. \end{aligned}$$

Here, $H \subseteq \mathrm{Gal}(\mathbb{Q}(\zeta_p)/\mathbb{Q}) \cong (\mathbb{Z}/p\mathbb{Z})^\times$ is unique subgroup of index 2.

□

CHAPTER 4

FINITENESS THEOREMS

4.1. Finiteness of class numbers

A subset $\Lambda \subseteq \mathbb{R}^n$ is called a lattice, if Λ is a free abelian subgroup of rank n containing a $\mathbb{R}$-basis of $\mathbb{R}^n$. For a lattice $\lambda \subseteq \mathbb{R}^n$, a $\mathbb{Z}$-basis of Λ is necessarily a $\mathbb{R}$-basis of $\mathbb{R}^n$, and Λ is discrete for the natural topology on $\mathbb{R}^n$. We define $\mathrm{Vol}(\mathbb{R}^n/\Lambda)$ as the volume of the parallelogram spanned by a basis of Λ.

We have the following elementary

Lemma 4.1.1 **(Minkowski)**. — *Let $\Lambda \subset \mathbb{R}^n$ be a lattice, and $X \subseteq \mathbb{R}^n$ be a centrally symmetric convex connected region of finite measure $\mu(X)$. Assume that $\mu(X) > 2^n \mathrm{Vol}(\mathbb{R}^n/\Lambda)$. Then there exists $\alpha \neq 0$ in $\Lambda \cap X$.*

Proof. — Let P_e be the parallelogram spanned by a $\mathbb{Z}$-basis of the lattice 2Λ so that

$$\mu(P_e) = \mathrm{Vol}(\mathbb{R}^n/2\Lambda) = 2^n \mathrm{Vol}(\mathbb{R}^n/\Lambda).$$

Note that $\mathbb{R}^n = \sqcup_{\lambda \in 2\Lambda}(\lambda + P_e)$, hence

$$\mu(X) = \sum_{\lambda \in 2\Lambda} \mu((\lambda + P_e) \cap X)$$

by the additivity of Lebesgue measure. As μ is invariant under translation, we have

$$\mu((\lambda + P_e) \cap X) = \mu((-\lambda + X) \cap P_e).$$

As $\mu(X) > \mu(P_e)$, there exists two $\lambda_1, \lambda_2 \in 2\Lambda$ such that $(-\lambda_1 + X) \cap (-\lambda_2 + X) \neq \emptyset$ (otherwise one would have $\mu(X) < \mu(P_e)$). Let $x, y \in X$ such that $-\lambda_1 + x = -\lambda_2 + y$, it follows that $x - y \in 2\Lambda$. Since X is symmetric and convex, one gets $\alpha = \frac{x-y}{2} \in \Lambda \cap X$. □

Let $K/\mathbb{Q}$ be a number field of degree $n = [K : \mathbb{Q}]$. Denote by $\sigma_1, \cdots, \sigma_{r_1} : K \to \mathbb{R}$ the real embeddings of K, and $\sigma_{r_1+1}, \sigma_{r_1+2}, \cdots, \sigma_{r_1+2r_2-1}, \sigma_{r_1+2r_2} : K \hookrightarrow \mathbb{C}$ be the complex embeddings such that $\sigma_{r_1+2i} = \bar{\sigma}_{r_1+2i-1}$ and $n = r_1 + 2r_2$. Consider another embedding

$$\lambda : K \longrightarrow \mathbb{R}^{r_1} \times \mathbb{C}^{r_2} \cong \mathbb{R}^n$$

sending x to $((\sigma_i(x))_{1\leq i\leq r_1}, (\sigma_{r_1+2j}(x))_{1\leq j\leq r_2})$. Here, the identification $\mathbb{R}^{r_1}\times\mathbb{C}^{r_2}\cong\mathbb{R}^n$ is given by

$$((y_i)_{1\leq i\leq r_1}, (z_j)_{1\leq j\leq r_2})\mapsto (y_1,\cdots,y_{r_1},\Re(z_1),\Im(z_1),\cdots,\Re(z_{r_2}),\Im(z_{r_2})).$$

Let I be a fractional ideal of $\mathcal{O}_K$. Then I is a free abelian group of rank n. Denote by $\alpha_1,\cdots,\alpha_n$ a $\mathbb{Z}$-basis of I, and we define

$$\mathrm{Disc}(I)=\mathrm{Disc}(\alpha_1,\cdots,\alpha_n)=\det(\sigma_i(\alpha_j))^2.$$

We see easily that the definition is independent of the choice of the basis $(\alpha_i)_{1\leq i\leq n}$, and $\mathrm{Disc}(I)=\Delta_K\mathrm{N}(I)^2$, where Δ_K denotes the discriminant of K.

Lemma 4.1.2. — *For any fractional ideal I, $\lambda(I)$ is a lattice of $\mathbb{R}^n$ with*

$$\mathrm{Vol}(\mathbb{R}^n/\lambda(I))=\frac{1}{2^{r_2}}\sqrt{|\Delta_K|}\mathrm{N}(I)$$

Proof. — It is clear that $\lambda(I)$ is a $\mathbb{Z}$-lattice of rank n. To compute $\mathrm{Vol}(\mathbb{R}^n/\lambda(I))$, we choose a basis $(\alpha_1,\cdots,\alpha_n)$ of I over $\mathbb{Z}$. Denote by $\lambda(\alpha_i)\in\mathbb{R}^n$ the column vector given by α_i. Then we have

$$\mathrm{Vol}(\mathbb{R}^n/\lambda(I))=|\det(\lambda(\alpha_1),\lambda(\alpha_2),\cdots,\lambda(\alpha_n))|.$$

Then we have

$$(\sigma_i(\alpha_j))_{1\leq i,j\leq n}=\begin{pmatrix} I_{r_1} & 0 & 0 & \cdots & 0\\ 0 & \begin{smallmatrix}1&-i\\1&i\end{smallmatrix} & 0 & \cdots & 0\\ 0 & 0 & \begin{smallmatrix}1&-i\\1&i\end{smallmatrix} & \cdots & 0\\ 0 & 0 & 0 & \cdots & \begin{smallmatrix}1&-i\\1&i\end{smallmatrix}\end{pmatrix}(\lambda(\alpha_1),\lambda(\alpha_2),\cdots,\lambda(\alpha_n)).$$

It follows that

$$\det(\sigma_i(\alpha_j))=(2i)^{r_2}\det(\lambda(\alpha_1),\lambda(\alpha_2),\cdots,\lambda(\alpha_n)).$$

But $\mathrm{Disc}(I)=\det(\sigma_i(\alpha_j))^2$, one obtains that

$$\mathrm{Vol}(\mathbb{R}^n/\lambda(I))=2^{-r_2}\sqrt{|\mathrm{Disc}(I)|}=2^{-r_2}\sqrt{|\Delta_K|}\mathrm{N}(I).$$

□

Theorem 4.1.3. — *Let K be a number field of degree n, Δ_K denote the absolute discriminant of K, and r_1,r_2 be the integers defined above. Let I be a fractional ideal of $\mathcal{O}_K$.*

(1) *Given arbitrary constants $c_1,\cdots,c_{r_1+r_2}>0$ with*

$$\prod_{i=1}^{r_1+r_2}c_i>(\frac{2}{\pi})^{r_2}|\Delta_K|^{1/2}\mathrm{N}(I),$$

there exists a nonzero $\alpha\in I$ with $|\sigma_i(\alpha)|<c_i$ for $1\leq i\leq r_1$, and $|\sigma_{r_1+2j}(\alpha)|^2<c_{r_1+j}$ for $1\leq j\leq r_2$.

(2) *There exists a non-zero $x\in I$ such that*

$$\mathrm{N}_{K/\mathbb{Q}}(x)\leq(\frac{4}{\pi})^{r_2}\frac{n!}{n^n}|\Delta_K|^{1/2}\mathrm{N}(I).$$

We need the following

Lemma 4.1.4. — *For $t \in \mathbb{R}_{\geq 0}$, let B_t denote the subset of all $(y_1, \cdots, y_{r_1}, z_1, \cdots, z_{r_2}) \in \mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$ such that*

$$\sum_{i=1}^{r_1} |y_i| + 2\sum_{j=1}^{r_2} |z_j| \leq t.$$

Then the Lesbegue measure of B_t is

$$\mu(B_t) = 2^{r_1}\big(\frac{\pi}{2}\big)^{r_2}\frac{t^n}{n!}$$

Proof. — Put $V(r_1, r_2, t) = \mu(B_t)$. We will prove the formula by double induction on r_1 and r_2. It is clear that $V(1, 0, t) = 2t$ and $V(0, 1, t) = (\pi/2)t^2$. Now assume the formula for $V(r_1, r_2, t)$ is true, and we deduce from it the formula of $V(r_1+1, r_2, t)$ and $V(t_1, t_2+1, t)$.

Note that

$$V(r_1 + 1, r_2, t) = \int_{-t}^{t} V(t_1, t_2, t - |y|)dy = 2\int_0^t V(r_1, r_2, t - y)dy.$$

Using induction hypothesis, one gets

$$V(r_1 + 1, r_2, t) = 2\int_0^t 2^{r_1}\big(\frac{\pi}{2}\big)^{r_2}\frac{(t-y)^n}{n!}dy = 2^{r_1+1}\big(\frac{\pi}{2}\big)^{r_2}\frac{t^{n+1}}{(n+1)!}.$$

For $V(r_1, r_2 + 1, t)$, we have similarly

$$V(r_1, r_2 + 1, t) = \int_{|z|\leq t/2} V(r_1, r_2, t - |z|)d\mu(z),$$

where $d\mu(z)$ denotes the Lesbegue measure on $\mathbb{C}$. Using the polar coordinates and induction hypothesis, one gets

$$V(r_1, r_2 + 1, t) = \int_{\rho=0}^{t/2}\int_{\theta=0}^{2\pi} 2^{r_1}\big(\frac{\pi}{2}\big)^{r_2}\frac{(t-2\rho)^n}{n!}\rho d\rho d\theta.$$

An easy computation shows that $V(r_1, r_2 + 1, t) = 2^{r_1}\big(\frac{\pi}{2}\big)^{r_2+1}\frac{t^{n+2}}{(n+2)!}$. □

Proof of Theorem 4.1.3. — (1) Consider the region

$$W(c) = \{x = (y, z) \in \mathbb{R}^n \cong \mathbb{R}^{r_1} \times \mathbb{C}^{r_2} | |y_i| < c_i \text{ for } 1 \leq i \leq r_1,\ |z_j|^2 < c_{r_1+j} \text{ for } 1 \leq j \leq r_2\}.$$

It is clear that $W(c)$ is symmetric convex with

$$\mu(W(c)) = 2^{r_1}\pi^{r_2}\prod_{i=1}^{r_1+r_2} c_i > 2^n\frac{1}{2^{r_2}}\sqrt{|\Delta_K|}\mathrm{N}(I) = 2^n\mathrm{Vol}(\mathbb{R}^n/\lambda(I))$$

Statement (1) now follows easily from Minkowski's Lemma 4.1.1.

(2) To prove (2), we consider the region

$$B_t = \{(y, z) \in \mathbb{R}^{r_1} \times \mathbb{C}^{r_2} | \sum_{i=1}^{r_i} |y_i| + 2\sum_{j=1}^{r_2} |z_{r_1+j}| \leq t\}$$

for some $t \in \mathbb{R}_{>0}$. Let

$$t_0 = \left(n!(\frac{4}{\pi})^{r_2} \sqrt{|\Delta_K|}\mathrm{N}(I) \right)^{1/n}$$

By Lemmas 4.1.4 and 4.1.2, we have $\mu(B_t) > 2^n \times 2^{-r_2}\sqrt{|\Delta_K|}\mathrm{N}(I) = 2^n \mathrm{Vol}(\mathbb{R}^n/\lambda(I))$ for $t > t_0$. By Minkowski's Lemma 4.1.1, $B_t \cap \lambda(I)$ contains a non-zero element for any $t > t_0$. As $B_{t_0+1/2}$ is compact and $\lambda(I)$ is discrete, $B_{t_0+1/2} \cap \lambda(I)$ is finite. Therefore, there exists a nonzero α which belong to $B_{t_0+1/2^m} \cap \lambda(I)$ for infinitely many (hence for all) $m \geq 1$. But $B_{t_0} = \bigcap_{m\geq 1} B_{t_0+1/2^m}$, it follows that $\alpha \in B_{t_0} \cap \lambda(I)$. Then

$$\begin{aligned} |\mathrm{N}(\alpha)| &= \prod_{i=1}^{r_1} |\sigma_i(\alpha)| \prod_{j=1}^{r_2} |\sigma_{r_1+2j}(\alpha)|^2 \\ &\leq n^{-n} \big(\sum_{i=1}^{r_1} |\sigma_i(\alpha)| + 2\sum_{j=1}^{r_2} |\sigma_{r_1+2j}(\alpha)| \big)^n \\ &= t_0^n/n^n = (\frac{4}{\pi})^{r_2} \frac{n!}{n^n} \sqrt{\Delta_K}\mathrm{N}(I). \end{aligned}$$

where the second step is the arithmetic-geometric mean inequality. □

Corollary 4.1.5. — *For a number field K of degree n, we have*

$$|\Delta_K|^{1/2} \geq (\frac{\pi}{4})^{n/2} \frac{n^n}{n!},$$

where the sequence $a_n := (\frac{\pi}{4})^{n/2} \frac{n^n}{(n!)}$ is strictly increasing with $a_n \to \infty$ and $a_2 > 1$. In particular, $|\Delta_K| > 1$ if $K \neq \mathbb{Q}$; in other words, if K is a number field in which all prime p is unramified, then $K = \mathbb{Q}$.

Proof. — Applying Theorem 4.1.3(2) to the case $I = \mathcal{O}_K$, one see that there exists $\alpha \in \mathcal{O}_K$ such that

$$1 \leq \mathrm{N}_{K/\mathbb{Q}}(\alpha) \leq (\frac{4}{\pi})^{r_2} \frac{n!}{n^n} \sqrt{|\Delta_K|}.$$

So one obtains $\sqrt{|\Delta_K|} \geq (\frac{\pi}{4})^{r_2} \frac{n^n}{n!} \geq (\frac{\pi}{4})^{n/2} \frac{n^n}{n!} = a_n$. Note that

$$\frac{a_{n+1}}{a_n} = \sqrt{\frac{\pi}{4}}(1 + \frac{1}{n})^n > \sqrt{\frac{\pi}{4}}(1 + \frac{1}{2})^2 > 1.$$

Therefore, a_n is strictly increasing and one has $\sqrt{|\Delta_K|} \geq a_2 > 1$. □

Corollary 4.1.6 **(Hermite)**. — *For a fixed integer Δ, there exist only finitely many number fields with discriminant Δ.*

Proof. — By the previous Corollary, if K is a number field with discriminant Δ, then its degree $n = [K : \mathbb{Q}]$ is bounded by in terms of $|\Delta|$. It suffices to prove that there are only finitely many number fields K of given discriminant Δ, and whose number of real and non-real embeddings are respectively r_1 and r_2. We construct an algebraic integer $\alpha \in \mathcal{O}_K$ as follows.

– Consider first the case $r_1 > 0$, i.e. K admits real embeddings. Choose real numbers c_i for $1 \leq i \leq r_1 + r_2$ such that $c_1 > 1$, $c_i < 1$ for $i > 1$, and

$$\prod_{i=1}^{r_1+r_2} c_i > (\frac{2}{\pi})^{r_2}\sqrt{|\Delta|}.$$

Then Theorem 4.1.3(1) implies that there exists a nonzero $\alpha \in \mathcal{O}_K$ such that $|\sigma_i(\alpha)| < c_i$ for $1 \leq i \leq r_1$ and $|\sigma_{r_1+2j}(\alpha)|^2 < c_{r_1+j}$ for $1 \leq j \leq r_2$. Since

$$1 \leq |\mathrm{N}_{K/\mathbb{Q}}(\alpha)| = |\sigma_1(\alpha)| \prod_{i=2}^{r_1} |\sigma_i(\alpha)| \prod_{j=1}^{r_2} |\sigma_{r_1+j}(\alpha)|^2,$$

it follows that $|\sigma_1(\alpha)| > 1$ and $|\sigma_i(\alpha)| < 1$ for $\sigma_i \neq \sigma_1$. In particular, one has $\sigma_1(\alpha) \neq \sigma_i(\alpha)$ if $\sigma_i \neq \sigma_1$.

– If $r_1 = 0$, consider the centrally symmetric convex region X of $\mathbb{C}^{r_2}$ given by

$$X = \{z \in \mathbb{C}^{r_2} || \Re(z_1)| < 1/2, |\Im(z_1)| < c_1, |z_j|^2 < c_j = 1/2, \forall 2 \leq j \leq r_2\},$$

where c_1 is some constant such that $\mu(X) > 2^n 2^{-r_2}\sqrt{|\Delta_K|}$. Applying Minkowski's lemma 4.1.1 to X and $\lambda(\mathcal{O}_K)$, one sees that there exists nonzero $\alpha \in X \cap \lambda(\mathcal{O}_K)$. Similarly to the previous case, one has $|\sigma_j(\alpha)| < 1$ for $\sigma_j \neq \sigma_1, \bar{\sigma}_1$ and $|\sigma_1(\alpha)| > 1$. But $|\Re(\sigma_1(\alpha))| < 1/2$ by construction so that $|\Im(\sigma_1(\alpha))| > \frac{\sqrt{3}}{2}$. In particular, one has $\sigma_i(\alpha) \neq \sigma_1(\alpha)$ for all $\sigma_i \neq \sigma_1$.

In both cases, α must have degree n over $\mathbb{Q}$; otherwise, by Proposition 1.2.3, there will be some $\sigma_i \neq \sigma_1$ such that $\sigma_1(\alpha) = \sigma_i(\alpha)$ by the existence of $[K : \mathbb{Q}(\alpha)]$ complex embeddings of K extending $\sigma_1|_{\mathbb{Q}(\alpha)}$. Hence, $\mathbb{Q}(\alpha) = K$. If $f(X)$ denotes the monic minimal polynomial of α over $\mathbb{Q}$, then $f(X) \in \mathbb{Z}[X]$ and its coefficients are clearly bounded above in terms of some functions of c_i. Therefore, there are only finitely many possibilities for $f(X)$. □

***Corollary 4.1.7* (Minkowski bound).** — *Let K be a number field of degree n and with r_2 pairs of complex embeddings, Δ_K be the absolute discriminant of K. Then every ideal class of K contains an integral ideal $\mathfrak{a}$ with norm*

$$\mathrm{N}(\mathfrak{a}) \leq (\frac{4}{\pi})^{r_2} \frac{n!}{n^n}\sqrt{|\Delta_K|}.$$

Proof. — Let J be an arbitrary fractional ideal, $I = J^{-1}$. Then by Theorem 4.1.3, there exists a nonzero $\alpha \in I$ such that

$$|\mathrm{N}_{K/\mathbb{Q}}(\alpha)| \leq (\frac{4}{\pi})^{r_2} \frac{n!}{n^n}\sqrt{|\Delta_K|}\mathrm{N}(I).$$

Put $\mathfrak{a} = \alpha I^{-1} = \alpha J$. Then $\mathfrak{a}$ is an integral ideal in the same ideal class as J and satisfies the required property. □

Theorem 4.1.8. — *For any number field K, its ideal class group $\mathcal{C}l_K$ is a finite abelian group.*

Proof. — This follows immediately from Corollary 4.1.7 and the fact that the number of integral ideals of $\mathcal{O}_K$ with a given norm is finite (Proposition 3.1.2). □

Using Corollary 4.1.7, one can compute effectively the ideal class group of a given number field.

Example 4.1.9. — Let $K = \mathbb{Q}(\sqrt{-14})$. Then we have $n = 2$, $r_2 = 1$ and $\Delta_K = -56$. The Minkowski bound is

$$(\frac{4}{\pi})^{r_2}\frac{n!}{n^n}\sqrt{|\Delta_K|} = \frac{4}{\pi}\sqrt{14} \approx 4.765 < 5.$$

By Corollary 4.1.7, every ideal class of K contains an integral ideal of norm ≤ 4. Note that $(2) = \mathfrak{m}_2^2$ with $\mathfrak{m}_2 = (2, \sqrt{-14})$ and $\mathrm{N}(\mathfrak{m}_2) = 2$. Since $\mathrm{N}_{K/\mathbb{Q}}(a + \sqrt{-14}b) = a^2 + 14b^2 = 2$ has no integral solutions, $\mathfrak{m}_2$ is not principal. Hence, $\mathfrak{m}_2$ has order 2 in the ideal class group. Consider the integral ideals of norm 3. We have

$$(3) = \mathfrak{p}_3\bar{\mathfrak{p}}_3, \quad \text{with } \mathfrak{p}_3 = (3, \sqrt{-14} + 1).$$

Note that $\mathfrak{p}_3^2 = (9, -2 + \sqrt{-14}) = (\frac{-2+\sqrt{-14}}{2})\mathfrak{m}_2$. Note also that (2) is the only integral ideal of $\mathcal{O}_K$ with norm 4. It follows that $\mathfrak{p}_3$ has order 4 in $\mathcal{C}l_K$, and $\mathcal{C}l_K \cong \mathbb{Z}/4\mathbb{Z}$.

Example 4.1.10. — For $K = \mathbb{Q}(\sqrt[3]{2})$, we have $n = 3$, $r_2 = 1$ and $\Delta_K = -2^2 3^3$. The Minkowski bound for K is

$$(\frac{4}{\pi})\frac{3!}{3^3}\sqrt{3^3 2^2} \approx 2.94 < 3.$$

But the only integral ideal of $\mathcal{O}_K$ with norm 2 is $(\sqrt[3]{2})$. It follows that $\mathbb{Q}(\sqrt[3]{2})$ has class number 1, hence $\mathcal{O}_K = \mathbb{Z}[\sqrt[3]{2}]$ is a principal ideal domain.

4.2. Dirichlet's unit theorem

Let K be a number field. Denote by $U_K = \mathcal{O}_K^\times$ the group of units (i.e. invertible elements) of $\mathcal{O}_K$. It is clear that an element $x \in \mathcal{O}_K$ is a unit, if and only if $|\mathrm{N}_{K/\mathbb{Q}}(x)| = 1$. The torsion subgroup of U_K, denoted by W_K, is the group of roots of unity contained in K.

Lemma 4.2.1. — *The group W_K is a finite cyclic group. Moreover, an element $u \in \mathcal{O}_K$ belongs to W_K if and only if $|\sigma_i(u)|_{\mathbb{C}} = 1$ for every complex embedding $\sigma_i : K \to \mathbb{C}$.*

Proof. — It is clear that W_K is a finite group, since $K/\mathbb{Q}$ is a finite extension. If W_K were not cyclic, there would exist a prime p such that the p-torsion of W_K is isomorphic to $(\mathbb{Z}/p\mathbb{Z})^r$ for some $r \geq 2$. But this is impossible since $x^p = 1$ has at most p solutions in K. It is clear that $|\sigma_i(u)|_{\mathbb{C}} = 1$ for $u \in W_K$ and any $\sigma_i : K \hookrightarrow \mathbb{C}$. Conversely, assume $u \in \mathcal{O}_K$ is an element with $|\sigma_i(u)| = 1$ for all complex embeddings σ_i. Let $f(X) \in \mathbb{Z}[X]$ denote the monic minimal polynomial of u. Then the coefficients of x^i in $f(X)$ is bounded by $\binom{n}{i}$. Denote by $\mathcal{S}$ the finite subset of $\mathbb{Z}[X]$ consisting of polynomials of degree n and such that the coefficients of X^i is bounded by $\binom{n}{i}$. Then the roots of some polynomial in $\mathcal{S}$ form a finite set. For all $n \in \mathbb{Z}$, u^n satisfies the same condition, thus u^n is a root of some polynomial in $\mathcal{S}$. There exist integers $m > n$ with $m \neq n$ and $u^m = u^n$, that is $u^{m-n} = 1$. □

The main result of this section is the following

***Theorem 4.2.2* (Dirichlet's Unit Theorem).** — *Let K be a number field of degree n with r_1 real embeddings and r_2 pairs of non-real complex embeddings. Then there exists a free abelian group V_K of rank $r_1 + r_2 - 1$ such that $U_K = W_K \times V_K$.*

Note that W_K is canonically determined by K, but V_K is not. A $\mathbb{Z}$-basis of V_K is usually called a *system of fundamental units* of K. If $\{\eta_1, \cdots, \eta_{r_1+r_2-1}\}$ is such a basis, then every $u \in U_K$ writes uniquely as

$$u = w\eta_1^{a_1} \cdots \eta_{r_1+r_2-1}^{a_{r_1+r_2-1}}, \quad \text{with } w \in W_K, a_i \in \mathbb{Z}.$$

Proof. — Let $\sigma_1, \cdots, \sigma_{r_1}$ denote the real embeddings of K, and $\sigma_{r_1+j}, \bar{\sigma}_{r_1+j}$ with $1 \leq j \leq r_2$ be the non-real embeddings. Let $\lambda : K \to \mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$ be the Minkowski embedding given by $x \mapsto (\sigma_i(x))_{1 \leq i \leq r_1+r_2}$. Then the image of $\mathcal{O}_K$ is a lattice in $\mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$, and $\lambda(\mathcal{O}_K \backslash \{0\})$ is contained in $\mathbb{R}^{\times,r_1} \times \mathbb{C}^{\times,r_2}$. Define the map $\ell : U_K \to \mathbb{R}^{r_1+r_2}$ as the composite of the inclusion

$$U_K \subseteq \mathcal{O}_K \backslash \{0\} \xrightarrow{\lambda} \mathbb{R}^{\times,r_1} \times \mathbb{C}^{\times,r_2}$$

with the logarithmic map

$$\mathrm{Log} : \mathbb{R}^{\times,r_1} \times \mathbb{C}^{\times,r_2} \to \mathbb{R}^{r_1+r_2} = \mathbb{R}^{r_1} \times \mathbb{R}^{r_2}$$

given by

$$(y_1, \cdots, y_{r_1}, z_1, \cdots, z_{r_2}) \mapsto (\log|y_1|, \cdots, \log|y_{r_1}|, 2\log|z_1|, \cdots, 2\log|z_{r_2}|).$$

Then ℓ is homomorphism of abelian groups. By Lemma 4.2.1, the kernel of ℓ is W_K, and the image of ℓ is contained in the hyperplane $H \subseteq \mathbb{R}^{r_1+r_2}$ defined by $\sum_{i=1}^{r_1+r_2} x_i = 0$ since

$$\sum_{i=1}^{r_1} \log|\sigma_i(u)| + 2\sum_{j=1}^{r_2} \log|\sigma_{r_1+j}(u)| = \log|\mathrm{N}_{K/\mathbb{Q}}(u)| = 0, \quad \forall u \in U_K.$$

We will prove that $\ell(U_K)$ is actually a full lattice in H, hence of rank $r_1 + r_2 - 1$. Then if V_K is the image of a section of the quotient $U_K \to \ell(U_K)$, we have $U_K \cong W_K \times V_K$. If $r_1 + r_2 = 1$, the statement is trivial. Thus we assume that $r_1 + r_2 > 1$.

First, we show that $\ell(U_K)$ is a discrete subgroup in H. For any $\delta \in \mathbb{R}$, let B_δ denote the subset consisting of $(y, z) \in \mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$ such that $|y_i|, |z_j|^2 \leq e^\delta$ for all i, j, and $\bar{B}_\delta$ be the closure of B_δ. Assume $\delta > 0$, and put $C_\delta = \bar{B}_\delta \backslash B_{-\delta}$. Since $\lambda(\mathcal{O}_K)$ is discrete in $\mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$, $\lambda(\mathcal{O}_K) \cap C_\delta$ is a finite set. Thus for sufficiently small $\delta > 0$, we have

$$\lambda(\mathcal{O}_K) \cap C_\delta = \lambda(\mathcal{O}_K) \cap C_0 = \lambda(W_K).$$

We fix such a δ, and put $D_\delta = \{x \in \mathbb{R}^{r_1+r_2} | |x_i| < \delta\}$. Then $\ell(U_K) \cap D_\delta$ is contained in the image of $\lambda(\mathcal{O}_K) \cap C_\delta$ under the map Log, hence $\ell(U_K) \cap D_\delta \subseteq \ell(W_K) = \{0\}$. This proves the discreteness of $\ell(U_K)$ in $\mathbb{R}^{r_1+r_2}$, and hence in H.

To finish the proof, it remains to show that $\ell(U_K)$ has rank $r_1 + r_2 - 1$. We need the following

Lemma 4.2.3. — *For each integer k with $1 \leq k \leq r_1 + r_2$, there exists $u_k \in U_K$ such that $|\sigma_k(u_k)| > 1$ and $|\sigma_i(u_k)| < 1$ for all $i \neq k$.*

Proof. — We fix a k as in the statement, and a constant $A > (\frac{2}{\pi})^{r_2}|\Delta_K|^{1/2}$. Let $c_1, \cdots, c_{r_1+r_2} > 0$ be such that $c_i < 1$ for all $i \neq k$ and $c_k = A/\prod_{i\neq k} c_i$. By Theorem 4.1.3, there exists a non-zero $a_1 \in \mathcal{O}_K$ such that $|\sigma_i(a_1)| < c_i$ for all $1 \leq i \leq r_1$ and $|\sigma_i(a_1)|^2 < c_i$ for $r_1 + 1 \leq i \leq r_1 + r_2$. Put $c_i^{(1)} := |\sigma_i(a_1)|$ for $i \neq k$ and $1 \leq i \leq r_1$, $c_i^{(1)} = |\sigma_i(a_1)|^2$ for $i \neq k$ and $r_1 + 1 \leq i \leq r_1 + r_2$, and $c_k^{(1)} = A/\prod_{i\neq k} c_i^{(1)}$. Replacing c_i by $c_i^{(1)}$ and applying Theorem 4.1.3 again, one gets a nonzero $a_2 \in \mathcal{O}_K$ such that $|\sigma_i(a_2)| < |\sigma_i(a_1)|$ for $i \neq k$ and

$$|\mathrm{N}_{K/\mathbb{Q}}(a_2)| = \prod_{i=1}^{r_1} |\sigma_i(a_2)| \prod_{j=1}^{r_2} |\sigma_{r_1+j}(a_2)|^2 < \prod_{i=1}^{r_1+r_2} c_i^{(1)} = A.$$

Repeating this process, one gets a sequence $a_1, a_2, \cdots, a_n, \cdots$, such that $|\sigma_i(a_{n+1})| < |\sigma_i(a_n)|$ for all $i \neq k$ and $|\mathrm{N}_{K/\mathbb{Q}}(a_n)| < A$. But there exist only finitely many integral ideals of $\mathcal{O}_K$ with norm strictly less than A. Therefore, there are integers $m > n$ such that $(a_m) = (a_n)$. Then $u_k = a_m/a_n$ satisfies the requirement of the Lemma.

□

We come back to the proof of Theorem 4.2.2 as follows. Let u_k with $1 \leq k \leq r_1 + r_2$ be as in the Lemma, and view $\ell(u) \in \mathbb{R}^{r_1+r_2}$ as a column vector. Then the entries on the main diagonal of the $(r_1 + r_2) \times (r_1 + r_2)$-matrix

$$(\ell(u_1), \cdots, \ell(u_{r_1+r_2}))$$

are positive, and all entries off the main diagonal are negative. Then it follows from the following elementary Lemma that this matrix has rank $r_1 + r_2 - 1$. □

Lemma 4.2.4. — *Let $A = (a_{i,j})_{1\leq i,j\leq n}$ be an $n \times n$ real matrix. Assume that $\sum_{i=1}^n a_{i,j} = 0$ for all j, $a_{i,i} > 0$ for all i and $a_{i,j} < 0$ if $i \neq j$. Then the rank of A is $n - 1$.*

Proof. — It suffices to show that the first $n-1$ rows of A are linearly independent. Assume in contrary that there exist $x_1, \cdots, x_{n-1} \in \mathbb{R}$ not all equal to 0 such that $\sum_{i=1}^{n-1} x_i a_{i,j} = 0$ for all $1 \leq j \leq n$. Putting $j = n$, we see that the x_i's can not be all positive or all negative. Let $1 \leq j_0 \leq n - 1$ be such that $x_{j_0} = \max_{1\leq j\leq n-1}\{x_j\} > 0$. Then one has

$$0 = \sum_i x_i a_{i,j_0} = x_{j_0} \sum_{i=1}^{n-1} a_{i,j_0} + \sum_{i\neq j_0} (x_i - x_{j_0}) a_{i,j} > 0,$$

which is absurd.

□

CHAPTER 5

BINARY QUADRATIC FORMS AND CLASS NUMBER

In this chapter, we will discuss an interesting relation between the ideal class group and integral binary quadratic forms. These results are very classical, and go back to Gauss.

5.1. Binary quadratic forms

Definition 5.1.1. — An (integral) *binary quadratic form* is a homogenous polynomial of the form $F(x,y) = ax^2+bxy+cy^2$ with $a,b,c \in \mathbb{Z}$. We say F is *primitive* if $\gcd(a,b,c) = 1$. The discriminant of F is defined as

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

An immediate remark is that, a binary quadratic form F can be written as a product of linear functions with rational coefficients if and only if d is a square integer. From now on, we assume that *the discriminants of all binary quadratic forms involved are not square integers.*

We say that an integral binary quadratic form $F(x,y)$ is

- *indefinite* if $F(x,y)$ takes both positive and negative values, or equivalently its discriminant $d > 0$;
- *positive definite* if $F(x,y) > 0$ for all $(x,y) \neq 0$, or equivalently its discriminant $d < 0$ and the coefficient of x^2 is postive.

Definition 5.1.2. — We say two binary forms $F(x,y)$ and $G(x,y)$ are *equivalent* if there exists $\gamma = \begin{pmatrix} r & s \\ u & v \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ such that $G(x,y) = F(rx+sy, ux+vy)$.

