FMA51052EP

Algebraic Number Theory

Diego Izquierdo

École Polytechnique

March 27, 2025

Contents

1 Ring arithmetic **4**
- 1.1 Definitions and examples . . . 4
- 1.2 Ring morphisms . . . 6
- 1.3 Subrings and ideals . . . 7
- 1.4 Divisibility and arithmetic in a domain . . . 14
- 1.5 Prime and maximal ideals . . . 15
- 1.6 Euclidean and principal ideal domains . . . 17
- 1.7 Unique factorization domains . . . 18
- 1.8 Polynomials over unique factorization domains . . . 20

2 Number fields and their rings of integers **25**
- 2.1 Rings of algebraic integers . . . 25
- 2.2 How to compute rings of integers? . . . 28
- 2.3 Algebraic properties of rings of integers . . . 37
- 2.4 Fractional ideals in Dedekind domains . . . 41
- 2.5 Factorization of ideals in Dedekind domains . . . 43
- 2.6 Minkowski's Theorem on lattices . . . 44
- 2.7 Geometry of numbers . . . 48
- 2.8 Finiteness of the ideal class group . . . 51

3 p-adic rings and fields **54**
- 3.1 p-adic numbers . . . 54
- 3.2 The p-adic valuation . . . 56
- 3.3 The p-adic absolute value . . . 60
- 3.4 The Weak Approximation Theorem . . . 63
- 3.5 p-adic equations and Hensel's Lemma . . . 65

4 Riemann zeta function, L-functions and Dirichlet's Theorem on arithmetic progressions **69**
- 4.1 Dirichlet series and Euler products . . . 69
- 4.2 Riemann's zeta function . . . 71

4.3 Dirichlet L-functions 73
4.4 Behaviour of L-functions at $s = 1$ 77
4.5 Dirichlet's Theorem on arithmetic progressions 81

5 Quadratic forms over $\mathbb{Q}$ **83**
5.1 Short reminders on quadratic forms 83
5.2 Quadratic forms over finite fields 85
5.3 Quadratic forms over p-adic fields 91
5.4 Quadratic forms over $\mathbb{Q}$ 102

Bibliography **109**

Foreword

This course is an introduction to algebraic number theory for Master 1 students. The main goal that we will reach at the end is the celebrated Hasse-Minkowski Theorem, which can be stated in very elementary terms as follows:

Theorem 0.0.1 (Hasse-Minkowski)

Let $n \geq 3$ and let $q = a_1 X_1^2 + \cdots + a_n X_n^2$ be a quadratic form with $a_1, \ldots, a_n \in \mathbb{Z} \setminus \{0\}$. Let $p_1, \ldots, p_r$ be the odd prime divisors of the product $a_1 \ldots a_n$ and set:

$$N := 8 p_1 \ldots p_r \cdot (a_1 \vee \cdots \vee a_n).$$

Then the following assertions are equivalent:

(i) the equation $q = 0$ has a non-zero solution in $\mathbb{Q}$;

(ii) not all the a_i's have the same sign and there exist $x_1, \ldots, x_n \in \mathbb{Z}$ such that $x_1 \wedge \cdots \wedge x_n \wedge N = 1$ and $q(x_1, \ldots, x_n) \equiv 0 \mod N$.

In order to prove this theorem, we will need to explore quite diverse areas of algebra and number theory:

- in Chapter 1, we will introduce a certain number of basic algebraic notions about **rings**;
- in Chapter 2, we will introduce **number fields** and their **rings of integers**, and we will study their arithmetic properties;
- Chapter 3 will be devoted to the study of p-**adic rings** and p-**adic fields**;
- in Chapter 4, we will introduce a certain number of tools of **analytic number theory** in order to prove **Dirichlet's Theorem on arithmetic progressions**;
- in Chapter 5, we will use the theories developed in the previous chapters to settle the **Hasse-Minkowski Theorem**.

As prerequisites, it will be useful to have some notions in group theory, field theory and Galois theory. However, we will make reminders on the tools we will need so that the course can be followed with as little background as prossible.
All pictures of mathematicians in these course notes are taken from Wikipedia.

Chapter 1

Ring arithmetic

1.1 Definitions and examples

Definition 1.1.1

A **ring** is a set A endowed with two binary operations:

$$\begin{aligned} +: A \times A &\to A \\ (a,b) &\mapsto a+b \end{aligned} \qquad \begin{aligned} \cdot: A \times A &\to A \\ (a,b) &\mapsto a \cdot b, \end{aligned}$$

which are respectively called **addition** and **multiplication**, such that:

(R1) the pair $(A,+)$ is an abelian group, whose neutral element is usually denoted by 0.

(R2) multiplication is *associative*:

$$\forall (a,b,c) \in A^3, \quad (a \cdot b) \cdot c = a \cdot (b \cdot c).$$

(R3) multiplication has a *neutral element*, usually denoted by 1:

$$\forall a \in A, \quad 1 \cdot a = a \cdot 1 = a.$$

(R4) multiplication is *distributive* with respect to addition, that is:

$$\begin{aligned} \forall (a,b,c) \in A, \quad (a+b) \cdot c &= a \cdot b + a \cdot c \\ a \cdot (b+c) &= a \cdot b + a \cdot c. \end{aligned}$$

A ring A is said to be **commutative** if the multiplication is commutative. In other words:

$$\forall (a,b) \in A^2, ab = ba.$$

Example 1.1.2.

(i) The sets $\mathbb{Z}$, $\mathbb{Z}/n\mathbb{Z}$, $\mathbb{Q}$ and $\mathbb{C}$ endowed with the usual addition and multiplication are commutative rings.

(ii) If A is a commutative ring, the set of polynomials $A[X]$ endowed with the usual addition and multiplication is also a commutative ring.

(iii) If A is a ring, the set of $n \times n$-matrices $\mathscr{M}_n(A)$ endowed with the usual addition and multiplication is a ring. It can be non-commutative even if A is commutative.

Unless otherwise stated, in the sequel of this course, all rings will be assumed to be commutative.

Definition 1.1.3

Let A be a ring.

1. A non-zero element $a \in A$ is said to be a **zero-divisor** if there exists $b \in A \setminus \{0\}$ such that $ab = 0$. The ring A is said to be a **domain** if it has no zero-divisors.

2. An element $a \in A$ is said to be an **invertible element** (or a **unit**) if there exists $b \in A$ such that $ab = 1$. The set of invertible elements in A is a group, usually denoted $A^\times$. The ring A is said to be a **field** if it is non-zero and every non-zero element in A is invertible.

Example 1.1.4.

(i) Every field is a domain.

(ii) The rings $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ are fields.

(iii) The rings $\mathbb{Z}$ and $\mathbb{C}[T]$ are domains but not fields. We have $\mathbb{Z}^\times = \{-1,1\}$ and $\mathbb{C}[T]^\times = \mathbb{C}^\times$.

(iv) Let A be a domain. Let K be the set of fractions $\frac{a}{b}$ with $a \in A$ and $b \in A \setminus \{0\}$ modulo the equivalence relation given by:

$$\frac{a}{b} \sim \frac{a'}{b'} \Leftrightarrow ab' = a'b.$$

If we endow K with the usual addition and multiplication for fractions, one can easily check that K is a field, called the **fraction field** of A.

(v) Consider the ring $\mathbb{Z}/n\mathbb{Z}$ for some $n > 1$. According to Bézout's Theorem:

$$(\mathbb{Z}/n\mathbb{Z})^\times = \{a \mod n \mid 1 \leq a \leq n, \gcd(a,n) = 1\}.$$

Hence $\mathbb{Z}/n\mathbb{Z}$ is a field if, and only if, n is a prime number. When n is not prime, one can say a bit more: indeed, one can write $n = ab$ for some $a, b > 1$, so that the classes of a and b modulo n are zero-divisors in $\mathbb{Z}/n\mathbb{Z}$. It follows that $\mathbb{Z}/n\mathbb{Z}$ is not a domain.

Summing up, we have just checked that the following three assertions are equivalent:

a. $\mathbb{Z}/n\mathbb{Z}$ is a field;

b. $\mathbb{Z}/n\mathbb{Z}$ is a domain;

c. n is a prime number.

1.2 Ring morphisms

Definition 1.2.1

Let A and B be two rings. A map $f : A \to B$ is called a **ring homomorphism** if f is a group homomorphism $(A,+) \to (B,+)$ such that:

$$f(1) = 1,$$
$$\forall (a,a') \in A^2, f(aa') = f(a)f(a').$$

A bijective ring homomorphism is said to be an **isomorphism**.

Example 1.2.2.

(i) Let A be a ring. The map $\mathbb{Z} \to A, n \mapsto n \cdot 1$ is the unique ring morphism from $\mathbb{Z}$ to A.

(ii) Let $\varphi : A \to B$ be a ring homomorphism (for instance, one could take $A = B$ and $\varphi = \mathrm{Id}_A$). By abuse of notation, we also denote by φ the induced morphism $A[T] \to B[T]$. Given $b \in B$, we can define a ring morphism by:

$$\begin{aligned} \mathrm{ev}_b : A[T] &\to B \\ P &\mapsto \varphi(P)(b). \end{aligned}$$

Every ring homomorphism $f : A[T] \to B$ such that $f|_A = \varphi$ is of this form.

1.3 Subrings and ideals

Definition 1.3.1

Let A be a ring. A **subring** of A is a subset B of A such that the ring operations $+$ and $\cdot$ of A can be restricted to B so that $(B,+,\cdot)$ is a ring. In other words:

$$\begin{cases} (B,+) \text{ is a subgroup of } (A,+), \\ \forall (b_1,b_2) \in B, b_1 b_2 \in B, \\ 1 \in B. \end{cases}$$

Example 1.3.2.

(i) The ring of integers $\mathbb{Z}$ is a subring of $\mathbb{Q}$. The only subring of $\mathbb{Z}$ is $\mathbb{Z}$.

(ii) If A is a ring, it is a subring of $A[X]$.

(iii) When A is a subring of a ring B and S is a subset of B, we denote by $A[S]$ the intersection of all the subrings of B containing A and S. It is the smallest such subring, and it is called the **subring spanned by** A **and** S. Its elements are the elements of A that can be written as $P(s_1,...,s_n)$ with $s_1,...,s_n \in S$ and $P \in A[X_1,...,X_n]$.

(iv) If $f : A \to B$ is a ring homomorphism, then the image $\mathrm{im}(f)$ is a subring of B. However, the **kernel** $\ker(f) := \{a \in A | f(a) = 0\}$ is not a subring of A in general.

Definition 1.3.3

Let A be a ring. An **ideal** of A is a subgroup I of $(A,+)$ such that:

$$\forall a \in A, \forall i \in I, \quad ai \in I.$$

Example 1.3.4.

(i) The ideals of $\mathbb{Z}$ are the $n\mathbb{Z}$ for $n \in \mathbb{N}$.

(ii) The kernel of a ring homomorphism $f : A \to B$ is an ideal of A.

(iii) Let I and J be two ideals in a ring A. Let $I+J$ be the subset of A given by elements of the form $x+y$ with $x \in I$ and $y \in J$ and let IJ be the subset of A given by elements of the form $\sum_{k=1}^{n} x_i y_i$ for some $n \in \mathbb{N}$, some $x_1,...,x_n \in I$ and some $y_1,...,y_n \in J$. The sets $I+J$ and IJ are ideals in A. They are respectively called the **sum** and the **product** of I and J.

When S is a subset of a ring A, we denote by (S) the intersection of all the ideals of A containing S. It is the smallest ideal of A containing S, and it is called the **ideal spanned by** S. Its elements are the elements of A that can be written as a finite sum $\sum_{i=1}^{n} a_i s_i$ with $a_1, ..., a_n \in A$ and $s_1, ..., s_n \in S$.

An ideal in A is said to be **principal** if it is of the form (a) for some $a \in A$. A domain in which every ideal is principal is called a **principal ideal domain**.

Example 1.3.5.

(i) The ring $\mathbb{Z}$ is principal. Moreover, if S is a subset of $\mathbb{Z}$, we have:

$$(S) = \gcd(S)\mathbb{Z}.$$

(ii) Similarly, if K is a field, the ring $K[X]$ is principal. If S is a subset of $K[X]$, then $(S) = \gcd(S)K[X]$.

(iii) The ring $\mathbb{C}[X, Y]$ is not principal, since the ideal (X, Y) is not principal.

(iv) The ring $\mathbb{Z}[X]$ is not principal, since the ideal $(2, X)$ is not principal.

We finish this section by discussing quotients of rings by ideals. Given a ring A and an ideal I in A, one can endow the quotient of abelian groups A/I with a ring structure by:

$$\forall (a, b) \in A, (a + I) \cdot (b + I) = (ab) + I.$$

The ring A/I is then called the **quotient** of A by I. It has the following universal property:

Theorem 1.3.6 – Universal property of quotient

Let $f : A \to B$ be a ring homomorphism. Let I be an ideal of A and denote by $p : A \to A/I$ the natural projection. Then there exists a unique ring homomorphism $\hat{f} : A/I \to B$ such that $f = \hat{f} \circ p$ if, and only if, $I \subseteq \ker(f)$. In that case:

$$\operatorname{im}(\hat{f}) = \operatorname{im}(f), \quad \ker(\hat{f}) = \ker(f)/I.$$

In particular, if $f : A \to B$ is a ring homomorphism, then f induces an isomorphism between $A/\ker(f)$ and $\operatorname{im}(f)$.

Example 1.3.7.

(i) The ring homomorphism $f : \mathbb{R}[T] \to \mathbb{C}, P \mapsto P(i)$ is surjective and has kernel (T^2+1). Hence it induces an isomorphism:

$$\mathbb{R}[T]/(T^2+1) \cong \mathbb{C}.$$

(i) The ring homomorphism $f : \mathbb{Z}[T] \to \mathbb{Z}[i], P \mapsto P(i)$ is surjective and one can check that its kernel is $(T^2+1)\mathbb{Q}[T] \cap \mathbb{Z}[T] = (T^2+1)\mathbb{Z}[T]$. Hence it induces an isomorphism:

$$\mathbb{Z}[T]/(T^2+1) \cong \mathbb{Z}[i].$$

(ii) Let $\varphi : A \to B$ be a ring homomorphism. By abuse of notation, we also denote by φ the induced morphism $A[T] \to B[T]$. Given an ideal I of $A[T]$ and and element $b \in B$, the ring morphism

$$\mathrm{ev}_b : A[T] \to B$$

induces a ring homomorphism:

$$\overline{\mathrm{ev}}_b : A[T]/I \to b$$

if, and only if, $\varphi(P)(b) = 0$ for each $P \in I$. Moreover, every ring homomorphism

$$f : A[T]/I \to B$$

such that $f|_A = \varphi$ is of this form.

An important consequence of the universal property of the quotient is the so-called Chinese Remainder Theorem. For the ring $\mathbb{Z}$, it states that $\mathbb{Z}/(mn)\mathbb{Z}$ is isomorphic to $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ whenever m and n are coprime. More generally:

Theorem 1.3.8 (Chinese Remainder Theorem)

Let $I_1, ..., I_n$ be ideals of a ring A such that $I_i + I_j = A$ for all $i \neq j$. Then there is a natural isomorphism:

$$A/(I_1 \cap ... \cap I_n) \cong A/I_1 \times ... \times A/I_n.$$

Proof. Consider the morphism:

$$f : A \to A/I_1 \times A/I_2 \times ... \times A/I_n,\ a \mapsto (a \mod I_1, ..., a \mod I_n).$$

Its kernel is $I_1 \cap ... \cap I_n$. By the universal property of the quotient, it only remains to prove that f is surjective. For that purpose, it suffices to check that $(1,0,0,\dots,0),(0,1,0,\dots,0),\dots,(0,0,\dots,0,1)$ all belong to the image of f. Without loss of generality, let us prove that $(1,0,0,\dots,0)$ is in the image of f. To do so, for each $2 \leq i \leq n$, we write $a_i + b_i = 1$ for some $a_i \in I_1$ and $b_i \in I_i$. We can then easily check that the image of $\prod_{i=2}^{n}(1-a_i) = \prod_{i=2}^{n} b_i$ by f is precisely $(1,0,0...0)$. □

Example 1.3.9. For $n \geq 2$, we have already seen that:

$$(\mathbb{Z}/n\mathbb{Z})^\times = \{[a] | 1 \leq a \leq n, \gcd(a,n) = 1\}.$$

In particular, if we introduce *Euler's totient function*:

$$\begin{aligned} \varphi : \mathbb{N} &\to \mathbb{N} \\ n &\mapsto |\{a \in \mathbb{N} | a \leq n, \gcd(a,n) = 1\}|, \end{aligned}$$

then the order of $(\mathbb{Z}/n\mathbb{Z})^\times$ is $\varphi(n)$. To compute the number $\varphi(n)$, let us write the prime factorization of n:

$$n = p_1^{r_1} ... p_s^{r_s}.$$

The Chinese Remainder Theorem then gives a ring isomorphism:

$$\mathbb{Z}/n\mathbb{Z} \cong \prod_{i=1}^{s} \mathbb{Z}/p_i^{r_i}\mathbb{Z},$$

and hence we get a group isomorphism:

$$(\mathbb{Z}/n\mathbb{Z})^\times \cong \prod_{i=1}^{s} (\mathbb{Z}/p_i^{r_i}\mathbb{Z})^\times. \tag{1.1}$$

Since:

$$|(\mathbb{Z}/p_i^{r_i}\mathbb{Z})^\times| = |\{a \in \mathbb{N} | a \leq p_i^{r_i}, p_i \nmid a\}| = p_i^{r_i - 1}(p_i - 1),$$

we deduce that:

$$\varphi(n) = \prod_{i=1}^{s} p_i^{r_i - 1}(p_i - 1).$$

By the isomorphism (1.1), the description of the group $(\mathbb{Z}/n\mathbb{Z})^\times$ can be reduced to the case where n is a power of a prime number. Hence the following theorem rounds off the description of the unit group of $\mathbb{Z}/n\mathbb{Z}$:

Theorem 1.3.10

Let r be a positive integer. Then the group $(\mathbb{Z}/2^r\mathbb{Z})^\times$ is isomorphic to:

$$\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2^{r-2}\mathbb{Z},$$

and for each odd prime number p, the group $(\mathbb{Z}/p^r\mathbb{Z})^\times$ is isomorphic to:

$$\mathbb{Z}/p^{r-1}(p-1)\mathbb{Z}.$$

The rest of this section is devoted to the proof of this result. We start with the case $r = 1$:

Proposition 1.3.11

Let K be a field. Any finite subgroup H of $K^\times$ is cyclic.