For a binary quadratic form $F(x,y) = ax^2 + bxy + cy^2$, we call $Q = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}$ the matrix associated with F. Then a binary quadratic form G is equivalent to F if and only if there exists a $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ such that $\gamma^t Q \gamma$ is the matrix associated to G. As $d = -4\det(Q)$, equivalent binary quadratic forms have the same discriminant.

It is also clear that if F and G are equivalent binary quadratic forms, then the number of solutions to $F(x,y) = n$ is the same as $G(x,y) = n$ for any integer n.

We are interested in classifying binary quadratic forms up to equivalent classes.

Lemma 5.1.3. — *Every binary quadratic form is equivalent to a form* $ax^2 + bxy + cy^2$ *with*

$$|b| \leq |a| \leq |c|.$$

Proof. — Let a be the integer such that $F(x, y) = a$ has integral solutions and $|a|$ is minimal. Let $r, s \in \mathbb{Z}$ such that $F(r, s) = a$. Then $\gcd(r, s) = 1$; otherwise one has $F(r/q, s/q) = a/q^2$ with $q = \gcd(r, s)$, which contradicts with the minimality of $|a|$. Then there exist $u, v \in \mathbb{Z}$ such that $rv - us = 1$, and

$$F(rx + uy, sx + vy) = ax^2 + b'xy + c'y^2 \quad \text{for some } b', c' \in \mathbb{Z}.$$

Note that

$$a(x + hy)^2 + b'(x + hy)y + c'y^2 = ax^2 + (b' + 2ah)xy + (ah^2 + b'h + c')y^2.$$

Then one can choose $h \in \mathbb{Z}$ such that $|b' + 2ah| \leq |a|$. Put $b = (b' + 2ah)$ and $c = ah^2 + b'h + c'$. Then $|c| \geq |a|$ by the minimality of $|a|$.

□

Theorem 5.1.4. — *For a fixed non-square integer* d*, there exists only finite equivalent classes of binary quadratic forms with discriminant* d*.*

Proof. — By the previous lemma, every equivalent class of binary quadratic forms of discriminant d contains a representative of the form

$$ax^2 + bxy + cy^2, \quad \text{with } d = b^2 - 4ac \text{ and } |b| \leq |a| \leq |c|.$$

We have the following two cases:

– If $d > 0$, then it follows from $|ac| \geq b^2 = d + 4ac \geq 4ac$ that $ac < 0$. Hence, one has

$$d \geq 4|ac| \geq 4a^2, \quad \text{i.e. } a \leq \frac{\sqrt{d}}{2}.$$

Once a is fixed, there are only finitely many possible choices for b as $|b| \leq |a|$, hence for $c = (b^2 - d)/4a$.

– If $d < 0$, then one has

$$|d| = 4ac - b^2 \geq 4a^2 - a^2 = 3a^2, \quad \text{i.e. } |a| \leq \sqrt{\frac{|d|}{3}}.$$

□

Theorem 5.1.5. — *Every positive definite equivalent class of primitive binary quadratic forms contains a unique form* $ax^2 + bxy + cy^2$ *with*

$$|b| \leq a \leq c \quad \textit{and } b \geq 0 \textit{ if } |b| = a \textit{ or } a = c.$$

Remark 5.1.6. — A positive definite binary quadratic form of the form in the Theorem is called *reduced*.

Proof. — We have already seen in Lemma 5.1.3 that any binary quadratic form is equivalent to a form $F(x,y)$ with $|b| \leq a \leq c$. Such a form is already reduced unless $b = -a$ or $a = c$ and $b < 0$. In the these cases, we make the following substitutions to make F reduced:

$$F(x,y) = ax^2 - axy + cy^2 \Longrightarrow F(x+y,y) = ax^2 + axy + cy^2;$$
$$F(x,y) = ax^2 + bxy + ay^2 \Longrightarrow F(-y,x) = ax^2 - bxy + ay^2.$$

We verify now that any two reduced primitive binary forms can not be equivalent. Let $F(x,y) = ax^2 + bxy + cy^2$ be a reduced form. Then we claim that

$$F(x,y) \geq (a + c - |b|) \min\{x^2, y^2\}, \quad \forall x, y \in \mathbb{Z}. \tag{5.1.6.1}$$

Indeed, without loss of generality, we may assume that $|x| \geq |y|$. Then

$$F(x,y) \geq (a - |b|)|x||y| + cy^2 \geq (a + c - |b|)y^2.$$

In particular, one has $F(x,y) \geq a + c - |b|$ if $xy \neq 0$, and the equality holds only if $(x,y) = \pm(1, -\mathrm{sign}(b))$. The smallest three integers represented by F are

$$a \leq c \leq a + c - |b|. \tag{5.1.6.2}$$

Assume now $G(x,y)$ is another reduced form equivalent to $F(x,y)$. Then one has $G(x,y) = ax^2 + b'xy + c'y^2$. We distinguish several cases:

- If $a = c = b \geq 0$, then the equality in $-d = 4ac' - b'^2 \geq 4a^2 - b^2$ holds. Therefore, $c' = a$ and $|b'| = a$. Since G is also reduced, then $b' = a$, i.e. $F = G$.
- If $a = c > b \geq 0$, then one has either $c' = a$ or $c' = 2a - b$. But $F(x,y) = a$ has 4 solutions, namely $(\pm 1, 0)$ and $(0, \pm 1)$. It follows that $c' = 2a - b$ is impossible, because otherwise $G(x,y) = a$ would have only 2 solutions. It follows also from $b' = \sqrt{4ac + d}$ that $b = b'$.
- If $c > a = |b|$, then $b = a$ and $c = a + c - |b|$ is the second smallest integer represented by n. Thus one have $c' = c$ or $c' = a$. But the second case can not be true because of the discussions on the previous two cases. It follows that $c' = c$ and hence $b' = \sqrt{4ac' + d} = b$.
- If $c > a > |b|$, then one has $c' > a > |b'|$ by applying the previous discussion to G. Since the inequalities in (5.1.6.2) are strict, we see that $c' = c$ and $|b'| = |b|$. Using the fact that the only solutions to $F(x,y) = a$ (resp. to $F(x,y) = c$) are $(\pm 1, 0)$ (resp. $(0, \pm 1)$), one checks easily that $ax^2 + bxy + cy^2$ is not equivalent to $ax^2 - bxy + cy^2$. Therefore, one concludes that $b' = b$.

□

Using Theorem 5.1.5, it is easy to list all the equivalent classes of positive definite binary quadratic forms with given discriminant $d < 0$. We have the following table for small $-d > 0$ (note that one has always $d \equiv 0, 3 \mod 4$):

values of d	reduced binary forms with discriminant d
-3	x^2+xy+y^2
-4	x^2+y^2
-7	$x^2+xy+2y^2$
-8	x^2+2y^2
-11	$x^2+xy+3y^2$
-12	x^2+3y^2
-20	$x^2+5y^2, 2x^2+2y^2+3y^2$
-23	$x^2+xy+6y^2, 2x^2 \pm xy+3y^2,$
-52	$x^2+13y^2, 2x^2+2xy+7y^2$
-56	$x^2+14y^2, 2x^2+7y^2, 3x^2 \pm 2xy+5y^2$

5.2. Representation of integers by binary quadratic forms

Definition 5.2.1. — We say that an integer n is *represented* by a binary quadratic form $F(x,y)$, if $F(x,y)=n$ has solutions in $\mathbb{Z}^2$, and n is *properly represented* by $F(x,y)$, if one can choose (x,y) so that $\gcd(x,y)=1$.

Lemma 5.2.2. — *A binary quadratic form $F(x,y)$ properly represents an integer n if and only if $F(x,y)$ is equivalent to the form $nx^2+b'xy+c'y^2$ for some $b',c'\in\mathbb{Z}$.*

Proof. — The condition is clearly sufficient. Suppose that $F(u,v)=n$ with $\gcd(u,v)=1$. Then one chooses $r,s\in\mathbb{Z}$ such that $us-rv=1$. Then $F(x,y)$ is equivalent to

$$F(ux+ry,vx+sy)=nx^2+(2aur+bus+brv+2cvs)xy+F(r,s)y^2.$$

□

Proposition 5.2.3. — *Let $n\neq 0$ and d be integers. Then the following are equivalent:*

1. *There exists a binary quadratic form of discriminant d that properly represents n.*
2. *d is square modulo $4n$.*

Proof. — If F is a binary quadratic form that properly represents n, then F is equivalent to $nx^2+bxy+cy^2$ for some $b,c\in\mathbb{Z}$ by the previous Lemma. Hence, one gets $d=b^2-4nc$ and $d\equiv b^2 \mod 4n$.

Conversely, suppose that $d\equiv b^2 \mod 4n$ so $d=b^2-4nc$ for some $c\in\mathbb{Z}$. Then the binary form $F(x,y)=nx^2+bxy+cy^2$ has discriminant d and represents n. □

Corollary 5.2.4. — *Let n be an integer represented by a binary quadratic form with discriminant d, and p be a prime with $\left(\frac{d}{p}\right)=-1$. Then the exponent of p in n is even.*

Proof. — Indeed, if the exponent of p in n is odd, d would be a quadratic residue modulo p by the Proposition. □

Example 5.2.5. — When $n=1,2,3$, the only positive definite reduced form of discriminant $d=-4n$ is x^2+ny^2. Thus by Proposition 5.2.3, an integer m can be properly represented by x^2+ny^2 if and only if $-4n$ is a square modulo $4m$, i.e. $-n$ is a square

modulo m. In particular, if p is a prime coprime with n with $\left(\frac{-n}{p}\right) = 1$, then it is represented by $x^2 + ny^2$. In view of $\mathrm{N}_{\mathbb{Q}(\sqrt{-n})/\mathbb{Q}}(x + \sqrt{-n}y) = x^2 + ny^2$ and the multiplicativity of norms, any product of such primes is represented by $x^2 + ny^2$. By quadratic reciprocity law, we have

$$\left(\frac{-1}{p}\right) = -1 \Leftrightarrow p \equiv 3 \mod 4$$
$$\left(\frac{-2}{p}\right) = -1 \Leftrightarrow p \equiv 5, 7 \mod 8$$
$$\left(\frac{-3}{p}\right) = -1 \Leftrightarrow p \equiv 2 \mod 3$$

After checking the representability of $m = 2^k$ by hand, one obtains by Corollary 5.2.4 the following

m is represented by	iff the following primes have even exponent in m
$x^2 + y^2$	$p \equiv 3 \mod 4$
$x^2 + 2y^2$	$p \equiv 5, 7 \mod 8$
$x^2 + 3y^2$	$p \equiv 2 \mod 3$

Example 5.2.6. — A positive integer n is represented by $x^2 + 5y^2$ if and only if

(1) any prime $p \equiv 11, 13, 17, 19 \mod 20$ appears in n with even exponent;
(2) the total number of prime divisors $p \equiv 2, 3, 7 \mod 20$ (counted with multiplicity) is even.

Note that there are no restrictions on the number of primes $p \equiv 1, 5, 9 \mod 20$.

First, there are only two binary quadratic forms with discriminant -20, namely

$$f(x, y) = x^2 + 5y^2, \quad g(x, y) = 2x^2 + 2xy + 3y^2.$$

By Proposition 5.2.3, a prime p coprime to -20 is represented by f or g if and only if $\left(\frac{-5}{p}\right) = 1$. By quadratic reciprocity law, we have

$$\left(\frac{-5}{p}\right) = \begin{cases} 1 & p \equiv 1, 3, 7, 9 \mod 20 \\ -1 & p \equiv 11, 13, 17, 19 \mod 20. \end{cases}$$

Then the first statement follows from Corollary 5.2.4. By checking modulo 4, we see easily that primes $p \equiv 1, 9 \mod 20$ can not be represented by g, thus they must be represented by f; on the other hand, primes $p \equiv 3, 7 \mod 20$ can not be represented by f, thus they are represented by g. One notes also that 2 is presented by g not by f, and 5 is represented by f by not by g. Therefore, any power of primes $p \equiv 1, 5, 9 \mod 20$ can appear in n. Finally, statement (2) comes from the magical identity

$$(2x^2 + 2xy + 3y^2)(2z^2 + 2zw + 3w^2) = (2xz + xy + yz + 3yw)^2 + 5(xw - yz)^2.$$

What does the magical identity come from? Actually, if $\mathfrak{m}_2 = (2, 1+\sqrt{-5})$ denotes the unique prime of $K = \mathbb{Q}(\sqrt{-5})$ above 2, then

$$2x^2 + 2xy + 3y^2 = \frac{\mathrm{N}_{K/\mathbb{Q}}(2x + y(1+\sqrt{-5}))}{\mathrm{N}(\mathfrak{m}_2)} = \frac{\mathrm{N}_{K/\mathbb{Q}}(2x + y(1-\sqrt{-5}))}{\mathrm{N}(\mathfrak{m}_2)}.$$

Now the magical formula follows from

$$(2x + y + \sqrt{-5}y)(2z + w - \sqrt{-5}w) = 2[(2xz + yz + xw + 3yw) + (yz - xw)\sqrt{-5}]$$

by taking norms.

5.3. Ideal class groups and binary quadratic forms

Let $K = \mathbb{Q}(\sqrt{d})$ be a quadratic field with discriminant d. Denote by $x \mapsto \bar{x}$ the non-trivial automorphism of $K/\mathbb{Q}$. We will consider a slightly different ideal class group in the real quadratic case.

Definition 5.3.1. —
- We say an element $x \in K$ is *totally positive* if $\sigma(x) > 0$ for all real embedding σ of K (so that the condition is empty if K is imaginary).
- Let $\mathcal{I}_K$ be the group of fractional ideals of K. We denote by $\mathcal{P}_K^+ \subseteq \mathcal{I}_K$ the subgroup consisting of fractional ideals generated by a totally positive element. We define the *strict (or narrow) ideal class group* of K as

$$\mathcal{C}l_K^+ = \mathcal{I}_K/\mathcal{P}_K^+.$$

If K is imaginary quadratic, $\mathcal{C}l_K^+$ is the usual ideal class group $\mathcal{C}l_K$. If K is real quadratic, $\mathcal{C}l_K$ is a quotient of $\mathcal{C}l_K^+$ with kernel $\mathcal{P}_K/\mathcal{P}_K^+$.[(1)]

Definition 5.3.2. — Let α_1, α_2 be two $\mathbb{Q}$-linearly independent elements of K. We say that (α_1, α_2) is *positively oriented* if

$$\frac{\det\begin{pmatrix}\alpha_1 & \alpha_2 \\ \bar{\alpha}_1 & \bar{\alpha}_2\end{pmatrix}}{\sqrt{d}} > 0.$$

Note that exactly one of pairs (α_1, α_2) and (α_2, α_1) is positively oriented. So the notion of positive orientation gives a way to choose the order of any two linearly independent elements in K.

Let I be a fractional ideal of K, and (ω_1, ω_2) be a positively oriently basis of I over $\mathbb{Z}$. We put

$$f_{\omega_1,\omega_2}(x, y) = \frac{\mathrm{N}_{K/\mathbb{Q}}(x\omega_1 + y\omega_2)}{\mathrm{N}(I)}.$$

Lemma 5.3.3. — *The quadratic form f_{ω_1,ω_2} has coefficients in $\mathbb{Z}$ and has discriminant d, and it is positive definite if K is imaginary quadratic. Moreover, the equivalent class of f_{ω_1,ω_2} depends only on the class of I in $\mathcal{C}l_K^+$.*

(1) Actually, $\mathcal{P}_K/\mathcal{P}_K^+$ is trivial if K has a unit of norm -1, and it is of order 2 otherwise.

Proof. — It is clear that $f_{\omega_1,\omega}(x,y) \in \mathbb{Z}$ for any $x, y \in \mathbb{Z}$. The coefficients of x^2, xy and y^2 are respectively given by $f_{\omega_1,\omega_2}(1,0)$, $f_{\omega_1,\omega_2}(1,1) - f_{\omega_1,\omega_2}(1,0) - f_{\omega_1,\omega_2}(0,1)$ and $f_{\omega_1,\omega_2}(0,1)$. This shows that f_{ω_1,ω_2} is integral. A direct computation also shows that the discriminant of f_{ω_1,ω_2} is given by

$$\frac{(\omega_1\bar{\omega}_2 - \bar{\omega}_1\omega_2)^2}{\mathrm{N}(I)^2} = \frac{\mathrm{Disc}(\omega_1,\omega_2)}{\mathrm{N}(I)^2} = d.$$

If K is imaginary quadratic, f_{ω_1,ω_2} is clearly positive definite, since the norm of any element in K is positive definite.

Now if (ω_1', ω_2') is another positively oriented basis of I, then there exists $\gamma \in \mathrm{GL}_2(\mathbb{Z})$ such that $(\omega_1', \omega_2') = (\omega_1, \omega_2)\gamma$. As both (ω_1', ω_2') and (ω_1, ω_2) are positively oriented, we have $\gamma \in \mathrm{SL}_2(\mathbb{Z})$. Therefore, $f_{\omega_1',\omega_2'}$ is equivalent to f_{ω_1,ω_2} under the action of $\mathrm{SL}_2(\mathbb{Z})$. Let J be a fractional ideal in the same class in $\mathcal{C}l_K^+$ as I. Then there exists $\alpha \in K$ such that $J = I(\alpha)$, where α is totally positive if K is quadratic real. Then $(\alpha\omega_1, \alpha_2\omega_2)$ is positively oriented basis of J, and one has $f_{\alpha\omega_1,\alpha\omega_2} = f_{\omega_1,\omega_2}$. □

For an integral binary quadratic form f, let $[f]$ denote the equivalent class of f under the action of $\mathrm{SL}_2(\mathbb{Z})$.

Theorem 5.3.4. — *The above construction $I \mapsto [f_{\omega_1,\omega_2}]$ induces a bijection between the set $\mathcal{C}l_K^+$ and the set of equivalent classes of binary quadratic forms with discriminant d, which are positive definite if $d < 0$.*

Proof. — We prove first the surjectivity of the morphism. Given a binary quadratic form $f(x,y) = ax^2 + bxy + cy^2$ with discriminant d, which is not negative definite, we have to show that there exists a fractional ideal I and a positive oriented basis (ω_1, ω_2) of I such that $[f] = [f_{\omega_1,\omega_2}]$. Up to replacing f by a form equivalent to it, we may assume that $a > 0$. Let τ denote the root of $ax^2 - bx + c = 0$ such that $(1, \tau)$ is positively oriented. Consider the lattice $I = \mathbb{Z} + \mathbb{Z}\tau \subseteq K$. We verify that I is a fractional ideal of K. We distinguish two cases:

1. $d \equiv 0 \mod 4$. Then $2|b$, and $\mathcal{O}_K = \mathbb{Z} + \mathbb{Z}\frac{\sqrt{d}}{2}$. It suffices to show that I is stable under multiplication by $\sqrt{d}$. If $\tau = \frac{b \pm \sqrt{d}}{2a}$, then
$$\frac{\sqrt{d}}{2}(1,\tau) = (1,\tau) \pm \begin{pmatrix} -\frac{b}{2} & -c \\ a & -\frac{b}{2} \end{pmatrix}.$$

2. $d \equiv 1 \mod 4$. According to $\tau = \frac{b \pm \sqrt{d}}{2a}$, we take correspondingly $\omega_d = \frac{1 \pm \sqrt{d}}{2}$. Then we have $\mathcal{O}_K = \mathbb{Z} + \mathbb{Z}\omega_d$ and
$$\omega_d(1,\tau) = (1,\tau)\begin{pmatrix} \frac{1-b}{2} & -c \\ a & \frac{1+b}{2} \end{pmatrix}.$$
As b is odd, this implies that I is stable under the multiplication by ω_d (hence a fractional ideal).

Next, we compute $\mathrm{N}(I)$. Note that $\mathrm{Disc}(1,\tau) = \det \begin{pmatrix} 1 & \tau \\ 1 & \bar{\tau} \end{pmatrix}^2 = d/a^2$. It follows that $\mathrm{N}(I) = a^{-1}$. Therefore, we see easily that

$$f_{1,\tau} = \frac{\mathrm{N}_{K/\mathbb{Q}}(x - y\tau)}{\mathrm{N}(I)} = ax^2 + bxy + cy^2 = f.$$

Now we prove the injectivity of the morphism. Let $f(x,y)$ and τ be as above. Suppose that J is a fractional ideal with positive oriented basis (ω_1, ω_2) such that $[f_{\omega_1,\omega_2}] = [f]$. We have to show that the class of J in $\mathcal{C}l_K^+$ is the same as $I = (1,\tau)$. There exists $\gamma = \begin{pmatrix} r & s \\ u & v \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ such that

$$f_{\omega_1,\omega_2}(rx + sy, ux + vy) = f(x,y).$$

Then up to replacing (ω_1, ω_2) by $(\omega_1', \omega_2') = (\omega_1, \omega_2)\gamma$, we may assume that $f_{\omega_1,\omega_2} = f$. Thus one gets

$$\mathrm{N}_{K/\mathbb{Q}}(\omega_1 x + y\omega_2) = (\omega_1 x + y\omega_2)(\bar{\omega}_1 x + \bar{\omega}_2 y) = \mathrm{N}(J)(ax^2 + bxy + cy^2).$$

Note that $\mathrm{N}_{K/\mathbb{Q}}(\omega_1) = a\mathrm{N}(J) > 0$. Therefore, up to replacing (ω_1, ω_2) by $(-\omega_1, -\omega_2)$, one may assume that ω_1 is totally positive. Putting $(x,y) = (-\tau, 1)$, we see that if $\tau' = \omega_2/\omega_1$, then either $\tau' = \tau$ or $\tau' = \bar{\tau}$. Note that

$$\det \begin{pmatrix} \omega_1 & \omega_2 \\ \bar{\omega}_1 & \bar{\omega}_2 \end{pmatrix} = \mathrm{N}_{K/\mathbb{Q}}(\omega_1) \det \begin{pmatrix} 1 & \tau' \\ 1 & \bar{\tau}' \end{pmatrix}.$$

Since (ω_1, ω_2) and $(1,\tau)$ are both positively oriented, the determinant above has the same sign as that of $\begin{pmatrix} 1 & \tau \\ 1 & \bar{\tau} \end{pmatrix}$. Thus one gets $\tau' = \tau$, and hence $J = (\omega_1)I$.

□

Remark 5.3.5. — (1) Note that d being the discriminant of a quadratic field is equivalent to the following conditions:

- d the exponent of any odd primes in d is at most one;
- $d \equiv 1 \mod 4$ or $d \equiv 8, 12 \mod 16$.

Such a d is usually called a *fundamental discriminant*. The discriminant of a general binary quadratic form writes uniquely as $d = f^2 d_K$, where d_K is a fundamental discriminant of a quadratic field K, and $f > 0$ is an integer, called the conductor of d. There exists a similar bijection between the equivalent classes of non-negative definite binary quadratic forms with discriminant d and the strict ideal classes of the subring $\mathcal{O} = \mathbb{Z} + f\mathcal{O}_K$ of $\mathcal{O}_K$.

(2) Combining with Theorem 5.1.4, this theorem gives another proof of the finiteness of the class number of a quadratic field. Actually, Theorem 5.1.5 even gives a very efficient algorithm to compute the class number of an imaginary quadratic field.

(3) Another interesting consequence of Theorem 5.3.4 is that there exists a natural abelian group structure on the set of equivalent classes of non-negative binary quadratic forms with discriminant d. The multiplication law of two equivalence classes in this group is usually called "Gauss composition law", which was discovered by Gauss around 1800. It is quite remarkable that, at the time of Gauss, the ideal class group for a general number

field had not been defined yet. For a different approach to Gauss composition law using cubes, see Bhargava's paper [**Bh04**].

CHAPTER 6

DISTRIBUTION OF IDEALS AND DEDEKIND ZETA FUNCTIONS

6.1. Distribution of ideals in a number field

Let K be a number field of degree n, and C be an ideal class of K. Given a positive real number t, we denote by $N_C(t)$ the number of ideals I of $\mathcal{O}_K$ in the given ideal class C with norm $\mathrm{N}(I) \leq t$. The aim of this section is to prove the following

Theorem 6.1.1. — *There exists a positive number κ, which depends on K but is independent of C, such that*

$$N_C(t) = \kappa t + O(t^{1-1/n}).$$

Here, the error term $O(t^{1-1/n})$ means that, there exists a real positive number A, depending on K and C but independent of t, such that $|N_C(t) - \kappa t| \leq At^{1-1/n}$ for all $t \geq 1$.

Remark 6.1.2. — We will give later a formula for κ after we define the regulator of K.

We now explain how to prove Theorem 6.1.1. First of all, instead of counting (integral) ideals in the ideal class, we reduce the problem to counting the elements in a fractional ideal.

Lemma 6.1.3. — *Let J be a fractional ideal in the ideal class C^{-1}. Let S_t be the equivalent class of $x \in J$ with norm $|\mathrm{N}_{K/\mathbb{Q}}(x)| \leq t\mathrm{N}(J)$ modulo the action of units of K. Then $\alpha \mapsto (\alpha)J^{-1}$ induces a bijection between S_t with the set of integral ideals I of $\mathcal{O}_K$ in C with $\mathrm{N}(I) \leq t$.*

Proof. — Indeed, the set S_t is in natural bijection with the set of principal ideals (α) contained in J with $|\mathrm{N}_{K/\mathbb{Q}}(\alpha)| \leq t\mathrm{N}(J)$. The multiplication by J^{-1} induces a bijection between such principal ideals with the set of integral integral ideals I contained in C with norms $\mathrm{N}(I) \leq t$. □

In the rest of this section, we fix such a fractional ideal J as in the Lemma.

6.1.4. The case of quadratic fields. — To illustrate the ideas of the proof, let us consider first the case when $K/\mathbb{Q}$ is quadratic.

(1) Assume that $K/\mathbb{Q}$ is imaginary quadratic field. We fix a complex embedding

$$\lambda : K \to \mathbb{C} \cong \mathbb{R}^2.$$

The image of J is a lattice in $\mathbb{C}$. Choose a basis (α_1, α_2) of $\lambda(J)$. Then a fundamental domain of $\mathbb{C}/\lambda(J)$ is given by the parallelogram D with vertex points $\frac{1}{2}(\pm\alpha_1, \pm\alpha_2)$. One has

$$\mu(D) = \mathrm{Vol}(\mathbb{C}/\lambda(J)) = 2^{-1}\sqrt{|\Delta_K|}\mathrm{N}(J).$$

For any $\rho > 0$, denote

$$B_\rho = \{z \in \mathbb{C} : |z| \leq \rho\}.$$

Let $n(t)$ denote the cardinality of $\lambda(J) \cap B_{\sqrt{t\mathrm{N}(J)}}$, $n_-(t)$ denote the number of $\alpha \in J$ such that $\alpha + D$ is contained in $B_{\sqrt{t\mathrm{N}(J)}}$, and $n_+(t)$ the number of $\alpha \in J$ such that $\alpha + D$ has non-empty intersection with $B_{\sqrt{t\mathrm{N}(J)}}$. It is clear that

$$n_-(t) \leq n(t) \leq n_+(t).$$

Let δ denote the maximal length of two elements in D, then one has

$$n_-(t) \geq \frac{\mu(B_{\sqrt{t\mathrm{N}(J)}-\delta})}{\mu(D)}, \quad n_+(t) \leq \frac{\mu(B_{\sqrt{t\mathrm{N}(J)}+\delta})}{\mu(D)}.$$

Therefore, one gets

$$n(t) = \frac{\mu(B_{\sqrt{t\mathrm{N}(J)}})}{\mu(D)} + O(\sqrt{t}) = \frac{2\pi}{\sqrt{|\Delta_K|}}t + O(\sqrt{t}).$$

Modulo the unit group U_K, we get

$$N_C(t) = \#S_t = \frac{2\pi}{w|\Delta_K|}t + O(\sqrt{t}),$$

where w denotes the cardinality of U_K. This finishes the proof of Theorem 6.1.1, with $\kappa = \frac{2\pi}{w|\Delta_K|}$.

(2) Assume now $K/\mathbb{Q}$ is real quadratic. The situation is complicated by the existence of free part of the unit group U_K. By Dirichlet's unit theorem, one has

$$U_K \cong \{\pm 1\} \times \varepsilon^{\mathbb{Z}},$$

where $\varepsilon \in U_K$ is a fundamental unit. Let

$$\lambda : K \to \mathbb{R}^2$$

denote the embedding given by $\alpha \mapsto (\sigma_1(\alpha), \sigma_2(\alpha))$, where σ_1, σ_2 are the two real embeddings of K. We may assume that $\sigma_1(\varepsilon) > 1$. Then $\lambda(J)$ is a lattice in $\mathbb{R}^2$ with $\mathrm{Vol}(\mathbb{R}^2/J) = \sqrt{|\Delta_K|}\mathrm{N}(J)$. We consider the orbit of $\mathbb{R}^2$ under the action of $\lambda(\varepsilon)^{\mathbb{Z}}$. Then for every $y_0 \in \mathbb{R}^2$, there exists $n \in \mathbb{Z}$ such that $y_0\lambda(\varepsilon^n)$ lies in the subset

$$\{y = (y_1, y_2) \in \mathbb{R}^2 : 1 < \frac{|y_1|}{|y_2|} \leq \frac{\sigma_1(\varepsilon)}{|\sigma_2(\varepsilon)|} = \sigma_1(\varepsilon)^2.\}$$

Therefore, if we put

$$D_{t\mathrm{N}(J)} = \{(y_1, y_2) \in \mathbb{R}^2 : |y_1 y_2| \leq t\mathrm{N}(J), 1 < \frac{|y_1|}{|y_2|} \leq \sigma_1(\varepsilon)^2\},$$

then every element in J with norm less than $t\mathrm{N}(J)$ is uniquely, under the action of $\varepsilon^{\mathbb{Z}}$, equivalent to an element in $\lambda(J) \cap D_{t\mathrm{N}(J)}$. Therefore, by the same arguments as in the imaginary quadratic case, we have

$$N_C(t) = \frac{\#\lambda(J) \cap D_{t\mathrm{N}(J)}}{w} = \frac{\mu(D_{t\mathrm{N}(J)})}{w\mathrm{Vol}(\mathbb{R}^2/\lambda(J))} + O(\sqrt{t}),$$

where $w = 2$ is the cardinality of $W_K = \{\pm 1\}$. So the problem is reduced to computing $\mu(D_{t\mathrm{N}(J)})$. We have

$$\begin{aligned}
\mu(D_{t\mathrm{N}(J)}) &= 4 \int_{\substack{y_1, y_2 > 0 \\ 0 < y_1 y_1 \leq t\mathrm{N}(J) \\ 1 < y_1/y_2 \leq \sigma_2(\varepsilon)^2}} dy_1 dy_2 \\
&= 4 \int_{\substack{x_1 + x_2 \leq \log(t\mathrm{N}(J)) \\ 0 < x_1 - x_2 \leq 2\log(\sigma_1(\varepsilon))}} e^{x_1 + x_2} dx_1 dx_2 \quad (\text{letting } y_i = e^{x_i} \text{ for } i = 1, 2) \\
&= 2 \int_{u=-\infty}^{\log(t\mathrm{N}(J))} \int_{v=0}^{2\log(\sigma_1(\varepsilon))} e^u du dv \quad (u = x_1 + x_2 \text{ and } v = x_1 - x_2) \\
&= 4t\mathrm{N}(J)\log(\sigma_1(\varepsilon)).
\end{aligned}$$

So finally, one gets

$$N_C(t) = \frac{2\log(\sigma_1(\varepsilon))}{\sqrt{|\Delta_K|}} t + O(\sqrt{t}),$$

where $\log(\sigma_1(\varepsilon))$ is usually called the *regulator* of K.

6.1.5. A formula for the number of lattice points. — In the discussion above, we have used an estimation for the number of lattice points contained in a bounded region D in $\mathbb{R}^2$ in terms of the area of D. It is reasonable to expect that, if $\Lambda \subseteq \mathbb{R}^n$ is a lattice and $B \subseteq \mathbb{R}^n$ is a bounded region, then $\#(\Lambda \cap B)$ can be estimated in terms of the ratio $\mu(B)/\mathrm{Vol}(\mathbb{R}^n/\Lambda)$, once the boundary of B is "not too bad". In order to put this in a rigorous form, we need the following

Definition 6.1.6. — (1) Let $[0,1]^{n-1}$ denote the $(n-1)$-dimensional unit cube. A function

$$f : [0,1]^{n-1} \to \mathbb{R}^n$$

is called Lipschitz, if the ratio

$$\frac{|f(x) - f(y)|}{|x - y|}$$

is uniformly bounded as x and y range over $[0,1]^{n-1}$, where $|\cdot|$ means the length in $\mathbb{R}^{n-1}$ or $\mathbb{R}^n$.

(2) Let B be a bounded region in $\mathbb{R}^n$. We define the boundary of B as $\partial B = \bar{B} - B^{\mathrm{int}}$, where $\bar{B}$ denotes the closure of B in $\mathbb{R}^n$ and B^{int} the interior of B. We say that ∂B is

$(n-1)$-*Lipschitz parametrizable*, if it is covered by the images of finitely many Lipschitz functions: $f : [0,1]^{n-1} \to \mathbb{R}^n$.

Lemma 6.1.7 **([Ma77] Chap. 6, Lemma 2).** — *Let B be a bounded region in $\mathbb{R}^n$ such that the boundary of B is $(n-1)$-Lipschitz parametrizable, and $\Lambda \subseteq \mathbb{R}^n$ be a full lattice. Then for $a > 1$, we have*

$$\#(\Lambda \cap aB) = \frac{\mu(B)}{\mathrm{Vol}(\mathbb{R}^n/\Lambda)} a^n + O(a^{n-1}).$$

Proof. — Let $L : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation such that $L(\Lambda) = \mathbb{Z}^n$. If B' denotes the image of B, then

$$\mu(B') = \mu(B)|\det(L)| = \frac{\mu(B)}{\mathrm{Vol}(\mathbb{R}^n/\Lambda)}.$$

Clearly, we have $\#(\mathbb{Z}^n \cap aB') = \#(\Lambda \cap aB)$, so the statement for (Λ, B) follows immediately from that for $(\mathbb{Z}^n, B')$. Thus we may assume that $\Lambda = \mathbb{Z}^n$.