Proof. Let n be the order of H. For $d|n$, let H_d be the set of elements of H of order d. Assume that $H_d \neq \emptyset$ and take an element $x \in H_d$. The group $\langle x \rangle$ is then cyclic of order d. Since K is a field, we deduce that $\langle x \rangle = \{x \in K : x^d = 1\}$. In particular, H_d is the set of elements of order d in the cyclic group $\langle x \rangle$ and hence $|H_d| = \varphi(d)$. Therefore we get:

$$n = |H| = \sum_{d|n} |H_d| \leq \sum_{d|n} \varphi(d).$$

By observing that $\varphi(d)$ is the number of elements in $\mathbb{Z}/n\mathbb{Z}$ of order d for each $d|n$, we get $\sum_{d|n} \varphi(d) = n$, and hence the previous inequality is in fact an equality. In particular, $|H_n| = \varphi(n) \neq 0$, so that H is cyclic. □

In particular, we get that $(\mathbb{Z}/p\mathbb{Z})^\times \cong \mathbb{Z}/(p-1)\mathbb{Z}$ for each prime number p. We now settle two preliminary lemmas that will allow us to prove Theorem 1.3.10.

Lemma 1.3.12. *(i) Let p be an odd prime number. For every $s \geq 0$:*

$$(1+p)^{p^s} \equiv 1 + p^{s+1} \mod p^{s+2}.$$

(ii) For every $s \geq 0$:

$$5^{2^s} \equiv 1 + 2^{s+2} \mod 2^{s+3}.$$

Proof. (i) We proceed by induction on s. For $s = 0$, the congruence is obvious. Now consider an integer $s \geq 1$ and assume that:

$$(1+p)^{p^{s-1}} \equiv 1 + p^s \mod p^{s+1}.$$

Take $k \in \mathbb{Z}$ such that:

$$(1+p)^{p^{s-1}} = 1 + p^s + kp^{s+1}.$$

Then, by Newton's binomial expansion:

$$\begin{aligned}(1+p)^{p^s} &= (1+p^s+kp^{s+1})^p \\ &\equiv (1+p^s)^p \mod p^{s+2} \\ &\equiv 1+p^{s+1} \mod p^{s+2}.\end{aligned}$$

(ii) We proceed again by induction on s. For $s = 0$, the congruence is obvious. Now consider an integer $s \geq 1$ and assume that:

$$5^{2^{s-1}} \equiv 1 + 2^{s+1} \mod 2^{s+2}.$$

Take $k \in \mathbb{Z}$ such that:

$$5^{2^{s-1}} = 1 + 2^{s+1} + k2^{s+2}.$$

Then, by Newton's binomial expansion:

$$\begin{aligned}5^{2^s} &= (1+2^{s+1}+k2^{s+2})^2 \\ &\equiv 1+2^{s+2} \mod 2^{s+3}.\end{aligned}$$

□

Lemma 1.3.13. *Let G be a group and let g and h be elements of G of respective orders r ans s. Assume that $gh = hg$ and that* $\gcd(r,s) = 1$. *Then gh has order rs.*

Proof. Of course, we have $(gh)^{rs} = (g^r)^s(h^s)^r = 1$. Conversely, let n be an integer such that $(gh)^n = 1$ and let's prove that n is divisible by rs. To do so, observe that the order of $g^n = h^{-n}$ divides $\gcd(r,s) = 1$. Hence $xg^n = h^{-n} = 1$, and n has to be divisible both by r and s. By using again the coprimality of r and s, the integer n has to be divisible by rs. □

We are now ready to prove Theorem 1.3.10.

Proof of Theorem 1.3.10.

(i) On the one hand, by lemma 1.3.12(i), we have:

$$\begin{aligned}(1+p)^{p^{r-2}} &\not\equiv 1 \mod p^r, \\ (1+p)^{p^{r-1}} &\equiv 1 \mod p^r.\end{aligned}$$

Hence $1+p$ has order p^{r-1} in $(\mathbb{Z}/p^r\mathbb{Z})^\times$.

On the other hand, consider the projection $\pi : (\mathbb{Z}/p^r\mathbb{Z})^\times \to (\mathbb{Z}/p\mathbb{Z})^\times$ and fix a generator x of $(\mathbb{Z}/p\mathbb{Z})^\times$. Let $\tilde{x}$ be a lifting of x in $(\mathbb{Z}/p^r\mathbb{Z})^\times$. Since $x = \pi(\tilde{x})$ has order $p-1$, the order of $\tilde{x}$ is necessarily divisible by $p-1$, so that we can find a power y of $\tilde{x}$ that has order exactly $p-1$.

By lemma 1.3.13, the order of $(1+p)y$ is $(p-1)p^{r-1}$. This finishes the proof since $|(\mathbb{Z}/p^r\mathbb{Z})^\times| = (p-1)p^{r-1}$.

(ii) By lemma 1.3.12(ii), we have:

$$\begin{aligned} 5^{2^{r-3}} &\not\equiv 1 \mod 2^r, \\ 5^{2^{r-2}} &\equiv 1 \mod 2^r. \end{aligned}$$

Hence 5 has order 2^{r-2} in $(\mathbb{Z}/2^r\mathbb{Z})^\times$.

Now consider the projection $\pi : (\mathbb{Z}/2^r\mathbb{Z})^\times \to (\mathbb{Z}/4\mathbb{Z})^\times$ and the group homomorphism:

$$\begin{aligned} f : \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2^{r-2}\mathbb{Z} &\to (\mathbb{Z}/2^r\mathbb{Z})^\times \\ (a,b) &\mapsto (-1)^a 5^b. \end{aligned}$$

Let's prove that f is isomorphism. Since

$$|\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2^{r-2}\mathbb{Z}| = |(\mathbb{Z}/2^r\mathbb{Z})^\times|,$$

it is enough to check that f is injective. To do so, fix $(a,b) \in \ker(f)$. Then $(-1)^a 5^b = 1$. In particular:

$$1 = \pi((-1)^a 5^b) = (-1)^a$$

in $\mathbb{Z}/4\mathbb{Z}$, so that $a = 0$. But then $5^b = 1$ in $\mathbb{Z}/2^r\mathbb{Z}$ and hence $b = 0$.

□

1.4 Divisibility and arithmetic in a domain

Definition 1.4.1

Let A be a domain and let a and b be elements in A. ‘

(i) We say that a **divides** b, denoted $a|b$, if there exists $c \in A$ such that $b = ac$. Equivalently, $(b) \subseteq (a)$.

(ii) We say that a and b are **associated** if $a|b$ and $b|a$. Equivalently, there exists a unit $u \in A^\times$ such that $b = ua$.

(iii) The element a is said to be **irreducible** if it is not a unit and if, for any $x, y \in A$ with $xy = a$, either x or y is a unit. In particular, the divisors of a in A are the units and the elements of A that are associated to a.

(iv) The elements a and b are **coprime** if any common divisor of a and b is a unit.

Example 1.4.2.

(i) Take $A = \mathbb{Z}$. Two integers a and b are associated if, and only if, $a = -b$. The irreducible elements in $\mathbb{Z}$ are the prime numbers and their opposites.

(ii) Take $A = K[T]$ with K a field. Two polynomials P and Q are associated if, and only if, $P = \lambda Q$ for some $\lambda \in K^\times$. An element $P \in K[T]$ is irreducible if, and only if, it is an irreducible polynomial (in the sense you already know).

The two previous examples deal with the rings $\mathbb{Z}$ and $K[X]$, which are both principal ideal domains. Take now any principal ideal domain A and consider two elements $a, b \in A$. The ideal (a, b) is principal, and hence we can find $c \in A$ such that $(a, b) = (c)$. The element c, which is well-defined up to multiplication by a unit, is called the **gcd** of a and b. It divides both a and b, and any common divisor of a and b divides c. The elements a and b are coprime if, and only if, one of the following equivalent conditions is satisfied:

(a) the gcd of a and b is a unit;

(b) $(a, b) = A$;

(c) $1 \in (a, b)$;

(d) there exist u and v in A such that $au + bv = 1$.

Remark 1.4.3. The assertions (b), (c) and (d) are still equivalent in any domain. But they might not be equivalent to the coprimality of a and b. For instance, in the domain $A = K[X, Y]$, the elements X and Y are coprime, but $(X, Y) \neq A$.

1.5 Prime and maximal ideals

Definition 1.5.1

Let A be a ring.

(i) A **maximal ideal** in A is an ideal $\mathfrak{m}$ of A that is maximal among the strict ideals of A. Equivalently, it is an ideal such that the quotient $A/\mathfrak{m}$ is a field.

(ii) A **prime ideal** in A is an ideal $\mathfrak{p}$ such that the quotient $A/\mathfrak{p}$ is a domain. Equivalently, it is an ideal such that:

$$\forall (a, b) \in A^2, \quad ab \in \mathfrak{p} \Rightarrow (a \in \mathfrak{p} \text{ or } b \in \mathfrak{p}).$$

In particular, any maximal ideal is a prime ideal.

Proof.

(i) Assume that $\mathfrak{m}$ is a maximal ideal in A and let $\pi : A \to A/\mathfrak{m}$ be the natural projection. Take any non-zero element $\overline{a} \in A/\mathfrak{m}$. We can then find $a \in A \setminus \mathfrak{m}$ such that $\overline{a} = \pi(a)$. By maximality of $\mathfrak{m}$, we have $\mathfrak{m} + (a) = A$. Hence we can find $b \in \mathfrak{m}$ and $c \in A$ such that $b + ac = 1$. By applying π, we get $\overline{a}\pi(c) = \pi(a)\pi(c) = 1$. Thus $\overline{a} \in (A/\mathfrak{m})^\times$, and $A/\mathfrak{m}$ is indeed a field.

Conversely, assume that $A/\mathfrak{m}$ is a field and let I be an ideal in A that strictly contains $\mathfrak{m}$. Take $a \in I \setminus \mathfrak{m}$. Since $A/\mathfrak{m}$ is a field, $\pi(a) \in (A/\mathfrak{m})^\times$, and hence we can find $b \in A$ such that $\pi(a)\pi(b) = 1$. In other words, $ab - 1 \in \mathfrak{m}$, so that $1 \in I$ and $I = A$.

(ii) Assume that $A/\mathfrak{p}$ is a domain and take $(a, b) \in A^2$ such that $ab \in \mathfrak{p}$. If $\pi : A \to A/\mathfrak{p}$ is the natural projection, then $\pi(a)\pi(b) = 0$, and hence $\pi(a) = 0$ or $\pi(b) = 0$ since $A/\mathfrak{p}$ is a domain. We deduce that $a \in \mathfrak{p}$ or $b \in \mathfrak{p}$.

Conversely, assume that:

$$\forall (a, b) \in A^2, \quad ab \in \mathfrak{p} \Rightarrow (a \in \mathfrak{p} \text{ or } b \in \mathfrak{p}).$$

Take $\overline{a}$ and $\overline{b}$ in $A/\mathfrak{p}$ such that $\overline{a}\overline{b} = 0$. One can find a and b in A such that $\pi(a) = \overline{a}$ and $\pi(b) = \overline{b}$. We get $\pi(ab) = \pi(a)\pi(b) = 0$. Hence $ab \in \mathfrak{p}$. By assumption, we deduce that $a \in \mathfrak{p}$ or $b \in \mathfrak{p}$, so that $\overline{a} = 0$ or $\overline{b} = 0$. The ring $A/\mathfrak{p}$ is therefore a domain.

□

Proposition 1.5.2

Let A be a ring and let $a \in A \setminus \{0\}$. Consider the following assertions:

(i) the element a is irreducible in A.

(ii) the ideal (a) is prime.

(iii) the ideal (a) is maximal.

We have $(iii) \Rightarrow (ii) \Rightarrow (i)$. Moreover, if A is a principal ideal domain, then $(i) \Rightarrow (iii)$.

Proof. The implications $(iii) \Rightarrow (ii) \Rightarrow (i)$ follow immediately from the definitions.

Assume now that A is a principal domain and that a is irreducible. We are going to prove that (a) is a maximal ideal in A. To do so, let I be an ideal strictly containing (a), and take $b \in I \setminus (a)$. Since A is a principal ideal domain, we can find $c \in A$ such that $(a, b) = (c)$. But a is irreducible. Hence either c is a unit or c is associated to a. This second case is impossible because $b \notin (a)$. Hence c is a unit and $I \supseteq (c) = A$. This establishes the maximality of (a). □

Remark 1.5.3. The implications $(i) \Rightarrow (iii)$ and $(ii) \Rightarrow (iii)$ are false in general: for instance, the ideal (X) is a prime ideal in $\mathbb{C}[X, Y]$, but it is not maximal.

Theorem 1.5.4 (Krull)

Let A be a ring. For any strict ideal I in A, there exists a maximal ideal in A containing I.

Proof. Let $\mathscr{E}$ be the family of strict ideals in A containing I. Endow $\mathscr{E}$ with the order given by inclusion. The family $\mathscr{E}$ is not empty since it contains I, and if $(I_j)_{j\in J}$ is a totally ordered family of elements in $\mathscr{E}$, then $\bigcup_{j\in J} I_j \in \mathscr{E}$. By Zorn's lemma, we deduce that $\mathscr{E}$ has a maximal element. □

In particular, any non-zero ring has a maximal ideal.

1.6 Euclidean and principal ideal domains

Recall that the principality of $\mathbb{Z}$ and $K[X]$ follows from the existence of a Euclidean division in those rings. That is the most common method to prove that a ring is a principal ideal domain. To formalize this, we introduce the following definition:

Definition 1.6.1

(i) Let A be a domain. A **Euclidean function** on A is a map $f : A\setminus\{0\} \to \mathbb{N}_{\geq 0}$ satisfying the following property: for any a and b in A with $b \neq 0$, one can write $a = bq + r$ with $r = 0$ or $r \neq 0$ and $f(r) < f(b)$.

(ii) A **Euclidean domain** is a domain that has a Euclidean function.

Example 1.6.2.

(i) The domain $\mathbb{Z}$ is Euclidean with respect to the Euclidean function $f(n) = |n|$.

(ii) If K is a field, the ring $K[X]$ is Euclidean with respect to the Euclidean function $f(P) = \deg P$.

(iii) Consider the domain $A = \mathbb{Z}[i]$ and define the map:

$$\begin{aligned} f : A &\to \mathbb{N}_{\geq 0} \\ z &\mapsto |z|^2. \end{aligned}$$

Take a and b in A with $b \neq 0$, and write $\frac{a}{b} = x + iy$ with $x, y \in \mathbb{Q}$. We can find $x', y' \in \mathbb{Z}$ such that $|x - x'| \leq \frac{1}{2}$ and $|y - y'| \leq \frac{1}{2}$. Hence:

$$f\left(\frac{a}{b} - x' - iy'\right) = f\left((x - x') + i(y - y')\right) \leq \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 < 1.$$

Hence:

$$f(a - b(x' + iy')) < f(b),$$

and A is a Euclidean domain.

Theorem 1.6.3

Any Euclidean domain is a principal ideal domain.

Proof. Let A be a Euclidean domain with respect to some Euclidean function f. Let I be a non-zero ideal of A and consider an element $b \in I \setminus \{0\}$ that minimizes $f(b)$. Of course, $(b) \subseteq I$. Conversely, given $a \in I$, we can find q and r in A such that $a = bq + r$ and either $r = 0$ or $r \neq 0$ and $f(r) < f(b)$. Since $r = a - bq \in I$, the minimality of $f(b)$ implies that the case $r \neq 0$ and $f(r) < f(b)$ is excluded. Hence $r = 0$ and $a \in (b)$. As a consequence, $I = (a)$, and A is indeed a principal ideal domain. □

Corollary 1.6.4

The ring $\mathbb{Z}[i]$ is a principal ideal domain.

Remark 1.6.5. There are principal ideal domains that are not Euclidean. That is for instance the case of the domain $\mathbb{Z}\left[\frac{1+i\sqrt{19}}{2}\right]$.

1.7 Unique factorization domains

Definition 1.7.1

Let A be a ring. We say that A is a **unique factorization domain** if it is a domain, each element $a \in A \setminus \{0\}$ can be written as:

$$a = up_1 \ldots p_r \tag{1.2}$$

where $u \in A^\times$ and $p_1, \ldots, p_r$ are irreducible elements in A, and this decomposition is unique in the following sense: if $a = vq_1 \ldots q_s$ is any other decomposition of a in which $v \in A^\times$ and $q_1, \ldots q_s$ are irreducible elements, then $r = s$ and there exist a permutation $\sigma \in \mathcal{S}_r$ and units $u_1, \ldots, u_r \in A^\times$ such that $q_i = u_i p_{\sigma(i)}$ for each i.