Consider translates of n-cubes $[0,1]^n$ with centers at points of $\mathbb{Z}^n$. We call simply such a translate a unit n-cube. The number of unit n-cubes contained in aB is roughly $\mu(aB) = a^n\mu(B)$, and the difference is controlled by the number of unit n-cubes which intersect with the boundary $\partial(aB)$.

Let us call *small* n-cubes the translates of $[1, a^{-1}]^n$ with centers at points of $a^{-1}\mathbb{Z}^n$. The number of unit n-cubes intersecting with $\partial(aB)$ equals to the number of small n-cubes intersecting with ∂B. Let $f : [0,1]^{n-1} \to \partial B$ be a $(n-1)$-Lipschitz function. Since ∂B is covered by the image of finitely many such functions, we just need to show that the number of small n-cubes intersecting with image of f is $O(a^{n-1})$. Let $\lambda > 0$ be such that

$$|f(x) - f(y)| \leq \lambda |x - y|, \quad \text{for all } x, y \in [0,1]^{n-1}.$$

Consider all the points $x \in [0,1]^{n-1}$ whose coordinates are of the form $x_i = \frac{b}{[a]}$ for some integer b with $0 \leq b \leq [a] - 1$. Here $[a]$ denotes the maximal integer less or equal to a. Then there are $[a]^{n-1}$ such points, and we label them as x_i for $1 \leq i \leq [a]^{n-1}$. Now assume that Δ is a small n-cube intersecting with the image of f. Let $y = f(x)$ be an intersection point. Then there exists some x_i as above such that $|x - x_i| \leq \sqrt{n-1}/(2a)$, so that

$$|y - f(x_i)| \leq \lambda\sqrt{n-1}/(2a).$$

As the diameter of Δ is $\sqrt{n}/a$, then there exists $c > 0$, independent of a, such that C is completely contained in the ball D_i with center $f(x_i)$ and radius c/a. Clearly, all the small cubes intersecting with $\mathrm{Im}(f)$ are contained in the union of the D_i's. But the volume of each D_i is of the form c'/a^n for some $c' > 0$ independent of a. Therefore, the volume of the union of all D_i is bounded by $c'/a^n[a]^{n-1} \leq c'/a$, and the number of small n-cubes intersecting with $\mathrm{Im}(f)$ is bounded above by

$$(c'/a)/a^{-n} = c'a^{n-1}.$$

□

6.1.8. Start of the proof of Theorem 6.1.1. — We now turn to the proof of Theorem 6.1.1 in the general case. Let $\sigma_1, \cdots, \sigma_{r_1}$ denote the real embeddings of K, and $\sigma_{r_1+j}, \bar{\sigma}_{r+j}$ for $1 \leq j \leq r_2$ denote the non-real embeddings. We have a Minkowski embedding

$$\lambda : K \hookrightarrow \mathbb{R}^{r_1} \times \mathbb{C}^{r_2} \cong \mathbb{R}^n,$$

under which the image of J is lattice such that (Lemma 4.1.2)

$$\mathrm{Vol}(\mathbb{R}^n/\lambda(J)) = 2^{-r_2}\sqrt{|\Delta_K|}\mathrm{N}(J). \tag{6.1.8.1}$$

Let J be the fixed fractional ideal in C^{-1}. For any real number $t > 0$, put

$$X_t = \{(y, z) \in \mathbb{R}^{r_1} \times \mathbb{C}^{r_2} : \prod_{i=1}^{r_1} |y_i| \prod_{j=1}^{r_2} |z_{r_1+j}|^2 \leq t\mathrm{N}(J)\}.$$

The elements of J with norm less than $t\mathrm{N}(J)$ are exactly those in $\lambda(J) \cap X_t$. Note that $\lambda(J\backslash\{0\})$ lies in $(\mathbb{R}^\times)^{r_1} \times (\mathbb{C}^\times)^{r_2}$. We put

$$X_t^* = X_t \cap (\mathbb{R}^{\times, r_1} \times \mathbb{C}^{\times, r_2}).$$

The unit group U_K acts naturally on X_t^* and J via multiplication, and we need to count the number of orbits of $\lambda(J) \cap X_t^*$ modulo the action of U_K. For this, we need to find a fundamental domain of X_t^* for the action of U_K. Choose a fundamental system of units $u_1, \cdots, u_{r_1+r_2-1}$ so that

$$U_K = W_K \times \prod_{i=1}^{r_1+r_2-1} u_i^{\mathbb{Z}},$$

where W_K is the subgroup of roots of unity in K. Consider the following commutative diagram

$$\begin{array}{ccc} \mathbb{R}^{\times, r_1} \times \mathbb{C}^{\times, r_2} & \xrightarrow{\mathrm{Log}} & \mathbb{R}^{r_1+r_2} \\ {\scriptstyle \lambda|_{U_K}}\uparrow & & \uparrow \\ U_K & \xrightarrow{\ell} & H \end{array}$$

where Log is given by

$$(y, z) \mapsto (\log|y_1|, \cdots, \log|y_{r_1}|, 2\log|z_1|, \cdots, 2\log|z_{r_2}|),$$

and H is the hyperplane of $\mathbb{R}^{r_1+r_2}$ defined by $\sum_{i=1}^{r_1+r_2} x_i = 0$. In the proof of Dirichlet's unit theorem 4.2.2, we have seen that ℓ is a group homomorphism with kernel W_K, and its image is a full lattice in H generated by $\ell(u_1), \cdots, \ell(u_{r_1+r_2-1})$. Note that the image of X_t^* under Log is the region $X_t^{*,\log}$ defined by $\sum_{i=1}^{r_1+r_2} x_i \leq \log(t\mathrm{N}(J))$, and $\ell(U_K)$ acts naturally by translation on $X_t^{*,\log}$. Put

$$\mathbf{n} = \frac{1}{r_1 + r_2}(1, \cdots, 1) \in \mathbb{R}^{r_1+r_2}. \tag{6.1.8.2}$$

Then $(\mathbf{n}, \ell(u_1), \cdots, \ell(u_{r_1+r_2-1}))$ form a basis of $\mathbb{R}^{r_1+r_2}$, and a fundamental domain for $X_t^{*,\log}$ under the action of $\ell(U_K)$ is given by

$$D_t^{\log} := \{t_0\mathbf{n} + t_1\ell(u_1) + \cdots + t_{r_1+r_2-1}\ell(u_{r_1+r_2-1}) : \\ t_0 \in (-\infty, \log(t\mathrm{N}(J))],\ t_i \in [0,1), i = 1, \cdots, r_1+r_2-1\}.$$

Let D_t denote the inverse image of $D_t^{\log}$ under Log. Note that $D_t \subseteq X_t^*$. Then we have the following

Lemma 6.1.9. — *Let V_K denote the subgroup $\prod_{i=1}^{r_1+r_2-1} u_i^{\mathbb{Z}}$ of U_K. Then D_t is a fundamental domain for X_t^* under the action of the subgroup V_K.*

Proof. — Let $\mathbf{y} \in X_t^*$ be a point. Since $D_t^{\log}$ is a fundamental domain for $X_t^{*,\log}$ under the action of $\ell(U_K) = \ell(V_K)$, there exist a unique $u \in V_K$ such that $\mathrm{Log}(\mathbf{y}) - \ell(u) \in D_t^{\log}$. Then $\mathbf{y}' = \mathbf{y}/u \in D_t$ by definition. □

Note that D_t still keeps the action of W_K. It follows immediately from the Lemma above that, the number of orbits of $\lambda(J) \cap X_t^*$ under U_K is the same as the number of orbits of $\lambda(J) \cap D_t$ under the action of W_K. In summary, the number of integral ideals in the ideal class C with norm less or equal to t is given by

$$N_C(t) = \frac{\#(\lambda(J) \cap D_t)}{w}.$$

Note that $D_t = t^{1/n} D_1$ and the boundary of D_1 is clearly $(n-1)$-Lipschitz. Hence by Lemma 6.1.7, we have

$$N_C(t) = \frac{\mu(D_1)}{w\mathrm{Vol}(\mathbb{R}^n/\lambda(J))} t + O(t^{1-1/n}). \tag{6.1.9.1}$$

To compute $\mu(D_1)$, we need to introduce the following

Definition 6.1.10. — Let $(u_1, \cdots, u_{r_1+r_2-1})$ be a fundamental system of the unit group U_K, and $\mathbf{n} \in \mathbb{R}^n$ be the vector defined in (6.1.8.2). We define the *regulator* of K as

$$R_K = |\det(\mathbf{n}, \ell(u_1), \cdots, \ell(u_{r_1+r_2-1}))|.$$

Note that, in the definition of R_K, one can replace $\mathbf{n}$ by any vector in $\mathbb{R}^{r_1+r_2}$ whose components sum up to 1.

Lemma 6.1.11. — *We have*

$$\mu(D_1) = 2^{r_1}\pi^{r_2} R_K \mathrm{N}(J).$$

Proof. — Let dy_i for $1 \le i \le r_1$ (resp. μ_{z_j} for $1 \le j \le r_2$) denote the Lebesgue measure on $\mathbb{R}$ (resp. on $\mathbb{C}$). Using polar coordinates $z_j = \rho_j e^{i\theta_j}$, then we have $\mu_{z_j} = \rho_j d\rho_j d\theta_j$. It follows that

$$\mu(D_1) = \int_{D_1} dy_1 \cdots dy_{r_1} \rho_1 \cdots \rho_{r_2} d\rho_1 \cdots d\rho_{r_2} d\theta_1 \cdots \theta_{r_2}.$$

By definition, if we put $x_i = \log|y_i|$ for $1 \le i \le r_1$ and $x_{r_1+j} = 2\log|z_j| = 2\log\rho_j$, then D_1 is defined by $\mathbf{x} = (x_1, \cdots, x_{r_1+r_2}) \in D_1^{\log}$. Changing the variables to x_i's, one gets

$$\mu(D_1) = 2^{r_1}\pi^{r_2} \int_{D_1^{\log}} e^{\sum_{i=1}^{r_1+r_2} x_i}\, dx_1 \cdots dx_{r_1+r_2}.$$

If $t_0, t_1, \cdots, t_{r_1+r_2-1}$ are new variables defined by

$$(x_1, x_2, \cdots, x_{r_1+r_2})^t = (\mathbf{n}, \ell(u_1), \cdots, \ell(u_{r_1+r_2-1}))(t_0, t_1, \cdots, t_{r_1+r_2-1})^t,$$

where $(x_1, \cdots, x_{r_1+r_2})^t$ means the column vector in $\mathbb{R}^{r_1+r_2}$, then $D_1^{\log}$ is defined by

$$-\infty < t_0 \le \log(\mathrm{N}(J)) \quad \text{and} \quad 0 < t_i \le 1, \quad i = 1, \cdots, r_1 + r_2 - 1.$$

Note also $\sum_{i=1}^{r_1+r_2} x_i = t_0$, since the components of each $\ell(u_i)$ sum up to 0. Then by Jacobian's rule, we have

$$dx_1 \cdots dx_{r_1+r_2} = R_K\, dt_0 dt_1 \cdots dt_{r_1+r_2-1},$$

and hence

$$\begin{aligned}\mu(D_1) &= 2^{r_1}\pi^{r_2} R_K \int_{-\infty}^{\log(\mathrm{N}(J))} e^{t_0} dt_0 \prod_{i=1}^{r_1+r_2-1} \int_0^1 dt_i \\ &= 2^{r_1}\pi^{r_2} R_K \mathrm{N}(J)\end{aligned}$$

□

We can now prove a more precise form of Theorem 6.1.1:

Theorem 6.1.12. — *Let K be a number field of degree n with r_1 real embeddings and r_2 non-real embeddings. For a fixed ideal class C of K and a real number $t > 0$, let $N_C(t)$ denote the number of integral ideals of K with norms less than t. Then we have*

$$N_C(t) = \frac{2^{r_1}(2\pi)^{r_2} R_K}{w\sqrt{|\Delta_K|}} t + O(t^{1-1/n}),$$

where w is the number of roots of unity in K and Δ_K is the discriminant of K.

Proof. — Indeed, this follows immediately from (6.1.9.1), Lemma 6.1.11 and (6.1.8.1). □

We have the following immediate

Corollary 6.1.13. — *Under the notation of the Theorem, let h denote the class number of K. Then the number of integral ideals of K with norm less than t is given by*

$$N(t) = \frac{2^{r_1}(2\pi)^{r_2} R_K h}{w\sqrt{|\Delta_K|}} t + O(t^{1-1/n}).$$

6.2. Residue formula of Dedekind Zeta functions

Recall that the Riemann zeta function is defined by

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

which absolutely converges for $\Re(s) > 1$.

Lemma 6.2.1. — *The function $\zeta(s)$ can be analytically extended to a meromorphic function in s on $\Re(s) > 0$ with a simple pole at $s = 1$ with residue 1.*

Proof. — Note that

$$\int_1^{\infty} t^{-s} dt = \frac{1}{s-1} \quad \text{for } \Re(s) > 1.$$

Then we have

$$\begin{aligned}
\zeta(s) - \int_1^{\infty} t^{-s} dt &= \sum_{n=1}^{\infty} \int_n^{n+1} (\frac{1}{n^s} - \frac{1}{t^s}) dt \\
&= \sum_{n=1}^{\infty} \int_n^{n+1} s \int_{x=n}^{t} \frac{1}{t^{s+1}} dx dt \\
&= \sum_{n=1}^{\infty} s \int_n^{n+1} \frac{n+1-x}{x^{s+1}} dx.
\end{aligned}$$

If σ denotes the real part of s, then

$$|\sum_{n=1}^{\infty} s \int_n^{n+1} \frac{n+1-x}{x^{s+1}} dx| \leq |s| \int_1^{\infty} \frac{dx}{x^{\sigma+1}} = \frac{|s|}{\sigma}, \quad \text{for } \sigma > 0.$$

Thus the sum in $|\cdot|$ on the left hand side defines a holomorphic function in $\Re(s) > 0$. Thus, one may define the analytic continuation of $\zeta(s)$ as

$$\zeta(s) = \frac{1}{s-1} + \sum_{n=1}^{\infty} s \int_n^{n+1} \frac{n+1-x}{x^s} dx, \quad \text{for } \Re(s) > 0. \tag{6.2.1.1}$$

□

Proposition 6.2.2. — *Let $f(s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s}$ be a Dirichlet series, which converges for $\Re(s)$ sufficiently large. Let $S_t = \sum_{n \leq t} a_n$ for any $t > 0$. Assume that there exists some $\kappa \in \mathbb{C}$ and δ with $0 < \delta \leq 1$ such that*

$$S_t = \kappa t + O(t^{1-\delta}) \quad \textit{when } t \to +\infty.$$

Then $f(s)$ can be analytically extended to a meromorphic function on $\Re(s) > 1 - \delta$ with at most a simple pole at $s = 1$ with residue κ.

Proof. — Put

$$g(s) = f(s) - \kappa\zeta(s) = \sum_{n=1}^{\infty} \frac{b_n}{n^s}, \quad \text{with } b_n = a_n - \kappa.$$

Then if $S'_t = \sum_{n\leq t} b_n$, then $S'_t = O(t^{1-\delta})$. By Lemma 6.2.1, $\kappa\zeta(s)$ can be analytically extended to a meromorphic function on $\Re(s) > 0$ with a simple pole at $s = 1$ with residue κ. Therefore, the Proposition will be proved if one shows that $g(s)$ has an analytic continuation to a holomorphic function on $\Re(s) > 1 - \delta$. For $\Re(s) >> 0$, we have

$$\begin{aligned} g(s) &= \sum_{n=1}^{\infty} \frac{S'_n - S'_{n-1}}{n^s} = \sum_{n=1}^{\infty} \frac{S'_n}{n^s} - \sum_{n=1}^{+\infty} \frac{S'_n}{(n+1)^s} \\ &= \sum_{n=1}^{+\infty} S'_n \left(\frac{1}{n^s} - \frac{1}{(n+1)^s}\right) \\ &= \sum_{n=1}^{+\infty} S'_n \int_n^{n+1} st^{-s-1} dt. \end{aligned}$$

Let $C > 0$ be such that $|S'_n| \leq Cn^{1-\delta}$, and $\sigma = \Re(s)$. Then we have

$$\begin{aligned} \left|\sum_{n=1}^{+\infty} S'_n \int_n^{n+1} st^{-s-1} dt\right| &\leq C|s| \sum_{n=1}^{\infty} n^{1-\delta} \int_n^{n+1} t^{-\sigma-1} dt \\ &= C|s| \sum_{n=1}^{+\infty} \int_n^{n+1} t^{-\sigma-\delta} dt \quad (\text{since } n^{1-\delta} \leq t^{1-\delta}) \\ &= C|s| \frac{1}{\sigma + \delta - 1} \quad \text{for } \sigma > 1 - \delta. \end{aligned}$$

This shows that $g(s)$ can be extended to a holomorphic function on $\Re(s) > 1 - \delta$.

□

Now let K be a number field of degree n.

Lemma 6.2.3. — *The series*

$$\zeta_K(s) := \sum_{\mathfrak{a}\subset\mathcal{O}_K} \frac{1}{\mathrm{N}(\mathfrak{a})^s}$$

absolutely converges for $\Re(s) > 1$, where $\mathfrak{a}$ runs through all the integral ideals of $\mathcal{O}_K$. Moreover, we have the Euler product

$$\zeta_K(s) = \prod_{\mathfrak{p}} \frac{1}{1 - \mathrm{N}(\mathfrak{p})^{-s}} \quad \textit{for } \Re(s) > 1, \tag{6.2.3.1}$$

where $\mathfrak{p}$ runs through all non-zero prime ideals of $\mathcal{O}_K$.

Proof. — Write $\zeta_K(s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s}$, where a_n is the number of integral ideals of $\mathcal{O}_K$ with norm equal to n. If $n = \prod_{i=1}^{r} p_i^{e_i}$ is the prime factorization of n, then $a_n \leq [K : \mathbb{Q}] \sum_i e_i \leq [K : \mathbb{Q}] \log_2 n$, since there are at most $[K : \mathbb{Q}]e_i$ primes with norm equal to p^{e_i}. In

particular, one has $a_n = O(n^\epsilon)$ for any positive $\epsilon > 0$. Therefore, for any $\epsilon > 0$, there exists a $C_\epsilon > 0$ such that

$$|\zeta_K(s)| \leq \sum_{n=1}^{+\infty} \frac{a_n}{|n^s|} \leq C_\epsilon \sum_{n=1}^{\infty} \frac{1}{n^{\Re(s)-\epsilon}},$$

which uniformly absolutely converges for $\Re(s) \geq 1 + 2\epsilon$. Next, we note that

$$\begin{aligned} |\prod_{\mathfrak{p}} \frac{1}{1 - \mathrm{N}(\mathfrak{p})^{-s}}| &\leq \prod_{\mathfrak{p}} \frac{1}{1 - \mathrm{N}(\mathfrak{p})^{-\Re(s)}} \\ &\leq \prod_{p} \frac{1}{(1 - p^{-\Re(s)})^{[K:\mathbb{Q}]}} = \zeta(\Re(s))^{[K:\mathbb{Q}]}, \end{aligned}$$

which proves the absolute convergence of the Euler product. Then the equality (6.2.3.1) is an immediate consequence of the unique factorization law of ideals in $\mathcal{O}_K$. □

The complex analytic function $\zeta_K(s)$ is usually called the *Dedekind zeta function* of K.

Theorem 6.2.4. — *The function $\zeta_K(s)$ has a meromorphic continuation to $\Re(s) > 1 - \frac{1}{n}$ with a simple pole at $s = 1$ with residue*

$$\kappa = \frac{2^{r_1}(2\pi)^{r_2} R_K h}{w\sqrt{|\Delta_K|}},$$

where the meanings of $r_1, r_2, R_K, h, w, \Delta_K$ are the same as in Corollary 6.1.13.

Proof. — Write $\zeta_K(s) = \sum_{n=1}^{+\infty} a_n n^{-s}$, where a_n is the number of integral ideals of $\mathcal{O}_K$ with norm exactly equal to n. Then $S_t = \sum_{n \leq t} a_n$ equals to the number of integral ideals of $\mathcal{O}_K$ with norm less than or equal to t. Now the Theorem follows immediately from Proposition 6.2.2 and Corollary 6.1.13. □

Let $f(s)$ and $g(s)$ be complex valued functions in a neighborhood of 1. We write $f(s) \sim g(s)$ when $s \to 1^+$, if the limit of $\frac{f(s)}{g(s)}$ when s approaches 1 along the real axis from the right.

Corollary 6.2.5. — *Let K be a number field. Then we have*

$$\sum_{\mathfrak{p}} \frac{1}{\mathrm{N}(\mathfrak{p})^s} \sim \sum_{\mathfrak{p}, \deg(\mathfrak{p})=1} \frac{1}{\mathrm{N}(\mathfrak{p})^s} \sim \log \frac{1}{s-1} \quad \textit{when } s \to 1^+,$$

where $\mathfrak{p}$ runs through all the prime ideals of $\mathcal{O}_K$ in the first summation, and through the primes of degree 1 in the second.

Proof. — By the Euler product of $\zeta_K(s)$, we have

$$\begin{aligned} \log \zeta_K(s) &= \sum_{\mathfrak{p}} -\log(1 - \mathrm{N}(\mathfrak{p})^{-s}) = \sum_{\mathfrak{p}} \sum_{n=1}^{+\infty} \frac{1}{n\mathrm{N}(\mathfrak{p})^{ns}} \\ &= \sum_{\mathfrak{p}, \deg(\mathfrak{p})=1} \frac{1}{\mathrm{N}(\mathfrak{p})^s} + \sum_{\mathfrak{p}, \deg(\mathfrak{p}) \geq 2} \frac{1}{\mathrm{N}(\mathfrak{p})^s} + \sum_{n \geq 2} \frac{1}{n\mathrm{N}(\mathfrak{p})^{ns}}, \quad \text{for } \Re(s) > 1. \end{aligned}$$

By Theorem 6.2.4, we have $\log \zeta_K(s) \sim \log \frac{1}{s-1}$ when $s \to 1$. Therefore, to finish the proof, it suffices to show that the second and the third terms are bounded when $s \to 1^+$. Let $\sigma = \Re(s)$. Then

$$|\sum_{\mathfrak{p},\deg(\mathfrak{p})\geq 2} \frac{1}{\mathrm{N}(\mathfrak{p})^s}| \leq \sum_p \frac{[K:\mathbb{Q}]}{p^{2\sigma}} \leq [K:\mathbb{Q}] \sum_{n\geq 1} \frac{1}{n^{2\sigma}},$$

which is convergent for $\sigma > 1/2$, hence bounded when $s \to 1^+$. Similarly, one has

$$\begin{aligned} |\sum_{\mathfrak{p},n\geq 2} \frac{1}{n\mathrm{N}(\mathfrak{p})^{ns}}| &\leq + \sum_{\mathfrak{p}} \frac{1}{\mathrm{N}(\mathfrak{p})^\sigma(\mathrm{N}(\mathfrak{p})^\sigma - 1)} \\ &\leq 2[K:\mathbb{Q}] \sum_p \frac{1}{p^\sigma(p^\sigma - 1)} \\ &\leq [K:\mathbb{Q}] \sum_{n\geq 2} \frac{1}{n^\sigma(n^\sigma - 1)} \end{aligned}$$

which is convergent when $\sigma > 1/2$. □

CHAPTER 7

DIRICHLET L-FUNCTIONS AND ARITHMETIC APPLICATIONS

A good reference for this chapter is [**Wa96**, Chap. 3, 4].

7.1. Dirichlet characters

Let G be a finite abelian group. Recall that a *character* of G is a group homomorphism

$$\chi : G \to \mathbb{C}^\times.$$

We say χ is trivial, if $\chi(g) = 1$ for all $g \in G$. If χ_1 and χ_2 are two characters, we define their product by the formula $\chi_1\chi_2(g) = \chi_1(g)\chi_2(g)$. The set of characters of G form an abelian group, which we denote by $\widehat{G}$.

Lemma 7.1.1. — *There exists a non-canonical isomorphism $G \cong \widehat{G}$.*

Proof. — Since every finite abelian group is a direct sum of cyclic groups, we may assume that $G \cong \mathbb{Z}/n\mathbb{Z}$. Then a character χ of G is determined by its value at $1 \in \mathbb{Z}/n\mathbb{Z}$, which is necessarily an n-th root of unity, and vice versa. Thus $\widehat{G}$ is canonically isomorphic to the group of n-th roots of unity, which is isomorphic to $\mathbb{Z}/n\mathbb{Z}$. □

A group homomorphism $f : G_1 \to G_2$ induces a natural map $\hat{f} : \widehat{G}_2 \to \widehat{G}_1$ given by $\chi \mapsto \chi \circ f$.

Corollary 7.1.2. — *If $0 \to G_1 \to G \to G_2 \to 0$ is an exact sequence of finite abelian groups, then the induced sequence $0 \to \widehat{G}_2 \to \widehat{G} \to \widehat{G}_1 \to 0$ is also exact.*

Proof. — Let f denote the injection $G_1 \to G$. It is easy to see that $\mathrm{Ker}(\hat{f}) = \widehat{G}_2$, so that only the surjectivity of $\widehat{f} : \widehat{G} \to \widehat{G}_1$ is non-trivial. By the Lemma, $\widehat{G}/\widehat{G}_2$ has the same cardinality as $\widehat{G}_1$. It follows that $\widehat{f}$ induces an isomorphism $\widehat{G}/\widehat{G}_2 \cong \widehat{G}_1$, that is $\widehat{f}$ is surjective.

□

For any finite abelian group G, there is a natural morphism $G \to \widehat{\widehat{G}}$ sending $g \in G$ to the character $\chi \mapsto \chi(g)$ on $\widehat{G}$.

Proposition 7.1.3. — *The canonical morphism $G \to \widehat{\widehat{G}}$ is an isomorphism.*

Proof. — Since G and $\widehat{\widehat{G}}$ have the same cardinality. It suffices to show that $G \to \widehat{\widehat{G}}$ is injective, that is, for any non-trivial $g \in G$, we have to construct a $\chi \in \widehat{G}$ such that $\chi(g) \neq 1$. Let $H \subseteq G$ be the subgroup generated by g. Then $\widehat{H} \neq 1$. By Corollary 7.1.2, there exists $\chi \in \widehat{G}$ with non-trivial image in H. Then $\chi(g) \neq 1$ since g is a generator of H.

□

Proposition 7.1.4. — *Let G be a finite abelian group. We have*

1. $\sum_{g\in G} \chi(g) = 0$ *for all non-trivial* $\chi \in \widehat{G}$,
2. $\sum_{\chi\in \hat{G}} \chi(g) = 0$ *for all* $g \neq 1$ *in* G.

Proof. — We just prove statement 1, the second follows from the first by Proposition 7.1.3. Since $\chi \in G$ is non-trivial, there exists $h \in G$ such that $\chi(h) \neq 1$. Then

$$\chi(h) \sum_{g\in G} \chi(g) = \sum_{g\in G} \chi(gh) = \sum_{g\in G} \chi(g).$$

□

7.1.5. Dirichlet characters. — A Dirichlet character is a character χ of the group $(\mathbb{Z}/N\mathbb{Z})^\times$ for some integer $N \geq 1$. Note that for $N|M$, χ induces a character of $(\mathbb{Z}/M\mathbb{Z})^\times$ by the natural surjection $(\mathbb{Z}/M\mathbb{Z})^\times \to (\mathbb{Z}/N\mathbb{Z})^\times$. We say $\chi : (\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{C}^\times$ is primitive, if it is not induced by any characters of $(\mathbb{Z}/d\mathbb{Z})^\times$ for $d|N$ and $d \neq N$; in that case, we say χ has conductor N, and write $f_\chi = N$. We say χ is even if $\chi(-1) = 1$ and odd if $\chi(-1) = -1$.

Many times, it is convenient to regard a Dirichlet character $\chi : (\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{C}^\times$ as a function $\mathbb{Z} \to \mathbb{C}$ by setting $\chi(a) = 0$ if $\gcd(a, f_\chi) \neq 1$. Note that $\chi(nm) = \chi(n)\chi(m)$ for all $n, m \in \mathbb{Z}$.

Example 7.1.6. — (1) Let $\chi : (\mathbb{Z}/8\mathbb{Z})^\times \to \mathbb{C}^\times$ be defined by $\chi(1) = 1$, $\chi(3) = -1$, $\chi(5) = 1$ and $\chi(7) = -1$. Then it is clear that $\chi(a + 4) = \chi(a)$, and thus $f_\chi = 4$.

(2) Let p be an odd prime. Then Legendre symbol $a \mapsto \left(\frac{a}{p}\right)$ defines a Dirichlet character of conductor p.

Let χ and ψ be two Dirichilet characters with conductors f_χ and f_ψ. Consider their the homomorphism

$$\gamma : (\mathbb{Z}/\mathrm{lcm}(f_\chi, f_\psi)\mathbb{Z})^\times \to \mathbb{C}^\times.$$

We define $\chi\psi$ to be the the primitive character associated to γ. Note that, in general, the conductor of $\chi\psi$ is smaller than $\mathrm{lcm}(f_\chi, f_\psi)$.

Example 7.1.7. — Define χ modulo 12 by $\chi(1) = 1$, $\chi(5) = -1$, $\chi(7) = -1$, $\chi(11) = 1$ and define ψ modulo 3 by $\psi(1)$ and $\psi(2) = -1$. Then $\chi\psi$ on $(\mathbb{Z}/12\mathbb{Z})^\times$ has values $\chi\psi(1) = 1$, $\chi\psi(5) = \chi(5)\psi(2) = 1$, $\chi\psi(7) = \chi(7)\psi(1) = -1$ and $\chi\psi(11) = \chi(11)\psi(11) = -1$. One sees easily that $\chi\psi$ has conductor 4, and $\chi\psi(1) = 1$ and $\chi\psi(3) = -1$. Note that $\chi\psi(3) = -1 \neq \chi(3)\psi(3) = 0$.

Let χ be a Dirichlet character modulo N. We put

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

Proposition 7.1.8. — 1. *The series $L(\chi, s)$ absolutely converges in $\Re(s) > 1$;*
2. *We have the Euler product:*

$$L(\chi, s) = \prod_p \frac{1}{1 - \chi(p)p^{-s}}, \quad \textit{for } \Re(s) > 1.$$

3. *If χ is non-trivial, then $L(\chi, s)$ has an analytic continuation into a holomorphic function in $\Re(s) > 0$.*

Proof. — Statement 1 follows from the fact that $|\sum_n \frac{\chi(n)}{n^s}| \leq \zeta(\sigma)$ with $\sigma = \Re(s)$. Statement 2 follows from the fact that $\chi(mn) = \chi(m)\chi(n)$. For statement 3, it follows easily from Proposition 7.1.4(1) that $S_t := \sum_{n \leq t} \chi(n) = O(1)$ when $t \to +\infty$. Then we conclude by Proposition 6.2.2. □

7.2. Factorization of Dedekind zeta functions of abelian number fields

We fix an integer $N \geq 3$. Consider the N-th cyclomotic field $\mathbb{Q}(\zeta_N)$. We recall that

$$\mathcal{G} = \mathrm{Gal}(\mathbb{Q}(\zeta_N)/\mathbb{Q}) \xrightarrow{\sim} (\mathbb{Z}/N\mathbb{Z})^\times$$

such that for $a \in (\mathbb{Z}/N\mathbb{Z})^\times$, the corresponding $\sigma_a \in \mathrm{Gal}(\mathbb{Q}(\zeta_N)/\mathbb{Q})$ is defined by $\sigma_a(\zeta_N) = \zeta_N^a$. Therefore, we can identify the set of Dirichlet characters modulo N with $\widehat{\mathcal{G}}$.

Let K be a subfield of $\mathbb{Q}(\zeta_N)$, denote $H = \mathrm{Gal}(\mathbb{Q}(\zeta_N)/K)$ and $G = \mathrm{Gal}(K/\mathbb{Q}) = \mathcal{G}/H$. By Corollary 7.1.2, there is an exact sequence

$$0 \to \widehat{G} \to \widehat{\mathcal{G}} \to \widehat{H} \to 0.$$

Thus $\widehat{G}$ is identified with the set of Dirichlet characters modulo N which are trivial on H.

Example 7.2.1. — Let p be an odd prime, and $p^* = (-1)^{\frac{p-1}{2}}p$. Then $K = \mathbb{Q}(\sqrt{p^*})$ is the unique quadratic field contained in $\mathbb{Q}(\zeta_p)$. The Galois group $G = \mathrm{Gal}(K/\mathbb{Q})$ is isomorphic to $\mathbb{Z}/2\mathbb{Z}$, and the non-trivial element of $\widehat{G}$ is the Dirichlet character with conductor p given by $\chi(a) = \left(\frac{a}{p}\right)$ for $a \in (\mathbb{Z}/p\mathbb{Z})^\times$. Thus, for an odd prime q, $\chi(q) = 1$ if and only if q splits in K.

Proposition 7.2.2. — *Under the above notation, we have*

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

Remark 7.2.3. — The assumption that K is contained some cyclotomic field is equivalent to saying that $K/\mathbb{Q}$ is a Galois extension with abelian Galois group. Actually, it is famous Kronecker-Weber Theorem. We refer the reader to [**Wa96**, Chap. 14] for a complete proof using class field theory. In Section 7.5, we will give an elementary proof when K is a quadratic field.

Now we turn to the proof of Proposition 7.2.2.

Proof. — In view of the Euler products for $\zeta_K(s)$ and $L(\chi, s)$ (cf. Lemma 6.2.3 and Proposition 7.1.8), it suffices to prove that, for every rational prime p, one has

$$\prod_{\mathfrak{p}|p}(1-\mathrm{N}(\mathfrak{p})^{-s}) = \prod_{\chi\in\widehat{G}}(1-\chi(p)p^{-s}), \tag{7.2.3.1}$$

where $\mathfrak{p}$ runs through the primes of K above p. Recall that K is a subfield of $\mathbb{Q}(\zeta_N)$ with subgroup $H = \mathrm{Gal}(\mathbb{Q}(\zeta_N)/K) \subseteq \mathrm{Gal}(\mathbb{Q}(\zeta_N)/\mathbb{Q})$. We distinguish two cases:

– Consider first the case $p \nmid N$. Then p is unramified in $\mathbb{Q}(\zeta_N)$ by Proposition 3.5.1, and hence in K. Denote by $\sigma_p \in G$ the Frobenius element of p. If we regard G as a quotient of $\mathrm{Gal}(\mathbb{Q}(\zeta_N)/\mathbb{Q}) \cong (\mathbb{Z}/N\mathbb{Z})^\times$, then σ_p is given by the image of $p \mod N$. Let $D_p = \langle\sigma_p\rangle \subseteq G$ denote the decomposition group at p, $f = \#D_p$ be the order of σ_p so that $D_p \cong \mathbb{Z}/f\mathbb{Z}$, and put $g = \#G/\#D_p$. Then p splits into g primes in $\mathcal{O}_K$ and each of them has residue degree f. Hence, one has

$$\prod_{\mathfrak{p}|p}(1-\mathrm{N}(\mathfrak{p})^{-s}) = (1-p^{-fs})^g.$$

On the other hand, recall that $\widehat{D}_p \xrightarrow{\sim} \mu_f$, where μ_f is the group of f-th roots of unity, and the isomorphism is given by sending a character ψ to its value at σ_p. By Corollary 7.1.2, each character of D_p lifts to exactly g characters of G. Hence, when χ runs through $\hat{G}$, $\chi(p)$ will take every f-th root of unity exactly g times. One get thus

$$\prod_{\chi\in\widehat{G}}(1-\chi(p)p^{-s}) = \prod_{\xi\in\mu_f}(1-\xi p^{-s})^g = (1-p^{-fs})^g,$$

and (7.2.3.1) is proved. Here, the last step used the equality

$$\prod_{\xi\in\mu_f}(X-\xi a) = X^f - a^f,$$

for the variables X and a.

– Assume now $p|N$. Write $N = p^k m$ with $\gcd(p, m) = 1$. Then $\mathbb{Q}(\zeta_N)$ is the composite of $\mathbb{Q}(\zeta_{p^k})$ and $\mathbb{Q}(\zeta_m)$. Since p is totally ramified in $\mathbb{Q}(\zeta_{p^k})$ and unramified in $\mathbb{Q}(\zeta_m)$,

$$\mathrm{Gal}(\mathbb{Q}(\zeta_N)/\mathbb{Q}(\zeta_m)) \cong \mathrm{Gal}(\mathbb{Q}(\zeta_{p^k})/\mathbb{Q}) \cong (\mathbb{Z}/p^k\mathbb{Z})^\times$$

is the inertia subgroup of $\mathrm{Gal}(\mathbb{Q}(\zeta_N)/\mathbb{Q})$ at p. Let I_p denote the image of $\mathrm{Gal}(\mathbb{Q}(\zeta_N)/\mathbb{Q}(\zeta_m))$ in G. Then I_p is the inertia subgroup of G at p, and $K_0 = K \cap \mathbb{Q}(\zeta_m) = K^{I_p}$ is the maximal sub-extension of K in which p is unramified. Assume that

$$p\mathcal{O}_{K_0} = \mathfrak{p}_{0,1}\cdots\mathfrak{p}_{0,g},$$

where each $\mathfrak{p}_i$ has residue degree f over p. Then, for each $\mathfrak{p}_{0,i}$, there exists a unique prime $\mathfrak{p}_i$ such that $\mathfrak{p}_{0,i}\mathcal{O}_K = \mathfrak{p}_i^e$ and $\mathrm{N}(\mathfrak{p}_i) = \mathrm{N}(\mathfrak{p}_{0,i}) = p^f$, where $e = [K : K_0]$.

Therefore,

$$\prod_{\substack{\mathfrak{p}\subseteq\mathcal{O}_K\\ \mathfrak{p}|p}}(1-\mathrm{N}(\mathfrak{p})^{-s})=\prod_{\substack{\mathfrak{p}_0\subseteq\mathcal{O}_{K_0}\\ \mathfrak{p}_0|p}}(1-\mathrm{N}(\mathfrak{p}_0)^{-s}).$$

Put $G_0=G/I_p=\mathrm{Gal}(K_0/\mathbb{Q})$. Then $\widehat{G}_0$ is identified with the subgroup of $\chi\in\widehat{G}$ that factorizes through G_0, and a character $\chi\in\widehat{G}$ if and only if it is induced from a Dirichlet character of $(\mathbb{Z}/m\mathbb{Z})^\times$. Hence, for $\chi\in\widehat{G}$, we have $\chi\in\widehat{G}_0$ if and only if $\chi(p)\neq 0$. It follows that

$$\prod_{\chi\in\widehat{G}}(1-\chi(p)p^{-s})=\prod_{\chi\in\widehat{G}_0}(1-\chi(p)p^{-s}).$$

Now the equality (7.2.3.1) follows immediately from the previous case with K replaced by K_0 and $\mathbb{Q}(\zeta_N)$ replaced by $\mathbb{Q}(\zeta_m)$.

□

One deduces easily from Proposition 7.2.2 the following important:

Theorem 7.2.4. — *Let K and G be as above, and $\chi_0\in\widehat{G}$ be the trivial character. Then*

$$\prod_{\substack{\chi\in\widehat{G}\\ \chi\neq\chi_0}}L(\chi,1)=\frac{2^{r_1}(2\pi)^{r_2}R_Kh}{w\sqrt{|\Delta_K|}},$$

where r_1 and r_2 denote respectively the number of real embeddings and non-real complex embeddings of K, R_K the regulator of K, h the class number, w the number of roots of unity in K, and Δ_K the discriminant of K. In particular, $L(\chi,1)\neq 0$ if $\chi\neq\chi_0$.

Proof. — By Proposition 7.2.2, we have

$$\lim_{s\to 1}(s-1)\zeta_K(s)=\lim_{s\to 1}\Big((s-1)\zeta(s)\Big)\prod_{\substack{\chi\in\widehat{G}\\ \chi\neq\chi_0}}L(\chi,1).$$

Since both $\zeta_K(s)$ and $\zeta(s)$ have a simple zero at $s=1$, the Theorem follows immediately from Theorem 6.2.4. □

7.3. Density of primes in arithmetic progressions

We now deduce from Theorem 7.2.4 Dirichlet's famous theorem on primes in arithmetic progressions. We have seen in Corollary 6.2.5 that

$$\sum_p\frac{1}{p^s}\sim\log\frac{1}{s-1},\quad\text{when } s\to 1.$$

For a subset T of rational primes, if there exists a real $\rho\in[0,1]$ such that

$$\sum_{p\in T}\frac{1}{p^s}\sim\rho\log\frac{1}{s-1}\quad\text{when } s\to 1,$$

we say that T has Dirichlet density ρ.

***Theorem 7.3.1* (Dirichlet, 1837).** — *Let $N \geq 1$ and a be integers with $\gcd(a, N) = 1$. Then the subset of primes p with $p \equiv a \mod N$ has Dirichlet density $\frac{1}{\varphi(N)}$, where $\varphi(N)$ denotes Euler function. In particular, there are infinitely many primes p with $p \equiv a \mod N$.*

Proof. — For a Dirichlet character χ, the Euler product for $L(\chi, s)$ implies that

$$\log L(\chi, s) = \sum_p \sum_{m=1}^{+\infty} \frac{\chi(p^m)}{mp^{ms}}, \quad \text{for } \Re(s) > 1.$$

Taking sums, one gets

$$\sum_\chi \chi(a^{-1}) \log L(\chi, s) = \sum_\chi \sum_p \sum_{m \geq 1} \frac{\chi(p^m a^{-1})}{mp^{ms}} \quad \text{for } \Re(s) > 1,$$

where χ runs through all the Dirichlet characters $(\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{C}^\times$. By the orthogonality of characters (Proposition 7.1.4), we have

$$\sum_\chi \chi(p^m a^{-1}) = \begin{cases} \varphi(N) & \text{if } p^m \equiv a \mod N, \\ 0 & \text{otherwise.} \end{cases}$$

Therefore, one has

$$\sum_\chi \chi(a^{-1}) \log L(\chi, s) = \varphi(N) \sum_{p \equiv a \mod N} \frac{1}{p^s} + \varphi(N) \sum_{m \geq 2} \left(\sum_{p^m \equiv a \mod N} \frac{1}{mp^{ms}} \right).$$

Now if χ is trivial, then $\log L(\chi, s) \sim \log \frac{1}{s-1}$ when $s \to 1$; if χ is non-trivial, then the non-vanishing of $L(\chi, 1)$ by Theorem 7.2.4 implies that $\log L(\chi, s)$ is bounded in a neighborhood of $s = 1$. On the other hand, the term of summation on $m \geq 2$ is bounded for $\Re(s) > 1/2$ as in the proof of Corollary 6.2.5. It follows immediately that

$$\sum_{p \equiv a \mod N} \frac{1}{p^s} \sim \frac{1}{\varphi(N)} \log \frac{1}{s-1}, \quad \text{when } s \to 1.$$

□

Remark 7.3.2. — (1) We can also define the natural density for a subset T of all prime numbers as follows. Let $\pi(x)$ denotes the number of primes less than x for any $x > 0$, and $\pi_T(x)$ be the number of primes numbers in T less than x. Then the natural density of T is defined to be the limit

$$\lim_{x \to +\infty} \frac{\pi_T(x)}{\pi(x)}$$

whenever it exists. In general, a subset T has a natural density $\rho \in [0, 1]$, then its Dirichlet density must exist and equal to ρ. Conversely, if T has Dirichlet density ρ, then it is possible that the natural density of T does not exist at all.

However, for the subset of primes p with $p \equiv a \mod N$ for $\gcd(a,N)=1$, this "pathology" does not occur. Actually, the proof of Dirichlet's Theorem uses essentially the non-vanishing of $L(\chi,1)$ for any non-trivial Dirichlet character χ. It is true that $L(\chi,s)$ does not vanishes on the whole line $\Re(s)=1$. Using this fact (together with Tauberian Theorem), one can prove that the natural density of primes $p \equiv a \mod N$ exists and equals to $\frac{1}{\varphi(N)}$. We refer the reader to [**La94**, Chap. XV] for a proof.

(2) It is easy to generalize the notion of Dirichlet density and natural density to subsets of an arbitrary number field. The generalization of Dirichlet's Theorem to this case is called Chebotarev density theorem:

Theorem 7.3.3 **([La94] Chap. VIII, Theorem 10).** — *Let L/K be a finite Galois extension of number fields with Galois group G. Let $\sigma \in G$ and c be the conjugacy class of σ in G. Then the subset of primes $\mathfrak{p}$ of K which are unramified in L and for which there exists $\mathfrak{P}|\mathfrak{p}$ such that*

$$\sigma = \left(\frac{L/K}{\mathfrak{P}}\right)$$

has a density, and this density equals to $\rho = \frac{|c|}{|G|}$. Here, for a finite set S, $|S|$ denotes its cardinality.

It is clear that when L/K is $\mathbb{Q}(\zeta_N)/\mathbb{Q}$ and $\sigma = \sigma_a$ with $a \in (\mathbb{Z}/N\mathbb{Z})^\times$, Chebatarev density theorem is equivalent to Dirichlet's theorem.

7.4. Values of $L(\chi,1)$ and class number formula

7.4.1. Gauss sums. — Let χ be a Dirichlet character of conductor $f \geq 3$, and $\bar{\chi}$ denote the complex conjugate of χ, i.e. $\bar{\chi}(x) = \overline{\chi(x)}$. Note that $\bar{\chi}(x) = \chi(x^{-1})$ for $x \in (\mathbb{Z}/f\mathbb{Z})^\times$. Put $\zeta_f = e^{\frac{2\pi i}{f}}$, and we write $\zeta = \zeta_f$ when there is no ambiguity. We define the Gauss sum associated to χ to be

$$\tau(\chi) = \sum_{x \in \mathbb{Z}/f\mathbb{Z}} \chi(x)\zeta^x$$

and put $\tau(\chi) = \tau_1(\chi)$. More generally, for any $a \in \mathbb{Z}/N\mathbb{Z}$, we put

$$\tau_a(\chi) = \sum_{x \in \mathbb{Z}/f\mathbb{Z}} \chi(x)\zeta^{ax}.$$

Lemma 7.4.2. — *The following statements hold:*

1. *$\tau_a(\tau) = \bar{\chi}(a)\tau(\chi)$ for any $a \in (\mathbb{Z}/f\mathbb{Z})$. In particular, $\tau_a(\chi) = 0$ if $\gcd(a,f) \neq 1$.*
2. *$\tau(\chi)\tau(\bar{\chi}) = \chi(-1)f$.*
3. *$|\tau(\chi)| = \sqrt{f}$.*

Proof. — (1) We consider first the case $\gcd(a,f)=1$. Then

$$\tau_a(\chi) = \sum_{x \in \mathbb{Z}/f\mathbb{Z}} \zeta^{ax}\chi(x) = \bar{\chi}(a) \sum_{x \in \mathbb{Z}/f\mathbb{Z}} \zeta^{ax}\chi(ax) = \bar{\chi}(a)\tau(\chi).$$

Assume now that $\gcd(a,f)=d>1$. Write $a=a'd$, $f=f'd$, and $\zeta_{f'}=\zeta^d$. Then

$$\begin{aligned}\tau_a(\chi)&=\sum_{x\in\mathbb{Z}/f\mathbb{Z}}\zeta_{f'}^{a'x}\chi(x)=\sum_{s=0}^{f'-1}\sum_{t=0}^{d-1}\zeta_{f'}^{a'(s+tf')}\chi(s+tf')\\&=\sum_{s=0}^{d-1}\zeta_{f'}^{s}\Big(\sum_{t=0}^{f'-1}\chi(s+tf')\Big)\end{aligned}$$

We claim that $\sum_{t=0}^{f'-1}\chi(s+tf')=0$. Actually, If $\gcd(s,f')>1$, then every term in the summation is 0. If $\gcd(s,f')=1$, let H denote the kernel of the natural reduction map $(\mathbb{Z}/f\mathbb{Z})^\times\to(\mathbb{Z}/f'\mathbb{Z})^\times$. Then the sum is the same as

$$\sum_{\substack{x\in(\mathbb{Z}/f\mathbb{Z})^\times\\ x\equiv s \mod f'}}\chi(x)=\chi(s)\sum_{x\in H}\chi(x)=0.$$

Here, the last equality uses Proposition 7.1.4 and the fact that $\chi|_H$ is non-trivial (as χ has conductor f).

For (2), we have

$$\begin{aligned}\tau(\chi)\tau(\bar\chi)&=\sum_{a\in(\mathbb{Z}/f\mathbb{Z})}\tau(\chi)\bar\chi(a)\zeta^a=\sum_{a\in\mathbb{Z}/f\mathbb{Z}}\tau_a(\chi)\zeta^a\\&=\sum_{x\in\mathbb{Z}/f\mathbb{Z}}\chi(x)\Big(\sum_{a\in\mathbb{Z}/f\mathbb{Z}}\zeta^{a(1+x)}\Big)=\chi(-1)f.\end{aligned}$$

For (3), it suffices to show that $|\tau(\bar\chi)|=|\tau(\chi)|$. Indeed,

$$\overline{\tau(\bar\chi)}=\overline{\big(}\sum_{x\in\mathbb{Z}/f\mathbb{Z}}\bar\chi(x)\zeta^x\big)=\sum_{x\in\mathbb{Z}/f\mathbb{Z}}\chi(x)\zeta^{-x}=\chi(-1)\tau(\chi).$$

□

Theorem 7.4.3. — *Let χ be a Dirichlet character of conductor $f\geq 3$. Then*

$$L(\chi,1)=-\frac{\tau(\chi)}{f}\sum_{a=1}^{f-1}\bar\chi(a)\log|1-\zeta^a|=-\frac{\tau(\chi)}{f}\sum_{a=1}^{f-1}\bar\chi(a)\log(\sin\frac{\pi a}{f})$$

if $\chi(-1)=1$, and

$$L(\chi,1)=\frac{\tau(\chi)\pi i}{f^2}\sum_{a=1}^{f-1}\bar\chi(a)a$$

if $\chi(-1)=-1$.

Proof. — We consider for any $a\in(\mathbb{Z}/f\mathbb{Z})^\times$ the Dirichlet series

$$\sum_{n=1}^{+\infty}\frac{\zeta^{an}}{n^s},$$

which absolutely converges in $\Re(s) > 1$. Using the same method as Proposition 6.2.2, we see that it has limit

$$\sum_{n=1}^{+\infty} \frac{\zeta^{an}}{n} = -\log(1-\zeta^a) \quad \text{when } s \to 1^+,$$

where we take the branch of the multiple valued function $\log(z)$ on $z \in \mathbb{C} - \{0\}$ that takes real values on $z \in \mathbb{R}_{>0}$. Multiplying with $\bar{\chi}(a)$ and taking sums, we get

$$\begin{aligned}\sum_{a=1}^{f-1} \bar{\chi}(a) \sum_{n=1}^{+\infty} \frac{\zeta^{an}}{n^s} &= \sum_{n=1}^{+\infty} \frac{1}{n^s} \Big(\sum_{a \in (\mathbb{Z}/f\mathbb{Z})^\times} \bar{\chi}(a)\zeta^{an} \Big) \\ &= \tau(\bar{\chi}) \sum_{n=1}^{+\infty} \frac{\chi(n)}{n^s} = \frac{f}{\chi(-1)\tau(\chi)} L(\chi, s).\end{aligned}$$

Here, the last two equalities uses Lemma 7.4.2. Hence, it follows that

$$L(\chi,1) = -\frac{\chi(-1)\tau(\chi)}{f} \sum_{a=1}^{f-1} \bar{\chi}(a) \log(1-\zeta^a).$$

Note that

$$\log(1-\zeta^a) = \log|1-\zeta^a| + \pi i(\frac{a}{f} - \frac{1}{2})$$

We now distinguish the two cases on the parity of χ.

– χ is odd, i.e. $\chi(-1) = -1$. In this case, since $|1-\zeta^a| = |1-\zeta^{-a}|$, we get

$$\begin{aligned}\sum_{a=1}^{f-1} \bar{\chi}(a) \log|1-\zeta^a| &= \frac{1}{2} \sum_{a=1}^{f-1} \Big(\bar{\chi}(a) \log|1-\zeta^a| + \bar{\chi}(a) \log|1-\zeta^{-a}| \Big) \\ &= \frac{1}{2} \sum_{a=1}^{f-1} (\bar{\chi}(a) + \bar{\chi}(-a)) \log|1-\zeta^a| = 0.\end{aligned}$$

Thus we get

$$L(\chi,1) = \frac{\tau(\chi)}{f} \sum_{a=1}^{f-1} \bar{\chi}(a)\pi i(\frac{a}{f} - \frac{1}{2}) = \frac{\tau(\chi)\pi i}{f^2} \sum_{a=1}^{f-1} \bar{\chi}(a)a.$$

– χ is even, i.e. $\chi(-1) = 1$. Then

$$\sum_{a=1}^{f-1} \bar{\chi}(a)(\frac{a}{f} - \frac{1}{2}) = \frac{1}{2f} \sum_{a=1}^{f-1} (\bar{\chi}(a) + \bar{\chi}(-a))a = 0.$$

It follows immediately that

$$L(\chi,1) = -\frac{\tau(\chi)}{f} \sum_{a=1}^{f-1} \bar{\chi}(a) \log|1-\zeta^a| = -\frac{\tau(\chi)}{f} \sum_{a=1}^{f-1} \bar{\chi}(a) \log(\sin\frac{\pi a}{f}).$$

□

7.5. Class number formula for quadratic fields

Let $K = \mathbb{Q}(\sqrt{d_K})$ be a quadratic field of discriminant d_K, and $G = \mathrm{Gal}(K/\mathbb{Q})$ denote the Galois group.

Proposition 7.5.1. — *The quadratic field K is a subfield of $\mathbb{Q}(\zeta_{d_K})$. Moreover, the non-trivial Dirichlet character χ_{d_K} of G has conductor $|d_K|$ and is determined by the following rules:*

(a) $\chi_{d_K}(-1) = \frac{d_K}{|d_K|}$.

(b) $\chi_{d_K}(2) = (-1)^{\frac{d_K^2-1}{8}}$ *if* $d_K \equiv 1 \mod 4$ *and* $\chi_{d_K}(2) = 0$ *otherwise.*

(c) $\chi_{d_K}(p) = \left(\frac{d_K}{p}\right)$ *if p is an odd prime; in particular,* $\chi_{d_K}(p) = 0$ *if* $p \mid d_K$.

Proof. — To show that $K \subseteq \mathbb{Q}(\zeta_{d_K})$, we proceed by induction on the number of distinct prime factors of d_K. Assume first d_K has only one prime factor. If $|d_K| = p$ is odd, then $d_K = (-1)^{\frac{p-1}{2}}p$ and K is the unique quadratic subfield of $\mathbb{Q}(\zeta_p)$ by Lemma 3.5.5. If d_K is even, then the possible values for d_K are -4 and ± 8. It is clear that $K \subseteq \mathbb{Q}(\zeta_8)$ in all three cases. Assume now d_K has $r \geq 2$ distinct primes factors, and that the assertion is true for $K' = \mathbb{Q}(\sqrt{d'})$ with $d'|d_K$ and $d' \neq d_K$. Write $d_K = mp^*$, where $p^* = (-1)^{\frac{p-1}{2}}p$ for some odd prime p. By induction hypothesis, we have $\mathbb{Q}(\sqrt{m}) \subseteq \mathbb{Q}(\zeta_m)$ and $\mathbb{Q}(\sqrt{p^*}) \subseteq \mathbb{Q}(\zeta_p)$. Therefore, we get

$$K \subseteq \mathbb{Q}(\sqrt{m}, \sqrt{p^*}) \subseteq \mathbb{Q}(\zeta_p, \zeta_m) = \mathbb{Q}(\zeta_{d_K}).$$

This completes the proof of the first assertion of the Proposition.

Let f denote the conductor of χ_{d_K}. Then it is the minimal positive integer such that $\chi_{d_K} : (\mathbb{Z}/d_K\mathbb{Z})^\times \to \mathbb{C}^\times$ factors through $(\mathbb{Z}/f\mathbb{Z})^\times$. Since the quotient $(\mathbb{Z}/f\mathbb{Z})^\times$ of $(\mathbb{Z}/d_K\mathbb{Z})^\times$ corresponds to the subfield $\mathbb{Q}(\zeta_f)$ of $\mathbb{Q}(d_K)$. Thus, $\mathbb{Q}(\zeta_f)$ is the minimal cyclotomic field that contains K. Because of ramification, it is clear that f must contain all the prime factors of d_K. Therefore, it remains to exclude the case where $d_K = 8d'$ with d' odd and $f = 4|d'|$. We need to show that it is impossible that $\sqrt{2d'} \in \mathbb{Q}(\zeta_{4d'})$. Write $2^* = (-1)^{\frac{d'-1}{2}}2$ and $d'^* = (-1)^{\frac{d'-1}{2}}d'$. Then $d'^* \equiv 1 \mod 4$, and there exist distinct odd primes p_i such that $d'^* = \prod_{i=1}^r p_i^*$, where $p_i^* = (-1)^{\frac{p_i-1}{2}}p_i$. Note that $\mathbb{Q}(\sqrt{p_i^*}) \subseteq \mathbb{Q}(\zeta_{p_i}) \subseteq \mathbb{Q}(\zeta_{4d'})$. Thus, if $\sqrt{2d'} \in \mathbb{Q}(\zeta_{4d'})$, then one would have

$$\sqrt{2^*} = \sqrt{2d'}\prod_i^r \frac{1}{\sqrt{p_i^*}} \in \mathbb{Q}(\zeta_{4d'}),$$

and hence $\zeta_8 = \frac{\sqrt{2}}{2} + \frac{\sqrt{-2}}{2} \in \mathbb{Q}(\zeta_{4d'})$ and $\mathbb{Q}(\zeta_{8d'}) = \mathbb{Q}(\zeta_{4d'})$. This is clearly absurd.

It remains to prove that χ_{d_K} satisfies the rules (a), (b), (c), which clearly determine χ_K. We regard G as a quotient of $(\mathbb{Z}/d_K\mathbb{Z})^\times$. The complex conjugate of K is the image of -1. Then rule (a) follows immediately, since χ_{d_K} sends always the non-trivial element of G to -1. Similarly, if p is a prime not dividing d_K, then p is unramified in K and the Frobenius element σ_p of $K/\mathbb{Q}$ is given by the image of $p \mod d_K$. Thus, $\chi_{d_K}(p) = -1$ if and only if p is inert in K. If $p \mid d_K$, then $\chi_{d_K}(p) = 0$ by definition. Now rules (b) and (c) follow immediately from Theorem 3.2.5. □

Example 7.5.2. — If $|d_K| = p$ is an odd prime, then $\chi_{d_K}(x) = \left(\frac{x}{p}\right)$ is the Legendre symbol by quadratic reciprocity law.

We have seen that Gauss sums enter into the computation of $L(\chi, 1)$ for a Dirichlet character χ. In general, it is hard to give an explicit value for $\tau(\chi)$. However, in the our case, we have the following

Theorem 7.5.3. — *Let χ_{d_K} denote the non-trivial Dirichlet character associated to a quadratic field K. Then*

$$\tau(\chi_{d_K}) = \begin{cases} \sqrt{|d_K|} & \text{if } \chi(-1) = 1, \\ i\sqrt{|d_K|} & \text{if } \chi(-1) = -1, \end{cases}$$

where $\sqrt{|d_K|}$ is the positive square root of $|d_K|$.

We will postpone the proof of this Theorem to the end of this section. We now state the main result of this section.

***Theorem 7.5.4* (Dirichlet's class number formula)**. — *The class number of a quadratic field K of discriminant d_K is given by*

$$h = -\frac{1}{|d_K|} \sum_{a=1}^{|d_K|-1} \chi_{d_K}(a) a$$

if $d_K < -4$, and

$$h = -\frac{1}{\log(\varepsilon)} \sum_{a=1}^{[\frac{d_K}{2}]} \chi_{d_K}(a) \log(\sin \frac{\pi a}{d_K}),$$

if $d_K > 0$, where $\varepsilon > 1$ is the fundamental unit of K.

Proof. — By Theorem 7.2.4, we have

$$h = \frac{\sqrt{|d_K|}}{\pi} L(\chi_{d_K}, 1) \quad \text{if } d_K < -4,$$

and

$$h = \frac{\sqrt{|d_K|}}{2\log(\varepsilon)} L(\chi_{d_K}, 1)$$

if $d_K > 0$. By Proposition 7.5.1, χ has conductor $|d_K|$. It follows from Theorems 7.4.3 and 7.5.3 that

$$h = -\frac{\tau(\chi_{d_K}) i}{|d_K|^{3/2}} \sum_{a=1}^{|d_K|} \chi_{d_K}(a) a = -\frac{1}{|d_K|} \sum_{a=1}^{|d_K|-1} \chi_{d_K}(a) a$$

if $d_K < -4$, and

$$h = -\frac{\tau(\chi_{d_K})}{2|d_K|^{1/2}\log(\varepsilon)} \sum_{a=1}^{d_K-1} \chi_{d_K}(a) \log(\sin \frac{\pi a}{d_K}) = -\frac{1}{\log(\varepsilon)} \sum_{a=1}^{[\frac{d_K}{2}]} \chi_{d_K}(a) \log(\sin \frac{\pi a}{d_K})$$

if $d_K > 0$. Here, the last step uses the symmetry $\chi_{d_K}(a) = \chi(d_K - a)$ and $\sin \frac{\pi(d_K-a)}{d_K} = \sin \frac{\pi a}{d_K}$.

□

Corollary 7.5.5. — *Under the notation of the Theorem, and assume that $|d_K| = p$ is an odd prime. Let R (resp. N) denote the subset of quadratic residues (resp. quadratic non-residues) modulo p of $\{1, 2, \cdots, p-1\}$.*

1. *If $d_K = -p$, then*
$$h = \frac{1}{p}(\sum_{b\in N} b - \sum_{a\in R} a) = \frac{p-1}{2} - \frac{2}{p}\sum_{a\in R} a.$$
In particular, h is always odd.
2. *If $d_K = p$, then*
$$\varepsilon^h = \frac{\prod_b \sin \frac{\pi b}{p}}{\prod_a \sin \frac{\pi a}{p}},$$
where a (resp. b) runs through the elements of $R \cap [1, \frac{p-1}{2}]$ (resp. of $N \cap [1, \frac{p-1}{2}]$).

Proof. — Indeed, if $|d_K| = p$ is an odd prime, then $\chi_{d_K}(x) = \left(\frac{x}{p}\right)$. Thus $\chi_{d_K}(x) = -1$ (resp. $\chi_{d_K}(x) = 1$) if and only if x is a quadratic non-residue (resp. quadratic residue) modulo p. The Corollary follows immediately from the Theorem. □

Remark 7.5.6. — In the case $d_K = -p$, we have in particular that $\sum_{b\in N} b > \sum_{a\in N} a$. Despite of its elementary appearance, an elementary proof of this fact has not been found yet.

Corollary 7.5.7. — *Assume that $d_K < 0$ and $2|d_K$. Then the class number of K is*
$$h = -\frac{2}{|d_K|}\sum_{0<a<|d_K|/2} \chi_{d_K}(a)a + \sum_{0<a<|d_K|/2} \chi_{d_K}(a).$$

Proof. — Indeed, one has
$$\begin{aligned} h &= -\frac{1}{|d_K|}\sum_{0<a<|d_K|/2} (\chi_{d_K}(a)a + (|d_K| - a)\chi_{d_K}(|d_K| - a)a) \\ &= -\frac{1}{|d_K|}\sum_{0<a<|d_K|/2} \Big(\chi_{d_K}(a)a - (|d_K| - a)\chi_{d_K}(a)\Big), \end{aligned}$$

where the last equality used the fact that $\chi_{d_K}(|d_K| - a) = -\chi_{d_K}(a)$. Now the corollary follows immediately. □

Example 7.5.8. — Let $K = \mathbb{Q}(\sqrt{-56})$. For an odd prime p, one has
$$\chi_{-56}(p) = \left(\frac{-56}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{2}{p}\right)\left(\frac{7}{p}\right) = \left(\frac{2}{p}\right)\left(\frac{p}{7}\right).$$

Now it is easy to see that for $0 < p < 28$, one has

$$\chi_{-56}(p) = \begin{cases} 1 & \text{if } p = 3, 5, 13, 19, 23 \\ -1 & \text{if } p = 11, 17. \end{cases}$$

Therefore, it follow that

$$\sum_{0<a<28} \chi_{-56}(a) = 8, \qquad \sum_{0<a<28} \chi_{-56}(a)a = 112.$$

Hence, one gets $h = 4$, which coincides with the computation in Chapter 5.

We now turn to the proof of Theorem 7.5.3 by starting with the following

Proposition 7.5.9. — *If d_1 and d_2 are discriminants of some quadratic fields with* $\gcd(d_1, d_2) = 1$, *then*

$$\chi_{d_1}(|d_2|) = \varepsilon_{d_1,d_2}\chi_{d_2}(|d_1|),$$

where $\varepsilon_{d_1,d_2} = -1$ if both d_1 and d_2 are negative, and $\varepsilon_{d_1,d_2} = 1$ otherwise.

Proof. — Indeed, if $d_1 = d_3d_4$ where d_3 and d_4 are coprime discriminants of some quadratic fields, then it is easy to see that $\chi_{d_3d_4} = \chi_{d_3}\chi_{d_4}$. One checks also easily that the right hand side is multiplicative in d_1. The same thing holds for d_2 by symmetry. By induction, we may assume that both d_1 and d_2 are powers of primes. The possible values for d_1 and d_2 are $-4, \pm 8$, p and $-q$, where p is some prime with $p \equiv 1 \mod 4$ and q is a prime with $q \equiv 3 \mod 4$. By a direct check, we see easily that the statement is equivalent to the quadratic reciprocity law. For instance, if both $d_1 = -q_1$ and $d_2 = -q_2$ where q_1 and q_2 are primes congruent to 3 modulo 4. Then it follows that

$$\begin{aligned} \chi_{d_1}(|d_2|) = \chi_{-q_1}(q_2) &= \left(\frac{-q_1}{q_2}\right) \\ &= \left(\frac{q_2}{q_1}\right) = -\chi_{d_2}(q_1) = -\chi_{d_2}(d_1), \end{aligned}$$

where the third equality uses quadratic reciprocity law. □

Proof of Theorem 7.5.3. — First, we reduce the problem to the case when d_K has only one distinct prime factor. Indeed, if d_K has more than one distinct prime factors, one can write $d_K = d_1d_2$ with $\gcd(d_1, d_2) = 1$, where d_1, d_2 are discriminants of some quadratic fields. Then by Proposition 7.5.1, it is easy to see that $\chi_{d_K} = \chi_{d_1}\chi_{d_2}$. Since $\{x|d_1| + y|d_2| :$

$1 \leq x \leq |d_2|, 1 \leq y \leq |d_1|\}$ form a set of representatives of $\mathbb{Z}/d_K\mathbb{Z}$, we have

$$\begin{aligned}\tau(\chi_{d_K}) &= \sum_{x=1}^{|d_2|}\sum_{y=1}^{|d_1|}\chi_{d_K}(x|d_1|+y|d_2|)\zeta_{|d_K|}^{x|d_1|+y|d_2|}, \quad \text{with } \zeta_{|d_K|} = e^{\frac{2\pi i}{|d_K|}} \\ &= \chi_{d_1}(|d_2|)\chi_{d_2}(|d_1|)\sum_{x=1}^{|d_2|}\sum_{y=1}^{|d_1|}\chi_{d_2}(x)\chi_{d_1}(y)\zeta_{|d_2|}^{x}\zeta_{|d_1|}^{y} \\ &= \chi_{d_1}(|d_2|)\chi_{d_2}(|d_1|)\tau(\chi_{d_1})\tau(\chi_{d_2}). \\ &= \varepsilon_{d_1,d_2}\tau(\chi_{d_1})\tau(\chi_{d_2}),\end{aligned}$$

where we have used Proposition 7.5.9 in the last step. Clearly, if Theorem 7.5.3 holds for $\tau(\chi_{d_1})$ and $\tau(\chi_{d_2})$, then it holds for $\tau(\chi_{d_1d_2})$.