Let A be a unique factorization domain, and fix a set of irreducible elements $\mathcal{P}$ in A such that each irreducible element in A is associated to exactly one element in $\mathcal{P}$. Every non-zero element $a \in A$ can then be written as:

$$a = u \prod_{p \in \mathcal{P}} p^{\alpha_p}$$

with $u \in A^\times$ and $\alpha_p \in \mathbb{N}_{\geq 0}$ for each p. The unit u and the α_p's are uniquely determined, and hence we can set $v_p(a) := \alpha_p$.

Now take two non-zero elements a and b in A. Then:

(i) a divides b if, and only if, $v_p(a) \leq v_p(b)$ for all $p \in \mathscr{P}$;

(ii) a and b are associated if, and only if, $v_p(a) = v_p(b)$ for all $p \in \mathscr{P}$;

(iii) a and b are coprime if, and only if, $v_p(a) = 0$ or $v_p(b) = 0$ for each $p \in \mathscr{P}$.

Moreover the gcd and the lcm of a and b can be defined by:

$$\gcd(a,b) := \prod_{i=1}^{r} p_i^{\min\{v_p(a),v_p(b)\}},$$
$$\operatorname{lcm}(a,b) := \prod_{i=1}^{r} p_i^{\max\{v_p(a),v_p(b)\}}.$$

Even if these definitions depend on the choice of the family of irreducible elements $\mathscr{P}$, the gcd and the lcm are well-defined up to multiplication by a unit in A.

The gcd divides both a and b and every common divisor of a and b divides it. Similarly, the lcm is a multiple of both a and b, and it divides every common multiple of a and b. Note however that, contrary to what happens in a principal ideal domain, the ideal (a,b) is not necessarily spanned by $\gcd(a,b)$.

Proposition 1.7.2

Let A be a unique factorization domain and let p be an irreducible element in A. The ideal (p) is prime.

Proof. Take two elements $b, c \in A \setminus \{0\}$ such that $bc \in (p)$. Since p divides bc, we have:

$$v_p(bc) = v_p(b) + v_p(c) \geq 1.$$

Hence $v_p(b) \geq 1$ or $v_p(c) \geq 1$, so that p divides b or c, as wished. □

Theorem 1.7.3

A principal ideal domain is a unique factorization domain.

Proof. Let A be a principal ideal domain, and take a non-zero element $a \in A$. Assume that a cannot be written in the form (1.2). Then a is not a unit and is not irreducible. We can therefore write $a = a_1 b_1$, with a_1 and b_1 in $A \setminus A^\times$. Of course, at least one of the elements a_1 and b_1 cannot be written in the form (1.2). Hence, by repeating this argument, we can construct an infinite sequence $a_0 = a, a_1, a_2, \ldots$ such that, for each i, we have $a_{i+1} | a_i$, but a_i and a_{i+1} are not

associated. Let I be the ideal spanned by all the a_i's. Since A is a principal ideal domain, we can find $c \in A$ such that $I = (c)$. But $I = \bigcup_i (a_i)$. Hence $c \in (a_n)$ for some $n \geq 0$, so that $a_n | c | a_{n+1}$: this contradicts the fact that a_n and a_{n+1} are not associated! We deduce that a can be written in the form (1.2).

Consider now two units $u, v \in A^\times$, pairwise non-associated irreducible elements $p_1, ..., p_r$, and non-negative integers $\alpha_1, ..., \alpha_r, \beta_1, ..., \beta_r$ such that:

$$a = u p_1^{\alpha_1} ... p_r^{\alpha_r} = v p_1^{\beta_1} ... p_r^{\beta_r}.$$

Assume by contradiction that $\alpha_1 > \beta_1$. Then p_1 should divide $v p_2^{\beta_2} ... p_r^{\beta_r}$. But since A is a principal ideal domain, the ideal (p_1) is prime, and hence p_1 should divide v or one of the p_i's with $i \neq 1$: contradiction! We deduce that $\alpha_1 \leq \beta_1$, and by symmetry that $\alpha_1 = \beta_1$. The same argument can be used to prove that $\alpha_i = \beta_i$ for all i. □

Corollary 1.7.4

The ring $\mathbb{Z}[i]$ is a unique factorization domain.

Remark 1.7.5. There are domains that do not have unique factorization. That is for instance the case of the domain $\mathbb{Z}\left[\sqrt{5}\right]$. One has:

$$2 \times 2 = (3 + \sqrt{5})(3 - \sqrt{5})$$

and one can check that 2, $3 + \sqrt{5}$ and $3 - \sqrt{5}$ are non-associated irreducible elements in $\mathbb{Z}\left[\sqrt{5}\right]$.

1.8 Polynomials over unique factorization domains

In this section, we are interested in studying irreducible elements in polynomial rings. We focus on the most favorable situation where our polynomials have coefficients in a unique factorization domain. For instance, the results of this section can be applied to the polynomial ring $\mathbb{Z}[T]$

Definition 1.8.1

Let A be a unique factorization domain with fraction field K and let $P \in K[T]$. Given $a \in A \setminus \{0\}$ such that $aP \in A[T]$, the **content** of P, denoted $\mathrm{ct}(P)$, is the class of $\frac{d}{a}$ in the quotient $K^\times / A^\times$ where d is the gcd of the coefficients of aP. It is independent of a.

Proposition 1.8.2 (Gauss)

Let A be a unique factorization domain. For any P and Q in $K[T]$, we have:

$$\mathrm{ct}(PQ) = \mathrm{ct}(P) \cdot \mathrm{ct}(Q).$$

Carl Friedrich Gauss (Brunswick 1777 - Göttingen 1855)

Proof. Up to replacing P and Q by $\mathrm{ct}(P)^{-1}P$ and $\mathrm{ct}(Q)^{-1}Q$, we can assume that $\mathrm{ct}(P) = \mathrm{ct}(Q) = 1$. By contradiction, assume that $\mathrm{ct}(PQ) \neq 1$. We can then find an irreducible element π that divides $\mathrm{ct}(PQ)$. By denoting by $\overline{P}$ and $\overline{Q}$ the reductions of P and Q modulo π, we get the following equality in $(A/(\pi))[T]$:

$$\overline{P} \cdot \overline{Q} = 0.$$

But $A/(\pi)$ is a domain by proposition 1.7.2, hence so is $(A/(\pi))[T]$ and we get that $\overline{P} = 0$ or $\overline{Q} = 0$. This means that π divides $\mathrm{ct}(P)$ or $\mathrm{ct}(Q)$: contradiction! Hence $\mathrm{ct}(PQ) = 1$, as wished. □

Theorem 1.8.3

Let A be a unique factorization domain with fraction field K. The irreducible elements of $A[T]$ are:

(a) the constant polynomials p with p irreducible in A;

(b) the polynomials of degree at least 1 that have content 1 and that

are irreducible in $K[T]$.

Proof. Since $(A[T])^\times = A^\times$, it is clear that the constant polynomials that are irreducible in $A[T]$ are the polynomials stated in (a).

Let now $P \in A[T]$ be a polynomial of degree at least 1 that is irreducible in $K[T]$ and such that $\mathrm{ct}(P) = 1$. Let $Q, R \in A[T]$ such that $P = QR$. By Proposition 1.8.2, we have $\mathrm{ct}(Q) = \mathrm{ct}(R) = 1$. Moreover, since P is irreducible in $K[T]$, we have $Q \in K^\times$ or $R \in K^\times$. Hence $Q \in A^\times$ or $R \in A^\times$, and P is irreducible in $A[T]$.

Conversely, assume that P is a non-constant irreducible polynomial in $A[T]$. Since $\mathrm{ct}(P)$ divides P, the content of P is necessarily 1. Let us prove that P is irreducible in $K[T]$. To do so, take $Q, R \in K[T]$ such that $P = QR$. Then $1 = \mathrm{ct}(P) = \mathrm{ct}(Q)\mathrm{ct}(R)$. Hence:

$$P = \frac{Q}{\mathrm{ct}(Q)} \cdot \frac{R}{\mathrm{ct}(R)}$$

with $\frac{Q}{\mathrm{ct}(Q)} \in A[T]$ and $\frac{R}{\mathrm{ct}(R)} \in A[T]$. Since P is irreducible in $A[T]$, we deduce that $\frac{Q}{\mathrm{ct}(Q)} \in A^\times$ or $\frac{R}{\mathrm{ct}(R)} \in A^\times$. Hence $Q \in K^\times$ or $R \in K^\times$. This establishes that P is irreducible in $K[T]$. □

Theorem 1.8.4

Let A be a unique factorization domain. Then so is $A[T]$.

Proof. Let K be the fraction field of A and let P be a non-zero polynomial in $A[T]$. If P is a constant polynomial, then it can be written in the form:

$$P = up_1 \ldots p_r$$

where $u \in A^\times$ and $p_1, \ldots, p_r$ are irreducible elements in A. But then $u \in A[T]^\times$ and, by Theorem 1.8.3, $p_1, \ldots p_r$ are irreducible in $A[T]$. Hence P is either a unit or a product of irreducible elements in $A[T]$.

Assume now that P has degree at least 1. Since $K[T]$ is a unique factorization domain, we can write P in the form:

$$P = aP_1 \ldots P_r$$

with $P_1, \ldots, P_r$ irreducible polynomials in $K[T]$ and $a \in A$. By Proposition 1.8.2, we have:

$$a \cdot \mathrm{ct}(P_1) \ldots \mathrm{ct}(P_r) = \mathrm{ct}(P) \in A.$$

Hence, by setting $\tilde{P}_i = \mathrm{ct}(P_i)^{-1} P_i$ for each i, we get:

$$P = \lambda \tilde{P}_1 ... \tilde{P}_r$$

with $\lambda = a \cdot \mathrm{ct}(P_1)...\mathrm{ct}(P_r) \in A$. We already know that λ is either a unit or a product of irreducible elements in $A[T]$. Moreover, by Theorem 1.7.2, each $\tilde{P}_i$ is irreducible in $A[T]$ since it has content 1 and it is irreducible in $K[T]$. Hence P is either a unit or a product of irreducible elements in $A[T]$.

Assume now that:

$$P = u P_1^{\alpha_1} ... P_r^{\alpha_r} = v P_1^{\beta_1} ... P_r^{\beta_r}$$

where u, v are units in $A[T]^\times = A^\times$, $P_1, ..., P_r$ are pairwise non-associated irreducible elements in $A[T]$, and $\alpha_1, ..., \alpha_r, \beta_1, ..., \beta_r$ are non-negative integers. We can assume without loss of generality that there exists $k \in \{1, ..., r+1\}$ such that $P_1, ..., P_{k-1}$ are constant polynomials and $P_k, ..., P_r$ have degree at least 1. Since $K[T]$ has unique factorization and since $P_k, ..., P_r$ are all irreducible in $K[T]$ by Theorem 1.7.2, we deduce that $\alpha_i = \beta_i$ for each $i \in \{k, ..., r\}$. As a consequence, we get:

$$u P_1^{\alpha_1} ... P_{k-1}^{\alpha_{k-1}} = v P_1^{\beta_1} ... P_{k-1}^{\beta_{k-1}}.$$

But A has unique factorization and $P_1, ..., P_{k-1}$ are irreducible in A by Theorem 1.7.2. Hence $\alpha_i = \beta_i$ for each $i \in \{1, ..., k-1\}$. It follows that $A[T]$ has unique factorization. □

Example 1.8.5. For any $n \geq 1$, the rings $\mathbb{C}[X_1, ..., X_n]$ and $\mathbb{Z}[X_1, ..., X_n]$ are unique factorization domains.

Theorem 1.8.6 – Eisenstein's criterion

Let A be a unique factorization domain with field of fractions K and let π be an irreducible element in A. Let $P(T) = \sum_{k=0}^{n} a_k T^k$ be a polynomial in $A[T]$. Assume that:

(a) π does not divide a_n;

(b) π divides a_k for each $k \in \{0, ..., n-1\}$;

(c) π^2 does not divide a_0.

Then P is irreducible in $K[T]$.

Proof. Let $Q, R \in K[T]$ be non-constant polynomials such that $P = QR$. Let $m \in \{1, ..., n-1\}$ be the degree of Q. By Lemma 1.8.2, we have $\mathrm{ct}(P) = \mathrm{ct}(Q)\mathrm{ct}(R)$.

Hence, by replacing P, Q and R by $\frac{P}{\mathrm{ct}(P)}$, $\frac{Q}{\mathrm{ct}(Q)}$ and $\frac{R}{\mathrm{ct}(R)}$, we can assume that $\mathrm{ct}(P) = \mathrm{ct}(Q) = \mathrm{ct}(R) = 1$.

Let now $\overline{P}$, $\overline{Q}$ and $\overline{R}$ be the reductions of P, Q and R modulo π. We then have the following equality in $(A/(\pi))[T]$:

$$\overline{a_n}T^n = \overline{P} = \overline{Q} \cdot \overline{R}$$

where $\overline{a_n} \in A/(\pi)$ is the reduction of a_n modulo π. But $\overline{a_n} \neq 0$ and $A/(\pi)$ is a domain by Proposition 1.7.2. Hence $\overline{Q} = \alpha T^m$ and $\overline{R} = \beta T^{n-m}$ for some $\alpha, \beta \in (A/(\pi)) \setminus \{0\}$. In particular, π divides both $Q(0)$ and $R(0)$. But by assumption, π^2 does not divide $a_0 = P(0) = Q(0)R(0)$: contradiction! Hence P is irreducible in $K[T]$. □

Example 1.8.7. The polynomial $T^{10} + 4T^3 + 2T + 6$ is irreducible in $\mathbb{Z}[T]$ and in $\mathbb{Q}[T]$.

Gotthold Eisenstein (Berlin 1823 - 1852)

Chapter 2

Number fields and their rings of integers

2.1 Rings of algebraic integers

Definition 2.1.1

A complex number is said to be an **algebraic integer** if it is a root of a monic polynomial with integer coefficients. The set of algebraic integers is denoted $\overline{\mathbb{Z}}$.

Example 2.1.2. We have $\mathbb{Z} \subset \overline{\mathbb{Z}}$. For $n \geq 1$, the complex numbers $\sqrt{n}$, $i\sqrt{n}$ and $e^{2i\pi/n}$ are algebraic integers.

Proposition 2.1.3

A rational number is an algebraic integer if, and only if, it is an integer. In other words, $\overline{\mathbb{Z}} \cap \mathbb{Q} = \mathbb{Z}$.

Proof. Take a rational number $x = p/q$, with $p \in \mathbb{Z}$, $q \in \mathbb{Z} \setminus \{0\}$ and $p \wedge q = 1$. Assume it is a root of some monic polynomial $f = X^n + \sum_{i=0}^{n-1} a_i X^i \in \mathbb{Z}[X]$. Then:

$$p^n + q \sum_{i=0}^{n-1} a_i p^i q^{n-i-1} = 0$$

Hence q divides p^n. But p and q are coprime. Hence $q = 1$ and $x \in \mathbb{Z}$. □

Let us now provide a useful criterion to prove that a complex number is an algebraic integer, and more generally to prove that an element of some ring satisfies some monic polynomial equation.

Proposition 2.1.4

Let $A \subset B$ be an inclusion of domains. Fix a finite number of elements $b_1, \ldots, b_n \in B$, assumed to be not all zero, and set $\Gamma := b_1 A + \cdots + b_n A \subset B$. Consider an element $b \in B$ such that $b\Gamma \subset \Gamma$, and for $1 \leq i \leq n$, write:

$$bb_i = \sum_{j=1}^{n} a_{ij} b_i$$

with $a_{ij} \in A$. Introduce the monic polynomial:

$$\chi(X) := \det(X\mathrm{I}_n - (a_{ij})_{1 \leq i,j \leq n}) \in A[X].$$

Then $\chi(b) = 0$.

Remark 2.1.5. In the proof, we will use without further explanation the following basic result of linear algebra: for any matrix $M \in \mathscr{M}_n(A)$ with coefficients in a ring A, we have ${}^t(\mathrm{Com}(M))M = \det M \cdot \mathrm{I}_n$ where $\mathrm{Com}(M)$ stands for the comatrix of M.

Proof. Consider the matrix $M := b\mathrm{I}_n - (a_{ij})_{1 \leq i,j \leq n} \in \mathscr{M}_n(B)$. We then have:

$$M \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = 0,$$

and hence:

$$\det M \cdot \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = {}^t(\mathrm{Com}(M))M \cdot \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = 0.$$

We deduce that $\det M = 0$, and hence that $\chi(b) = 0$. □

Example 2.1.6. Take $A = \mathbb{Z}$, $B = \mathbb{C}$ and $b_1 = 1$, $b_2 = \sqrt[3]{2}$, $b_3 = \sqrt[3]{4}$ and $b = 1 + 3\sqrt[3]{2} - \sqrt[3]{4}$. We have $b\Gamma \subset \Gamma$, and we can compute:

$$\chi = \begin{vmatrix} X-1 & 2 & -6 \\ -3 & X-1 & 2 \\ 1 & -3 & X-1 \end{vmatrix} = X^3 - 3X^2 + 21X + 31.$$

We deduce that b is a root of χ, and hence that b is an algebraic integer.

Theorem 2.1.7

The set $\overline{\mathbb{Z}}$ is a subring of $\mathbb{C}$.

Proof. Consider two elements $x, y \in \overline{\mathbb{Z}}$ and let $A = \mathbb{Z}[x, y]$ be the subring of $\mathbb{C}$ spanned by x and y. Let P and Q be monic polynomials with integer coefficients such that $P(x) = 0$ and $Q(y) = 0$, and let m and n be their respective degrees. Observe that x^m (resp. y^n) is a linear combination with integers coefficients of $1, x, x^2, \dots x^{m-1}$ (resp. of $1, y, y^2, \dots y^{n-1}$). A simple induction shows that the same holds for all powers of x (resp. y). In particular, A is spanned as an abelian group by the finite family $(x^i y^j)_{0 \leq i \leq m-1, 0 \leq j \leq n-1}$. But $(x+y)A \subset A$ and $xyA \subset A$. Hence, by Proposition 2.1.4, both $x+y$ and xy belong to $\overline{\mathbb{Z}}$. □

Corollary 2.1.8

An algebraic complex number $z \in \mathbb{C}$ is an algebraic integer if, and only if, its minimal polynomial has integer coefficients.