Thus we may assume that d_K has only one prime factor. When d_K is even, then $d_K = -4, \pm 8$ and the formula for $\tau(\chi_{d_K})$ can be checked easily by hand. We assume hence that d_K is odd, and let $p = |d_K|$. Then $d_K = p$ if K is real, and $d_K = -p$ if K is imaginary. Let R (resp. N) denote the set of integers a with $1 \leq a \leq p-1$ which are quadratic residues (resp. non-residues) modulo p. Then

$$\tau(\chi_{d_K}) = \sum_{a=1}^{p-1}\chi_{d_K}(a)\zeta^a = \sum_{a\in R}\zeta^a - \sum_{b\in N}\zeta^b$$

where $\zeta = e^{\frac{2\pi i}{p}}$. Since $\sum_{a\in R}\zeta^a + \sum_{b\in N}\zeta^b = -1$, one obtains

$$\tau(\chi_{d_K}) = 1 + 2\sum_{a\in R}\zeta^a = \sum_{x=0}^{p-1} e^{\frac{2\pi i x^2}{p}}.$$

Then the Theorem follows immediately from Proposition 7.5.10 below. □

Proposition 7.5.10. — *Let $N \geq 1$ be an integer, and $\sqrt{N}$ denote its positive square root. Then we have*

$$S_N := \sum_{x=0}^{N-1} e^{2\pi i x^2/N} = \begin{cases}(1+i)\sqrt{N} & \text{if } N \equiv 0 \mod 4, \\ \sqrt{N} & \text{if } N \equiv 1 \mod 4, \\ 0 & \text{if } N \equiv 2 \mod 4, \\ i\sqrt{N} & \text{if } N \equiv 3 \mod 4.\end{cases}$$

Proof. — Let $f(x)$ denote the periodic function with period 1 and

$$f(x) = \sum_{n=0}^{N-1} e^{2\pi i(x+n)^2/N}, \quad \text{for } x \in [0,1).$$

It is continuously differentiable except at $x \in \mathbb{Z}$ and continuous everywhere. Hence, its Fourier series converges to $f(x)$ pointwise. We obtain therefore

$$f(x) = \sum_{m=-\infty}^{+\infty} a_m e^{-2\pi i x},$$

with

$$a_m = \int_0^1 f(t)e^{2\pi imt} = \sum_{n=0}^{N-1} \int_0^1 e^{2\pi i(mt+(t+n)^2/N)}dt = \int_0^N e^{2\pi i(mt+t^2/N)}dt.$$

Taking $x = 0$, we get

$$\begin{aligned} S_N = f(0) &= \sum_{m=-\infty}^{+\infty} \int_0^N e^{2\pi i(mt+t^2/N)}dx \\ &= N \sum_{m=-\infty}^{+\infty} \int_0^1 e^{2\pi iN(t^2+mt)}dt \quad (\text{set } x = Nt) \\ &= N \sum_{m=-\infty}^{+\infty} e^{-\frac{\pi i}{2}Nm^2} \int_{\frac{m}{2}}^{1+\frac{m}{2}} e^{2\pi iNy^2}dy \quad (\text{set } y = t + \frac{m}{2}). \end{aligned}$$

Note that $e^{\frac{\pi i}{2}Nm^2}$ equals to 1 if m is even, and to i^{-N} if m is odd. Thus, we divide the sum over m into two parts according to the parity of m, and we put $m = 2k$ and $m = 2k - 1$ in the two cases respectively. Then we get

$$\begin{aligned} S_N &= N \sum_{k=-\infty}^{+\infty} \int_k^{k+1} e^{2\pi iNy^2}dy + Ni^{-N} \int_{k-1/2}^{k+1/2} e^{2\pi iNy^2}dy \\ &= N(1 + i^{-N}) \int_{-\infty}^{+\infty} e^{2\pi iNy^2}dy \\ &= \sqrt{N}(1 + i^{-N}) \int_{-\infty}^{+\infty} e^{2\pi iz^2}dz \quad (\text{set } z = \sqrt{N}y). \end{aligned}$$

Now the value of $C := \int_{-\infty}^{+\infty} e^{2\pi iz^2}dz$ is a independent of N. Letting $N = 1$, one finds easily that $C = \frac{1}{1+i^{-1}}$. Therefore, one finds

$$S_N = \sqrt{N}\frac{1 + i^{-N}}{1 + i^{-1}},$$

from which Proposition 7.5.10 follows immediately. □

CHAPTER 8

NONARCHMEDEAN VALUATION FIELDS

8.1. The introduction of p-adic fields

The idea of introducing p-adic numbers comes from solving polynomial equations modulo arbitrary powers of a prime number p. There are two approaches to define p-adic numbers:

1. We define first $\mathbb{Z}_p$ as an inverse limit of $\mathbb{Z}/p^n\mathbb{Z}$, and consider $\mathbb{Q}_p$ as the fraction field of $\mathbb{Z}_p$.
2. We equip first $\mathbb{Q}$ with a p-adic absolute value, define $\mathbb{Q}_p$ as the completion of $\mathbb{Q}$ under this absolute value, and $\mathbb{Z}_p$ as the valuation ring of $\mathbb{Q}_p$ with p-adic absolute values ≤ 1.

8.1.1. p-adic numbers as inverse limit. — Let p be a prime. We put

$$\mathbb{Z}_p = \varprojlim_n (\mathbb{Z}/p^n\mathbb{Z}) = \{(x_n)_n \in \prod_{n\geq 1} \mathbb{Z}/p^n\mathbb{Z} : x_{n+1} \mod p^n = x_n, \}$$

that is $\mathbb{Z}_p$ is the subset of $\prod_{n\geq 1}(\mathbb{Z}/p^n\mathbb{Z})$ such that its n-th component is the reduction modulo p^n of its $n+1$-th component. We equip $\mathbb{Z}_p$ the componentwise ring structure induced from $\prod_{n\geq 1}\mathbb{Z}/p^n\mathbb{Z}$.

Proposition 8.1.2. — *The ring $\mathbb{Z}_p$ is an integral domain, and it is a local ring with unique maximal ideal $p\mathbb{Z}_p$.*

Proof. — We prove first that $\mathbb{Z}_p$ is integral. Let $x = (x_n)_n, y = (y_n)_n \in \mathbb{Z}_p$ be non-zero elements. Let m_0, n_0 be the minimal integers such that $x_{m_0+1}, y_{n_0+1} \neq 0$. Then for any $m \geq m_0+1$ and $n \geq n_0+1$, one has $x_m = p^{m_0}u_m$ and $y_n = p^{n_0}v_n$ with u_m, v_n not divisible by p. Then

$$(xy)_{m_0+n_0+2} = p^{m_0+n_0}u_{m_0+n_0+2}v_{m_0+n_0+2}$$

is nonzero in $\mathbb{Z}/p^{m_0+n_0}\mathbb{Z}$. This proves that $\mathbb{Z}_p$ is integral. It is clear that $p\mathbb{Z}_p$ is a maximal ideal of $\mathbb{Z}_p$. To show that it is the unique maximal ideal, we have to show that every element of $\mathbb{Z}_p\backslash p\mathbb{Z}_p$ is invertible in $\mathbb{Z}_p$. Actually, if $x \in \mathbb{Z}_p\backslash p\mathbb{Z}_p$, then there exists a unique $y_n \in \mathbb{Z}/p^n$ such that $x_n y_n = 1$. So $y = (y_n)_{n\geq 1}$ well defines an element of $\mathbb{Z}_p$ by the uniqueness, and $xy = 1$ in $\mathbb{Z}_p$.

□

We equip $\mathbb{Z}_p$ with the topology such that, for every $a \in \mathbb{Z}_p$, $(a + p^n\mathbb{Z}_p)_{n\geq 1}$ form a fundamental system of open neighborhood of a. It is clear that the topology is invariant under translation, and $\mathbb{Z}_p$ becomes a topological ring, i.e. the addition and multiplication on $\mathbb{Z}_p$ are both continuous under this topology.

Proposition 8.1.3. — *The topological ring $\mathbb{Z}_p$ is complete in the sense that every Cauch sequence $(a_n)_{n\geq 1} \in \mathbb{Z}_p$ has a limit in $\mathbb{Z}_p$. Moreover, $\mathbb{Z}$ is dense in $\mathbb{Z}_p$.*

Proof. — For any integer $m \geq 1$, there exists $N(m)$ such that for any $n_1, n_2 \geq N(m)$ one has $a_{n_1} - a_{n_2} \in p^m\mathbb{Z}_p$. Therefore, the image of a_n in $\mathbb{Z}/p^{m\mathbb{Z}}$ is independent of $n \geq N$, and we denote it by b_m. Then $(b_m)_{m\geq 1}$ well defines an element of $\mathbb{Z}_p$, and $a_n \to b$ when $n \to \infty$. The subring $\mathbb{Z} \subseteq \mathbb{Z}_p$ is clearly dense in $\mathbb{Z}_p$, since the natural map $\mathbb{Z} \to \mathbb{Z}/p^n\mathbb{Z}$ is surjective. □

We define $\mathbb{Q}_p$ as the fraction field of $\mathbb{Z}_p$. The topology on $\mathbb{Z}_p$ extends naturally to a topology on $\mathbb{Q}_p$ so that $\mathbb{Q}_p$ is a complete topological field. Then every element of $\mathbb{Q}_p$ writes uniquely as $x = p^n u$ with $n \in \mathbb{Z}$ and $u \in \mathbb{Z}_p^\times$. Since $\mathbb{Z} \subseteq \mathbb{Z}_p$ is dense, $\mathbb{Q}$ is also dense in $\mathbb{Q}_p$.

8.1.4. p-adic numbers as completions. — We now explain an alternative way to define the p-adic field $\mathbb{Q}_p$. We define a p-adic norm $|\cdot|_p : \mathbb{Q} \to \mathbb{R}_{\geq 0}$ by the following rules:

1. $|0|_p = 0$;
2. for $x = p^n \frac{a}{b} \in \mathbb{Q}$ with $p \nmid a, b$, we put $|x|_p = p^{-n}$.

Then one verify easily that $|\cdot|_p$ satisfies the ultra-metric inequality

$$|x + y|_p \leq \max\{|x|_p, |y|_p\}, \quad \text{for } x, y \in \mathbb{Q}.$$

Hence, $|\cdot|_p$ defines a metric on $\mathbb{Q}$. We can define alternatively $\mathbb{Q}_p$ as the completion of $\mathbb{Q}$ under the norm $|\cdot|_p$.

The existence of the completion follows from the general theorems in analysis. Then $\mathbb{Z}_p$ can be defined as the subring of $\mathbb{Q}_p$ consisting of $x \in \mathbb{Q}_p$ with $|x|_p \leq 1$. The equivalence of these two definitions follows from Proposition 8.1.3.

A series $\sum_{n=1}^{+\infty} a_n$ with $a_n \in \mathbb{Q}_p$ converges in $\mathbb{Q}_p$, if and only if $|a_n|_p \to 0$ as $n \to +\infty$. Every element $x \in \mathbb{Q}_p$ writes uniquely of the form

$$x = \sum_{n>>-\infty} a_n p^n, \quad \text{with } a_n \in \{0, 1, \cdots, p-1\}.$$

If $x \in \mathbb{Q}$, we call such an expression the p-adic expansion of x. In particular, we have an equality in $\mathbb{Q}_p$:

$$\frac{1}{1-p} = \sum_{n=0}^{+\infty} p^n.$$

Example 8.1.5. — In $\mathbb{Q}_5$, we have the expansions:

$$\frac{1}{2} = 3 + 2 \times 5 + 2 \times 5^2 + \cdots + 2 \times 5^n + \cdots,$$
$$\frac{1}{3} = 2 + 3 \times 5 + 5^2 + \cdots + 3 \times 5^{2n-1} + 5^{2n} + \cdots$$

8.2. Absolute values and completion

Definition 8.2.1. — A *valuation field* $(K, |\cdot|)$ is a field K together with an *absolute value* (or a *multiplicative valuation*) $|\cdot| : K \to \mathbb{R}_{\geq 0}$ such that

(1) $|x| = 0$ if and only if $x = 0$;
(2) $|xy| = |x||y|$ for any $x, y \in K$;
(3) $|x + y| \leq |x| + |y|$ for any $x, y \in K$.

If condition (3) is replaced by the ultra metric inequality

$$|x + y| \leq \max\{|x|, |y|\},$$

we say that $|\cdot|$ is *non-archimedean*; otherwise, we say $|\cdot|$ is *archimedean*. We say two norms $|\cdot|_1$ and $|\cdot|_2$ are equivalent if there exists $r > 0$ such that $|\cdot|_2 = |\cdot|_1^r$.

Note that condition 2 implies that $|\zeta| = 1$ for any root of unity ζ contained in K.

A norm $|\cdot|$ on K makes K a metric space, and hence defines a topology on K: the subsets

$$U(a, \epsilon) = \{x \in K : |x - a| < \epsilon\} \quad \text{for } a \in K \text{ and } \epsilon > 0$$

form a topological basis. It is clear that equivalent valuations define the same topology on K.

Proposition 8.2.2. — *Let $(K, |\cdot|)$ be a valued field. Then there exists a unique field $(\widehat{K}, |\cdot|_{\widehat{K}})$ such that*

(1) *K is a subfield of $\widehat{K}$, and the restriction of $|\cdot|_{\widehat{K}}$ to K is $|\cdot|$;*
(2) *K is dense in $\widehat{K}$ for topology defined by $|\cdot|_{\widehat{K}}$;*
(3) *$\widehat{K}$ is complete under $|\cdot|_{\widehat{K}}$;*
(4) *if $f : (K, |\cdot|) \hookrightarrow (L, |\cdot|_L)$ is an embedding of normed fields with L complete, then f extends uniquely to an embedding $\hat{f} : (\widehat{K}, |\cdot|_{\widehat{K}}) \hookrightarrow (L, |\cdot|_L)$ of normed fields.*

Proof. — This is a standard abstract nonsense in analysis. We recall briefly the arguments. Recall that a sequence $(a_n)_{n\geq 1}$ in K is called a Cauchy sequence, if for any $\epsilon > 0$, there exists an integer $N \geq 1$ such that $|a_n - a_m| < \epsilon$ for any $n, m \geq N$. Two Cauchy sequences $(a_n)_{n\geq 1}$ and $(b_n)_{n\geq 1}$ are called equivalent if for any $\epsilon > 0$, there exists an integer $N \geq 1$ such that $|a_n - b_n| < \epsilon$ for any $n \geq N$. As a set, $\widehat{K}$ is the equivalence class of Cauchy sequences in K, and K embeds naturally into $\widehat{K}$ via $a \mapsto (a, a, \cdots, a, \cdots)$. For a Cauchy sequence $x = (x_n)_{n\geq 1}$ in K, one put

$$|x|_{\widehat{K}} = \lim_n |x_n|.$$

It is easy to check that if $x = (x_n)_{n\geq 1}$ and $y = (y_n)_{n\geq 1}$ are equivalent Cauchy sequences, then $|x|_{\widehat{K}} = |y|_{\widehat{K}}$. This implies that $|\cdot|_{\widehat{K}}$ is well defined, and $|\cdot|_{\widehat{K}}$ restricts to $|\cdot|$ on K.

It is clear that K is dense in $\widehat{K}$, since every Cauchy sequence $(a_n)_{n\geq 1}$ in K is approximated by the constant Cauchy sequences $(a_n, \cdots, a_n, \cdots, a_n, \cdots)$ when $n \to +\infty$. It is also clear that the addition and multiplication on K extend naturally to $\widehat{K}$. We verify now that $\widehat{K}$ is a field. Let $a = (a_n)_{n\geq 1}$ be a non-zero Cauchy sequence. Then there exists an integer $N \geq 1$ and a constant $C > 0$ such that $a_n \neq 0$ and $|a_n| \geq C$ for all $n \geq N$. Define a sequence $b = (b_n)_{n\geq 1}$ with $b_n = 1$ for $1 \leq n \leq N-1$ and $b_n = a_n^{-1}$ for $n \geq N$. Then, for any $n, m \geq N$, one has

$$|b_n - b_m| = |a_n|^{-1}|a_m|^{-1}|a_n - a_m| \leq C^{-2}|a_n - a_m|.$$

Thus it follows that $(b_n)_{n\geq 1}$ is a Cauchy sequence. It is clear that $ab = 1$ in $\widehat{K}$. This proves that $\widehat{K}$ is a field.

Now we prove that $\widehat{K}$ is complete. Suppose given a Cauchy sequence $(a_n)_{n\geq 1}$ in $\widehat{K}$. Let $(a_{n,m})_{m\geq 1}$ with $a_{n,m} \in K$ be a Cauchy sequence in K that represents a_n. Then, one checks easily that the diagonal Cauchy sequence $(a_{n,n})_{n\geq 1}$ is the limit of $(a_n)_{n\geq 1}$ under $|\cdot|$ on $\widehat{K}$. This shows that $\widehat{K}$ is complete.

Let $f : (K, |\cdot|) \hookrightarrow (L, |\cdot|_L)$ be as in (4). For a Cauchy sequence $a = (a_n)_{n\geq 1}$ in K, we define $\hat{f}(a) = \lim_{n\to+\infty} f(a_n)$. It is clear that equivalent Cauchy sequences have the same image under $\hat{f}$. Hence, this defines the unique extension of f to $(\widehat{K}, |\cdot|_{\widehat{K}})$.

□

We call $\widehat{K}$ as above the completion of K, and we still denote $|\cdot|_{\widehat{K}}$ by $|\cdot|$ for simplicity.

Now we introduce another way to define a non-archimedean norm on K.

Definition 8.2.3. — An *additive valuation* (or simply a *valuation*) on K is a map $v : K \to \mathbb{R} \cup \{+\infty\}$ such that

(1) $v(x) = +\infty$ if and only if $x = 0$;
(2) $v(xy) = v(x) + v(y)$;
(3) $v(x + y) \geq \min\{v(x), v(y)\}$.

If $v(K^\times)$ is a discrete subgroup of $\mathbb{R}$, we say that v is a *discrete* valuation; if $v(K^\times) = \mathbb{Z}$, then we say that v is a *normalized discrete valuation*.

Two additive valuations v_1 and v_2 are called equivalent if there exists $r > 0$ such that $v_1(x) = rv_2(x)$ for any $x \in K$.

Given an additive valuation v on K and any real number $q > 1$, we define a non-archimedean norm on K by $|x| = q^{-v(x)}$. Equivalent additive valuations or different choices of q will give rise to equivalent norms. Conversely, if $|\cdot|$ is a non-archimedean absolute value on K and $q > 1$ be a fixed real number, then $v(x) := -\log_q(|x|)$ is an additive valuation on K. Thus there is a natural one-to-one correspondence between the equivalence classes of additive valuations and those of non-archimedean norms.

Example 8.2.4. — (1) Let K be any field. We put $|x| = 1$ if $x \neq 0$. Such a norm $|\cdot|$ is called the trivial absolute value on K. The topology defined by K is the discrete topology on K, i.e. every element of K is both open and closed in K.

(2) Let K be a number field, and σ be a complex embedding of K. Then $|x|_\sigma := |\sigma(x)|_{\mathbb{C}}$ defines a archimedean valuation on K. Later on, we will see that every archimedean absolute value of K arises in this way. The completion of K under the norm $|\cdot|_\sigma$ is $\mathbb{R}$ if σ is a real embedding, and is $\mathbb{C}$ if σ is non-real. Moreover, K admits also non-archimedean valuations. For every prime ideal $\mathfrak{p}$ of $\mathcal{O}_K$, let $v_\mathfrak{p}(x) \in \mathbb{Z}$ denote the exponent of $\mathfrak{p}$ appearing in the fractional ideal (x). Then $x \mapsto v_\mathfrak{p}(x)$ defines an additive valuation on K. The corresponding non-archimedean absolute value $|x|_\mathfrak{p} = \mathrm{N}(\mathfrak{p})^{-v_\mathfrak{p}(x)}$ is called the normalized $\mathfrak{p}$-adic norm on K.

(3) Let k be a field, and $k(x)$ be the rational function field over k. Let $p(x) \in k[x]$ be an irreducible polynomial. For any $f(x) \in k(x)$, one can write uniquely

$$f(x) = p(x)^e \frac{a(x)}{b(x)}, \quad \text{with } a(x), b(x) \in k[x] \text{ and } p(x) \nmid a(x)b(x).$$

Then $v_{p(x)}(f) = e$ defines an additive valuation on $k(x)$.

Proposition 8.2.5. — *Let K be a field, and $|\cdot|$ be a norm on K. Then $|\cdot|$ is non-archimedean if and only if $|\cdot|$ is bounded above on the image of $\mathbb{Z}$ in K. In particular, if K has characteristic $p > 0$, then every norm on K is non-archimedean.*

Proof. — If $|\cdot|$ is non-archimedean, then

$$|n| = |1 + \cdots + 1| \leq \max\{|1|\} = 1$$

Therefore, $|\cdot|$ is bounded on the image of $\mathbb{Z}$ in K. Conversely, suppose that $|\cdot|$ is bounded on the image of $\mathbb{Z}$ by a constant C. Then for any $x, y \in K$ and integer $n \geq 1$, we have

$$|(x+y)^n| = |\sum_{i=0}^{n} \binom{n}{i} x^{n-i} y^i| \leq \max_i \{|\binom{n}{i}||x|^{n-i}|y|^i\} \leq C \max\{|x|, |y|\}^n.$$

Hence, $|x+y| \leq C^{1/n} \max\{|x|, |y|\}$. Letting $n \to +\infty$, one obtains $|x+y| \leq \max\{|x|, |y|\}$. □

In the sequel, we will mainly focus on the non-archimedean norms.

Lemma 8.2.6. — *Let $|\cdot|$ be a non-archimedean norm on a field K. Then one has*

$$|x+y| = \max\{|x|, |y|\} \quad \textit{for } |x| \neq |y|.$$

Proof. — We may assume that $|x| > |y|$. Then by the ultra-metric equality, we have

$$|y| < |x| = |x + y - y| \leq \max\{|x+y|, |-y|\} = \max\{|x+y|, |y|\}.$$

It follows immediately that $|x| = |x+y|$.

□

Definition 8.2.7. — Let K be a field equipped with a non-archimedean absolute value $|\cdot|$. We define the valuation ring of $(K, |\cdot|)$ as the subring $\mathcal{O}_K$ consisting of $x \in K$ with $|x| \leq 1$. Equivalently, if $v : K \to \mathbb{R} \cup \{+\infty\}$ is an additive valuation such that $|\cdot| = q^{-v(\cdot)}$ for some $q > 1$, then $\mathcal{O}_K = \{x \in \mathcal{O}_K : v(x) \geq 0\}$. Moreover, if the valuation v is discrete, then we call $\mathcal{O}_K$ a *discrete valuation ring*.

Remark 8.2.8. — One may think of the valuation ring $\mathcal{O}_K$ as the closed unit ball $D(0,1)$ in the topological field K. Let $B = \{x \in K : |x| = 1\}$ denote the boundary, and $D(0,1^-) = \{x \in K : |x| < 1\}$ denote the open unit ball. Then Lemma 8.2.6 implies that the distance between each point of B and a point in $D(0,1^-)$ is always 1, and that the open unit ball $D(a,1^-)$ with center at a point $a \in B$ is contained in $D(0,1^-)$.

Example 8.2.9. — (1) The valuation ring of $(\mathbb{Q}, |\cdot|_p)$ is $\mathbb{Z}_{(p)}$, while that of $(\mathbb{Q}_p, |\cdot|_p)$ is $\mathbb{Z}_p$. Both $\mathbb{Z}_{(p)}$ and $\mathbb{Z}_p$ are discrete valuation rings, and $\mathbb{Z}_p$ is the completion of $\mathbb{Z}_{(p)}$ under the p-adic topology (i.e. the topology defined by $|\cdot|_p$.)

(2) Consider the field $\mathbb{C}(x)$ equipped with the additive valuation v_0 defined by the irreducible polynomial $p(x) = x$ as in Example 8.2.4(3). Then the valuation ring of v_0 is $\mathbb{C}[x]_{(x)}$, and the its completion under this valuation is $\mathbb{C}[[x]]$.

(3) Let $K = \mathbb{C}\{\{z\}\}$ denote the subset of $f(z) = \sum_{n>>-\infty} a_n z^n \in \mathbb{C}((z))$ such that $f(z)$ defines a meromorphic function in a neighborhood of $z = 0$. Then $f(z) \mapsto \mathrm{ord}_z(f)$ defines an additive valuation on $\mathbb{C}\{\{z\}\}$. The valuation ring $\mathcal{O}_K$ is the subring of $\mathbb{C}\{\{z\}\}$ consisting of holomorphic functions in a neighborhood at $z = 0$. The completion of $\mathcal{O}_K$ under this valuation is $\mathbb{C}[[x]]$.

Proposition 8.2.10. — *Let $(K, |\cdot|)$ be non-archimedean valuation field, and $\mathcal{O}_K$ be its valuation field. Then*

(1) *$\mathcal{O}_K$ is an integrally closed local ring with maximal ideal $\mathfrak{m}_K = \{x \in K : |x| < 1\}$.*

(2) *If $\widehat{K}$ denotes the completion of K under $|\cdot|$, then the valuation ring $\mathcal{O}_{\widehat{K}}$ is the completion of $\mathcal{O}_K$ under $|\cdot|$. Moreover, if $\pi \in \mathfrak{m}$ is a non-zero element, then one has a canonical isomorphism*

$$\mathcal{O}_{\widehat{K}} := \varprojlim_n \mathcal{O}_K/\pi^n = \{(x_n)_n \in \prod_{n\geq 1} \mathcal{O}_K/\pi^n : (x_{n+1} \mod \pi^n) = x_n\}.$$

(3) *If $|\cdot|$ is a discrete valuation, then $\mathfrak{m}_K$ is principal and all the non-zero ideals of $\mathcal{O}_K$ are of the form $\mathfrak{m}_K^n$ with some $n \in \mathbb{Z}_{\geq 0}$; in particular, the only prime ideals of $\mathcal{O}_K$ are 0 and $\mathfrak{m}_K$, and $\mathfrak{m}_K^n/\mathfrak{m}_K^{n+1}$ is a one-dimensional vector space over $k := \mathcal{O}_K/\mathfrak{m}_K$.*

Proof. — (1) Let $x \in K$ be a nonzero integral element over $\mathcal{O}_K$. Assume that

$$x^n + a_1 x^{n-1} + \cdots + a_n = 0$$

with $a_i \in \mathcal{O}_K$. Assume that $|x| > 1$. Then $|a_i x^{n-i}| = |a_i||x|^{n-i} < |x|^n$ for any $i \geq 1$. Therefore, we have $0 = |x^n + a_1 x^{n-1} + \cdots + a_n| = |x|^n > 1$ by Lemma 8.2.6, which is absurd. This shows that $|x| \leq 1$, i.e. $x \in \mathcal{O}_K$. To prove that $\mathfrak{m}$ is the unique maximal ideal of $\mathcal{O}_K$, it suffices to note that for any $u \in \mathcal{O}_K - \mathfrak{m}$, one has $|u^{-1}| = |u|^{-1} = 1$, hence $u^{-1} \in \mathcal{O}_K$.

(2) Since K is dense in $\widehat{K}$, the closure of $\mathcal{O}_K$ in $\widehat{K}$ is the exactly the subset $\{x \in \widehat{K} : |x| \leq 1\}$, that is $\mathcal{O}_{\widehat{K}}$. Denote temporarily $R = \varprojlim_n (\mathcal{O}_K/\pi^n)$. We define first a map $\phi : R \to \mathcal{O}_{\widehat{K}}$ as follows: for $x = (x_n)_{n\geq 1} \in R$, let $\tilde{x}_n \in \mathcal{O}_K$ be an arbitrary lift of $x_n \in \mathcal{O}_K/\pi^n$ for any $n \geq 1$. Then, one has

$$|\tilde{x}_n - \tilde{x}_m| \leq |\pi|^m \quad \text{for all } n \geq m.$$

As $|\pi| < 1$, we see that $(\tilde{x}_n)_{n\geq 1}$ is a Cauchy sequence in $\mathcal{O}_K$. We define $\phi(x) \in \mathcal{O}_{\widehat{K}}$ to be the equivalence class of $(\tilde{x}_n)_{n\geq 1}$. Conversely, one has a map $\psi : \mathcal{O}_{\widehat{K}} \to R$ given as follows. Let $a = (a_n)_{n\geq 1}$ be a Cauchy sequences whose equivalence class is in $\mathcal{O}_{\widehat{K}}$. Up to replacing $(a_n)_{n\geq 1}$ by a subsequence, we may assume that $a_n \in \mathcal{O}_K$ and $a_{n+1} \equiv a_n \mod \pi^n$ for all $n \geq 1$. Let $\bar{a}_n$ denote the image of a_n in $\mathcal{O}_K/\pi^n$. We put $\psi(a) = \prod_{n\geq 1}(\bar{a}_n) \in R$, which depends only on the equivalence class of the Cauchy sequence $(a_n)_{n\geq 1}$. Therefore, one gets a morphism $\psi : \mathcal{O}_{\widehat{K}} \to R$. It is easy to check that ϕ and ψ are inverse of each other. This proves that $\mathcal{O}_{\widehat{K}} \xrightarrow{\sim} \varprojlim_n (\mathcal{O}_K/\pi^n)$.

(3) Assume that $|\cdot|$ is discrete, and let $v : K \to \mathbb{Z}\cup\{+\infty\}$ denote the normalized additive valuation attached to $|\cdot|$. The maximal ideal of $\mathcal{O}_K$ is given by $\mathfrak{m}_K = \{x \in K : v(x) \geq 1\}$. If $\pi \in K$ is an element $v(\pi) = 1$, then one has $v(x\pi^{-1}) \geq 0$ for every $x \in \mathfrak{m}_K$, i.e. $x\pi^{-1} \in \mathcal{O}_K$. This implies that $\mathfrak{m}_K = (\pi)$. Let $I \subseteq \mathcal{O}_K$ be a non-zero ideal, and $n \geq 0$ be the minimal integer such that there exists $x \in I$ with $v(x) = n$. Then one sees that $\pi^n x^{-1} \in \mathcal{O}_K$, hence $\pi^n = (\pi^n x^{-1})x \in I$. Moreover, for any $y \in I$, we have $v(y\pi^{-n}) \geq 0$ by the minimality of n, and thus $y \in \pi^n \mathcal{O}_K$. It follows immediately $I = (\pi^n) = \mathfrak{m}_K^n$.

□

Remark 8.2.11. — One can give a simple "algebraic" characterization of discrete valuation ring: *for an integral domain R to be a discrete valuation ring, it is necessary and sufficient that R is noetherian, integrally closed and has only one non-zero prime ideal.* For a proof of this statement, see [**Se68**, Chap. I, §2 Proposition 3].

Definition 8.2.12. — Let $(K, |\cdot|)$ be a non-archimedean valuation field, and $\mathcal{O}_K$ be its valuation ring and $\mathfrak{m}_K \subseteq \mathcal{O}_K$ the maximal ideal.

- We call $k := \mathcal{O}_K/\mathfrak{m}_K$ the *residue field* of K (or of $\mathcal{O}_K$).
- If the valuation $|\cdot|$ is discrete with normalized additive valuation $v : K \to \mathbb{Z}\cup\{+\infty\}$, a uniformizer of K (or of $\mathcal{O}_K$) is an element $\pi \in \mathcal{O}_K$ such that $v(\pi) = 1$.

By Proposition 8.2.10, the residue field of K is stable under completion. It is clear that $\mathfrak{m}_K = (\pi)$ for any uniformizer π of K, and an element $x \in \mathcal{O}_K$ is invertible in $\mathcal{O}_K$ if and only if $\pi \nmid x$. Every element of K writes uniquely as $x = \pi^n u$ with $n = v(x) \in \mathbb{Z}$ and $u \in \mathcal{O}_K^\times$.

The following Proposition is a very useful criterion for the equivalence of two non-archimedean absolute values on a field K.

Proposition 8.2.13. — *Two nontrivial non-archimedean absolute values $|\cdot|_1$ and $|\cdot|_2$ on a field K are equivalent if and only if their valuation rings are the same.*

Proof. — Assume first that $|\cdot|_1$ and $|\cdot|_2$ are equivalent. They it is clear that $|x|_1 \leq 1$ if and only if $|x|_2 \leq 1$. Therefore, the valuation rings of K for $|\cdot|_1$ and $|\cdot|_2$ are the same. Assume now that $\mathcal{O}_K$ is the common valuation ring of K for both $|\cdot|_1$ and $|\cdot|_2$. Let $b \in K$ such that $|b|_1 > 1$. Then $b \notin \mathcal{O}_K$, so that $|b|_2 > 1$ and there exists a constant $c > 0$ such that $|b|_2^c = |b|_1$. Up to replacing $|\cdot|_2$ by $|\cdot|_2^c$, we may assume that $|b|_2 = |b|_1$. We have to prove that $|x|_1 = |x|_2$ for all $x \in K$. For any $x \in K$, there exists a real number $\rho > 0$ such that $|x|_1 = |b|_1^\rho$. For any rational number $r/s \geq \rho$, we have

$$|x|_1 \leq |b|_1^{r/s} \Leftrightarrow |x^s b^{-r}|_1 \leq 1 \Leftrightarrow x^s b^{-r} \in \mathcal{O}_K.$$

It follows that $|x^s b^{-r}|_2 \leq 1$, i.e. $|x|_2 \leq |b|_2^{r/s}$. Similarly, for any rational $m/n \leq \rho$, we also have $|x|_2 \geq |b|_2^{m/n}$. Letting m/n and r/s tend to ρ, we get

$$|x|_2 = |b|_2^\rho = |x|_1.$$

□

8.3. Structure of complete discrete valuation fields

Let K be a complete discrete valuation field with normalized additive valuation v. We will study in this section the structure of abelian groups K and $K^\times$. We fix a uniformizer $\pi \in \mathcal{O}_K$, and a set of representatives $S \subseteq \mathcal{O}_K$ of the residue field k with $0 \in S$.

Proposition 8.3.1. — *Every element of K writes uniquely as a Laurent series*

$$x = \sum_{n \gg -\infty} a_n \pi^n, \quad \text{with } a_n \in S.$$

Proof. — Since every element of K writes uniquely as $x = \pi^n u$ with $u \in \mathcal{O}_K^\times$. It suffices to prove the Proposition for $x \in \mathcal{O}_K$. Let $\bar{x}$ be the image of x in k, and $a_0 \in S$ be the unique lift of $\bar{x}$. Then one has $y_1 = x - a_0 \in \mathfrak{m}_K = (\pi)$. Thus one may write $x = a_0 + \pi x_1$ with $x_1 \in \mathcal{O}_K$. Repeating this process, one gets, for any integer $N \geq 1$,

$$x = \sum_{n=0}^{N-1} a_n \pi^n + \pi^N x_N$$

with $a_n \in S$, $x_N \in \mathcal{O}_K$. Letting $N \to +\infty$, one proves the Proposition by Proposition 8.2.10.

□

8.3.2. Case of finite residue field. — Assume now that k is a finite field with cardinality $q = p^a$ for some prime p. We will construct a canonical choice for S. We start with

Lemma 8.3.3. — *For any integer $n \geq 1$ and $x \in \mathcal{O}_K$, one has*

$$(1 + \pi^n x)^p \in 1 + \pi^{\gamma(n)} \mathcal{O}_K, \quad \text{with } \gamma(n) = \min\{n + v(p), pn\}.$$

Here, if K has characteristic p, we put $v(p) = +\infty$.

Proof. — Note that the image of p (hence q) in $\mathcal{O}_K$ lies in $\mathfrak{m}_K$. Then

$$(1+\pi^n x)^p = 1 + px\pi^n + \cdots + p(x\pi^n)^{p-1} + (x\pi^n)^p.$$

Since the minimal valuation of $\binom{p}{i}(x\pi^n)^i$ for $1 \leq i \leq p$ is $\gamma(n)$, the Lemma follows immediately. □

Corollary 8.3.4. — *For any $n \geq 1$ and $x \in \mathcal{O}_K$, one has*

$$(1+\pi x)^{q^n} \in 1 + \pi^{n+1}\mathcal{O}_K.$$

Proof. — This follows immediately by applying repeatedly the Lemma above. □

Proposition 8.3.5. — *For any $a \in k$, there exists a unique lift $[a] \in \mathcal{O}_K$ such that a is the natural reduction of $[a]$ and $[a]^q = [a]$. In particular, if K has characteristic p, then $a \mapsto [a]$ gives rise to an embedding of k into K.*

Such a lift $[a] \in \mathcal{O}_K$ is called the Teichmüller lift of $a \in k$.

Proof. — The case $a = 0$ being trivial, assume that $a \neq 0$. We choose an arbitrary lift $\tilde{a} \in \mathcal{O}_K$, and consider the sequence $(\tilde{a}^{q^n})_{n\geq 1}$. Then, for any integers $m \geq n \geq 1$, one has

$$\tilde{a}^{q^m} - \tilde{a}^{q^n} = \tilde{a}^{q^n}(\tilde{a}^{q^n(q^{m-n}-1)} - 1).$$

Since $a^{q^{m-n}-1} = 1$ in k, one has $\tilde{a}^{q^{m-n}-1} \in 1+\pi\mathcal{O}_K$. By Corollary 8.3.4, one deduces that

$$\tilde{a}^{q^n(q^{m-n}-1)} - 1 \in \pi^{n+1}\mathcal{O}_K.$$

It follows immediately that

$$|\tilde{a}^{q^m} - \tilde{a}^{q^n}| \leq |\tilde{a}^{q^n}||\pi^{n+1}| = |\pi|^{n+1},$$

which tends to 0 when $n \to +\infty$. Thus, $(\tilde{a}^{q^n})_{n\geq 1}$ is a Cauchy sequence in $\mathcal{O}_K$. As K is complete, its limit, denoted by $[a]$, exists in $\mathcal{O}_K$. It is clear that $[a]^q = [a]$ by construction, and that the image of $[a]$ in k is a. When a runs over k, the $[a]$'s gives all the solutions to $x^q = x$ in K. Therefore, $[a]$ is the unique lift of a satisfying $[a]^q = [a]$.

□

Remark 8.3.6. — When $a \neq 0$, its Teichmüller lift $[a]$ is a $(q-1)$-th root of unity in $\mathcal{O}_K$. If K has characteristic 0, the map $[\cdot] : k \to \mathcal{O}_K$ is multiplicative, but not additive.

Corollary 8.3.7. — *Every element of $x \in K$ writes uniquely as*

$$x = \sum_{n \gg -\infty} [a_n]\pi^n, \quad \textit{with } a_n \in k.$$

In particular, if K has characteristic p, we have $K \cong k((x))$.

Proof. — This follows immediately from Proposition 8.3.1 and 8.3.5. □

8.3.8. Multiplicative structure. — We assume no longer that k is a finite field. We consider the structure of $K^\times$. Denote by $U_K = \mathcal{O}_K^\times$. We have an exact sequence

$$0 \to U_K \to K^\times \xrightarrow{v} \mathbb{Z} \to 0,$$

where the surjection is given by $x \mapsto v(x)$. The choice of a uniformizer π gives a (non-canonical) section of $v : K^\times \to \mathbb{Z}$. For any integer $n \geq 1$, we put

$$U_K^n = \{x \in U_K \mid x \equiv 1 \mod \pi^n\}.$$

One gets thus a decreasing filtration

$$U_K^0 := U_K \supsetneq U_K^1 \supsetneq \cdots \supsetneq U_K^n \supsetneq U_K^{n+1} \supsetneq \cdots .$$

The following properties for the filtration is easy to check:

- This filtration is separated in the sense that $\bigcap_{n\geq 0} U_K^n = 0$.
- By the completeness of K, one has $U_K = \varprojlim_n U_K/U_K^n$.
- One has isomorphisms of abelian groups $U_K^0/U_K^1 \cong k^\times$ and $U_K^n/U_K^{n+1} \cong k$.

8.4. Hensel's Lemma

Let $(K, |\cdot|)$ be a complete non-archimedean valuation field, $\mathcal{O}_K$ be its valuation ring with the maximal ideal $\mathfrak{m}_K \subseteq \mathcal{O}_K$, and $k = \mathcal{O}_K/\mathfrak{m}_K$. Consider a polynomial

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0 \in K[x].$$

Define the Gauss norm of f as

$$\|f\| = \max_{0\leq i\leq n}\{|a_i|\}. \tag{8.4.0.1}$$

We say that $f(x)$ is *primitive* if $\|f\| = 1$, or equivalent $f(x) \in \mathcal{O}_K[x]$ and $\bar{f}(x) \neq 0$, where $\bar{f}(x) \in k[x]$ denotes the reduction of $f(x)$ modulo $\mathfrak{m}_K$.

Proposition 8.4.1 **(Hensel's Lemma)**. — *Assume that $f(x)$ is primitive, and one has*

$$\bar{f}(x) = \bar{g}(x)\bar{h}(x)$$

in $k[x]$, where $\bar{g}(x)$ and $\bar{h}(x)$ are relatively prime. Then $f(x)$ admits a factorization

$$f(x) = g(x)h(x),$$

where $g(x), h(x) \in \mathcal{O}_K[x]$ and $\deg(g) = \deg(\bar{g})$ and $g(x) \mod \mathfrak{m}_K = \bar{g}(x)$ and $h(x) \mod \mathfrak{m}_K = \bar{h}(x)$. Moreover, g and h are unique up to a unit of $\mathcal{O}_K$.

Proof. — As g and h must be relatively prime in $K[x]$, the uniqueness of g and h follows easily from the unique factorization law in $K[x]$. Let $r = \deg(\bar{g})$ and $s = \deg(f) - \deg(\bar{g})$. First, we take g_0 and $h_0 \in \mathcal{O}_K[x]$ such that

- $g_0 \mod \mathfrak{m}_K = \bar{g}$ and $\deg(g_0) = r$,
- $h_0 \mod \mathfrak{m}_K = \bar{h}$ and $\deg(h_0) \leq s$,
- $g_0 h_0 \equiv f \mod \mathfrak{m}_K$.

Note that the leading coefficient of g_0 is a unit in $\mathcal{O}_K$. Since $\gcd(\bar{g}, \bar{h}) = 1$, there exists $a, b \in \mathcal{O}_K[x]$ such that

$$ag_0 + bh_0 \equiv 1 \mod \mathfrak{m}_K.$$

If $f = g_0h_0$, then we are done. Otherwise, let $\pi \in \mathfrak{m}_K$ be a non-zero element such that π divides all the coefficients of $f - g_0h_0$ and $ag_0 + bh_0 - 1$. We now construct by induction on $n \geq 1$ polynomials $g_n, h_n \in \mathcal{O}_K$ such that

- $g_n \equiv g_{n-1} \mod \pi^n$ and $\deg(g_n) = r$;
- $h_n \equiv h_{n-1} \mod \pi^n$ and $\deg(h_n) \leq s$;
- $f \equiv g_n h_n \mod \pi^{n+1}$.

Then are the limits of g_n and h_n as $n \to \infty$ are the desired polynomials g and h respectively. Assume now that g_{n-1} and h_{n-1} have been constructed so that

$$f = g_{n-1}h_{n-1} + \pi^n f_n$$

for some $f_n \in \mathcal{O}_K[x]$ and $\deg(f_n) \leq r + s$. Write $g_n = g_{n-1} + \pi^n p_n$ and $h_n = h_{n-1} + \pi^n q_n$, where $p_n, q_n \in \mathcal{O}_K[x]$ are polynomials to be determined later. Then one has

$$g_n h_n \equiv g_{n-1}h_{n-1} + \pi^n(p_n h_{n-1} + g_{n-1}q_n) \mod \pi^{n+1}$$
$$\equiv f + \pi^n(p_n h_0 + g_0 q_n - f_n) \mod \pi^{n+1}$$

Dividing π^n, we get an equation

$$p_n h_0 + g_0 q_n \equiv f_n \mod \pi. \tag{8.4.1.1}$$

From $ag_0 + bh_0 \equiv 1 \mod \pi$, one deduces that

$$af_n g_0 + bf_n h_0 \equiv f_n \mod \pi.$$

By Euclidean division, we get

$$bf_n = ug_0 + v$$

for some $u, v \in K[x]$ and $\deg(v) \leq \deg(g_0) - 1$. Since the leading coefficient of $g_0(x)$ is a unit in $\mathcal{O}_K[x]$ and $bf_n \in \mathcal{O}_K[x]$, one sees easily that $u, v \in \mathcal{O}_K[x]$. Then one get

$$vh_0 + (af_n + uh_0)g_0 \equiv f_n \mod \pi.$$

Since $\deg(f_n) \leq r + s$ and $\deg(vh_0) < r + s$, we can take $p_n = v$ and q_n to be the degree less than s part of $af_n + uh_0$ as the solutions to equation (8.4.1.1). □

Corollary 8.4.2. — *Let $f(x) \in \mathcal{O}_K[x]$ and $\alpha \in \mathcal{O}_K$ be such that $f(\alpha_0) \equiv 0 \mod \mathfrak{m}_K$ and $f'(\alpha_0) \equiv 0 \mod \mathfrak{m}_K$. Then there exists a unique $\alpha \in \mathcal{O}_K$ such that $f(\alpha) = 0$ and $\alpha \equiv \alpha_0 \mod \mathfrak{m}_K$.*

Proof. — This is the special case of Proposition 8.4.1 with $\bar{f}(x) = x - \bar{\alpha}_0$, where α_0 is the reduction modulo $\mathfrak{m}_K$ of α_0. □

Corollary 8.4.3. — *For an irreducible polynomial $f(x) = \sum_{i=0}^{n} a_i x^i \in K[x]$, then $\|f\| = \max\{|a_0|, |a_n|\}$.*

Proof. — Assume in contrary that there exists an integer j with $0 < j < n$ such that $|f| = |a_j|$. Up to multiplying a constant, we may assume that $\|f\| = |a_j| = 1$. Then $\bar{f}(x) = \bar{a}_j x^j$. Applying Proposition 8.4.1 with $\bar{g} = x^j$, we see that $f(x)$ is not irreducible, which contradicts with the assumption. □

8.5. Extensions of valuations

Let $(K, |\cdot|)$ be a complete non-archimedean valuation field, $\mathcal{O}_K$, $\mathfrak{m}_K$ and k be as in the previous section.

Theorem 8.5.1. — *Let L/K be an algebraic extension of fields. Then there exists a unique extension of $|\cdot|$ to a non-archimedean valuation on L. If L/K is finite of degree n, then this extension is given by*

$$|x|_L = |\mathrm{N}_{L/K}(x)|^{1/n}, \quad \textit{for any } x \in L$$

and L is complete with respect to this unique extension of $|\cdot|$.

Proof. — Since an algebraic extension is a union of finite extensions, we may assume that $n = [L : K]$ is finite. Let $|\cdot|_L$ be as in the statement. It is clear that $|x|_L = |x|$ for $x \in K$. We show first that $|\cdot|_L$ is indeed a non-archimedean valuation on L. Let $\mathcal{O}_L$ denote the integral closure of $\mathcal{O}_K$ in L. We claim that $\mathcal{O}_L$ is exactly the subring of $\alpha \in L$ with $|\alpha|_L \leq 1$. Actually, if $x \in \mathcal{O}_L$, then $\mathrm{N}_{L/K}(x) \in \mathcal{O}_L$, hence $|x|_L \leq 1$. Conversely, if $|x|_L \leq 1$, then $\mathrm{N}_{L/K}(x) \in \mathcal{O}_K$. Let $f(x)$ be the minimal monic polynomial of α over K. Then by Corollary 8.5.6, the Gauss norm of f is

$$\|f\| = \max\{1, |\mathrm{N}_{L/K}(\alpha)|\} = 1.$$

Therefore, one has $f(x) \in \mathcal{O}_K[x]$ and $\alpha \in \mathcal{O}_L$.

It is clear that $|xy|_L = |x|_L|y|_L$ and $|x|_L = 0$ if and only if $x = 0$. It remains to show that $|x + y|_L \leq \max\{|x|_L, |y|_L\}$. We may assume that $|x|_L \leq |y|_L$. Up to dividing by $|y|_L$, it suffices to show that $|x + 1|_L \leq 1$ for $|x|_L \leq 1$. By the discussion above, this is equivalent to saying that $x + 1 \in \mathcal{O}_L$ for $x \in \mathcal{O}_L$, which is obvious since $\mathcal{O}_L$ is ring.

Now assume that $|\cdot|'_L$ is another non-archimedean norm on L extending $|\cdot|$. We have to show that $|\alpha|_L = |\alpha|'_L$ for all $\alpha \in L$. We first prove that $|\alpha|_L \leq 1$ if and only if $|\alpha|'_L \leq 1$. Let $f(x) = x^d + a_{d-1}x^{d-1} + \cdots + a_0$ denote the minimal polynomial of α over K with some $d|n$. Assume $|\alpha| \leq 1$. Then one has $a_i \in \mathcal{O}_K$ for all i. If $|\alpha|'_L > 1$, then $|\alpha^d|'_L > |a_i\alpha^i|'_L \geq |\alpha|'_L$ for all $0 \leq i \leq d-1$. By Lemma 8.2.6, one has $0 = |f(\alpha)| = |\alpha|^d$, which is absurd. Hence, $|\alpha|_L \leq 1$ implies that $|\alpha|'_L \leq 1$. Assuming moreover that $|\alpha|_L = |a_0^{n/d}| < 1$, we prove that $|\alpha|'_L < 1$ as well. We claim that $a_i \in \mathfrak{m}_K$ for all $0 \leq i \leq n-1$. Otherwise, there would exist some j with $1 \leq j \leq n-1$ and $|a_j| = 1$. We may assume j is minimal with this property. Then the reduction of $f(x)$ is

$$\bar{f}(x) = x^n + \bar{a}_{n-1}x^{n-1} + \cdots \bar{a}_j x^j = x^j(x^{n-j} + \cdots + \bar{a}_j).$$

Now applying Hensel's Lemma 8.4.1 to $\bar{g} = x^j$, we get a factorization of $f(x)$ in $\mathcal{O}_K[x]$, which contradicts with the irreducibility of $f(x)$. By Lemma 8.2.6, we see that $|\alpha|'_L < 1$. Now if $|\alpha|_L > 1$, then applying the previous discussion to α^{-1}, we see that $|\alpha|'_L > 1$.

Therefore, this proves that $|\alpha|_L \leq 1$ if and only if $|\alpha|'_L \leq 1$, i.e. $|\cdot|_L$ and $|\cdot|'_L$ give rise to the same valuation ring on L. By Proposition 8.2.13, $|\cdot|_L$ and $|\cdot|'_L$ are equivalent. Since they both extend the absolute value $|\cdot|$ on K, these two valuations are actually the same.

Finally, the completeness of L with respect to the norm $|\cdot|_L$ follows from the Lemma 8.5.3 below.

□

Definition 8.5.2. — Let V be a vector space over $(K, |\cdot|)$. Then a (ultra-metric) *norm* on V is a map $\|\cdot\| : V \to \mathbb{R}_{\geq 0}$ such that

(1) $\|x\| = 0$ if and only if $x = 0$;
(2) $\|\lambda x\| = |\lambda|\|x\|$ for $\lambda \in K$ and $x \in V$;
(3) $\|x + y\| \leq \max\{\|x\|, \|y\|\}$.

We say two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on V are equivalent if there exist constants $C_1, C_2 > 0$ such that

$$C_1\|x\|_1 \leq \|x\|_2 \leq C_2\|x\|_1.$$

Equivalent norms on a vector space V define the same topology. Thus, V is complete with respect to a norm $\|\cdot\|_1$ if and only if so it is with respect to any norm equivalent to $\|\cdot\|_2$.

Lemma 8.5.3. — *Let V be a finite dimensional vector space over $(K, |\cdot|)$. Then any two norms on V are equivalent, and V is complete (with respect to any norm).*

Proof. — Let $(v_1, \cdots, v_n)$ be a basis of V over K. For $x = \sum_{i=1}^n a_i v_i$, we define

$$\|x\| = \max_{1 \leq i \leq n} \{|a_i|\}.$$

It is clear that V is complete under $\|\cdot\|$. It suffices to prove that every norm $\|\cdot\|'$ on V is equivalent to $\|\cdot\|$. Indeed, if $C_2 = \max_{1 \leq i \leq n} \|v_i\|'$, then

$$\|\sum_{i=1}^n a_i v_i\|' \leq C_2 \max\{|a_i|\} = C_2\|x\|.$$

To find $C_1 > 0$ such that $\|x\|' \geq C_1\|x\|$, we proceed by induction on $n \geq 1$. When $n = 1$, the assertion is trivial. Assume now $n \geq 2$ and the assertion is true for $n - 1$. Then, for $1 \leq i \leq n$, the vector space

$$V_i = Kv_1 + \cdots Kv_{i-1} + Kv_{i+1} + \cdots + Kv_n$$

is complete under the norm $\|\cdot\|'$ by induction hypothesis. Hence, $v_i + V_i$ is a closed in V with respect to the topology defined by $\|\cdot\|'$. As $0 \notin v_i + V_i$, there exists an $\epsilon > 0$ such that

$$U(0, \epsilon) = \{x \in K, \|x\|' < \epsilon\}$$

is disjoint with $v_i + V_i$, i.e. $\|x + v_i\|' \geq \epsilon$ for any $x \in V_i$. Now for $x = \sum_{i=1}^n a_i v_i$, suppose that $\|x\| = |a_r| \neq 0$ for some r. Then

$$\|x\|' = |a_r|\|\frac{a_1}{a_r}v_1 + \cdots + \frac{a_{r-1}}{a_r}v_{r-1} + v_r + \cdots + \frac{a_n}{a_r}v_n\|' \geq \epsilon|a_r| = \epsilon\|x\|.$$

We can take thus $C_1 = \epsilon$. □

We can also state the results of Theorem 8.5.1 in terms of additive valuation. If v denotes an additive valuation associated to $|\cdot|$, then for any finite extension L/K of degree n, then the unique extension of v is given by

$$v_L(x) = \frac{1}{n} v(\mathrm{N}_{L/K}(x)).$$

If v is a normalized discrete valuation, then v_L is also discrete, but not necessarily normalized. Note also that if x' is a Galois conjugate of $x \in L$, then $v_L(x') = v_L(x)$.

8.5.4. Newton Polygon. — Assume now K is a discrete valuation field with normalized additive valuation v. Let $\overline{K}$ be an algebraic closure of K. For simplicity, we still use v to denote the unique extension of v to $\overline{K}$. Consider a polynomial

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0 \in K[x]$$

with $a_n \neq 0$. We define the Newton polygon of $f(x)$, denoted by $\mathrm{NP}(f(x))$, as the lower convex envelop in $\mathbb{R}^2$ of the points

$$\{(i, v(a_i)) | 1 \leq i \leq n\}.$$

Assume that the breaking points of $\mathrm{NP}(f(x))$ are

$$(q_0, t_0) = (0, v(a_0)),\ (q_1, t_1),\ (q_2, t_2),\ \cdots,\ (q_r, t_r) = (n, v(a_n)) \in \mathbb{Z}^2.$$

For each j with $1 \leq j \leq r$, put

$$s_j = \frac{t_{j-1} - t_j}{q_j - q_{j-1}}$$

Then $s_1 > s_2 > \cdots > s_r$, and they are *negative* of the slopes of the Newton polygon of $f(x)$. We usually call the s_j's *slopes* of $f(x)$, and call $m_j := q_j - q_{j-1}$ the multiplicity of s_j.

Proposition 8.5.5. — *For each j, $f(x)$ has exactly m_j roots in $\overline{K}$ with valuation s_j.*

Proof. — Up to dividing $f(x)$ by a_0, we may assume that $a_0 = 1$. Let $\alpha_1, \cdots, \alpha_n \in \overline{K}$ be such that

$$f(x) = (1 - \alpha_1 x)(1 - \alpha_2 x) \cdots (1 - \alpha_n x).$$

Denote by $\rho_1 < \rho_2 < \cdots < \rho'_{r'}$ be the distinct valuations of the α_i's, and let m'_j with $1 \leq j \leq r'$ be the number of α_i's such that $v(\alpha_i) = \rho'_j$. Put $q'_0 = 0$ and $q'_j = \sum_{i=1}^{j} m'_j$ for $j \geq 1$ so that $m'_j = q'_j - q'_{j-1}$. We label the α_i's such that, for each j with $1 \leq j \leq r'$, we have

$$v(\alpha_i) = \rho'_j \quad \text{for } 1 + q'_{j-1} \leq i \leq q'_j.$$

Then one has

$$a_i = (-1)^i \sum_{1 \leq j_1 < \cdots < j_i \leq n} \alpha_{j_1} \alpha_{j_2} \cdots \alpha_{j_i}.$$

Then, one has

$$v(a_{q'_j}) = v(\alpha_1 \alpha_2 \cdots \alpha_{q_j}) = \sum_{\ell}^{j} \rho'_\ell m'_\ell,$$

and for $q_{j-1} < i \leq q_j$, one has

$$v(a_i) \geq \sum_{\ell=1}^{j-1} \rho'_\ell m'_\ell + (i - q_{j-1})\rho'_j.$$

Thus, $\mathrm{NP}(f(x))$ has breaking points $(q'_j, \sum_\ell^j \rho'_\ell m'_\ell)$, and slopes ρ'_j. Note that the roots of $f(x)$ are the α_i^{-1}'s. Hence, their valuations are $s_j = -\rho'_j$ for $1 \leq j \leq r$, and each s_j appear exactly $m_j = m'_j$ times. □

We have the following immediate

Corollary 8.5.6. — *If $f(x)$ is irreducible in $K[x]$, then $\mathrm{NP}(f(x))$ has only one slope. Conversely, if $f(x)$ has only one slope $s = \frac{t}{n}$ with $\gcd(t, n) = 1$ and $n = \deg(f)$, then $f(x)$ is irreducible in $K[x]$. In particular, if $f(x)$ is an Eisenstein polynomial, $f(x)$ is irreducible.*

Proof. — The first part follows from the Proposition 8.5.5 and the fact that all the Galois conjugate of an element $\alpha \in \overline{K}$ has the same valuation. For the second part, note that $NP(f(x))$ is given by the segment $y = -sx$ with $0 \leq x \leq n$ by assumption. If $f(x)$ were not irreducible, then $NP(f(x))$ will pass other integral points except the two ends. Since $\gcd(n, t) = 1$, this is not the case. For the second part, if $f(x)$ is Eisenstein, then $\frac{1}{n}$ is the only slope of $f(x)$; hence $f(x)$ is irreducible. □

Example 8.5.7. — We give an example on the application of Newton polygons to the irreducibility of rational polynomials. Consider the polynomial

$$f(x) = 1 + x + \frac{x^2}{2} + \frac{x^3}{3} + \frac{x^4}{4} + \frac{x^5}{5} + \frac{x^6}{6} \in \mathbb{Q}[x].$$

Regarding $f(x)$ as a polynomial in $\mathbb{Q}_5[x]$, then the 5-adic slopes of $f(x)$ are $1/5$ and -1. Therefore, $f(x)$ is a product of an irreducible polynomial of degree 5 with a factor of degree 1 in $\mathbb{Q}_5[x]$. If $f(x)$ were not irreducible in $\mathbb{Q}[x]$, then the only possible decomposition for $f(x)$ would be also of such form, that is, $f(x)$ would have a rational root. However, by checking the 2-adic slopes of $\mathbb{Q}_2[x]$, we see easily that $f(x)$ does not has any root in $\mathbb{Q}_2$. We thus conclude that $f(x)$ is irreducible in $\mathbb{Q}[x]$.

8.6. Krasner's Lemma and applications

In this section, let $(K, |\cdot|)$ be a complete non-archimedean valuation field. We fix an algebraic closure $\overline{K}$ of K, and extend $|\cdot|$ to $\overline{K}$ by Theorem 8.5.1.

***Lemma 8.6.1* (Krasner's Lemma)**. — *Let $\alpha, \beta \in \overline{K}$. Assume that*

$$|\beta - \alpha| < |\beta - \beta'|$$

for any Galois conjugate β' of β different from β. Then one has $\beta \in K(\alpha)$.

Proof. — By Galois theory, it suffices to show that for any K-automorphism σ of $\overline{K}$, if $\sigma(\alpha) = \alpha$, then $\sigma(\beta) = \beta$. Actually, if σ is such an automorphism, then

$$|\sigma(\beta) - \beta| = |\sigma(\beta) - \sigma(\alpha) + \alpha - \beta| \leq \max\{|\sigma(\beta - \alpha)|, |\beta - \alpha|\} = |\beta - \alpha|,$$

where the last step uses the invariance of norms under Galois conjugation. By hypothesis, one must have $\sigma(\beta) = \beta$. □

Let $\|\cdot\|$ denote the Gauss norm (8.4.0.1) on $K[x]$.

Theorem 8.6.2. — *Let $f(x) \in K[x]$ an irreducible monic polynomial of degree n. Put*

$$d_0 = \min_{\alpha \neq \alpha'}\{|\alpha - \alpha'|\},$$

where α, α' run through the distinct roots of $f(x)$. For any real number ϵ with $0 < \epsilon < d_0$, there exists $\delta > 0$ satisfying the following property: if $g(x) \in K[x]$ is a monic polynomial of degree n with $\|f - g\| < \delta$, then there exists an ordering of the roots $\alpha_1, \cdots, \alpha_n$ of $f(x)$ and $\beta_1, \cdots, \beta_n$ of $g(x)$ respectively, such that $|\alpha_i - \beta_i| < \epsilon$, and $K(\alpha_i) = K(\beta_i)$ as subfields of $\overline{K}$; in particular, $g(x)$ is irreducible.

Proof. — First, note that if $h(x) = x^n + \sum_{i=0}^{n-1} c_i x^n$ is a monic polynomial, then its roots can be bounded above in terms of $\|h\|$. Actually, if γ is a root of $h(x)$, then by Lemma 8.2.6, there exists j with $0 \leq j \leq n-1$ such that $|c_j \gamma^j| \geq |\gamma^n|$. Hence, one obtains

$$|\gamma| \leq \max_{0 \leq j \leq n-1} |c_j|^{1/(n-j)} \leq \max_{0 \leq j \leq n-1} \|h\|^{1/(n-j)}.$$

Now for any $\delta > 0$, if $g \in K[x]$ is a monic polynomial of degree n with $\|f - g\| < \delta$, then $\|g\| \leq \max\{\|f\|, \delta\}$ is bounded. Hence, if $\beta \in \overline{K}$ is a root of $g(x)$, then there exists a constant $C_0 > 0$ (depending only $\|f\|$) such that $|\beta| \leq C_0$. Hence, there exists thus a constant C_1 such that

$$\prod_{\alpha} |\beta - \alpha| = |f(\beta)| = |f(\beta) - g(\beta)| \leq C_1 \|f - g\| < C_1 \delta.$$

It follows that $\min_\alpha\{|\beta - \alpha|\}$, where α runs over all roots of f, tends to 0 when $\delta \to 0$. Thus, when δ is sufficiently small, one has

$$\min_{\alpha}\{|\beta - \alpha|\} < \epsilon.$$

Since $\epsilon < d_0$, Lemma 8.2.6 implies that there exists a unique root $\alpha(\beta)$ of $f(x)$ such that $|\alpha(\beta) - \beta| < \epsilon$. By Krasner's Lemma 8.6.1, we have $\alpha(\beta) \in K(\beta)$. Since $\alpha(\beta)$ has degree n, it follows that $K(\beta) = K(\alpha(\beta))$ and hence $g(x)$ is irreducible.

It remains to show that, when δ is sufficiently small, the map $\beta \mapsto \alpha(\beta)$ induces a bijection between the roots of $g(x)$ and those of $f(x)$. It suffices to show that this map is injective, i.e. $\alpha(\beta') \neq \alpha(\beta)$ for every root β' of $g(x)$ different from β. Suppose in contrary that $\alpha_0 = \alpha(\beta) = \alpha(\beta')$. Then

$$|\beta - \beta'| \leq \max\{|\beta - \alpha_0|, |\beta' - \alpha_0|\} < \epsilon.$$

It follows that

$$|g'(\beta)| = \prod_{\beta'' \neq \beta} |\beta - \beta''| = (\prod_{\beta'' \neq \beta', \beta} |\beta - \beta''|) \cdot (|\beta - \beta'|) < C_0^{n-1}\epsilon. \tag{8.6.2.1}$$

On the other hand, there exists a constant $C_2 > 0$ (depending only on $\|f\|$) such that

$$|g'(\beta) - f'(\beta)| \leq C_2\|f - g\| < C_2\delta.$$

Note that

$$|f'(\beta)| \leq \max\{|f'(\beta) - f'(\alpha(\beta))|, |f'(\alpha(\beta))|\},$$

where the first term in maximum is tending to 0 as $\delta \to 0$, and $|f'(\alpha(\beta))|$ is bounded below independent of δ. Therefore, when δ is sufficiently small, we see by Lemma 8.2.6 that

$$|g'(\beta)| = \max\{|g'(\beta) - f'(\beta)|, |f'(\beta) - f'(\alpha(\beta))|, |f'(\alpha(\beta))|\} = |f'(\alpha(\beta))|,$$

which contradicts with (8.6.2.1). This finishes the proof of Theorem 8.6.2.

□

Corollary 8.6.3. — *Let $L/\mathbb{Q}_p$ be a finite extension, then there exists a monic polynomial $g(x) \in \mathbb{Q}[x]$ which is irreducible in $\mathbb{Q}_p[x]$ and $L \cong \mathbb{Q}_p[x]/(g(x))$.*

Proof. — Assume $L \cong \mathbb{Q}_p[x]/(f(x))$ some monic irreducible polynomial $f \in \mathbb{Q}_p[x]$. By Theorem 10.1.5, if $g \in \mathbb{Q}[x]$ is sufficiently close to f in the Gauss norm, then $\mathbb{Q}_p[x]/(f(x)) \cong \mathbb{Q}_p[x]/(g(x))$. □

CHAPTER 9

FINITE EXTENSIONS OF COMPLETE DISCRETE VALUATION FIELDS

In this chapter, let K be a complete discrete valuation field, and $v_K : K \to \mathbb{Z} \cup \{\infty\}$ be the normalized additive valuation on K. Let $\mathcal{O}_K$ denote the valuation ring of K, $\mathfrak{m}_K \subseteq \mathcal{O}_K$ the maximal ideal and $k = \mathcal{O}_K/\mathfrak{m}_K$. We fix also a uniformizer π_K of K so that $\mathfrak{m}_K = (\pi_K)$. Fix an algebraic closure $\overline{K}$ of K, and denote still by v_K the unique extension of v_K to $\overline{K}$ by Theorem 8.5.1. All finite extensions of K are considered as subfields of $\overline{K}$.

If L/K is a finite extension, then L is also a complete discrete valuation field by Theorem 8.5.1. We denote usually by $\mathcal{O}_L$ the valuation ring of L, by $\mathfrak{m}_L$ its maximal ideal, and $k_L = \mathcal{O}_L/\mathfrak{m}_L$. Let v_L denote the normalized additive valuation on L.