Proof. The converse direction is obvious. Let us therefore assume that z is an algebraic integer and let $P \in \mathbb{Z}[X]$ be a monic polynomial such that $P(z) = 0$. Let $z_1 = z, \dots, z_r$ be the conjugates of z, that is the roots of the minimal polynomial μ of z. They are all roots of P, and hence they are all algebraic integers. Since $\overline{\mathbb{Z}}$ is a ring and the coefficients of μ belong to $\mathbb{Z}[z_1, \dots z_r]$, we deduce that the coefficients of μ belong to $\mathbb{Q} \cap \overline{\mathbb{Z}} = \mathbb{Z}$. □

Definition 2.1.9

A **number field** is a finite extension K of $\mathbb{Q}$. Its **ring of integers** is $\mathcal{O}_K := \overline{\mathbb{Z}} \cap K$.

Example 2.1.10. Let $d \in \mathbb{Z} \setminus \{0, 1\}$ be a square-free integer and set $K := \mathbb{Q}(\sqrt{d})$. Take an element $x = a + b\sqrt{d} \in K$ with $a \in \mathbb{Q}$ and $b \in \mathbb{Q}^\times$. Its minimal polynomial is $X^2 - 2aX + a^2 - db^2$, and hence x is an algebraic integer if, and only if, $2a$ and $a^2 - db^2$ both belong to $\mathbb{Z}$. The latter condition implies that $d(2b)^2 = (2a)^2 - 4(a^2 - db^2) \in \mathbb{Z}$, and hence that $2b \in \mathbb{Z}$ since d is square-free. Writing $a = m/2$ and $b = n/2$ with $m, n \in \mathbb{Z}$, the condition $a^2 - db^2 = \frac{m^2 - dn^2}{4} \in \mathbb{Z}$ means that:

- m and n have the same parity if $d \equiv 1 \mod 4$,
- m and n are both even if $d \equiv 2, 3 \mod 4$.

We deduce that:

$$\mathcal{O}_K = \begin{cases} \{a+b\sqrt{d} : (a,b) \in \mathbb{Z}^2 \cup (1/2+\mathbb{Z})^2\} = \mathbb{Z}[\frac{1+\sqrt{d}}{2}] & \text{if } d \equiv 1 \mod 4, \\ \{a+b\sqrt{d} : (a,b) \in \mathbb{Z}^2\} = \mathbb{Z}[\sqrt{d}] & \text{if } d \equiv 2,3 \mod 4. \end{cases}$$

In general, it can be quite hard to compute the ring of integers of a given number field. In the next section, we provide a theoretical method to do so.

2.2 How to compute rings of integers?

Take a finite field extension F/E. Given an element $x \in F$, Proposition 2.1.4 provides a simple way to find a non-zero polynomial $f \in E[T]$ such that $f(x) = 0$. Indeed, the multiplication-by-x map:

$$\begin{aligned} m_x : A &\to A \\ y &\mapsto xy \end{aligned}$$

is an F-linear map and x is a root of its characteristic polynomial:

$$\chi_{F/E,x}(T) := \det(T\mathrm{Id}_A - m_x).$$

Definition 2.2.1

The polynomial $\chi_{F/E,x}(T) := \det(T\mathrm{Id}_A - m_x)$ is called the **characteristic polynomial of** x. One also defines the **trace** and the **norm** of x as:

$$\mathrm{Tr}_{F/E}(x) := \mathrm{Tr}(m_x), \qquad N_{F/E}(x) := \det(m_x).$$

If we write $\chi_{F/E,x}(T) = T^n - c_1 T^{n-1} + \cdots + (-1)^n c_n$, we then have:

$$n = [F:E], \qquad c_1 = \mathrm{Tr}_{F/E}(x), \qquad c_n = N_{F/E}(x).$$

The trace defines an E-linear map $\mathrm{Tr}_{F/E} : F \to E$ and the norm defines a group homomorphism $N_{F/E} : F^\times \to E^\times$.

Let us now focus on the case where E and F are number fields. According to the Primitive Element Theorem[1], we can write $F = E(\alpha)$ for some $\alpha \in F$. Denote by $\mu_\alpha \in E[T]$ the minimal polynomial of α. The field F can then be identified with

[1] If F/E is a finite separable extension of fields (for instance a finite extension of fields of characteristic zero), then there exists $\alpha \in F$ such that $F = E(\alpha)$.

$E[T]/(\mu_\alpha(T))$, and hence, by the universal property of the quotient, the field embeddings of F into $\mathbb{C}$ fixing E can be described as:

$$\begin{aligned} \iota_\beta : F = E[T]/(\mu_\alpha(T)) &\to \mathbb{C} \\ P(T) &\mapsto P(\beta) \end{aligned}$$

for some root β of μ_α. This shows that the set of such embeddings, which we denote $\mathrm{Hom}_E(F,\mathbb{C})$, has exactly $n = [F:E] = \deg \mu_\alpha$ elements.

Proposition 2.2.2

With the previous notations, given an element $x \in F$ with minimal polynomial μ_x, we have:

$$\begin{aligned} \chi_{F/E,x}(T) = \mu_x(T)^{[F:E(x)]} &= \prod_{\sigma \in \mathrm{Hom}_E(F,\mathbb{C})} (T - \sigma(x)), \\ \mathrm{Tr}_{F/E}(x) &= \sum_{\sigma \in \mathrm{Hom}_E(F,\mathbb{C})} \sigma(x), \\ N_{F/E}(x) &= \prod_{\sigma \in \mathrm{Hom}_E(F,\mathbb{C})} \sigma(x). \end{aligned}$$

Proof. Set:

$$u := [E(x):E] = \deg \mu_x, \quad v := [F:E(x)], \quad n := uv = [F:E],$$

and write:

$$\mu_x(T) = T^u + c_1 T^{u-1} + \cdots + c_u \in E[T].$$

The family $(1, x, x^2, \ldots, x^{u-1})$ is an E-basis of $E(x)$. Hence, if $(\alpha_1, \ldots, \alpha_v)$ stands for some $E(x)$-basis of F, the family:

$$(\alpha_1, \alpha_1 x, \ldots, \alpha_1 x^{u-1}, \alpha_2, \alpha_2 x, \ldots, \alpha_2 x^{u-1}, \ldots, \alpha_v, \alpha_v x, \ldots, \alpha_v x^{u-1})$$

is an E-basis of F. In that E-basis, the matrix of the multiplication-by-x map m_x is a block diagonal matrix, that has exactly v diagonal blocks all of them equal to:

$$C(\mu_x) := \begin{pmatrix} 0 & 0 & 0 & \ldots & 0 & -c_u \\ 1 & 0 & 0 & \ldots & 0 & -c_{u-1} \\ 0 & 1 & 0 & \ldots & 0 & -c_{u-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & -c_{u-1} \end{pmatrix}.$$

Since $C(\mu_x)$ has characteristic polynomial μ_x, we deduce that $\chi_{F/E,x}(T) = \mu_x(T)^v$. But $\mu_x(T) = \prod_{\sigma \in \mathrm{Hom}_E(E(x),\mathbb{C})}(T - \sigma(x))$, and hence:

$$\chi_{F/E,x}(T) = \prod_{\sigma \in \mathrm{Hom}_E(E(x),\mathbb{C})} (T - \sigma(x))^v.$$

To conclude, it only remains to check that, for each $\sigma \in \mathrm{Hom}_E(E(x),\mathbb{C})$, there exist exactly v embeddings $\tilde{\sigma} \in \mathrm{Hom}_E(F,\mathbb{C})$ that extend σ. To do so, one can apply the Primitive Element Theorem to write $F = E(x,y)$ for some $y \in F$, so that $F \cong E(x)[T]/(\mu_y(T))$ with μ_y the minimal polynomial of y over $E(x)$. According to the universal property of quotient, the embeddings $\tilde{\sigma} \in \mathrm{Hom}_E(F,\mathbb{C})$ that extend σ can then be described as:

$$\begin{aligned} \iota_{\sigma,\beta} : F = E(x)[T]/(\mu_y(T)) &\to \mathbb{C} \\ P(T) &\mapsto \sigma(P)(\sigma(\beta)) \end{aligned}$$

for some root β of μ_y. This shows that there are exactly $v = [F : E(x)] = \deg \mu_y$ such embeddings. □

Corollary 2.2.3

Let K be a number field. An element $x \in K$ belongs to $\mathcal{O}_K$ if, and only if, the coefficients of $\chi_{K/\mathbb{Q},x}$ belong to $\mathbb{Z}$. In particular, if $x \in \mathcal{O}_K$, then $\mathrm{Tr}_{K/\mathbb{Q}}(x)$ and $N_{K/\mathbb{Q}}(x)$ both belong to $\mathbb{Z}$.

Proof. Assume that x belongs to $\mathcal{O}_K$. Then $\sigma(x)$ is an algebraic integer for every $\sigma \in \mathrm{Hom}_{\mathbb{Q}}(K,\mathbb{C})$. Hence, according to Proposition 2.2.2, all coefficients of $\chi_{K/\mathbb{Q},x}$ belong to $\mathcal{O}_K \cap \mathbb{Q} = \mathbb{Z}$. The converse and the statement about the trace and the norm are obvious. □

Let us now take a number field K and let $\mathbf{x} := (x_1,\dots,x_n)$ be a $\mathbb{Q}$-basis of K. Take some $x \in K$ and write $x = \sum_{i=1}^n \lambda_i x_i$ for some $\lambda_1,\dots,\lambda_n \in \mathbb{Q}$. We want to find some necessary conditions for x to be in $\mathcal{O}_K$ in terms of the coordinates $\lambda_1,\dots,\lambda_n$.

For that purpose, we start by reducing to the case where the $\mathbb{Q}$-basis $\mathbf{x}$ is contained in $\mathcal{O}_K$ thanks to the following Lemma:

Lemma 2.2.4. *Given any $y \in K$, there exists a positive integer n_y such that $n_y y \in \mathcal{O}_K$.*

Proof. Let $\mu_y = T^d + a_{d-1}T^{d-1} + \cdots + a_0 \in \mathbb{Q}[T]$ be the minimal polynomial y, and let n_y be the least common multiple of the denominators of the a_i's. Then $n_y y$ is a root of the polynomial $T^d + a_{d-1}n_y T^{d-1} + \cdots + a_0 n_y^d$, which is monic and has integer coefficients. We deduce that $n_y y$ is in $\mathcal{O}_K$. □

Up to multiplying the $\mathbb{Q}$-basis $\mathbf{x}$ by a suitable positive integer, we henceforth assume that $\mathbf{x}$ lies in $\mathscr{O}_K$. Now, by Corollary 2.2.3, for $x = \sum_{i=1}^{n} \lambda_i x_i$ to be in $\mathscr{O}_K$, it is necessary that, for each $1 \leq j \leq n$:

$$\mathrm{Tr}_{K/\mathbb{Q}}(xx_j) = \sum_{i=1}^{n} \lambda_i \mathrm{Tr}_{K/\mathbb{Q}}(x_i x_j) \in \mathbb{Z}.$$

Setting $M_{\mathbf{x}} := \left(\mathrm{Tr}_{K/\mathbb{Q}}(x_i x_j)\right)_{1\leq i,j\leq n} \in \mathscr{M}_n(\mathbb{Z})$, this system of equations can be rewritten as:

$$M_{\mathbf{x}} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \in \mathbb{Z}^n.$$

Hence:

$$\det(M_{\mathbf{x}}) \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = {}^t\mathrm{Com}(M_{\mathbf{x}}) \cdot M_{\mathbf{x}} \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \in \mathbb{Z}^n,$$

and for each i, the rational number $\det(M_{\mathbf{x}})\lambda_i$ is an integer. In other words, we have proved the following proposition:

Proposition 2.2.5

Let K be a number field and let $\mathbf{x} = (x_1, \ldots, x_n)$ be a $\mathbb{Q}$-basis of K that is contained in $\mathscr{O}_K$. Then:

$$d(\mathbf{x}) \cdot \mathscr{O}_K \subset \bigoplus_{i=1}^{n} \mathbb{Z} \cdot x_i,$$

where:

$$d(\mathbf{x}) := \det\left(\mathrm{Tr}_{K/\mathbb{Q}}(x_i x_j)\right)_{1\leq i,j\leq n}.$$

The integer $d(\mathbf{x})$ is called the **discriminant** of the basis $\mathbf{x}$. According to Proposition 2.2.2, if $\sigma_1, \ldots, \sigma_n$ stand for the field embeddings of K in $\mathbb{C}$, we have:

$$\left(\mathrm{Tr}_{K/\mathbb{Q}}(x_i x_j)\right)_{1\leq i,j\leq n} = \left(\sum_{r=1}^{n} \sigma_r(x_i)\sigma_r(x_j)\right)_{1\leq i,j\leq n} = {}^t\left(\sigma_i(x_j)\right)_{1\leq i,j\leq n} \cdot \left(\sigma_i(x_j)\right)_{1\leq i,j\leq n},$$

and hence the discriminant can be computed as:

$$d(\mathbf{x}) := \det\left(\mathrm{Tr}_{K/\mathbb{Q}}(x_i x_j)\right)_{1\leq i,j\leq n} = \left(\det\left(\sigma_i(x_j)\right)_{1\leq i,j\leq n}\right)^2.$$

A particularly interesting case is that of a basis $\mathbf{x}$ of the form $(1, x, x^2, \ldots, x^{n-1})$. In that case, we get:

$$d(1, x, x^2, \ldots, x^{n-1}) = \left(\det\left(\sigma_i(x)^{j-1}\right)_{1\leq i,j\leq n}\right)^2,$$

and hence, by recognizing a Vandermonde determinant:

$$d(1, x, x^2, \ldots, x^{n-1}) = \left(\prod_{i<j}(\sigma_i(x) - \sigma_j(x))\right)^2.$$

If $\mu_x = \prod_{i=1}^{n}(T - \sigma_i(x))$ is the minimal polynomial of x, this formula can be rewritten as:

$$d(1, x, x^2, \ldots, x^{n-1}) = (-1)^{n(n-1)/2} \cdot \prod_{i=1}^{n} \mu_x'(\sigma_i(x)). \tag{2.1}$$

Denoting by $\alpha_1, \ldots, \alpha_{n-1}$ the roots of μ_x', we get:

$$\begin{aligned} d(1, x, x^2, \ldots, x^{n-1}) &= (-1)^{n(n-1)/2} \cdot n^n \cdot \prod_{i=1}^{n}\prod_{j=1}^{n-1}(\sigma_i(x) - \alpha_j) \\ &= (-1)^{n(n-1)/2} \cdot n^n \cdot \prod_{j=1}^{n-1} \mu_x(\alpha_j). \end{aligned} \tag{2.2}$$

Equations (2.1) and (2.2) are often useful to compute $d(1, x, x^2, \ldots, x^{n-1})$.

Coming back to Proposition 2.2.5, for its statement to be not empty, the discriminant $d(\mathbf{x})$ has to be non-zero. This is in fact always the case, as the following proposition shows:

Proposition 2.2.6

Let K be a number field and let $\mathbf{x} = (x_1, \ldots, x_n)$ be a $\mathbb{Q}$-basis of K. The discriminant $d(\mathbf{x})$ is non-zero.

We start by proving the following lemma that shows that the non-vanishing of $d(\mathbf{x})$ is independent of the choice of the basis $\mathbf{x}$:

Lemma 2.2.7. *Let K be a number field and let $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$ be two $\mathbb{Q}$-bases of K. Denote by A the transition matrix from $\mathbf{x}$ to $\mathbf{y}$. Then:*

$$d(\mathbf{y}) = d(\mathbf{x}) \cdot (\det A)^2.$$

In particular, $d(\mathbf{x}) \neq 0$ if, and only if, $d(\mathbf{y}) \neq 0$

Proof. By definition, $d(\mathbf{x})$ is the determinant of the matrix $\left(\mathrm{Tr}_{K/\mathbb{Q}}(x_i x_j)\right)_{1\leq i,j\leq n}$. We recognize here the matrix of the bilinear form:

$$\begin{aligned} b : K \times K &\to \mathbb{Q} \\ (x, y) &\mapsto \mathrm{Tr}_{F/E}(xy) \end{aligned}$$

in the $\mathbb{Q}$-basis $\mathbf{x}$. The matrix of b in the $\mathbb{Q}$-basis $\mathbf{y}$ is then:

$$\left(\mathrm{Tr}_{K/\mathbb{Q}}(y_i y_j)\right)_{1\leq i,j\leq n} = {}^t A \left(\mathrm{Tr}_{K/\mathbb{Q}}(x_i x_j)\right)_{1\leq i,j\leq n} A$$

Hence:

$$d(\mathbf{y}) = d(\mathbf{x}) \cdot (\det A)^2.$$

In particular, $d(\mathbf{x}) \neq 0$ if, and only if, $d(\mathbf{y}) \neq 0$. □

Proof of Proposition 2.2.6. By the lemma, it suffices to find a $\mathbb{Q}$-basis $\mathbf{y} = (y_1, \ldots, y_n)$ of K for which we can easily compute $d(\mathbf{y})$. To do so, we can use the Primitive Element Theorem to write $K = \mathbb{Q}(x)$ for some $x \in K$. The family $(1, x, x^2, \ldots, x^{n-1})$ is then a $\mathbb{Q}$-basis of K and as we have already seen:

$$d(1, x, x^2, \ldots, x^{n-1}) = \left(\prod_{i<j} (\sigma_i(x) - \sigma_j(x)) \right)^2,$$

where $\sigma_1, \ldots, \sigma_n$ are the field embeddings of K in $\mathbb{C}$. We conclude by observing that the $\sigma_i(x)$'s are pairwise distinct. □

Now that we know that $d(\mathbf{x})$ is always non-zero, Proposition 2.2.5 provides a theoretical way to compute the ring of integers $\mathcal{O}_K$ of a number field K. Here is the step-by-step procedure:

1. Find a $\mathbb{Q}$-basis $\mathbf{x} = (x_1, \ldots, x_n)$ of K that is contained in $\mathcal{O}_K$. For instance, find an element $x \in \mathcal{O}_K$ such that $K = \mathbb{Q}(x)$ and set $\mathbf{x} := (1, x, x^2, \ldots, x^{n-1})$.

2. Compute $d(\mathbf{x})$ and deduce that:

$$\bigoplus_{i=1}^{n} \mathbb{Z} \cdot x_i \subset \mathcal{O}_K \subset \frac{1}{d(\mathbf{x})} \cdot \bigoplus_{i=1}^{n} \mathbb{Z} \cdot x_i.$$

3. Find representatives $\alpha_1, \ldots, \alpha_r \in \frac{1}{d(\mathbf{x})} \cdot \bigoplus_{i=1}^{n} \mathbb{Z} \cdot x_i$ of the finite quotient:

$$\left(\frac{1}{d(\mathbf{x})} \cdot \bigoplus_{i=1}^{n} \mathbb{Z} \cdot x_i \right) / \left(\bigoplus_{i=1}^{n} \mathbb{Z} \cdot x_i \right).$$

4. For each i, compute the characteristic polynomial $\chi_{K/\mathbb{Q},\alpha_i}$ and determine whether it has integer coefficients.

5. Deduce that:
$$\mathcal{O}_K = \bigsqcup_{\substack{i=1 \\ \chi_{K/\mathbb{Q},\alpha_i} \in \mathbb{Z}[T]}}^{r} \alpha_i + \left(\bigoplus_{i=1}^{n} \mathbb{Z} \cdot x_i \right).$$

Example 2.2.8. This method is often more theoretical than practical, because the integer r appearing in step 3 is often quite large. The following two examples illustrate this phenomenon.

(i) Consider the number field $K = \mathbb{Q}(x)$ with $x \in \mathbb{C}$ such that $x^3 + x + 1 = 0$. Introduce the $\mathbb{Q}$-basis $\mathbf{x} := (1, x, x^2)$ of K. Since the minimal polynomial of x is $\mu_x = T^3 + T + 1$, the roots of μ_x' are $\pm\frac{i}{\sqrt{3}}$ and the discriminant $d(\mathbf{x})$ can be computed as follows:
$$d(\mathbf{x}) = -3^3 \cdot \mu_x\left(\frac{i}{\sqrt{3}}\right) \cdot \mu_x\left(-\frac{i}{\sqrt{3}}\right) = -27\left(1 + \frac{2}{3\sqrt{3}}i\right)\left(1 - \frac{2}{3\sqrt{3}}i\right) = -31.$$
We deduce that:
$$\mathbb{Z} \oplus \mathbb{Z}x \oplus \mathbb{Z}x^2 \subset \mathcal{O}_K \subset \frac{1}{31} \cdot (\mathbb{Z} \oplus \mathbb{Z}x \oplus \mathbb{Z}x).$$
As representatives of the quotient:
$$\left(\frac{1}{31} \cdot \left(\mathbb{Z} \oplus \mathbb{Z}x \oplus \mathbb{Z}x^2\right)\right) / \left(\mathbb{Z} \oplus \mathbb{Z}x \oplus \mathbb{Z}x^2\right),$$
one can choose the $\alpha_{a,b,c} = \frac{a+bx+cx^2}{31}$ with $a, b, c \in \{0, 1, \ldots, 30\}$. For each such $\alpha_{a,b,c}$ we have to compute its characteristic polynomial and check whether it has integer coefficients. There are therefore $31^3 = 29791$ characteristic polynomials to compute!

(ii) Similarly, consider the number field $K = \mathbb{Q}(\sqrt[3]{2})$, choose $x = \sqrt[3]{2}$ and introduce the $\mathbb{Q}$-basis $\mathbf{x} := (1, \sqrt[3]{2}, \sqrt[3]{4})$ of K. Since the minimal polynomial of x is $\mu_x = T^3 - 2$, the discriminant $d(\mathbf{x})$ can be computed as follows:
$$d(\mathbf{x}) = \mu_x'(\sqrt[3]{2}) \cdot \mu_x'(e^{2i\pi/3}\sqrt[3]{2}) \cdot \mu_x'(e^{4i\pi/3}\sqrt[3]{2}) = 2^2 \times 3^3 = 108.$$
We deduce that:
$$\mathbb{Z} \oplus \mathbb{Z}\sqrt[3]{2} \oplus \mathbb{Z}\sqrt[3]{4} \subset \mathcal{O}_K \subset \frac{1}{108} \cdot \left(\mathbb{Z} \oplus \mathbb{Z}\sqrt[3]{2} \oplus \mathbb{Z}\sqrt[3]{4}\right).$$

As representatives of the quotient:

$$\left(\frac{1}{108}\cdot\left(\mathbb{Z}\oplus\mathbb{Z}\sqrt[3]{2}\oplus\mathbb{Z}\sqrt[3]{4}\right)\right)/\left(\mathbb{Z}\oplus\mathbb{Z}\sqrt[3]{2}\oplus\mathbb{Z}\sqrt[3]{4}\right),$$

one can choose the $\alpha_{a,b,c}=\frac{a+b\sqrt[3]{2}+c\sqrt[3]{4}}{108}$ with $a,b,c\in\{0,1,\dots,107\}$. There are therefore $108^3=1259712$ characteristic polynomials to compute!

In order to reduce the number of characteristic polynomials one needs to check in the previous procedure, one needs to strengthen Proposition 2.2.5. To do so, we start with the following linear algebra lemma:

Lemma 2.2.9. *Let V be a $\mathbb{Q}$-vector space of finite dimension n. Let $\mathbf{x}=(x_1,\dots,x_n)$ be a $\mathbb{Q}$-basis of V and set $\Gamma_{\mathbf{x}}:=\bigoplus_{i=1}^n\mathbb{Z}\cdot x_i$. Let Γ be a subgroup of V containing $\Gamma_{\mathbf{x}}$ and such that $[\Gamma:\Gamma_{\mathbf{x}}]<\infty$. Then V has a basis $\mathbf{y}=(y_1,\dots,y_n)$ such that:*

$$\Gamma=\Gamma_{\mathbf{y}}=\mathbb{Z}y_1\oplus\cdots\oplus\mathbb{Z}y_n.$$

Moreover, if $A_{\mathbf{y},\mathbf{x}}$ is the transition matrix from $\mathbf{y}$ to $\mathbf{x}$:

$$[\Gamma:\Gamma_{\mathbf{x}}]=|\det A_{\mathbf{y},\mathbf{x}}|.$$

Proof. We proceed by induction on $r:=[\Gamma:\Gamma_{\mathbf{x}}]$. If $r=1$, then $\Gamma=\Gamma_{\mathbf{x}}$, so we may take $\mathbf{y}=\mathbf{x}$, in which case we have $[\Gamma:\Gamma_{\mathbf{x}}]=1=|\det A_{\mathbf{y},\mathbf{x}}|$. Moreover, if $\mathbf{y}'$ is any other basis of V such that $\Gamma=\Gamma_{\mathbf{y}'}$, the transition matrix $A_{\mathbf{y}',\mathbf{x}}$ from $\mathbf{y}'$ to $\mathbf{x}$ is in $\mathrm{GL}_n(\mathbb{Z})$, and hence $|\det A_{\mathbf{y}',\mathbf{x}}|=1$.
Now assume that $r>1$. In that case there exists $z'\in\Gamma$ whose order in the quotient $\Gamma/\Gamma_{\mathbf{x}}$ is a prime number p. Write:

$$pz'=a_1x_1+\cdots+a_nx_n$$

with $a_1,\dots,a_n\in\mathbb{Z}$. Since $z'\notin\Gamma_{\mathbf{x}}$, there exists i such that p does not divide a_i. Without loss of generality, we may assume that p does not divide a_1. Then, by choosing $b\in\mathbb{Z}$ such that $a_1b\equiv 1 \mod p$, we may find $z\in bz'+\mathbb{Z}x_1$ such that:

$$pz=x_1+ba_2x_2+\cdots+ba_nx_n.$$

Set $\Gamma_{\mathbf{x},z}:=\Gamma_{\mathbf{x}}+\mathbb{Z}z\subset\Gamma$. One immediately sees that:

$$\Gamma_{\mathbf{x},z}=\Gamma_{\mathbf{z}}=\mathbb{Z}z\oplus\bigoplus_{i=2}^n\mathbb{Z}\cdot x_i\quad\text{with }\mathbf{z}=(z,x_2,\dots,x_n),$$
$$\Gamma_{\mathbf{z}}/\Gamma_{\mathbf{x}}\cong\mathbb{Z}/p\mathbb{Z},\quad[\Gamma:\Gamma_{\mathbf{z}}]=[\Gamma_{\mathbf{z}}:\Gamma_{\mathbf{x}}]^{-1}\cdot[\Gamma:\Gamma_{\mathbf{x}}]=p^{-1}\cdot[\Gamma:\Gamma_{\mathbf{x}}]<[\Gamma:\Gamma_{\mathbf{x}}].$$

By induction, there exists a basis $\mathbf{y}$ of V such that $\Gamma = \Gamma_{\mathbf{y}}$ and $[\Gamma : \Gamma_{\mathbf{z}}] = |\det A_{\mathbf{y},\mathbf{z}}|$. We then get:

$$[\Gamma : \Gamma_{\mathbf{x}}] = p \cdot [\Gamma : \Gamma_{\mathbf{z}}] = |\det A_{\mathbf{z},\mathbf{x}}| \cdot |\det A_{\mathbf{y},\mathbf{z}}| = |\det A_{\mathbf{y},\mathbf{x}}|.$$

Now if $\mathbf{y}'$ is any other basis of V such that $\Gamma = \Gamma_{\mathbf{y}'}$, the transition matrix $A_{\mathbf{y}',\mathbf{y}}$ from $\mathbf{y}'$ to $\mathbf{y}$ is in $\mathrm{GL}_n(\mathbb{Z})$, and hence:

$$|\det A_{\mathbf{y}',\mathbf{x}}| = |\det A_{\mathbf{y}',\mathbf{y}}| \cdot |\det A_{\mathbf{y},\mathbf{x}}| = |\det A_{\mathbf{y},\mathbf{x}}| = [\Gamma : \Gamma_{\mathbf{x}}].$$

□

One can apply the previous lemma to the situation where V is a number field. One then gets the following strengthening of Proposition 2.2.5:

Proposition 2.2.10

Let K be a number field. Let $\mathbf{x} = (x_1, \ldots, x_n)$ be a $\mathbb{Q}$-basis of K and set $\Gamma_{\mathbf{x}} := \bigoplus_{i=1}^n \mathbb{Z} \cdot x_i$. Let Γ be a subgroup of K containing $\Gamma_{\mathbf{x}}$ and such that $[\Gamma : \Gamma_{\mathbf{x}}] < \infty$. Then K has a basis $\mathbf{y} = (y_1, \ldots, y_n)$ such that:

$$\Gamma = \Gamma_{\mathbf{y}} = \mathbb{Z} y_1 \oplus \cdots \oplus \mathbb{Z} y_n.$$

Moreover, the discriminant $d(\mathbf{y})$ is independent of $\mathbf{y}$ and:

$$d(\mathbf{x}) = [\Gamma : \Gamma_{\mathbf{x}}]^2 \cdot d(\mathbf{y}).$$

Proof. One simply needs to combine Lemmas 2.2.7 and 2.2.9. □

Applying the previous proposition to some basis $\mathbf{x}$ of K contained in $\mathcal{O}_K$ and to $\Gamma = \mathcal{O}_K$, we get:

Corollary 2.2.11

Every number field K has a basis $\mathbf{y} = (y_1, \ldots, y_n)$ such that:

$$\mathcal{O}_K = \mathbb{Z} y_1 \oplus \cdots \oplus \mathbb{Z} y_n.$$

Moreover, the discriminant $\Delta_K := d(\mathbf{y})$ is independent of $\mathbf{y}$ and for every $\mathbb{Q}$-basis $\mathbf{x}$ of K contained in $\mathcal{O}_K$:

$$d(\mathbf{x}) = [\mathcal{O}_K : \Gamma_{\mathbf{x}}]^2 \cdot \Delta_K.$$

The discriminant Δ_K is called the *discriminant of* K.

In particular, if $\mathbf{x}$ is a $\mathbb{Q}$-basis of K contained in $\mathcal{O}_K$ such that $d(\mathbf{x})$ is square-free:

$$\mathcal{O}_K = \bigoplus_{i=1}^{n} \mathbb{Z} \cdot x_i.$$

Example 2.2.12. Let us come-back to examples 2.2.8.

(i) When $K = \mathbb{Q}(x)$ with $x^3 + x + 1 = 0$, the discriminant $d(1, x, x^2)$ is equal to -31, and hence is square-free. Hence:

$$\mathcal{O}_K = \mathbb{Z} \oplus \mathbb{Z}x \oplus \mathbb{Z}x^2.$$

(ii) The case of the number field $K = \mathbb{Q}(\sqrt[3]{2})$ is more complicated. Indeed, this time, if we set $x := \sqrt[3]{2}$, we have $d(1, x, x^2) = 108$, and hence:

$$[\mathcal{O}_K : \mathbb{Z}[\sqrt[3]{2}]] \mid 6.$$

In particular,

$$\mathcal{O}_K \subset \frac{1}{6}\mathbb{Z}[\sqrt[3]{2}],$$

and hence we only need to compute $6^3 = 216$ characteristic polynomials. This is still a lot. In the exercise sessions, we will develop a method based on Eisenstein polynomials that allows to prove that $\mathcal{O}_K = \mathbb{Z}[\sqrt[3]{2}]$.

2.3 Algebraic properties of rings of integers

In this section, we study some algebraic properties of rings of integers.

Definition 2.3.1

A ring A is **noetherian** if one of the following equivalent conditions holds

(i) every ideal of A is finitely generated;

(ii) every non-decreasing sequence of ideals in A is eventually constant;

(iii) every non-empty family of ideals in A has a maximal element.

Emmy Noether (Erlangen 1882 - Princeton 1935)

Proof. Assume (i) and consider a non-decreasing sequence of ideals $I_1 \subset I_2 \subset \ldots$ of A. Introduce the ideal $I := \bigcup_{n\geq 1} I_n$. By (i), it is finitely generated. Let $a_1, \ldots, a_r$ be generators of I and choose $N \geq 1$ such that I_N contains all of them. Then $I_N = I_{N+1} = \cdots = I$, and the sequence $(I_n)_n$ is eventually constant.

Assume now (ii) and let $\mathscr{E}$ be a non-empty family of ideals of A. Assume $\mathscr{E}$ has no maximal element. Then one can inductively construct a strictly increasing sequence $I_1 \subset I_2 \subset \ldots$ of ideals in $\mathscr{E}$, which constradicts assertion (ii). We deduce that $\mathscr{E}$ has a maximal ideal.

Assume finally (iii) and let I be an ideal of A. Let $\mathscr{E}$ be the family of finitely generated ideals of A contained in I. By (iii), it has a maximal element J. If $J \neq I$, then we can find $a \in I \setminus J$, so that $J + (a)$ is an element of $\mathscr{E}$ that is strictly larger than J: contradiction! We deduce that $J = I$, and hence that I is finitely generated. □

Proposition 2.3.2

The ring of integers $\mathscr{O}_K$ of a number field K is noetherian.

Proof. Take a non-zero ideal I of $\mathcal{O}_K$ and choose $a \in I \setminus \{0\}$. Then $n := N_{K/\mathbb{Q}}(a) \in (I \cap \mathbb{Z}) \setminus \{0\}$, and $n\mathcal{O}_K \subset I$. The quotient $I/n\mathcal{O}_K$ is a subgroup of $\mathcal{O}_K/n\mathcal{O}_K$, which is a finite group isomorphic to $(\mathbb{Z}/n\mathbb{Z})^{[K:\mathbb{Q}]}$. Hence $n\mathcal{O}_K$ has finite index in I. Recalling that $\mathcal{O}_K$ (and hence $n\mathcal{O}_K$) is a finitely generated abelian group, we deduce that so is I. In particular, $\mathcal{O}_K$ is noetherian. □

Proposition 2.3.3

Consider the ring of integers $\mathcal{O}_K$ of a number field K. Let I be a non-zero ideal in $\mathcal{O}_K$.

(i) The quotient $\mathcal{O}_K/I$ is finite.

(ii) If I is a (non-zero) prime ideal, then it is maximal.

Proof. (i) Take $x \in I \setminus \{0\}$. The norm $n := N_{K/\mathbb{Q}}(x)$ is non-zero and belongs to $\mathbb{Q} \cap \mathcal{O}_K = \mathbb{Z}$ and to I. Hence, as an abelian group, the quotient $\mathcal{O}_K/I$ is n-torsion. But it is also finitely generated. Hence it is finite.

(ii) This statement follows from (i) and from the following lemma.