9.1. Generalities

Let L/K be a finite extension of degree n, and $\mathcal{O}_L$ denote the integral closure of $\mathcal{O}_K$ in L. Then $\mathcal{O}_L$ is the valuation ring of L for the unique extension of v_K to L.

Lemma 9.1.1. — *Then ring $\mathcal{O}_L$ is a finite free $\mathcal{O}_K$-module of rank $n = [L : K]$.*

Proof. — For an element $x \in \mathcal{O}_L$, denote by $\bar{x}$ its image in $\mathcal{O}_L/\pi_K\mathcal{O}_L$. Choose a subset $\{b_i : i \in I\}$ of $\mathcal{O}_L$ such that $(\bar{b}_i)_{i\in I}$ is a basis of $\mathcal{O}_L/\pi_K\mathcal{O}_L$ over the field $k = \mathcal{O}_K/\pi_K\mathcal{O}_K$. First, we claim that the b_i's with $i \in I$ are linearly independent over K. Indeed, if $\sum_{i\in I} a_i b_i = 0$ is non-trivial linear relation with $a_i \in K$. Up to multiplying a power of π_K, we may assume that each $a_i \in \mathcal{O}_K$ and the reductions $\bar{a}_i \in k$ are not all zero. Then $\sum_{i\in I} \bar{a}_i\bar{b}_i = 0$ is a non-trivial linear relation of the $\bar{b}_i$'s, which contradicts with the choice of b_i. This proves the claim, which implies immediately that I is a finite set with cardinality at most n.

Secondly, we prove that the b_i's generate $\mathcal{O}_L$ as an $\mathcal{O}_K$-module, i.e. $\mathcal{O}_L$ is a free $\mathcal{O}_K$-module with basis $(b_i)_{i\in I}$. Indeed, for any $x \in \mathcal{O}_L$, there exists a $x_1 \in \mathcal{O}_L$ and $a_i^{(0)} \in \mathcal{O}_K$ such that

$$x = \sum_{i\in I} a_i^{(0)} b_i + \pi_K x_1,$$

since $(\bar{b}_i)_{i\in I}$ form a k-basis of $\mathcal{O}_L/\pi_K\mathcal{O}_L$. Repeating this process, we see that x writes as

$$x = \sum_{i\in I}(a_i^{(0)} + \pi_K a_i^{(1)} + \cdots + \pi_K^{n-1}a_i^{n-1})b_i + \pi_K^n x_n.$$

Since $\mathcal{O}_L$ is complete, we get $x \in \sum_{i\in I} b_i\mathcal{O}_K$ by letting $n \to +\infty$. Note that for every $y \in L$, there exists $m \geq 1$ such that $\pi_K^m y \in \mathcal{O}_L$. Therefore, $(b_i)_{i\in I}$ is actually a basis of L over K. Hence, I has cardinality n.

□

Remark 9.1.2. — In the proof of Lemma 9.1.1, we see that *n elements $b_1, \cdots, b_n \in \mathcal{O}_L$ form a basis of $\mathcal{O}_L$ over $\mathcal{O}_K$ if and only if their reduction modulo $\pi_K\mathcal{O}_L$ form a basis of $\mathcal{O}_L/\pi_K\mathcal{O}_L$ over k*. This gives a very convenient way to construct basis of $\mathcal{O}_L$ over $\mathcal{O}_K$.

9.1.3. Ramification index and residue degree. — By construction, for every $x \in L$, the unique extension of v to L is given by

$$v_K(x) := \frac{1}{n}v_K(\mathrm{N}_{L/K}(x)).$$

Therefore, there exists an integer $e = e(L|K) \geq 1$ dividing n such that

$$v_K(L^\times) = \frac{1}{e}\mathbb{Z},$$

or equivalently, e is the integer such that $\pi_K = u\pi_L^e$, where π_L denotes a uniformizer of K (resp. of L), and u is a unit in $\mathcal{O}_L$. We call e the *ramification index* of L/K. If $e = 1$, we say that the extension L/K is *unramified*.

Let k_L be the residue field of $\mathcal{O}_L$. Then k_L/k is a finite extension (since $\mathcal{O}_L/\pi_K\mathcal{O}_L$ has dimension n over k by Lemma 9.1.1). We call the integer

$$f(L|K) := [k_L : k]$$

the residue degree of the finite extension L/K. If $f(L|K) = 1$, we say that L/K is *totally ramified*.

Proposition 9.1.4. — (1) *We have $e(L|K)f(L|K) = n$.*

(2) *If $\alpha_1, \cdots, \alpha_f$ are elements of $\mathcal{O}_L$ such that their reduction in k_L form a basis of k_L over k, then $\{\alpha_i\pi_L^{j-1} : 1 \leq i \leq f = f(L|K), 1 \leq j \leq e = e(L|K)\}$ form a basis of $\mathcal{O}_L$ over $\mathcal{O}_K$.*

Proof. — (1) By Lemma 9.1.1, $\mathcal{O}_L/\pi_K\mathcal{O}_L$ has dimension n over k. On the other hand, we have a filtration

$$\mathcal{O}_L/\pi_K\mathcal{O}_L = \mathcal{O}_L/(\pi_L^e) \supsetneq (\pi_L)/(\pi_L^e) \supsetneq \cdots \supsetneq (\pi_L^{e-1})/(\pi_L^e) \supsetneq (0),$$

where each sub-quotient is one-dimensional k_L-vector space. Therefore,

$$n = \dim_k(\mathcal{O}_L/\pi_K\mathcal{O}_L) = e\dim_k k_L = ef.$$

(2) Fix a set of representatives $\{\lambda_i : i \in I\} \subseteq \mathcal{O}_K$ of k. Then $S = \{\lambda_i \alpha_j : i \in I, 1 \leq j \leq f\} \subseteq \mathcal{O}_L$ form a set of representatives of k_L. By Proposition 8.3.1, every element of $\mathcal{O}_L$ writes uniquely as

$$x = \sum_{n \geq 0} c_n \pi_L^n \quad \text{with } c_n \in S.$$

In particular, the reductions modulo $\pi_K \mathcal{O}_L = \pi_L^e \mathcal{O}_L$ of $\{\alpha_i \pi_L^{j-1} : 1 \leq i \leq f, 1 \leq j \leq e\}$ generate $\mathcal{O}_L/\pi_K \mathcal{O}_L$ over k. We conclude by Remark 9.1.2.

□

9.1.5. Ramified extension. — Assume that L/K is totally ramified of degree n. Let π_L denote a uniformizer of L, then one has $v_K(\pi_L) = \frac{1}{n}$. If

$$f(x) = \sum_{i=0}^{n} a_i x^i \in \mathcal{O}_K[x]$$

is the monic minimal polynomial of π_L, then $v_K(a_i) \geq 1$ and

$$v_K(a_0) = v_K(\mathrm{N}_{L/K}(\pi_L)) = 1,$$

i.e. $f(x)$ is an Eisenstein polynomial. In particular, $\mathrm{N}_{L/K}(\pi_L)$ is a uniformizer of K. By Proposition 9.1.4(2), we have $\mathcal{O}_L = \mathcal{O}_K[\pi_L]$.

9.2. Unramified extensions

If $\iota : L \hookrightarrow L'$ is a K-embedding of two finite extensions of K, then ι sends $\mathcal{O}_L$ into $\mathcal{O}_{L'}$ and $\mathfrak{m}_L$ into $\mathfrak{m}_{L'}$ respectively, hence it induces an embedding of residues fields $k_L \hookrightarrow k_{L'}$.

Theorem 9.2.1. — *Assume k'/k is a finite separable extension. Then the following assertions hold:*

(1) *There exists an unramified extension K'/K with residue field k'. Moreover, this extension is unique up to isomorphisms, and it is Galois if and only if k'/k is Galois.*
(2) *For any finite extension L/K with residue field k_L, there exists a natural bijection (induced by reduction) between the set of K-embeddings of K' into L and the set of k-embeddings of k' into k_L. In particular, if k'/k is Galois, we have* $\mathrm{Gal}(K'/K) \cong \mathrm{Gal}(k'/k)$.

Proof. — (1) We prove first the existence of K'. We may assume that $k' \cong k[x]/(\bar{f}(x))$ for some irreducible monic polynomial $\bar{f}(x) \in k[x]$ of degree n. Take a monic polynomial $f(x) \in \mathcal{O}_K[x]$ of degree n such that $f \mod \mathfrak{m}_K = \bar{f}$. Then $f(x)$ is necessarily irreducible, and we claim that $K' = K[x]/(f(x))$ satisfies the required property. Let α denote the image of x in K', and $\bar{\alpha}$ be its reduction modulo $\mathfrak{m}_K$. First, we note that $v(\alpha) = 0$, i.e. the image of α in the residue filed of $\mathcal{O}_{K'}$ is non-zero. Hence, the residue field of K' contains $k[\bar{\alpha}] = k'$. Thus, we get $f(K'|K) \geq n$. By Proposition 9.1.4, we see that $k_{K'} = k'$

and $e(K'|K) = 1$. Note that $1, \bar{\alpha}, \cdots, \bar{\alpha}^{n-1}$ are linearly independent over k. For rank reasons, we have

$$\mathcal{O}_{K'}/(\pi_K) \cong \sum_{i=1}^{n} k\bar{\alpha}^{i-1}.$$

By Remark 9.1.2, we have $\mathcal{O}_{K'} = \mathcal{O}_K[\alpha]$.

(2) We now prove that K' satisfies the property in (2) for any finite extension L/K. The rest part of the Theorem will be an easy consequence of this property. Let L/K be as in Statement (2). Let $S(L)$ denote the set of roots of $f(x)$ in L, and $S(k_L)$ the set of roots of $\bar{f}(x)$ in k_L. Then the set of K-embeddings of $K' = K[\alpha]$ into L is in natural bijection with S, while the set of k-embeddings of k' into k_L is in bijection with $S(k_L)$. Since $\bar{f}(x)$ is separable, Hensel's Lemma 8.4.1 implies that the reduction map from $S(L)$ to $S(k_L)$ is bijective. Now the assertion follows immediately. □

Corollary 9.2.2. — *Let L/K be a finite extension such that the residue extension k_L/k is separable. Then there exists a unique unramified sub-extension L_0/K of L/K with residue field k_L such that all unramified sub-extension K'/K of L/K is contained in L_0. In particular, if L/K is a normal extension, then L_0/K is also normal.*

Proof. — Let L_0 be the unramified extension of K given by Theorem 9.2.1 with residue extension k_L/k. We see that L_0/K is a sub-extension of L/K by applying the second part of Theorem 9.2.1 to the identity embedding $k_L \cong k_L$. The fact that L_0 contains all unramified sub-extension of L/K also follows easily. □

We call L_0 as in Corollary above *maximal unramified sub-extension of L*. It is clear that $[L_0 : K] = f(L|K)$, and L/L_0 is totally ramified of degree $e(L|K)$. Using L_0/K, the study of a general finite extension L/K can be reduced to the cases of totally ramified and unramified cases.

9.3. Different, discriminant and ramification

In this section, let L/K be a finite separable extension such that the residue extension k_L/k is also separable. Denote by π_L (resp. π_K) a uniformizer of L (resp. of K).

9.3.1. Norms of fractional ideals. — Note that both $\mathcal{O}_L$ and $\mathcal{O}_K$ are Dedekind domains. We have the notion of fractional ideals on L or on K. If $\mathfrak{a}$ is a fractional ideal of L, we define

$$v_L(\mathfrak{a}) = \min_{x \in \mathfrak{a}} \{v_L(x)\} \in \mathbb{Z},$$

and call it the *valuation* of $\mathfrak{a}$. Then it is clear that $\mathfrak{a} = (\pi_L^{v_L(\mathfrak{a})})$ We define the norm of $\mathfrak{a}$ as the fractional ideal of K given by

$$\mathrm{N}_{L/K}(\mathfrak{a}) := (\mathrm{N}_{L/K}(\pi_L))^{v_L(\mathfrak{a})}.$$

Lemma 9.3.2. — *We have $v_K(\mathrm{N}_{L/K}(\mathfrak{a})) = f(L|K)v_L(\mathfrak{a})$.*

Proof. — It suffices to show that $v_K(\pi_L) = f(L|K)$. Let L_0/K denote the maximal unramified extension of L/K. Then $\mathrm{N}_{L/K}(\pi_L) = \mathrm{N}_{L_0/K}(\mathrm{N}_{L/L_0}(\pi_L))$. Since L/L_0 is totally ramified, we see that $\mathrm{N}_{L/L_0}(\pi_L)$ is a uniformizer of L_0. Thus it suffices to show that for any uniformizer π_{L_0} of L_0, we have

$$v_K(\mathrm{N}_{L_0/K}(\pi_{L_0})) = f(L|K).$$

As L_0/K is unramified, one has $\pi_{L_0} = \pi_K u$ for some unit $u \in \mathcal{O}_{L_0}^\times$. Thus, we get

$$v_K(\mathrm{N}_{L_0/K}(\pi_{L_0})) = v(\mathrm{N}_{L_0/K}(\pi_K)) = f(L_0|K) = f(L|K).$$

□

9.3.3. Different and discriminant. — The theory of different and discriminant for number fields has an analog for L/K. Recall that the bilinear form $\mathrm{Tr}_{L/K}(xy)$ on L is non-degenerate by Theorem 1.2.4. We put

$$\mathcal{O}_L^* := \{x \in L : \mathrm{Tr}_{L/K}(xy) \in \mathcal{O}_K, \forall y \in \mathcal{O}_L\}.$$

Then $\mathcal{O}_L^*$ is a *fractional ideal* of L. It is clear that $\mathcal{O}_L \subseteq \mathcal{O}_L^*$. We define the *different* of L/K (or of $\mathcal{O}_L/\mathcal{O}_K$) as the ideal in $\mathcal{O}_L$

$$\delta_{L/K} := (\mathcal{O}_L^*)^{-1}. \tag{9.3.3.1}$$

and the discriminant of L/K (or of $\mathcal{O}_L/\mathcal{O}_K$) as the ideal in $\mathcal{O}_K$

$$\mathfrak{d}_{L/K} := \mathrm{N}_{L/K}(\delta_{L/K}). \tag{9.3.3.2}$$

Similar properties as in Section 3.3 hold in our case. In particular, we have

Proposition 9.3.4. — *Let K'/K be a sub-extension of L/K. Then we have*

$$\delta_{L/K} = (\delta_{K'/K}\mathcal{O}_L) \cdot \delta_{L/K'},$$

and

$$\mathfrak{d}_{L/K} = \mathrm{N}_{K'/K}(\mathfrak{d}_{L/K'})\mathfrak{d}_{K'/K}^{[L:K']}.$$

Proof. — The proof is exactly the same as Proposition 3.3.5 and Corollary 3.3.6. □

Applying this Proposition with K' equal to the maximal unramified sub-extension of L/K, we reduce the problem of computing $\delta_{L/K}$ to the case of L_0/K and L/L_0, i.e. it suffices to treat separately the unramified case and the totally ramified case.

Proposition 9.3.5. — *Assume that there exists an $\alpha \in \mathcal{O}_L$ such that $\mathcal{O}_L = \mathcal{O}_K[\alpha]$. Let $f(x) \in \mathcal{O}_K[x]$ be the monic minimal polynomial of α. Then we have $\delta_{L/K} = (f'(\alpha))$.*

We start with the following elementary

Lemma 9.3.6 **([Se68], Chap. III, §6 Lemme 2)**. — *If $f(x)$ has degree n, then we have*

$$\mathrm{Tr}_{L/K}\left(\frac{\alpha^i}{f'(\alpha)}\right) = \begin{cases} 0 & \text{if } 0 \leq i \leq n-2 \\ 1 & \text{if } i = n-1. \end{cases}$$

We now return to the proof of Proposition 9.3.5.

Proof of 9.3.5. — We have to prove that $\mathcal{O}_L^* = \frac{1}{f'(\alpha)}\mathcal{O}_L$, i.e. $\alpha^{i-1}/f'(\alpha)$ for $1 \leq i \leq n$ form a basis of $\mathcal{O}_L^*$ over $\mathcal{O}_K$. It suffices to show that the matrix $a_{i,j} = \mathrm{Tr}_{L/K}(\alpha^{i-1}\alpha^{j-1}/f'(\alpha))$ is invertible in $\mathrm{GL}_n(\mathcal{O}_K)$. But Lemma 9.3.6 implies that $a_{i,j} = 0$ if $i+j \leq n$, and $a_{i,j} = 1$ if $i+j = n+1$. A simple computation shows that $\det(a_{i,j}) = (-1)^{\frac{n(n-1)}{2}}$. □

Proposition 9.3.7. — (1) *Assume that L/K is totally ramified of ramification index e. Then*

$$v_L(\delta_{L/K}) \geq e-1,$$

and the equality holds if and only if e is prime to the characteristic of k.

(2) *The finite extension L/K is unramified if and only if $v_L(\delta_{L/K}) = 0$.*

Proof. — (1) Let π_L denote a uniformizer of L, and let

$$f(x) = x^e + a_{e-1}x^{e-1} + \cdots + a_0 \in \mathcal{O}_K[x]$$

be the minimal polynomial of π_L. By Subsection 9.1.5, $f(x)$ is an Eisenstein polynomial, $\mathcal{O}_L = \mathcal{O}_K[\pi_L]$. It follows thus from Proposition 9.3.5 that $\delta_{L/K} = (f'(\pi_L))$. Hence, we get

$$v_L(\delta_L) = v_L(f'(\pi_L)) = v_L(e\pi_L^{e-1} + a_{e-1}(e-1)\pi_L^{e-2} + \cdots + a_1).$$

As $\pi_K | a_i$ for all $0 \leq i \leq e-1$, we have

$$v_L(ia_i\pi_L^{i-1}) \geq i - 1 + v_L(ia_i) \geq e + i - 1, \quad \text{for } 1 \leq i \leq e-1,$$

and $v_L(e\pi_L^{e-1}) = e - 1 + v_L(e)$. Therefore, we have

$$v_L(f'(\pi_L)) \geq \min_{1 \leq i \leq e}\{v_L(ia_i\pi_L^{i-1})\} \geq e-1.$$

If e is prime to the residue characteristic, we have $v_L(e) = 0$ and $v_L(ia_i\pi_L^{i-1} > v_L(e\pi_L^{e-1}) = e-1$; hence, $v_L(f'(\pi)) = e-1$ by Lemma 8.2.6.

(2) Assume first that L/K is unramified. Then there exists a monic irreducible polynomial $f(x) \in \mathcal{O}_K[x]$ such that its reduction $\bar{f}(x) \in k[x]$ is also irreducible and $\mathcal{O}_L = \mathcal{O}_K[x]/(f(x))$. Let $\alpha \in \mathcal{O}_L$ denote the image of x. Then $\bar{f}'(\alpha) \neq 0$, and hence

$$v_L(\delta_{L/K}) = v_L(f'(\alpha)) = 0.$$

Suppose conversely that $v_L(\delta_{L/K}) = 0$. Let L_0 denote the maximal unramified subextension of L/K. By Proposition 9.3.4 and (1), we have

$$0 \geq v_L(\delta_{L/K}) = v_L(\delta_{L/L_0}) \geq e-1.$$

Thus, we get $e = 1$, i.e. L is unramified over K.

□

Definition 9.3.8. — We say the finite extension L/K is *tamely ramified* if its ramification index $e(L|K)$ is prime to the characteristic of k.

9.4. Galois extension of complete discrete valuation fields

We keep the notation of the previous section. Let L/K be a finite Galois extension of Galois group G such that the residue extension k_L/k is separable. Denote by L_0/K the maximal unramified sub-extension. We have seen in Corollary 9.2.2 that L_0/K is a Galois extension. By Theorem 9.2.1(1), the residue extension k_L/k is also Galois. The restriction to L_0 defines a natural surjective map

$$G \to \mathrm{Gal}(L_0/K) \xrightarrow{\sim} \mathrm{Gal}(k_L/k), \tag{9.4.0.1}$$

where the second isomorphism uses Theorem 9.2.1. We define the *inertia subgroup* of G, denoted by $I_{L/K}$ or simply I when no confusions arise, as the kernel of this map, or equivalently

$$I = \{\sigma \in G \mid \sigma(x) \equiv x \mod \mathfrak{m}_L\}. \tag{9.4.0.2}$$

It is clear that I is normal in G, and $I \cong \mathrm{Gal}(L/L_0)$ by (9.4.0.1).

In the rest of this section, we suppose that L/K is *totally ramified*, i.e. $G = I$. Denote by v_L the normalized additive valuation on L so that $v_L(L^\times) = \mathbb{Z}$.

Lemma 9.4.1. — *Let $\sigma \in G$. For any integer $n \geq 1$, the following two conditions are equivalent:*

(1) *For any $x \in \mathcal{O}_L$, we have $v_L(\sigma(x) - x) \geq n + 1$*
(2) *We have $v_L(\sigma(\pi_L) - \pi_L) \geq n + 1$, for any uniformizer π_L of L.*

Proof. — (1) $\Rightarrow$ (2) is trivial. We prove now that (2) $\Rightarrow$ (1). Indeed, every $x \in \mathcal{O}_L$ writes as

$$x = \sum_{i=0}^{+\infty} a_i \pi_L^i, \text{with } a_i \in \mathcal{O}_K.$$

It follows that

$$\sigma(x) - x = \sum_{i=1}^{+\infty} a_i(\sigma(\pi_L)^i - \pi_L^i),$$

which is divisible by $\sigma(\pi_L) - \pi_L$. Assertion (2) now follows immediately.

□

For any integer $n \geq 1$, we define G_n as the subgroup of $\sigma \in G$ which satisfies the equivalent conditions above. These groups are usually called *higher ramification subgroups* of G, and G_1 is also called the *wild inertia* subgroup of G. Note that

$$v_L(\tau^{-1}\sigma\tau(x) - x) = v_L(\sigma(\tau(x)) - \tau(x))$$

for any $\tau \in G$ and $x \in \mathcal{O}_L$. Hence, G_n is a normal subgroup of G by condition 9.4.1(1). We get thus a decreasing filtration

$$G_0 := G \supseteq G_1 \supseteq G_2 \supseteq \cdots \supseteq G_n \supseteq G_{n+1} \supseteq \cdots,$$

called the *(lower) ramification filtration on G*. We put $U_L^0 = \mathcal{O}_L^\times$, and

$$U_L^n = \{u \in \mathcal{O}_L^\times \mid u \equiv 1 \mod \pi_L^n\}$$

for an integer $n \geq 1$. Then the definition of G_n is equivalent to

$$G_n = \{\sigma \in G \mid \frac{\sigma(\pi_L)}{\pi_L} \in U_L^n\}.$$

Therefore, the map $\sigma \mapsto \sigma(\pi_L)/\pi_L \mod U_L^{n+1}$ induces a canonical injection

$$\theta_n : G_n/G_{n+1} \hookrightarrow U_n/U_{n+1} \cong \begin{cases} k_L^\times & \text{if } n = 0; \\ k_L & \text{if } n \geq 1. \end{cases} \tag{9.4.1.1}$$

In particular, the sub-quotient G_n/G_{n+1} is abelian.

Proposition 9.4.2. — (1) *If the characteristic of k_L is 0, then the wild inertia subgroup $G_1 = \{1\}$, and G_0/G_1 is a finite cyclic group.*
(2) *If the characteristic of k_L is $p > 0$, then G_1 is a finite group of p-power order, and G_0/G_1 is a cyclic group of order prime to p; in particular, the inertia group G is solvable.*

Proof. — Note that k_L has no finite subgroups if its characteristic is 0, and any finite subgroup of k_L must have the form $(\mathbb{Z}/p\mathbb{Z})^r$ if it has characteristic p. The assertions on G_1 in both (1) and (2) follow immediately. The assertion on G_0/G_1 is an immediate consequence of the fact that any finite subgroup of $k_L^\times$ is cyclic.

□

For more details on the ramification filtration on G, we refer the reader to [**Se68**, Chap. IV].

Example 9.4.3. — Let ζ_{p^n} be a primitive p^n-th root of unity. We consider the extension $L = \mathbb{Q}_p(\zeta_{p^n})$ over $\mathbb{Q}_p$. Then the minimal polynomial of $\zeta_{p^n} - 1$ over $\mathbb{Q}_p$ is

$$\begin{aligned} f(x) &= \frac{(x+1)^{p^n} - 1}{(x+1)^{p^{n-1}} - 1} = \sum_{i=0}^{p-1} (x+1)^{p^{n-1}i} \\ &= x^{p^{n-1}(p-1)} + \cdots + p \equiv x^{p^{n-1}(p-1)} \mod p, \end{aligned}$$

which is an Eisenstein polynomial over $\mathbb{Q}_p$. Therefore, $\mathbb{Q}_p(\zeta_{p^n})$ is totally ramified over $\mathbb{Q}_p$ of degree $p^{n-1}(p-1)$, and we have

$$v_p(\zeta_{p^n} - 1) = \frac{1}{p^{n-1}(p-1)}. \tag{9.4.3.1}$$

In particular, we see that $\zeta_{p^n} - 1$ is a uniformizer of $L = \mathbb{Q}(\zeta_{p^n})$. Here, v_p denotes the unique extension to L of the usual p-adic valuation on $\mathbb{Q}_p$.

The Galois group $G = \mathrm{Gal}(\mathbb{Q}_p(\zeta_{p^n})/\mathbb{Q}_p)$ is canonically isomorphic to $(\mathbb{Z}/p^n\mathbb{Z})^\times$, where, for each $a \mod p^n \in (\mathbb{Z}/p^n\mathbb{Z})^\times$, the corresponding element $\sigma_a \in G$ is defined by $\sigma_a(\zeta_{p^n}) = \zeta_{p^n}^a$. For each integer k with $1 \leq k \leq n-1$, let G^k denote the subgroup of G corresponding

to those $a \in (\mathbb{Z}/p^n\mathbb{Z})^\times$ with $a \equiv 1 \mod p^k$. We have $G^k \supseteq G^{k+1}$. If $\sigma_a \in G^k \backslash G^{k+1}$, then a writes as $a \equiv 1 + p^k b \mod p^n$ for some $b \in \mathbb{Z}_p^\times$ we have

$$v_p(\sigma_a(\zeta_{p^n} - 1) - (\zeta_{p^n} - 1)) = v_p(\zeta_{p^n}^a - \zeta_{p^n}) = v_p(\zeta_{p^{n-k}}^b - 1) = \frac{1}{p^{n-k-1}(p-1)},$$

where $\zeta_{p^{n-k}}^b = \zeta_{p^n}^{p^k b}$ is a primitive p^{n-k}-th root of unity, and the last equality follows from (9.4.3.1) with n replaced by $n-k$. If v_L denotes the normalized additive valuation on L, then $v_L = p^{n-1}(p-1)v_p$. Hence, it follows that

$$v_L(\sigma_a(\zeta_{p^n} - 1) - (\zeta_{p^n} - 1)) = p^k, \quad \text{for } \sigma_a \in G^k \backslash G^{k+1}.$$

Thus the lower ramification filtration of $G = \mathrm{Gal}(L/\mathbb{Q}_p)$ is given by

$$G_u = \begin{cases} G & \text{if } u = 0; \\ G^k & \text{if } p^{k-1} \leq u \leq p^k - 1 \text{ for some } k \text{ with } 1 \leq k \leq n-1; \\ \{1\} & \text{if } p^{n-1} \leq u. \end{cases}$$

CHAPTER 10

APPLICATIONS OF LOCAL METHODS TO NUMBER FIELDS

10.1. Norms and places on number fields

In this section, we will classify the norms on a number field. We start with $\mathbb{Q}$. Let $|\cdot|_\infty$ denote the usual real norm on $\mathbb{Q}$. For any rational prime p, let $|\cdot|_p$ denote the p-adic norm discussed in Subsection 8.1.4. A fundamental fact for norms on $\mathbb{Q}$ is the following

Theorem 10.1.1 **(Ostrowski)**. — *The norm $|\cdot|_p$ is not equivalent to $|\cdot|_q$ if $p \neq q$ with $p, q \leq \infty$. Every nontrivial norm $|\cdot|$ on $\mathbb{Q}$ is equivalent to $|\cdot|_p$ for some prime p or for $p = \infty$.*

Proof. — If one of p, q is ∞ (and the other one is finite), then it is clear that $|\cdot|_p$ is not equivalent to $|\cdot|_q$. If both p and q are finite primes, then $|p|_p = p^{-1}$ and $|p|_q = 1$. Therefore, $|\cdot|_p$ and $|\cdot|_q$ can not be equivalent.

Assume first that $|\cdot|$ is archimedean. By Proposition 8.2.5, $|\cdot|$ must be unbounded on $\mathbb{Z}$. Let $n_0 \geq 1$ be the first integer such that $|n_0| > 1$. Let $c \in \mathbb{R}_{>0}$ be such that $|n_0| = n_0^c$. We have to prove that $|n| = n^c$ for any positive integer n. Write

$$n = a_0 + a_1 n_0 + \cdots + a_s n_0^s, \quad \text{with } 0 \leq a_i < n_0, a_s \neq 0.$$

Then one has

$$\begin{aligned} |n| &\leq |a_0| + |a_1||n_0| + \cdots + |a_s||n_0|^s \\ &= |a_0| + |a_1| n_0^c + \cdots |a_s| n_0^{cs}. \end{aligned}$$

By our choice of n_0 and since $a_i < n_0$, we have $|a_i| \leq 1$. Hence,

$$|n| \leq 1 + n_0^c + \cdots + n_0^{cs} \leq n_0^{cs}(1 + n_0^{-c} + \cdots + n_0^{-cs}) \leq A n^c,$$

where A is some constant independent of n. Now replacing n by n^M and taking M-th radical, one gets

$$|n| \leq \sqrt[N]{A} n^c.$$

Letting $N \to +\infty$, one gets $|n| \leq n^c$. We now deduce an inequality in the other direction as follows. If one writes n in terms of n_0 as above, one has $n_0^{s+1} > n \geq n_0^s$. Thus, the

trigonometric inequality implies that

$$|n| \geq |n_0|^{s+1} - (|n_0^{s+1} - n|) \geq n_0^{c(s+1)} - (n_0^{s+1} - n)^c.$$

Since

$$(n_0^{s+1} - n)^c \leq (n_0^{s+1} - n_0)^c = n_0^{c(s+1)}(1 - \frac{1}{n_0})^c,$$

we get

$$|n| \geq n_0^{c(s+1)}(1 - (1 - \frac{1}{n_0})^c) \geq A' n^c$$

with $A' = 1 - (1 - \frac{1}{n_0})^c$. Replacing n by n^M and taking M-th root, we get

$$|n| \geq \sqrt[M]{A'} n^c.$$

Letting $M \to +\infty$, we get $|n| \geq n^c$. We conclude finally that $|n| = n^c$ if $|\cdot|$ is archimedean.

Now assume that $|\cdot|$ is non-archimedean. Then we have $|n| \leq 1$ for all $n \in \mathbb{Z}$. Let $\mathfrak{p} \subseteq \mathbb{Z}$ be the subset consisting of $n \in \mathbb{Z}$ with $|n| < 1$. Then one sees easily that $\mathfrak{p}$ is a prime ideal, and $\mathfrak{p} \neq 0$ since $|\cdot|$ is non-trivial. Therefore, there exists a prime number p such that $\mathfrak{p} = (p)$. By Proposition 8.2.13, $|\cdot|$ is equivalent to $|\cdot|_p$. □

Definition 10.1.2. — Let K be a number field. A *place* v of K is an equivalence class of norms on K. If these norms are archimedean (resp. non-archimedean), we say the place v is archimedean (resp. non-archimedean).

By Theorem 10.1.1 the set of places of $\mathbb{Q}$ is $\{\text{rational primes }\} \cup \{\infty\}$. This result can be generalized to any number field.

10.1.3. Places of an arbitrary number field. — Let K be a number field. Let σ_i with $1 \leq i \leq r_1$ denote the real embeddings of K, and $\sigma_{r_1+j}, \bar{\sigma}_{r_1+j}$ with $1 \leq j \leq r_2$ be the non-real complex embeddings of K. Then for each complex embedding σ_i with $1 \leq i \leq r_1 + r_2$, we have a norm $|\cdot|_\sigma$ on K given by

$$|x|_{\sigma_i} = |\sigma_i(x)|_{\mathbb{C}}.$$

On the other hand, for each prime ideal $\mathfrak{p}$ of the integral ring $\mathcal{O}_K$ of K, we have an additive valuation $v_\mathfrak{p}$ on K, which sends each $x \in K$ to the exponent of $\mathfrak{p}$ in the factorization of (x). We define the normalized $\mathfrak{p}$-norm on K by

$$|x|_\mathfrak{p} = \mathrm{N}(\mathfrak{p})^{-v_\mathfrak{p}(x)}.$$

Theorem 10.1.4. — (1) *Any two of the absolute values of* $|\cdot|_v$ *for* $v \in \{\sigma_i \mid 1 \leq i \leq r_1 + r_2\} \cup \{\text{prime ideals of } \mathcal{O}_K\}$ *are not equivalent to each other.*

(2) *Every absolute value on* K *is equivalent to some* $|\cdot|_{\sigma_i}$ *with* $1 \leq i \leq r_1 + r_2$ *or* $|\cdot|_\mathfrak{p}$ *for a prime ideal* $\mathfrak{p}$ *of* $\mathcal{O}_K$.

Proof. — It is obvious that a archimedean absolute value $|\cdot|_{\sigma_i}$ is not equivalent to any $|\cdot|_\mathfrak{p}$, which is non-archimedean. It is also clear that $|\cdot|_\mathfrak{p}$ is not equivalent to $|\cdot|_\mathfrak{q}$ if $\mathfrak{p} \neq \mathfrak{q}$, since they induces different prime ideals in $\mathcal{O}_K$. To prove that $|\cdot|_{\sigma_i}$ is not equivalent to $|\cdot|_{\sigma_j}$ for $i \neq j$ with $1 \leq i, j \leq r_1 + r_2$, we recall that, under the map

$$\lambda : K \to \mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$$

defined by $\lambda(x) = (\sigma_i(x))_{1\leq i\leq r_1+r_2}$, the image of $\mathcal{O}_K$ is a (full) lattice in $\mathbb{R}^{r_1}\times\mathbb{C}^{r_2}$. Then $\lambda(K)$ is dense in $\mathbb{R}^{r_1}\times\mathbb{C}^{r_2}$. In particular, for any i with $1\leq i\leq r_1+r_2$, there exists $x_i\in K$ such that $|x_i|_{\sigma_i}<1$ and $|x_i|_{\sigma_j}>1$ for any $j\neq i$. Hence, the sequence $(x_i^n)_{n\geq 1}$ converge to 0 for $|\cdot|_{\sigma_i}$ but diverges for $|\cdot|_{\sigma_j}$ with $j\neq i$. Hence, $|\cdot|_{\sigma_i}$ is not equivalent to $|\cdot|_{\sigma_j}$.