□

Lemma 2.3.4. *Every finite domain is a field.*

Proof. Let A be a finite domain and take $a \in A$. Since A is a domain, the multiplication-by-a map $A \to A$ is injective. But then it is a bijection since A is finite. Hence $a \in A^\times$ and A is a field. □

Definition 2.3.5

Let A be a domain with fraction field K. We say A is **integrally closed** if every root in K of a monic polynomial $f \in A[X]$ lies in A.

Proposition 2.3.6

The ring of integers $\mathcal{O}_K$ of a number field K is integrally closed.

Proof. Let f be a monic polynomial with coefficients in $\mathcal{O}_K$ and let n be its degree. Let $x \in K$ be a root of f and consider a finite system of generators $e_1, \ldots, e_r$

of $\mathcal{O}_K$ as an abelian group. Observe that $x^n \in \sum_{i=0}^{n-1} \mathcal{O}_K x^i$. A simple induction shows that $x^m \in \sum_{i=0}^{n-1} \mathcal{O}_K x^i$ for all $m \geq n$, and hence:

$$\mathcal{O}_K[x] = \sum_{i=0}^{n-1} \mathcal{O}_K x^i = \sum_{i=0}^{n-1} \sum_{j=1}^{r} \mathbb{Z} e_j x^i$$

In other words, $\mathcal{O}_K[x]$ is generated by $(x^i e_j)_{0 \leq i \leq n-1, 1 \leq j \leq r}$ as an abelian group. By Proposition 2.1.4, we deduce that x is an algebraic integer, and hence that $x \in K \cap \overline{\mathbb{Z}} = \mathcal{O}_K$. □

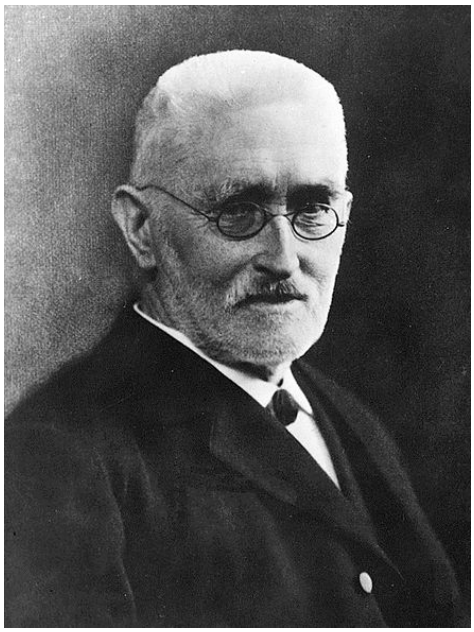
Definition 2.3.7

A domain A is said to be a **Dedekind domain** if it is noetherian, integrally closed and every non-zero prime ideal in A is maximal.

According to Propositions 2.3.2, 2.3.3 and 2.3.6, we get:

Theorem 2.3.8

The ring of integers $\mathcal{O}_K$ of a number field K is a Dedekind domain.



Richard Dedekind (Braunschweig 1831 - Braunschweig 1916)

2.4 Fractional ideals in Dedekind domains

Definition 2.4.1

Let A be a Dedekind domain with fraction field K and let I be a subset of K. We say that I is a **fractional ideal** of A if there exists an element $a \in A \setminus \{0\}$ such that aI is an ideal of A.

Example 2.4.2. If A is a principal ideal domain, its fractional ideals are the subsets of K of the form xA with $x \in K$. In particular, if A is not a field, one can check that K itself is not a fractional ideal.

As for usual ideals, one can define several arithmetical operations on fractional ideals. Indeed, if I and J are two fractional ideals of A, then:

$$I \cap J, \quad I + J := \{i + j : i \in I, j \in J\}, \quad IJ := \{\sum_{r=1}^{n} i_r j_r : n \geq 0, i_1, \ldots, i_n \in I, j_1, \ldots, j_n \in J\}$$

are all fractional ideals. A fractional ideal I is said to be **invertible** if there exists a fractional ideal J such that $IJ = A$.

Theorem 2.4.3

Let A be a Dedekind domain and let $\mathfrak{p}$ be a non-zero prime ideal in A. Then $\mathfrak{p}$ is an invertible fractional ideal.

Observe that, if J is a fractional ideal such that $\mathfrak{p}J = A$, then

$$J \subset \{x \in K : x\mathfrak{p} \subset A\}.$$

We therefore naturally define:

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

If y is a non-zero element in $\mathfrak{p}$, then $y\mathfrak{p}^{-1}$ is an ideal in A. Hence $\mathfrak{p}^{-1}$ is a fractional ideal in A. Theorem 2.4.3 is then an immediate consequence of the following proposition:

Proposition 2.4.4

Let A be a Dedekind domain, let $\mathfrak{p}$ be a prime ideal in A and let I be any non-zero ideal in A. Then $\mathfrak{p}^{-1}I$ strictly contains I. In particular, $\mathfrak{p}^{-1}\mathfrak{p} = A$.

Proof. We carry out the proof in three main steps.

Step 1: We start by proving that every non-zero ideal in A contains a product of prime ideals. For that purpose, consider the family $\mathscr{E}$ of non-zero ideals in A that do not satisfy this property and assume by contradiction that $\mathscr{E}$ is non-empty. Since A is noetherian, the family $\mathscr{E}$ contains a maximal element I_0. Since $I_0 \in \mathscr{E}$, the ideal I_0 cannot be prime. This means that there exist $a, b \in A \setminus I_0$ such that $ab \in I_0$. By maximality of I_0, the ideals $I_0 + (a)$ and $I_0 + (b)$ both contain a product of prime ideals. But then so does $(I_0 + (a))(I_0 + (b)) \subset I_0$. This contractics that $I_0 \in \mathscr{E}$.

Step 2: We now prove that $\mathfrak{p}^{-1}$ strictly contains A. By definition, $\mathfrak{p}^{-1}$ contains A. It remains to construct an element that lies in $\mathfrak{p}^{-1}$ but not in A. To do so, take a non-zero element $a \in \mathfrak{p}$ and apply Step 1 to find prime ideals $\mathfrak{p}_1, \dots, \mathfrak{p}_n$ such that :

$$\mathfrak{p}_1 \dots \mathfrak{p}_n \subset (a) \subset \mathfrak{p}. \tag{2.3}$$

Assume without loss of generality that n is as small as possible. By inclusion (2.3) and the primality of the ideal $\mathfrak{p}$, one of the $\mathfrak{p}_i$'s is contained in $\mathfrak{p}$. Assume without loss of generality that $\mathfrak{p}_n \subset \mathfrak{p}$. By maximality of $\mathfrak{p}_n$, we deduce that $\mathfrak{p}_n = \mathfrak{p}$. Moreover, by minimality of n, we can find an element $b \in \mathfrak{p}_1 \dots, \mathfrak{p}_{n-1}$ that does not lie in (a). This implies that $a^{-1}b$ belongs to $\mathfrak{p}^{-1}$ but not to A, as wished.

Step 3: We finally prove that the fractional ideal $\mathfrak{p}^{-1}I$ strictly contains I, and in particular that $\mathfrak{p}^{-1}\mathfrak{p} = A$. Assume by contradiction that $\mathfrak{p}^{-1}I = I$ and take $x \in \mathfrak{p}^{-1}$. We have $xI \subset I$, and since A is noetherian, the ideal I is finitely generated. By Proposition 2.1.4, we deduce that x is a root of a monic polynomial with coefficients in A. But A is integrally closed, hence $x \in A$ and $\mathfrak{p}^{-1} = A$. This contradicts Step 2, and therefore shows that $\mathfrak{p}^{-1}I$ strictly contains I. In particular, when $I = \mathfrak{p}$, we obtain $\mathfrak{p}\mathfrak{p}^{-1}$ is an ideal of A that strictly contains $\mathfrak{p}$. By maximality of $\mathfrak{p}$, we conclude that $\mathfrak{p}\mathfrak{p}^{-1} = A$. □

We will prove in the next section that in fact every non-zero fractional ideal in a Dedekind domain is invertible. To do so, we will first prove that, in that context, every ideal is a product of prime ideals

2.5 Factorization of ideals in Dedekind domains

Theorem 2.5.1

Let A be a Dedekind domain and let I be a non-zero ideal in A. Then there exist an integer $n \geq 0$ and some non-zero prime ideals $\mathfrak{p}_1, \ldots, \mathfrak{p}_n$ in A such that $I = \mathfrak{p}_1 \ldots \mathfrak{p}_n$. The integer n is uniquely determined, and the prime ideals $\mathfrak{p}_1, \ldots, \mathfrak{p}_n$ are uniquely determined up to the order.

Proof. We start by proving the existence. Let $\mathscr{E}$ be the family of ideals in A other than (0) and (1) that are not a product of prime ideals. By contradiction, assume that $\mathscr{E}$ is non-empty. Since A is noetherian, the family $\mathscr{E}$ contains a maximal element I_0. By Krull's Theorem, I_0 is contained in a maximal ideal $\mathfrak{p}$. By Proposition 2.4.4, the ideal I_0 is strictly contained in $\mathfrak{p}^{-1} I_0$, and therefore $\mathfrak{p}^{-1} I_0$ is a product of prime ideals. Hence by using again Proposition 2.4.4, the ideal $I_0 = \mathfrak{p}\mathfrak{p}^{-1} I_0$ is also a product of prime ideals. This contradicts that $I_0 \in \mathscr{E}$.

We now move on to the uniqueness. Consider non-zero prime ideals $\mathfrak{p}_1, \ldots, \mathfrak{p}_m$, $\mathfrak{q}_1, \ldots, \mathfrak{q}_n$ such that:

$$\mathfrak{p}_1 \ldots \mathfrak{p}_m = \mathfrak{q}_1 \ldots \mathfrak{q}_n.$$

We prove by induction on m that $m = n$ and that the $\mathfrak{p}_i$'s and the $\mathfrak{q}_j$'s coincide up to the order. Of course, the result is obvious for $m = 0$. We henceforth assume that $m > 0$. Since $\mathfrak{p}_1 \ldots \mathfrak{p}_m \subset \mathfrak{q}_n$ and $\mathfrak{q}_n$ is prime, one of the ideals $\mathfrak{p}_1, \ldots, \mathfrak{p}_m$ is contained in $\mathfrak{q}_n$. Without loss of generality, assume that $\mathfrak{p}_m \subset \mathfrak{q}_n$. Since A is a Dedekind domain, $\mathfrak{p}_m$ is a maximal ideal in A, and hence $\mathfrak{p}_m = \mathfrak{q}_n$. Therefore, by Proposition 2.4.4:

$$\mathfrak{p}_1 \ldots \mathfrak{p}_{m-1} = \mathfrak{p}_1 \ldots \mathfrak{p}_m \mathfrak{p}_m^{-1} = \mathfrak{q}_1 \ldots \mathfrak{q}_n \mathfrak{q}_n^{-1} = \mathfrak{q}_1 \ldots \mathfrak{q}_{n-1}.$$

By our inductive assumption, $m-1 = n-1$ and the $\mathfrak{p}_i$'s for $i \leq m-1$ and the $\mathfrak{q}_j$'s for $j \leq m-1$ coincide up to the order. We conclude that $m = n$ and that the $\mathfrak{p}_i$'s for $i \leq m$ and the $\mathfrak{q}_j$'s for $j \leq n$ coincide up to the order. □

Corollary 2.5.2

Every non-zero fractional ideal in a Dedekind domain is invertible.

Proof. Let I be a non-zero fractional ideal of a Dedekind domain A. There exists $x \in A \setminus \{0\}$ such that xI is an ideal in A. By Theorem 2.5.1, we can write:

$$xI = \mathfrak{p}_1 \ldots \mathfrak{p}_n$$

for some prime ideals $\mathfrak{p}_1, \ldots, \mathfrak{p}_n$ of A. But by Theorem 2.4.3, all the $\mathfrak{p}_i$'s are invertible. Hence so is I. □

Corollary 2.5.3

Let I and J be two non-zero ideals in a Dedekind domain A. Then $I \subseteq J$ if, and only if, J divides I.

Proof. Assume that $I \subseteq J$ and set $I' := J^{-1}I$. Then $I' \subseteq J^{-1}J = \mathcal{O}_K$ and $I = I'J$. The converse is obvious. □

2.6 Minkowski's Theorem on lattices

In this section, we fix an n-dimensional Euclidean vector space V. We endow it with the corresponding Lebesgue measure, so that the volume of the cube:

$$\{t_1 e_1 + \cdots + t_n e_n : 0 \leq t_1, \ldots, t_n < 1\}$$

spanned by an orthonormal basis $(e_1, \ldots, e_n)$ of V is always 1. More generally, if $(v_1, \ldots, v_n)$ is a basis of V, the parallepiped spanned by $(v_1, \ldots, v_n)$ is defined as:

$$\Pi_{v_1,\ldots,v_n} := \{t_1 v_1 + \cdots + t_r v_r : 0 \leq t_1, \ldots, t_r < 1\},$$

and its volume is the absolute value of the determinant of the transition matrix $A = \big(\langle e_i, v_j \rangle\big)_{1 \leq i,j \leq n}$ from some orthonormal basis $(e_1, \ldots, e_n)$ of V to $(v_1, \ldots, v_n)$. By observing that:

$$A^t A = \left(\sum_{k=1}^{n} \langle e_i, v_k \rangle \langle e_k, v_j \rangle \right)_{1 \leq i,j \leq n} = \big(\langle v_i, v_j \rangle\big)_{1 \leq i,j \leq n},$$

we get the orthonormal-basis-free formula:

$$\mathrm{Vol}(\Pi_{v_1,\ldots,v_n}) = \Big(\det\big(\langle v_i, v_j \rangle\big)_{1 \leq i,j \leq n}\Big)^{1/2}.$$

Definition 2.6.1

A **lattice** Γ in the n-dimensional Euclidean space V is a subgroup of the form:

$$\Gamma = \mathbb{Z}\gamma_1 + \cdots + \mathbb{Z}\gamma_r$$

where $\gamma_1, \ldots, \gamma_r$ are linearly independent vectors in V. We then say that $(\gamma_1, \ldots, \gamma_r)$ is a **basis** of Γ and that the set:

$$M := \{t_1\gamma_1 + \cdots + t_r\gamma_r : 0 \leq t_1, \ldots, t_r < 1\}$$

is a **fundamental parallepiped** of Γ.

The lattice Γ is said to be **complete** if $r = n$. In that case, the volume of a fundamental parallepiped M of Γ is independent of the choice of a basis of Γ. It is called the **covolume** of Γ and is denoted $\mathrm{Covol}(\Gamma)$.

Example 2.6.2. If K is a quadratic imaginary extension of $\mathbb{Q}$, its ring of integers $\mathcal{O}_K$ is a complete lattice in $\mathbb{C}$.

Quite often, the most practical way to check that a subgroup of $\mathbb{R}^n$ is a lattice is provided by the following topological characterization:

Proposition 2.6.3

A subgroup of $\mathbb{R}^n$ is a lattice if, and only if, it is discrete.

Proof. The direct implication is obvious. Let us prove the converse. Let Γ be a non-zero discrete subgroup in $\mathbb{R}^n$. We can then construct inductively a finite sequence $\gamma_1, \gamma_2, \gamma_3, \ldots$, of non-zero elements of Γ in the following way:

- for $i = 1$, we choose γ_1 to be a non-zero element of Γ with minimal norm;
- for $i > 1$, we set $V_i := \mathbb{R}\gamma_1 + \cdots + \mathbb{R}\gamma_{i-1}$, we choose some $\gamma_i' \in \Gamma \setminus V_i$ if Γ is not contained in V_i and we let γ_i be an element of $\overline{\Pi}_{\gamma_1,\ldots,\gamma_{i-1},\gamma_i'} \cap \Gamma$ such that the distance $d(\gamma_i, V_i)$ is non-zero and minimal. Such an element exists because $\overline{\Pi}_{\gamma_1,\ldots,\gamma_{i-1},\gamma_i'}$ contains a finite number of elements of Γ.

The vectors $\gamma_1, \gamma_2, \gamma_3, \ldots$ are linearly independent and hence the previous construction necessarily stops at some point. In other words, for some $r \geq 1$, we have $\Gamma \subset V_r$. Let us prove that $\Gamma = \bigoplus_{i=1}^r \mathbb{Z}\gamma_i$.

To do so, take some $\gamma \in \Gamma \setminus \left(\bigoplus_{i=1}^r \mathbb{Z}\gamma_i\right)$ and write $\gamma = x_1\gamma_1 + \cdots + x_r\gamma_r$ with $x_1, \ldots, x_r \in \mathbb{R}$. Let $s \in \{1, \ldots, r\}$ be the smallest index such that $x_s \notin \mathbb{Z}$ and set:

$$\gamma' := \lfloor x_1 \rfloor \gamma_1 + \cdots + \lfloor x_s \rfloor \gamma_s.$$

Then $x_1\gamma_1 + \cdots + x_s\gamma_s - \gamma' = (x_s - \lfloor x_s \rfloor)\gamma_s \in \overline{\Pi}_{\gamma_1,\ldots,\gamma_{s-1},\gamma_s'} \cap \Gamma$ and:

$$d(x_1\gamma_1 + \cdots + x_s\gamma_s - \gamma', V_s) = (x_s - \lfloor x_s \rfloor) \cdot d(\gamma_s, V_s) < d(\gamma_s, V_s).$$

This contradicts the definition of γ_s. We deduce that $\Gamma = \bigoplus_{i=1}^{r} \mathbb{Z}\gamma_i$, and hence Γ is a lattice. □

The following corollary turns out to be extremely useful to prove that some subgroup of $\mathbb{R}^n$ is a complete lattice and to compute its covolume:

Corollary 2.6.4

Let Γ be a complete lattive in $\mathbb{R}^n$. Let $\Gamma' \subset \Gamma$ be a finite index subgroup. Then Γ' is a complete lattice of $\mathbb{R}^n$ and:

$$\mathrm{Covol}(\Gamma') = \mathrm{Covol}(\Gamma) \cdot [\Gamma : \Gamma'].$$

Proof. Since Γ' is contained in Γ, it is discrete in $\mathbb{R}^n$, and hence it is a lattice by the previous proposition. The equality $\mathrm{Covol}(\Gamma') = \mathrm{Covol}(\Gamma) \cdot [\Gamma : \Gamma']$ follows immediately from the definition of the covolume and Lemma 2.2.9. □

We are now ready to move on to the so-called Minkowski's theory. The intuitive idea is that, if one takes a complete lattice Γ in $\mathbb{R}^n$ and some "sufficiently large" subset $X \subset \mathbb{R}^n$, then X has to contain some elements of Γ. We start with the following proposition:

Proposition 2.6.5 (Blichfeldt)

Let Γ be a complete lattice in the vector space $V = \mathbb{R}^n$, and let X be a measurable subset of V such that:

$$\mathrm{Vol}(X) > \mathrm{Covol}(\Gamma).$$

Then there exist two elements $x_1, x_2 \in X$ such that $x_2 - x_1 \in \Gamma \setminus \{0\}$.

Proof. Recall that M is a fundamental parallepiped. We have:

$$V = \bigcup_{\gamma \in \Gamma} (\gamma + M),$$

hence:

$$X = \bigcup_{\gamma \in \Gamma} \big((\gamma + M) \cap X\big).$$

We deduce that:

$$\mathrm{Vol}(M) < \mathrm{Vol}(X) \leq \sum_{\gamma \in \Gamma} \mathrm{Vol}\big((\gamma + M) \cap X\big) = \sum_{\gamma \in \Gamma} \mathrm{Vol}\big(M \cap (X - \gamma)\big).$$

Therefore there exist $\gamma_1, \gamma_2 \in \Gamma$ with $\gamma_1 \neq \gamma_2$ and:

$$M \cap (X - \gamma_1) \cap (X - \gamma_2) \neq \emptyset.$$

We can therefore find $x_1, x_2 \in X$ such that $x_1 - \gamma_1 = x_2 - \gamma_2$, and hence $x_2 - x_1 = \gamma_2 - \gamma_1 \in \Gamma$. □

Hans Frederick Blichfeldt (Illar 1873 - Palo Alto 1945)

Theorem 2.6.6 (Minkowski)

Let Γ be a complete lattice in the vector space $V = \mathbb{R}^n$, and let X be a (measurable) centrally symmetric convex subset of V such that:

$$\mathrm{Vol}(X) > 2^n \mathrm{Covol}(\Gamma).$$

Then X contains a non-zero element of Γ.

Proof. Set $Y := \frac{1}{2}X$, so that $\mathrm{Vol}(Y) > \mathrm{Covol}(\Gamma)$. By Proposition 2.6.5, there exist $y_1, y_2 \in Y$ such that $y_2 - y_1 \in \Gamma \setminus \{0\}$. Moreover, since X is centrally symmetric and convex:

$$y_2 - y_1 = \frac{(-2y_1) + (2y_2)}{2} \in X.$$

We have therefore found a non-zero element of Γ in X. □

Hermann Minkowski (Aleksotas 1864 - Göttingen 1909)

2.7 Geometry of numbers

Let K be a number field and consider the set of complex field embeddings $\Sigma := \mathrm{Hom}_{\mathbb{Q}}(K, \mathbb{C})$. This set has $n := [K : \mathbb{Q}]$ elements and it is naturally endowed with an action of $\mathrm{Gal}(\mathbb{C}/\mathbb{R})$ by complex conjugation. The subset of fixed points for this action is the set of real embeddings $\Sigma_{\mathbb{R}} := \mathrm{Hom}_{\mathbb{Q}}(K, \mathbb{R})$, whose cardinal will be denoted r. Its complement $\Sigma \setminus \Sigma_{\mathbb{R}}$ is partitioned into a certain number s of equivalence classes, each of them of cardinal 2. We let $\Sigma_{\mathbb{C}}$ be a set of representatives of those equivalence classes in $\Sigma \setminus \Sigma_{\mathbb{R}}$, so that:

$$\Sigma = \Sigma_{\mathbb{R}} \sqcup \Sigma_{\mathbb{C}} \sqcup \overline{\Sigma_{\mathbb{C}}}, \qquad n = r + 2s.$$

We introduce the embedding:

$$\begin{aligned} \iota : K &\to \mathbb{R}^{\Sigma_{\mathbb{R}}} \times \mathbb{C}^{\Sigma_{\mathbb{C}}} \\ x &\mapsto (\sigma(x))_{\sigma \in \Sigma_{\mathbb{R}} \sqcup \Sigma_{\mathbb{C}}}. \end{aligned}$$

Proposition 2.7.1

The image of $\mathcal{O}_K$ by ι through the embedding ι is a complete lattice in the n-dimensional real vector space $\mathbb{R}^{\Sigma_{\mathbb{R}}} \times \mathbb{C}^{\Sigma_{\mathbb{C}}}$.

Proof. According to Proposition 2.6.3, in order to check that $\iota(\mathcal{O}_K)$ is a lattice, we only need to prove that it is discrete. To do so, consider a sequence $(x_n)_{n \geq 0}$ of elements in $\mathcal{O}_K$ such that the sequence $(\iota(x_n))_{n \geq 0}$ converges to $\iota(x)$ for some

$x \in \mathcal{O}_K$. We prove that $(x_n)_{n\geq 0}$ is eventually constant equal to x. Indeed, for any $\sigma \in \mathrm{Hom}_{\mathbb{Q}}(K,\mathbb{C})$, we have $\sigma(x_n) \xrightarrow[n\to\infty]{} \sigma(x)$. Hence:

$$\chi_{K/\mathbb{Q},x_n} \xrightarrow[n\to\infty]{} \chi_{K/\mathbb{Q},x}$$

coefficient-wise. But all the x_n's and x belong to $\mathcal{O}_K$, so that all the $\chi_{K/\mathbb{Q},x_n}$'s and $\chi_{K/\mathbb{Q},x}$ have integer coefficients. We deduce that $\chi_{K/\mathbb{Q},x_n} = \chi_{K/\mathbb{Q},x}$, and hence that x_n is a root of $\chi_{K/\mathbb{Q},x}$, for n large enough. In particular, $(x_n)_{n\geq 0}$ is eventually constant equal to x, as wished. Since $\mathcal{O}_K$ has rank n and $\mathbb{R}^{\Sigma_{\mathbb{R}}} \times \mathbb{C}^{\Sigma_{\mathbb{C}}}$ has dimension n as a real vector space, we conclude that $\iota(\mathcal{O}_K)$ is a complete lattice in $\mathbb{R}^{\Sigma_{\mathbb{R}}} \times \mathbb{C}^{\Sigma_{\mathbb{C}}}$. □

Proposition 2.7.2

Endow the real vector space $\mathbb{R}^{\Sigma_{\mathbb{R}}} \times \mathbb{C}^{\Sigma_{\mathbb{C}}}$ with the scalar product:

$$\langle \left((x_\sigma)_{\sigma\in\Sigma_{\mathbb{R}}}, (z_\sigma)_{\sigma\in\Sigma_{\mathbb{C}}}\right), \left((x'_\sigma)_{\sigma\in\Sigma_{\mathbb{R}}}, (z'_\sigma)_{\sigma\in\Sigma_{\mathbb{C}}}\right)\rangle = \sum_{\sigma\in\Sigma_{\mathbb{R}}} x_\sigma x'_\sigma + 2\mathrm{Re}\left(\sum_{\sigma\in\Sigma_{\mathbb{C}}} z_\sigma \overline{z}'_\sigma\right).$$

The covolume of $\iota(\mathcal{O}_K)$ is then equal to $\sqrt{|\Delta_K|}$.

Proof. Consider a basis $(e_1,\dots,e_n)$ of K such that $\mathcal{O}_K = \mathbb{Z}e_1 + \cdots + \mathbb{Z}e_n$. We then have:

$$\begin{aligned}
\mathrm{Covol}(\iota(\mathcal{O}_K))^2 &= \det\left(\langle \iota(e_i), \iota(e_j)\rangle\right)_{1\leq i,j\leq n} \\
&= \det\left(\sum_{\sigma\in\Sigma_{\mathbb{R}}} \sigma(e_i e_j) + \sum_{\sigma\in\Sigma_{\mathbb{C}}} \sigma(e_i)\overline{\sigma}(e_j) + \sum_{\sigma\in\Sigma_{\mathbb{C}}} \overline{\sigma}(e_i)\sigma(e_j)\right)_{1\leq i,j\leq n} \\
&= \det\left(\sum_{\sigma\in\Sigma} \sigma(e_i)\overline{\sigma}(e_j)\right)_{1\leq i,j\leq n}.
\end{aligned}$$

Observing that :

$$\left(\sum_{\sigma\in\Sigma} \sigma(e_i)\overline{\sigma}(e_j)\right)_{1\leq i,j\leq n} = A^t\overline{A}$$

with $A = (\sigma(e_i))_{\sigma\in\Sigma, 1\leq i\leq n}$, we get:

$$\mathrm{Covol}(\iota(\mathcal{O}_K))^2 = |\det A|^2 = |\Delta_K|.$$

□

Corollary 2.7.3

Keeping the previous conventions and notations, if I is a non-zero ideal of $\mathcal{O}_K$, then $\iota(I)$ is a complete lattice in the n-dimensional real vector space $\mathbb{R}^{\Sigma_\mathbb{R}} \times \mathbb{C}^{\Sigma_\mathbb{C}}$ with covolume $[\mathcal{O}_K : I] \cdot \sqrt{|\Delta_K|}$.

Proof. By Proposition 2.3.3, the ideal I has finite index in $\mathcal{O}_K$. The result then follows by combining Propositions 2.7.1 and 2.7.2 with Corollary 2.6.4. □

Theorem 2.7.4

Let I be a non-zero ideal in $\mathcal{O}_K$. Fix $(c_\sigma)_{\sigma\in\Sigma} \in \mathbb{R}_+^\Sigma$ such that $c_{\overline{\sigma}} = c_\sigma$ for all $\sigma \in \Sigma$ and:

$$\prod_{\sigma\in\Sigma} c_\sigma > \left(\frac{2}{\pi}\right)^s \cdot [\mathcal{O}_K : I] \cdot \sqrt{|\Delta_K|}.$$

There exists a non-zero element $x \in I$ such that $|\sigma(x)| < c_\sigma$ for all $\sigma \in \Sigma$.

Proof. Consider the set:

$$X := \prod_{\sigma\in\Sigma_\mathbb{R}} (-c_\sigma, c_\sigma) \times \prod_{\sigma\in\Sigma_\mathbb{C}} D(0, c_\sigma) \subset \mathbb{R}^{\Sigma_\mathbb{R}} \times \mathbb{C}^{\Sigma_\mathbb{C}},$$

where $D(0, c_\sigma)$ stands for the disk of centre 0 and radius c_σ in $\mathbb{C}$. It is of course a centrally symmetric convex subset of $\mathbb{R}^{\Sigma_\mathbb{R}} \times \mathbb{C}^{\Sigma_\mathbb{C}}$, and:

$$\begin{aligned}
\mathrm{Vol}(X) &= \prod_{\sigma\in\Sigma_\mathbb{R}} (2c_\sigma) \cdot \prod_{\sigma\in\Sigma_\mathbb{C}} (2\pi c_\sigma^2) \\
&= 2^{r+s} \cdot \pi^s \cdot \prod_{\sigma\in\Sigma} c_\sigma \\
&> 2^{r+s} \cdot \pi^s \cdot \left(\frac{2}{\pi}\right)^s \cdot [\mathcal{O}_K : I] \cdot \sqrt{|\Delta_K|} \\
&> 2^n \cdot \mathrm{Covol}(\iota(I)).
\end{aligned}$$

By Theorem 2.6.6, we deduce that X contains some non-zero element of the lattice $\iota(I)$. In other words, there exists a non-zero element $x \in I$ such that $|\sigma(x)| < c_\sigma$ for all $\sigma \in \Sigma$. □

Corollary 2.7.5

For any non-zero ideal I in $\mathcal{O}_K$, there exists a non-zero element $x \in I$ such

that:

$$|N_{K/\mathbb{Q}}(x)| \leq \left(\frac{2}{\pi}\right)^s \cdot [\mathcal{O}_K : I] \cdot \sqrt{|\Delta_K|}.$$

Proof. Given $c > \left(\frac{2}{\pi}\right)^s \cdot [\mathcal{O}_K : I] \cdot \sqrt{|\Delta_K|}$, choose a family $(c_\sigma)_{\sigma\in\Sigma} \in \mathbb{R}_+^\Sigma$ such that $c_{\overline{\sigma}} = c_\sigma$ for all $\sigma \in \Sigma$ and:

$$\prod_{\sigma\in\Sigma} c_\sigma = c.$$

According to Theorem 2.7.4, there exists a non-zero element $x_c \in I$ such that $|\sigma(x_c)| < c_\sigma$ for all $\sigma \in \Sigma$. In particular:

$$|N_{K/\mathbb{Q}}(x_c)| = \prod_{\sigma\in\Sigma} |\sigma(x_c)| < \prod_{\sigma\in\Sigma} c_\sigma = c.$$

The corollary immediately follows since $N_{K/\mathbb{Q}}(x_c)$ always belongs to $\mathbb{Z}$ and since the previous inequality holds for any $c > \left(\frac{2}{\pi}\right)^s \cdot [\mathcal{O}_K : I] \cdot \sqrt{|\Delta_K|}$. □

2.8 Finiteness of the ideal class group

Definition 2.8.1

Let A be a Dedekind domain with fraction field K. We denote by $\mathcal{I}_K$ the set of non-zero fractional ideals of A and by $\mathcal{P}_K \subset \mathcal{I}_K$ the subset consisting of those fractional ideals that are **principal**, that is of the form xA for some $x \in K^\times$. According to Corollary 2.5.2, $\mathcal{I}_K$ is an abelian group, and $\mathcal{P}_K$ is a subgroup. The **ideal class group** of K is defined as $\mathrm{Cl}_K := \mathcal{I}_K/\mathcal{P}_K$.

In this section, we aim at proving that the ideal class group Cl_K of a number field is always finite. To do so, we will apply Corollary 2.7.5 to find in each class $c \in \mathrm{Cl}_K$ an ideal $I \in c$ contained in $\mathcal{O}_K$ and with bounded index $[\mathcal{O}_K : I]$. It will therefore be useful to understand how this index behaves:

Lemma 2.8.2. *Let K be a number field.*

(i) For any non-zero element $a \in \mathcal{O}_K$, we have $[\mathcal{O}_K : (a)] = |N_{K/\mathbb{Q}}(a)|$.

(ii) For any ideals I and J in $\mathcal{O}_K$, we have $[\mathcal{O}_K : IJ] = [\mathcal{O}_K : I] \cdot [\mathcal{O}_K : J]$.

Proof. (i) This is a direct consequence of Lemma 2.2.9.

(ii) We have to prove that, if $I = \mathfrak{p}_1^{e_1} \dots \mathfrak{p}_r^{e_r}$ is the prime decomposition of some ideal I of $\mathcal{O}_K$, then:

$$[\mathcal{O}_K : I] = \prod_{i=1}^{r} [\mathcal{O}_K : \mathfrak{p}_i]^{e_i}.$$

For that purpose, we first observe that, according to the Chinese remainder Theorem:

$$\mathcal{O}_K / I \cong \prod_{i=1}^{r} \mathcal{O}_K / \mathfrak{p}_i^{e_i},$$

and hence:

$$[\mathcal{O}_K : I] = \prod_{i=1}^{r} [\mathcal{O}_K : \mathfrak{p}_i^{e_i}].$$

Now fix some $i \in \{1, \dots, r\}$ and some $s \geq 0$. Since $\mathfrak{p}_i^s \neq \mathfrak{p}_i^{s+1}$, the quotient $\mathfrak{p}_i^s / \mathfrak{p}_i^{s+1}$ is a non-trivial $\mathcal{O}_K/\mathfrak{p}_i$-vector space. Choose some $x \in \mathfrak{p}_i^s \setminus \mathfrak{p}_i^{s+1}$, and consider the $\mathcal{O}_K/\mathfrak{p}_i$-linear map:

$$\begin{aligned} \varphi_s : \mathcal{O}_K/\mathfrak{p}_i &\to \mathfrak{p}_i^s/\mathfrak{p}_i^{s+1} \\ y &\mapsto xy. \end{aligned}$$

Since $x \in \mathfrak{p}_i^s \setminus \mathfrak{p}_i^{s+1}$, we have $(x) + \mathfrak{p}_i^{s+1} = \mathfrak{p}_i^s$, and hence φ_s is surjective. We deduce that φ_s is an isomorphism, so that:

$$[\mathcal{O}_K : \mathfrak{p}_i] = [\mathfrak{p}_i^s : \mathfrak{p}_i^{s+1}].$$

This being true for all s, we deduce that:

$$[\mathcal{O}_K : \mathfrak{p}_i^{e_i}] = [\mathcal{O}_K : \mathfrak{p}_i]^{e_i},$$

and hence:

$$[\mathcal{O}_K : I] = \prod_{i=1}^{r} [\mathcal{O}_K : \mathfrak{p}_i]^{e_i}.$$

□

Proposition 2.8.3

For any ideal class $c \in \mathrm{Cl}_K$, there exists a fractional ideal $I \in c$ that is contained in $\mathcal{O}_K$ and such that:

$$[\mathcal{O}_K : I] \leq \left(\frac{2}{\pi}\right)^s \cdot \sqrt{|\Delta_K|}.$$

Proof. Choose a fractional ideal $I_0 \in \mathscr{I}_K$ representing c and choose an element $a \in \mathscr{O}_K \setminus \{0\}$ such that $J_0 := aI_0^{-1}$ is an ideal in $\mathscr{O}_K$. According to Corollary 2.7.5, there exists a non-zero element $x \in J_0$ such that:

$$|N_{K/\mathbb{Q}}(x)| \leq \left(\frac{2}{\pi}\right)^s \cdot [\mathscr{O}_K : J_0] \cdot \sqrt{|\Delta_K|}.$$

Since $I_0 J_0 = (a)$, Lemma 2.8.2 implies that:

$$[\mathscr{O}_K : J_0] = \frac{[\mathscr{O}_K : (a)]}{[\mathscr{O}_K : I_0]} = \frac{|N_{K/\mathbb{Q}}(a)|}{[\mathscr{O}_K : I_0]},$$

and hence, noting that $xa^{-1}I_0$ is an ideal in $\mathscr{O}_K$:

$$[\mathscr{O}_K : xa^{-1}I_0] = \frac{|N_{K/\mathbb{Q}}(x)|}{|N_{K/\mathbb{Q}}(a)|} \cdot [\mathscr{O}_K : I_0] \leq \left(\frac{2}{\pi}\right)^s \cdot \sqrt{|\Delta_K|}.$$

We conclude by setting $I := xa^{-1}I_0$. □

Theorem 2.8.4

For any number field K, the ideal class group Cl_K is finite.