Suppose now $|\cdot|$ is an absolute value on K, we have to prove that $|\cdot|$ is equivalent to some $|\cdot|_{\sigma_i}$ or to some $|\cdot|_{\mathfrak{p}}$. Suppose first that $|\cdot|$ is non-archimedean. Since $\mathcal{O}_K$ is integral over $\mathbb{Z}$, we have $|x|\leq 1$ for all $x\in\mathcal{O}_K$. Let $\mathfrak{p}$ be the subset of $x\in\mathcal{O}_K$ such that $|x|<1$. One verifies easily that $\mathfrak{p}$ is a prime ideal of $\mathcal{O}_K$. Then $\mathcal{O}_{K,\mathfrak{p}}$ is the valuation ring of K for $|\cdot|$, which coincides that of K for $|\cdot|_{\mathfrak{p}}$. By Proposition 8.2.13, $|\cdot|$ is equivalent to $|\cdot|_{\mathfrak{p}}$.

Assume now that $|\cdot|$ is archimedean. Let K_v denote the completion of K under $|\cdot|$. Then K_v is a finite extension of $\mathbb{R}$, which is the completion of $\mathbb{Q}$ at its unique infinite place. Thus, K_v is either $\mathbb{R}$ or $\mathbb{C}$. In any case, the embedding $K\hookrightarrow K_v$ is one of σ_i or the complex conjugate of some σ_i.

□

Theorem 10.1.4 says that the non-archimedean (or finite) places of K are in natural bijection with non-zero prime ideals of $\mathcal{O}_K$, and the archimedean (or infinite) places of K are in natural bijection with the orbit of complex conjugate on the set of complex embeddings of K.

Theorem 10.1.5. — *Let $v_1,\cdots,v_r$ be distinct places of K. Then the diagonal embedding*

$$K\hookrightarrow\prod_{i=1}^{r}K_{v_i},$$

has dense image, where K_{v_i} denotes the completion of K at v_i.

Proof. — **Step 1**: Since K is dense in each K_{v_i}, it suffices to prove that, given $x_1,\cdots,x_r\in K$ and $\epsilon>0$, there exists $\xi\in K$ such that

$$|\xi-x_i|_{v_i}<\epsilon. \tag{10.1.5.1}$$

We claim that for each i with $1\leq i\leq r$ and any $\delta>0$, there exists a $\xi_i\in K$ such that

$$|\xi_i-1|_{v_i}<\epsilon,\quad |\xi_i|_j\leq\delta,\ \text{for } j\neq i.$$

Assuming the claim for a moment, then $\xi=\sum_{i=1}^{r}\xi_i x_i$ satisfies

$$\begin{aligned}|\xi-x_i|_{v_i}&=|x_i(\xi_i-1)+\sum_{j\neq i}x_j\xi_j|_{v_i}\\&\leq|x_i|_{v_i}\delta+\sum_{j\neq i}|x_j|_{v_i}\delta=\delta\sum_{j=1}^{r}|x_j|_{v_i}.\end{aligned}$$

One can choose appropriate δ so that (10.1.5.1) is satisfied. It remains to prove the claim.

Step 2. The proof of the claim can be reduced to showing that for any i with $1 \leq i \leq r$, there exists $\xi \in K$ such that $|\xi|_{v_j} < 1$ for all $j \neq i$, and $|\xi|_{v_i} > 1$. Since then, we have

$$|\frac{\xi^n}{1+\xi^n} - 1|_{v_i} = |\frac{1}{1+\xi^n}|_{v_i} \to 0, \quad |\frac{\xi^n}{1+\xi^n}|_{v_j} \to 0 \quad \text{for all } j \neq i$$

as $n \to +\infty$. We may assume that $i = 1$, and proceed by induction on r. For $r = 2$, the existence of ξ follows from the non-equivalence of v_1 and v_2. Assume now $r > 2$. By induction hypothesis, there exists $\xi \in K$ such that $|\xi|_{v_1} > 1$ and $|\xi|_{v_j} < 1$ for $j = 2, \cdots, r-1$. If $|\xi|_{v_r} < 1$, then the assertion is proved. Consider the cases $|\xi|_{v_r} \geq 1$. As v_1 and v_r are not equivalent, there exists $\alpha \in K$ such that $|\alpha|_{v_1} > 1$ and $|\alpha|_{v_r} < 1$. If $|\xi|_{v_r} = 1$, then $\xi^N \alpha$ for sufficiently large N will answer the question. If $|\xi|_{v_r} > 1$, we can take $\frac{\alpha\xi^N}{1+\xi^N}$ for N sufficiently. □

10.2. Tensor product and decomposition of primes

Let L/K be a finite extension of number fields. Let L_w be the completion of L at a place, and K_v be the closure of K in L_w. Then K_v is the completion of K at a place v, and L_w/K_v is a finite extension. We write $w|v$.

Assume now w is a non-archimedean place. Let $\mathfrak{P}_w \subseteq \mathcal{O}_L$ denote the prime ideal given by the place w. Then the prime ideal of $\mathcal{O}_K$ corresponding to v is $\mathfrak{p}_v = \mathfrak{P}_w \cap \mathcal{O}_K$. Recall that we defined in Section 3.2 the ramification index $e(\mathfrak{P}_w|\mathfrak{p}_v)$ and the residue degree $f(\mathfrak{P}_w|\mathfrak{p}_v)$. On the other hand, we defined in Subsection 9.1.3 the ramification index $e(L_w|K_v)$ and $f(L_w|K_v)$. Let $\mathcal{O}_{K_v}$ be the valuation ring of K_v and $\hat{\mathfrak{p}}_v \subseteq \mathcal{O}_{K_v}$ be the maximal ideal, and similar notation for $\mathcal{O}_{L_w}$ and $\hat{\mathfrak{P}}_w$. Then $k_v := \mathcal{O}_{K_v}/\hat{\mathfrak{p}}_v \cong \mathcal{O}_K/\mathfrak{p}_v$ is stable under completion, and similarly for the residue field k_w of L_w. It follows immediately that $f(L_w|K_v) = f(\mathfrak{P}_w|\mathfrak{p}_v)$. Besides, if π_v (resp. π_w) is a uniformizer of $\mathcal{O}_{K,\mathfrak{p}_v}$ (resp. of $\mathcal{O}_{L,\mathfrak{P}_w}$), then it is also a uniformizer of K_v (resp. L_w). Let v_{π_w} denote the normalized additive valuation on L_w. Then we have

$$e(\mathfrak{P}_w|\mathfrak{p}_v) = v_{\pi_w}(\pi_v) = e(L_w|K_v).$$

In the sequel, we will denote simply $e(w|v) = e(L_w|K_v)$ and $f(w|v) = f(L_w|K_v)$.

Theorem 10.2.1. — *Given a place v of K, we have a canonical isomorphism*

$$L \otimes_K K_v \cong \prod_{w|v} L_w.$$

Proof. — Let $f(x) \in K[x]$ be an irreducible polynomial such that $L = K[x]/(f(x))$. Assume that

$$f(x) = \prod_{i=1}^{r} g_i(x)$$

is a decomposition of $f(x)$ into irreducible factors in $K_v[x]$. Then we have

$$L \otimes_K K_v = K_v[x]/(f(x)) \cong \prod_{i=1}^{r} K_v[x]/(g_i(x)).$$

Each $L_i := K_v[x]/(g_i(x))$ is a finite extension of K_v, hence a complete discrete valuation ring. Note that L is dense in $L \otimes_K K_v$, hence in each factor L_i. Hence, L_i is the complete of L at a finite place w_i above v. Note that any two factors $g_i(x)$ and $g_j(x)$ are coprime with each other, so $L_i \neq L_j$. It follows thus that we have an injection:

$$L \otimes_K K_v \cong \prod_{i=1}^{r} L_i \hookrightarrow \prod_{w|v} L_w.$$

On the other hand, if L_w is the completion of L at a finite place w dividing v. Then by the universal property of $L \otimes_K K_v$, we have a homomorphism of K_v-algebras

$$L \otimes_K K_v \to L_w,$$

which is automatically continuous if we equip both sides the canonical topology for finite dimensional K_v-vector space. Since L is dense in L_w and $L \otimes_K K_v$ is complete as a finite dimensional K_v-vector space (c.f. Lemma 8.5.3 in the non-archimedean case), the map $L \otimes_K K_v \to L_w$ must be surjective. This shows that L_w is a quotient of $L \otimes_K K_v$, hence equals to one of L_i.

Now assume that v is non-archimedean. Note that the valuation ring $\mathcal{O}_{K_v}$ is the completion of $\mathcal{O}_K$ with respect to the norm $|\cdot|_v$. It follows that the residue field of K_v is exactly $k_v := \mathcal{O}_K/\mathfrak{p}_v$; similarly the residue field of L_w is $k_w := \mathcal{O}_L/\mathfrak{P}_w$ for $w|v$. It is now obvious that $f(\mathfrak{P}_w|\mathfrak{p}_v) = f(L_w|K_v)$.

□

Remark 10.2.2. — Theorem 10.2.1 gives another proof of the fundamental equality Proposition 3.2.2(2):

$$[L : K] = \dim_{K_v}(L \otimes_K K_v) = \sum_{w|v} [L_w : K_v] = \sum_{w|v} e(w|v)f(w|v),$$

where the last equality uses Proposition 9.1.4(1).

Corollary 10.2.3. — *Let v be a place of K. Then for any $x \in L$, we have*

$$\mathrm{Tr}_{L/K}(x) = \sum_{w|v} \mathrm{Tr}_{L_w/K_v}(x), \quad \mathrm{N}_{L/K}(x) = \prod_{w|v} \mathrm{N}_{L_w/K_v}(x).$$

Proof. — Indeed, $\mathrm{Tr}_{L/K}(x)$ equals to the trace of the K_v-linear endomorphism on $L \otimes_K K_v$ given by the multiplication by $x \otimes 1$. According to Theorem 10.2.1, this endomorphism is the direct sum of the multiplication by x on L_w for all $w|v$. Now the Corollary follows immediately. □

Example 10.2.4. — Consider the polynomial

$$f(x) = 1 + x + \frac{x^2}{2} + \frac{x^3}{3} + \frac{x^4}{4} + \frac{x^5}{5} \in \mathbb{Q}[x].$$

By looking at the 5-adic Newton polygon of $f(x)$, one sees easily that $f(x)$ is irreducible. The 3-adic Newton polygon of $f(x)$ have slopes $1/3$ with multiplicity 3 and $-1/2$ with multiplicity two. So $f(x)$ has two irreducible factors $g(x)$ and $h(x)$ in $\mathbb{Q}_3[x]$ with degree

3 and 2 respectively, and all the roots of $g(x)$ have 3-adic valuation $1/3$ and those of $h(x)$ have valuation $-1/2$. It follows by Theorem 10.2.1 that we have decomposition of primes in L:

$$3\mathcal{O}_L = \mathfrak{p}_1\mathfrak{p}_2$$

with $f(\mathfrak{p}_1|3) = f(\mathfrak{p}_2|3) = 1$ and $e(\mathfrak{p}_1|3) = 3$ and $e(\mathfrak{p}_2|3) = 2$.

10.3. Product formula

To state the result, we need to modify slightly the norm at a complex place. Let K be a number field, and v be a place of K. We define the *normalized* norm $|\cdot|_v$ of K at v as

$$|x|_v := \begin{cases} \frac{1}{\mathrm{N}(\mathfrak{p})^{v_\mathfrak{p}(x)}} & \text{if } v = \mathfrak{p} \text{ is non-archimedean;} \\ |\sigma(x)|_{\mathbb{R}} & \text{if } v \text{ is given by a real embedding } \sigma; \\ |\sigma(x)|_{\mathbb{C}}^2 & \text{if } v \text{ is given by a pair of complex embeddings } \sigma, \bar{\sigma}. \end{cases}$$

Lemma 10.3.1. — *For a place $p \leq \infty$ of $\mathbb{Q}$ and any $x \in K$, we have*

$$|\mathrm{N}_{K/\mathbb{Q}}(x)|_p = \prod_{v|p} |x|_v$$

Proof. — By Corollary 10.2.3, we have

$$|\mathrm{N}_{K/\mathbb{Q}}(x)|_p = \prod_{v|p} |\mathrm{N}_{K_v/\mathbb{Q}_p}(x)|_p.$$

To finish the proof, it suffices to show that $|\mathrm{N}_{K_v/\mathbb{Q}_p}(x)|_p = |x|_v$ for any $x \in K_v$. If v is an infinite place, then $\mathbb{Q}_\infty = \mathbb{R}$ and K_v is either $\mathbb{R}$ or $\mathbb{C}$; the assertion is obvious by our definition of $|\cdot|_v$. If p is finite, then by Lemma 9.3.2, we have

$$v_p(\mathrm{N}_{K_v/\mathbb{Q}_p}(x)) = f(v|p)v_{\mathfrak{p}_v}(x),$$

where $v_{\mathfrak{p}_v}$ denotes the normalized additive valuation on K_v. Thus, it follows that

$$|x|_v = \mathrm{N}_{\mathfrak{p}_v}^{-v_{\mathfrak{p}_v}(x)} = p^{-f(v|p)v_{\mathfrak{p}_v}(x)} = p^{-v_p(\mathrm{N}_{K_v/\mathbb{Q}_p}(x))} = |\mathrm{N}_{K_v/\mathbb{Q}_p}(x)|_p.$$

□

Proposition 10.3.2 **(Product formula)**. — *For every $x \in K$, we have*

$$\prod_v |x|_v = 1,$$

where v runs through all places of K, and $|\cdot|_v$ is the normalized norm of K at v.

Proof. — We treat first the case $K = \mathbb{Q}$. Since the formula is multiplicative in x, it suffices to prove the formula for $x = -1$ and $x = p$ with p a prime. The case is trivial for $x = -1$. When $x = p$, we have

$$\prod_{v \leq \infty} |x|_v = |p|_p \cdot |p|_\infty = 1.$$

For general number field K, it follows from Lemma 10.3.1 that

$$\prod_v |x|_v = \prod_{p\leq\infty} |\mathrm{N}_{K/\mathbb{Q}}(x)|_p = 1.$$

□

10.4. Comparison of local and global Galois groups

Let L/K be a finite Galois extension of number fields with Galois group G. Let v be a finite place of K, and w be a place of L above v. Let L_w (resp. K_v) denote the completion of L at w (resp K at v). The natural inclusion $L \hookrightarrow L_w$ induces a morphism of Galois group

$$i_w : \mathrm{Gal}(L_w/K_v) \to \mathrm{Gal}(L/K).$$

Let $\mathfrak{P}_w$ denote the prime ideal of $\mathcal{O}_L$ given by w, and $\mathfrak{p}_v$ be the prime ideal of $\mathcal{O}_K$ for v. Recall that we defined in 3.4.3 the decomposition subgroup $D_w := D(\mathfrak{P}_w|\mathfrak{p}_v) \subseteq G$. The following Proposition is fundamental for applying local methods to study the Galois group of number fields.

Proposition 10.4.1. — *The morphism i_w induces an isomorphism*

$$\mathrm{Gal}(L_w/K_v) \cong D_w.$$

Proof. — Note first that i_w is injective, because L is dense in L_w. Let $w = w_1, \cdots, w_g$ be the primes of L above v. Then G acts transitive on the set $\{w_1, \cdots, w_g\}$, and by definition, D_w is the stabilizer of w. Let $\sigma \in \mathrm{Gal}(L_w/K_v)$. Denote by $\mathfrak{m}_{L_w}$ the maximal ideal of $\mathcal{O}_{L_w}$. Then σ clearly stabilizes $\mathfrak{m}_{L_w}$, and $i_w(\sigma)$ stabilizes $\mathcal{O}_L$. Since $\mathcal{O}_L \cap \mathfrak{m}_{L_w} = \mathfrak{P}_w$, we see that $i_w(\sigma) \in D_w$. This shows that the image of i_w lies in D_w. On the other hand, if $e = e(w|v)$ and $f = f(w|v)$ denote the ramification and residue degree respectively, then both D_w and $\mathrm{Gal}(L_w|K_v)$ have cardinality ef. Hence, i_w induces an isomorphism between $\mathrm{Gal}(L_w/K_v)$ and D_w. □

Proposition 10.4.1 can be used to determine the Galois group of an irreducible polynomial over $\mathbb{Q}$.

Example 10.4.2. — Consider the polynomial $f(x) = x^5 - x - 1 \in \mathbb{Q}[x]$. Since $f(x)$ is irreducible modulo 5, we see that $f(x)$ is irreducible over $\mathbb{Q}$. We can interested in finding the Galois group G of $f(x)$. Let $K = \mathbb{Q}[x]/(f(x))$, and L be the Galois closure of K. Then $G = \mathrm{Gal}(L/\mathbb{Q})$ and can be viewed as a subgroup of the permutation group $\mathfrak{S}_5$. Since $5|[L : \mathbb{Q}]$, G must contain an element of order 5, which is necessarily a 5-cycle. On the other hand, we have the decomposition

$$f(x) \equiv (x^2 + x + 1)(x^3 + x^2 + 1) \mod 2.$$

By Hensel's Lemma 8.4.1, $f(x) = g(x)h(x)$ in $\mathbb{Q}_2[x]$, where $g(x)$ and $h(x)$ are monic irreducible polynomials in $\mathbb{Z}_2[x]$ lifting $x^2 + x + 1$ and $x^3 + x^2 + 1$. Hence, by the proof of Theorem 10.2.1, the prime 2 is unramified in K and splits into two places v_1 and v_2 in K with residue degrees 2 and 3 respectively. Let w be a place of L above v_2. Then L_w is the unique unramified extension of degree 6 over $\mathbb{Q}_2$, and has residue degree 2 over K_{v_2}.

Let $\sigma \in G$ denote the Frobenius element at w. Then σ^3 generates $\mathrm{Gal}(L_w/K_{v_w})$. If α_1, α_2 denote the two roots of $g(x)$ and $\alpha_3, \alpha_4, \alpha_5$ are the roots of $h(x)$, then σ^3 fixes $\alpha_3, \alpha_4, \alpha_5$ and interchanges α_1 and α_2. Hence, σ^3 is a transposition in $\mathfrak{S}_5$. It is well known that a subgroup of $\mathfrak{S}_5$ containing a 5-cycle and a transposition is necessarily $\mathfrak{S}_5$ itself. We conclude that $G \cong \mathfrak{S}_5$.

10.5. Local and global different

Let L/K be a finite extension of number fields, v be a finite place of K, and $\mathfrak{p}_v \subseteq \mathcal{O}_K$ be the prime ideal given by v. Denote by K_v the completion of K at v. We have the following series of Dedekind domains:

$$\mathcal{O}_K \hookrightarrow \mathcal{O}_{K,\mathfrak{p}_v} \hookrightarrow \hat{\mathcal{O}}_{K,\mathfrak{p}_v} = \varprojlim_n \left(\mathcal{O}_{K,\mathfrak{p}_v}/\mathfrak{p}_v^n \mathcal{O}_{K,\mathfrak{p}_v}\right) = \mathcal{O}_{K_v},$$

where $\mathcal{O}_K$ (resp. $\mathcal{O}_{K_v}$) denote the ring of integers of K (resp. K_v), $\mathcal{O}_{K,\mathfrak{p}_v}$ is the localization of $\mathcal{O}_K$ at $\mathfrak{p}_v$, and $\hat{\mathcal{O}}_{K,\mathfrak{p}_v}$ means the completion of $\mathcal{O}_{K,\mathfrak{p}_v}$ with respect to its maximal ideal.

Put $S_v = \mathcal{O}_K \backslash \mathfrak{p}_v$. This is a multiplicative subset of $\mathcal{O}_K$ or $\mathcal{O}_L$. We put $\mathcal{O}_{L,\mathfrak{p}_v} = S^{-1}\mathcal{O}_L = \mathcal{O}_L \cdot \mathcal{O}_{K,\mathfrak{p}_v}$. For any fractional ideal I of L, we put

$$I_{\mathfrak{p}_v} := I\mathcal{O}_{L,\mathfrak{p}_v} = \{x \in L \mid x = \frac{a}{s} \text{ for some } a \in I \text{ and } s \in S_v.\}$$

$$\hat{I}_{\mathfrak{p}_v} = \varprojlim_n \left(I_{\mathfrak{p}_v^n}/\mathfrak{p}_v^n I_{\mathfrak{p}_v}\right)$$

Lemma 10.5.1. — *If $I = \prod_w \mathfrak{P}_w^{a_w}$ is the prime decomposition of I, where w runs through finite places of L, then we have a canonical isomorphism*

$$\hat{I}_{\mathfrak{p}_v} \cong \prod_{w|v} \hat{\mathfrak{P}}_w^{a_w},$$

where $\hat{\mathfrak{P}}_w$ denotes the maximal ideal of $\mathcal{O}_{L_w}$. In particular, we have

$$\hat{\mathcal{O}}_{L,\mathfrak{p}_v} \cong \prod_{w|v} \hat{\mathcal{O}}_{L_w}.$$

Proof. — We have

$$I_{\mathfrak{p}_v} = \prod_{w|v} (\mathfrak{P}_w \mathcal{O}_{L,\mathfrak{p}_v})^{a_w}$$

$$\mathfrak{p}_v^n I_{\mathfrak{p}_v^n} = \prod_{w|v} \left(\mathfrak{P}_w^{a_w} \mathcal{O}_{L,\mathfrak{p}_v}\right).$$

Assume that $\mathfrak{p}_v\mathcal{O}_{L,\mathfrak{p}_v} = \prod_{w|v}(\mathfrak{P}_w^{e(w|v)}\mathcal{O}_{L,\mathfrak{p}_v})$. Then it follows that

$$\begin{aligned}\varprojlim_n \left(I_{\mathfrak{p}_v}/\mathfrak{p}_v^n I_{\mathfrak{p}_v}\right) &= \varprojlim_n \left(\mathfrak{P}_w^{a_w}\mathcal{O}_{L,\mathfrak{p}_v}/\mathfrak{P}_w^{a_w+ne(w|v)}\mathcal{O}_{L,\mathfrak{p}_v}\right)\\ &\cong \prod_{w|v}\varprojlim_n \left(\mathfrak{P}_w^{a_w}\mathcal{O}_{L,\mathfrak{p}_v}/\mathfrak{P}_w^{a_w+ne(w|v)}\mathcal{O}_{L,\mathfrak{p}_v}\right) = \prod_{w|v}\hat{\mathfrak{P}}_w^{a_w}.\end{aligned}$$

□

Theorem 10.5.2. — *Let $\delta_{L/K}$ be the relative different of L/K (cf. Definition 3.3.3), and for each place w dividing v, let δ_{L_w/K_v} be the different of the extension L_w/K_v. Then we have a canonical isomorphism*

$$\hat{\delta}_{L/K} \cong \prod_{w|v}\delta_{L_w/K_v}.$$

Proof. — By definition, $\delta_{L/K}$ is the ideal of $\mathcal{O}_L$ such that the trace map induces a perfect pairing

$$\mathrm{Tr}_{L/K} : \mathcal{O}_L \times \delta_{L/K}^{-1} \to \mathcal{O}_K.$$

Here, "perfect" means that

$$\delta_{L/K}^{-1} = \{y \in L \mid \mathrm{Tr}_{L/K}(xy) \in \mathcal{O}_K \forall x \in \mathcal{O}_L\},$$

or equivalently the map $y \mapsto (x \mapsto \mathrm{Tr}_{L/K}(xy))$ induces an isomorphism of $\mathcal{O}_L$-modules

$$\delta_{L/K}^{-1} \cong \mathrm{Hom}_{\mathcal{O}_K}(\mathcal{O}_L, \mathcal{O}_K).$$

Passing to localization at $\mathfrak{p}_v$, we get a pairing

$$\mathcal{O}_{L,\mathfrak{p}_v} \times \delta_{L/K,\mathfrak{p}_v}^{-1} \to \mathcal{O}_{L,\mathfrak{p}_v}. \tag{10.5.2.1}$$

We claim that this pairing is still perfect. Indeed, it is clear that

$$\delta_{L,\mathfrak{p}_v}^{-1} \subseteq \{y \in L \mid \mathrm{Tr}_{L/K}(xy) \in \mathcal{O}_{K,\mathfrak{p}_v} \forall x \in \mathcal{O}_{L,\mathfrak{p}_v}\}.$$

Conversely, given an element $y \in L$ such that $\mathrm{Tr}_{L/K}(xy) \in \mathcal{O}_{K,\mathfrak{p}_v}$ for all $x \in \mathcal{O}_{L,\mathfrak{p}_v}$, we need to prove that $y \in \delta_{L/K,\mathfrak{p}_v}^{-1}$, that is there exists $s \in S_v = \mathcal{O}_K \backslash \mathfrak{p}_v$ such that $sy \in \delta_{L/K}^{-1}$. Let $\{x_1, \cdots, x_r\}$ be a set of generators of $\mathcal{O}_L$ as an $\mathcal{O}_K$-module. Then they also form a set of generators of $\mathcal{O}_{L,\mathfrak{p}_v}$ over $\mathcal{O}_{K,\mathfrak{p}_v}$. Let

$$z_i = \mathrm{Tr}_{L/K}(x_i y) \in \mathcal{O}_{K,\mathfrak{p}_v}.$$

Then exists $s_i \in S$ such that $s_i z \in \mathcal{O}_K$. Put $s = \prod_{i=1}^r s_i$, then $\mathrm{Tr}_{L/K}(syx_i) \in \mathcal{O}_K$ for all i. It follows that $sy \in \mathcal{O}_K$. This finishes the proof of the claim.

Taking completion of (10.5.2.1), we get a paring

$$\hat{\mathrm{Tr}}_{L/K} : \hat{\mathcal{O}}_{L,\mathfrak{p}_v} \times \hat{\delta}_{L/K,\mathfrak{p}_v}^{-1} \to \hat{\mathcal{O}}_{K,\mathfrak{p}_v} = \mathcal{O}_{K_v}. \tag{10.5.2.2}$$

If $\delta_{L/K} = \prod_w \mathfrak{P}_w^{v_w(\delta_{L/K})}$ is the prime decomposition of $\delta_{L/K}$ for some $v_w(\delta_{L/K}) \in \mathbb{Z}$, then by Lemma 10.5.1 and Corollary 10.2.3, (10.5.2.2) is canonically isomorphic to the pairing

$$\sum_{w|v} \mathrm{Tr}_{L_w/K_v} : \prod_{w|v} \mathcal{O}_{L_w} \times \prod_{w|v} \hat{\mathfrak{P}}_w^{-v_w(\delta_{L/K})} \to \mathcal{O}_{K_v}.$$

To finish the proof of the Theorem, it suffices to show that (10.5.2.2) is perfect, since it will imply that $\delta_{L_w/K_v}^{-1} = \hat{\mathfrak{P}}_w^{-v_w(\delta_{L/K})}$. We note first that $\mathcal{O}_{K,\mathfrak{p}_v}$ is a principal ideal domain, and every finite generated torsion free module over $\mathcal{O}_{K,v}$ is free. Let $(\alpha_1, \cdots, \alpha_n)$ and $(\beta_1, \cdots, \beta_n)$ be respectively a basis of $\mathcal{O}_{L,\mathfrak{p}_v}$ and $\delta_{L/K,\mathfrak{p}_v}^{-1}$ over $\mathcal{O}_{K,\mathfrak{p}_v}$. Let $(\alpha_1^\vee, \cdots, \alpha_n^\vee)$ be the basis of $\mathrm{Hom}_{\mathcal{O}_{K,\mathfrak{p}_v}}(\mathcal{O}_{L,\mathfrak{p}_v}, \mathcal{O}_{K,\mathfrak{p}_v})$. If $\phi : \delta_{L/K,\mathfrak{p}_v}^{-1} \to \mathrm{Hom}_{\mathcal{O}_{K,\mathfrak{p}_v}}(\mathcal{O}_{L,\mathfrak{p}_v}, \mathcal{O}_{K,\mathfrak{p}_v})$ denotes the map induced by the pairing (10.5.2.1), then one has

$$\phi(\beta_j) = \sum_i \mathrm{Tr}_{L/K}(\alpha_i \beta_j) \alpha_i^\vee.$$

Therefore, the perfectness of (10.5.2.1) is equivalent to $\det(\mathrm{Tr}_{L/K}(\alpha_i\beta_j)) \in \mathcal{O}_{K,\mathfrak{p}_v}^\times$. Now, we view $(\alpha_i)_{1\le i\le n}$ and $(\beta_j)_{1\le j\le n}$ as basis of $\hat{\mathcal{O}}_{L,\mathfrak{p}_v}$ and $\hat{\delta}_{L/K,\mathfrak{p}_v}^{-1}$ over $\mathcal{O}_{K_v}$ respectively via the canonical injection from the non-completed modules to the completed ones, then similar arguments show that $\det(\mathrm{Tr}_{L/K}(\alpha_i\beta_j)) \in \mathcal{O}_{K,\mathfrak{p}_v}^\times \subseteq \hat{\mathcal{O}}_{K,\mathfrak{p}_v}^\times$ implies that (10.5.2.2) is perfect. □

10.6. Hermite-Minkowski's finitness theorem

In this section, we give a proof of Hermite-Minkowski's finiteness theorem using local methods. We start with a finiteness theorem on local fields.

Theorem 10.6.1. — *Let F be a finite extension of $\mathbb{Q}_p$. Given an integer $n \in \mathbb{Z}_{>0}$, there exists only finitely many extensions of F of degree n.*

Proof. — Since F has a unique unramified extension of degree d for each positive integer d dividing n, it suffices to prove that there exist only finitely many totally ramified extension of F of degree n. Let π_F be a uniformizer of F, and put

$$S = \{f(x) = x^n + a_{n-1}\pi_F x^{n-1} + \cdots + \pi_F a_0 \mid a_i \in \mathcal{O}_F \text{ for } i \neq 0, a_0 \in \mathcal{O}_F^\times\}.$$

Then we have an isomorphism of topological spaces

$$S \cong \mathcal{O}_F^{n-1} \times \mathcal{O}_F^\times.$$

Then a totally extension of F of degree n is generated by a root of $f(x)$ with $f(x) \in S$. By Theorem 8.6.2, for each $f(x) \in S$, there exists $\epsilon > 0$ such that $F[x]/(f(x)) \cong F[x]/(g(x))$ for any $g \in S$ with $\|f - g\| < \epsilon$. Such g's form an open neighborhood $U(f, \epsilon)$ of $f(x)$ in S. Then all such $U(f, \epsilon)$ cover the space S. Since S is compact, finitely many of $U(f, \epsilon)$, say $U(f_i, \epsilon_i)$ with $1 \le i \le r$, cover S. Then every totally ramified extension of F of degree n is isomorphic to one of $F[x]/(f_i(x))$'s.

□

The following Corollary is immediate from Theorem 10.6.1.

Corollary 10.6.2. — *Let E/F be a finite extension of degree n, and $v_E(\delta_{E/F})$ be the integer such that $\delta_{E/F} = (\pi_E^{v_E(\delta_{E/F})})$. Then $v_E(\delta_{E/F})$ is bounded above in terms of n.*

Theorem 10.6.3 **(Hermite-Minkowski)**. — *Let K be a number field, S be a finite set of places of K. Then for a fixed integer $n \in \mathbb{Z}_{>0}$, there exist only finitely many extensions of K unramified outside S of degree n.*

Proof. — By Corollary 4.1.6, it suffices to show that the disriminant of any extension L/K of degree n and unramified outside S is bounded above in terms of n and S. Indeed, if L/K is such an extension, it follows from Corollary 3.3.6 that

$$|\Delta_L| = |\Delta_K|^{[L:K]}\mathrm{N}_{K/\mathbb{Q}}(\mathrm{Disc}_{L/K}) = |\Delta_K|^{[L:K]}\mathrm{N}_{L/\mathbb{Q}}(\delta_{L/K}),$$

where Δ_L and Δ_K denote respectively the discriminant of L and K. By assumption, the order of $\delta_{L/K}$ at a finite place w of L is greater than 0 only if w divides a place $v \in S$. By Theorem 10.5.2 and Corollary 10.6.2, the exponent of $\mathfrak{P}_w$ in $\delta_{L/K}$ is equal to $v_{L_w}(\delta_{L_w/K_v})$, and is bounded above in terms of n. Hence, $\mathrm{N}_{L/\mathbb{Q}}(\delta_{L/K})$ is bounded above in terms of n and S. This finishes the proof. □

BIBLIOGRAPHY

[Bh04] M. Bhargava, Higher composition laws I: a new view on Gauss composition and quadratic generalizations, *Ann. Math.* **159**(1) (2004), 217-250.

[Da80] H. Davenport, *Multiplicative Number Theory*, Second Edition, *Graduate Texts in Mathematics*, **74**, Springer-Verlag, 1980.

[La94] S. Lang, *Algebraic Number Theory*, Second edition, *Graduate Texts in Mathematics*, **110**, Springer-Verlag, 1994.

[Ma77] D. Marcus, *Number Fields,* Springer-Verlag, 1977.

[Wa96] L. Washington, *Introduction to Cyclotomic Fields*, Second edition, Springer-Verlag, 1996.

[Se68] J. P. Serre, *Corps Locaux*, Publication de l'institut de Mathématique de l'Université de Nancago VIII, Hermann, Paris, 1968.