Proof. For any $M \geq 1$, there exist finitely many ideals I in $\mathscr{O}_K$ satisfying the inequality $[\mathscr{O}_K : I] \leq M$. The Theorem is therefore a direct application of Proposition 2.8.3. □

Chapter 3

p-adic rings and fields

3.1 p-adic numbers

Let p be a prime number. It is well-known that every positive integer n can be written in a unique way as:

$$n = \sum_{i=0}^{r} a_i p^i$$

for some $r \geq 0$ and some $a_0, \dots a_r \in \{0, 1, \dots, p-1\}$ with $a_r \neq 0$. Of course, negative integers and non integral rational numbers cannot be written in this way. To overcome this difficulty, one needs to allow infinite sums. For instance, we formally have:

$$-1 = (p-1) \cdot \frac{1}{1-p} = \sum_{i \geq 0} (p-1) p^i.$$

One way to understand this equality consists in interpreting it as the datum of an infinite number of congruences:

$$\forall r \geq 0, \quad -1 \equiv \sum_{i=0}^{r-1} (p-1) p^i \equiv \sum_{i \geq 0} (p-1) p^i \mod p^r.$$

More generally, an infinite power series $\sum_{i \geq 0} a_i p^i$ can be understood as a "number" n such that:

$$\forall r \geq 0, \quad n \equiv \sum_{i=0}^{r-1} a_i p^i \mod p^r.$$

In general, the "number" n does not necessarily exist as an integer nor even as a rational number (with prime-to-p denominator so that its congruence classes modulo powers of p make sense). This can be easily seen by observing that the set of formal power series $\sum_{i \geq 0} a_i p^i$ with the a_i's in $\{0, 1, \dots, p-1\}$ is uncountable while $\mathbb{Q}$ is countable. This motivates the following definition of p-adic integers.

Definition 3.1.1

A p-**adic integer** is a formal power series $\sum_{i\geq 0} a_i p^i$ with $a_0, a_1, \cdots \in \{0,1,\ldots,p-1\}$. The set of p-adic integers is denoted $\mathbb{Z}_p$.

We would now like to endow $\mathbb{Z}_p$ with a ring structure. However, with the previous definition, it does not seem that obvious to do so. For instance, it would seem natural to define the addition by:

$$\sum_{i\geq 0} a_i p^i + \sum_{i\geq 0} b_i p^i = \sum_{i\geq 0} (a_i + b_i) p^i,$$

but then the sums $a_i + b_i$ would not necessarily belong to $\{0,1,\ldots,p-1\}$.

In fact, as we have already explained, the set of positive integers $\mathbb{N}$ can be seen as a subset of $\mathbb{Z}_p$ by writing integers in base p. It is then natural to try to extend to $\mathbb{Z}_p$ the addition and the multiplication of integers in base p. A practical way to define this properly consists in understanding, as we already suggested, a p-adic integer as a "number" satisfying some congruences modulo powers of p. Consider the map:

$$\begin{aligned} \theta : \mathbb{Z}_p &\to \prod_{r\geq 0} \mathbb{Z}/p^r\mathbb{Z} \\ \sum_{i\geq 0} a_i p^i &\mapsto \left(\sum_{i=0}^{r-1} a_i p^i \mod p^r \right)_{r\geq 0}. \end{aligned}$$

We claim that θ is injective. By contradiction, take two distinct p-adic integers $\sum_{i\geq 0} a_i p^i$ and $\sum_{i\geq 0} b_i p^i$ that have the same image by θ, and let i_0 be the smallest integer i such that $a_i \neq b_i$. The congruence $\sum_{i=0}^{i_0} a_i p^i \equiv \sum_{i=0}^{i_0} b_i p^i \bmod p^{i_0+1}$ implies that $a_{i_0} p^{i_0} \equiv b_{i_0} p^{i_0} \bmod p^{i_0+1}$, and hence that $a_{i_0} \equiv b_{i_0} \bmod p$. But a_{i_0} and b_{i_0} belong to $\{0,1,\ldots,p-1\}$. Hence $a_{i_0} = b_{i_0}$, which is a contradiction.

Let us now compute the image of θ. It is obvious that it is contained in the set:

$$\mathbb{Z}'_p := \{(n_r)_{r\geq 0} : \forall r \geq 1, \pi_r(n_r) = n_{r-1}\} \subset \prod_{r\geq 0} \mathbb{Z}/p^r\mathbb{Z},$$

where $\pi_r : \mathbb{Z}/p^r\mathbb{Z} \to \mathbb{Z}/p^{r-1}\mathbb{Z}$ stands for the natural projection. Conversely, take an element $(n_r)_{r\geq 0} \in \mathbb{Z}'_p$. For each $r \geq 0$, consider a lifting $\tilde{n}_r$ of n_r in $\{0,1,\ldots,p^r-1\}$. Since $\pi_{r+1}(\tilde{n}_{r+1}) = \tilde{n}_r$, we have:

$$0 \leq \tilde{n}_{r+1} - \tilde{n}_r < p^{r+1}, \quad p^r \mid \tilde{n}_{r+1} - \tilde{n}_r,$$

and hence $a_r := \frac{\tilde{n}_{r+1}-\tilde{n}_r}{p^r} \in \{0,1,\dots,p-1\}$. Introduce the p-adic integer:

$$n := \sum_{i\geq 0} a_i p^i.$$

Then, for each $r \geq 0$:

$$\sum_{i=0}^{r} a_i p^i = \sum_{i=0}^{r} \frac{\tilde{n}_{i+1}-\tilde{n}_i}{p^i} p^i = \sum_{i=0}^{r} \frac{\tilde{n}_{i+1}-\tilde{n}_i}{p^i} p^i = \tilde{n}_r,$$

so that $\theta(n) = (n_r)_{r\geq 0}$.

We have therefore proved the following statement:

Proposition 3.1.2

The map θ induces a bijection between $\mathbb{Z}_p$ and the set:

$$\mathbb{Z}'_p := \{(n_r)_{r\geq 0} : \forall r \geq 1, \pi_r(n_r) = n_{r-1}\} \subset \prod_{r\geq 0} \mathbb{Z}/p^r\mathbb{Z}.$$

Since $\mathbb{Z}'_p$ is a subring of $\prod_{r\geq 0} \mathbb{Z}/p^r\mathbb{Z}$, one can endow $\mathbb{Z}_p$ with the unique ring structure for which θ is a ring isomorphism. The ring $\mathbb{Z}$ can then be seen as a subring of $\mathbb{Z}_p \cong \mathbb{Z}'_p$ through the diagonal embedding:

$$\begin{aligned} \iota : \mathbb{Z} &\to \mathbb{Z}_p \\ n &\mapsto (n \mod p^r)_{r\geq 0} \end{aligned}$$

In the sequel, we will systematically identify $\mathbb{Z}_p$ and $\mathbb{Z}'_p$, so the notation $\mathbb{Z}'_p$ will not be used anymore.

3.2 The p-adic valuation

Recall that, for each prime number p, there is a well-defined p-adic valuation on $\mathbb{Z}$:

$$\begin{aligned} v_p : \mathbb{Z} &\to \mathbb{N}_{\geq 0} \cup \{\infty\} \\ n &\mapsto v_p(n) := \sup\{r \geq 0 : p^r \mid n\}. \end{aligned}$$

This valuation satisfies the following basic properties:

$$\forall n \in \mathbb{Z}, \quad v_p(n) = +\infty \Leftrightarrow n = 0,$$
$$\forall m,n \in \mathbb{Z}, \quad v_p(mn) = v_p(m) + v_p(n), \quad v_p(m+n) \geq \min\{v_p(m), v_p(n)\}.$$

Proposition 3.2.1

The p-adic valuation $v_p : \mathbb{Z} \to \mathbb{N}_{\geq 0} \cup \{\infty\}$ can be extended in a unique way to $\mathbb{Z}_p$ so that:

$$\forall n \in \mathbb{Z}_p, \quad v_p(n) = +\infty \Leftrightarrow n = 0,$$
$$\forall m, n \in \mathbb{Z}_p, \quad v_p(mn) = v_p(m) + v_p(n), \quad v_p(m+n) \geq \min\{v_p(m), v_p(n)\}.$$

Proof. Consider the map:

$$v'_p : \mathbb{Z}_p \to \mathbb{N}_{\geq 0} \cup \{\infty\}$$
$$n := (n_r)_{r \geq 0} \mapsto v_p(n) := \sup\{r \geq 0 : n_r = 0\}.$$

We check the different axioms v'_p needs to satisfy.

- If $v'_p(n) = 0$, then $n_r = 0$ for arbitrarily large values of r, and hence, by definition of $\mathbb{Z}_p$, we get $n = 0$.
- If n is an integer, then $n \equiv 0 \mod p^{v_p(n)}$ and $n \not\equiv 0 \mod p^j$ for $j > v_p(n)$, hence $v'_p(n) = v_p(n)$. This shows that v'_p extends v_p.
- Take $m = (m_r)_{r \geq 0}$ and $n = (n_r)_{r \geq 0}$ in $\mathbb{Z}_p$ and set $a := v'_p(m)$ and $b := v'_p(n)$. If $m = 0$ or $n = 0$, the equality $v'_p(mn) = v'_p(m) + v'_p(n)$ is obvious. We henceforth assume that m and n are both non-zero. The images of $m_{a+b} \in \mathbb{Z}/p^{a+b}\mathbb{Z}$ in $\mathbb{Z}/p^a\mathbb{Z}$ and of $n_{a+b} \in \mathbb{Z}/p^{a+b}\mathbb{Z}$ in $\mathbb{Z}/p^b\mathbb{Z}$ are both zero, hence $m_{a+b}n_{a+b} = 0$. The images of $m_{a+b+1} \in \mathbb{Z}/p^{a+b+1}\mathbb{Z}$ in $\mathbb{Z}/p^{a+1}\mathbb{Z}$ and of $n_{a+b+1} \in \mathbb{Z}/p^{a+b+1}\mathbb{Z}$ in $\mathbb{Z}/p^{b+1}\mathbb{Z}$ are both non-zero, hence $m_{a+b+1}n_{a+b+1} \neq 0$. We deduce that $v'_p(mn) = v'_p(m) + v'_p(n)$.
- Take $m = (m_r)_{r \geq 0}$ and $n = (n_r)_{r \geq 0}$ in $\mathbb{Z}_p$ and set $a := v'_p(m)$ and $b := v'_p(n)$. If $m = 0$ or $n = 0$, the inequality $v_p(m+n) \geq \min\{v_p(m), v_p(n)\}$ is obvious. We henceforth assume that m and n are both non-zero. Assume moreover without loss of generality that $a \leq b$. Then $m_a = n_a = 0$, and hence $v'_p(a+b) \geq a = \min\{v'_p(m), v'_p(n)\}$.

□

Proposition 3.2.2

Fix $\alpha \geq 0$. For $n \in \mathbb{Z}_p$, we have $v_p(n) \geq \alpha$ if, and only if, $p^\alpha | n$.

Proof. If $n = 0$, there is nothing to prove. We henceforth assume that $n \neq 0$. Write $n = (n_r)_{r\geq 0}$. If p^α divides n in $\mathbb{Z}_p$, then p^α divides n_r in $\mathbb{Z}/p^r\mathbb{Z}$ for each $r \geq 0$. Hence $n_0 = n_1 = \cdots = n_\alpha = 0$ and $v_p(n) \geq \alpha$.

Conversely, assume that $v_p(n) \geq \alpha$. Then for each $r \geq 0$ we can find $m'_r \in \mathbb{Z}/p^r\mathbb{Z}$ such that $p^\alpha m'_r = n_r$. Set $m := (m_r)_{r\geq 0} \in \prod_{r\geq 0} \mathbb{Z}/p^r\mathbb{Z}$ where $m_r := \pi^\alpha_{r+\alpha}(m'_{r+\alpha})$ and $\pi^\alpha_{r+\alpha} : \mathbb{Z}/p^{r+\alpha}\mathbb{Z} \to \mathbb{Z}/p^r\mathbb{Z}$ stands for the natural projection. Let us check that $m \in \mathbb{Z}_p$ and that $p^\alpha m = n$.

- For $r \geq 1$, we have
$$p^\alpha \pi_{r+\alpha}(m'_{r+\alpha}) = \pi_{r+\alpha}(n_{r+\alpha}) = n_{r+\alpha-1} = p^\alpha m'_{r+\alpha-1} \in \mathbb{Z}/p^{r+\alpha-1}\mathbb{Z}.$$
Hence $p^{r-1} | \pi_{r+\alpha}(m'_{r+\alpha}) - m'_{r+\alpha-1}$ and:
$$\pi_r(m_r) = \pi_r(\pi^\alpha_{r+\alpha}(m'_{r+\alpha})) = \pi^\alpha_{r+\alpha-1}(\pi_{r+\alpha}(m'_{r+\alpha})) = \pi^\alpha_{r+\alpha-1}(m'_{r+\alpha-1}) = m_{r-1}.$$
We deduce that $m \in \mathbb{Z}_p$.
- We have:
$$p^\alpha m = (p^\alpha m_r)_{r\geq 0} = (\pi^\alpha_{r+\alpha}(p^\alpha m'_{r+\alpha}))_{r\geq 0} = (\pi^\alpha_{r+\alpha}(n_{r+\alpha}))_{r\geq 0} = (n_r)_{r\geq 0} = n.$$

Hence $p^\alpha | n$ in $\mathbb{Z}_p$. □

Corollary 3.2.3

For every $n \in \mathbb{Z}_p$, we have $v_p(n) = \sup\{r \geq 0 : p^r \mid n\}$.

Proposition 3.2.4

We have:
$$\mathbb{Z}_p^\times = \{n \in \mathbb{Z}_p : v_p(n) = 0\}.$$

Proof. If $n \in \mathbb{Z}_p^\times$, then there exists $m \in \mathbb{Z}_p$ such that $mn = 1$. Hence $v_p(m) + v_p(n) = v_p(1) = 0$, so that $v_p(n) = 0$.

Conversely, take $n \in \mathbb{Z}_p$ such that $v_p(n) = 0$ and write $n = (n_r)_{r\geq 0}$. Since $n_1 \neq 0$ and the image of each n_r in $\mathbb{Z}/p\mathbb{Z}$ is n_1, we deduce that $n_r \in (\mathbb{Z}/p^r\mathbb{Z})^\times$ for each $r \geq 1$. Set $m := (m_r)_{r\geq 0} \in \prod_{r\geq 0} \mathbb{Z}/p^r\mathbb{Z}$ where $m_r := n_r^{-1}$. Observing that $\pi_r(m_r) = \pi_r(n_r^{-1}) = \pi_r(n_r)^{-1} = n_{r-1}^{-1} = m_{r-1}$ for each $r \geq 1$, we see that $m \in \mathbb{Z}_p$. Since $mn = 1$, we deduce that $n \in \mathbb{Z}_p^\times$. □

Corollary 3.2.5

Every non-zero element $n \in \mathbb{Z}_p$ can be written uniquely as $n = up^r$ with $u \in \mathbb{Z}_p^\times$ and $r \geq 0$.

Proof. Take n a non-zero p-adic integer. By Proposition 3.2.2, we can write $n = up^r$ with $r = v_p(n)$ and $u \in \mathbb{Z}_p$. We then have $v_p(u) = v_p(n) - r\,v_p(p) = 0$, and hence, by Proposition 3.2.4, we get $u \in \mathbb{Z}_p^\times$. This settles the existence part of the statement.

Take now $u' \in \mathbb{Z}_p^\times$ and $r' \geq 0$ such that $n = u'p^{r'}$. Then

$$v_p(n) = v_p(u') + r'\,v_p(p) = r'$$

by Proposition 3.2.4. Hence $r' = r$ and $u' = u$, which shows the uniqueness. □

Theorem 3.2.6

The ring of p-adic integers $\mathbb{Z}_p$ is a principal ideal domain. Its ideals are the $p^r\mathbb{Z}_p$ for $r \in \mathbb{N}_{\geq 0}$ and for each $r \geq 0$, the quotient $\mathbb{Z}_p/p^r\mathbb{Z}_p$ is isomorphic to $\mathbb{Z}/p^r\mathbb{Z}$. In particular, $p\mathbb{Z}_p$ is the unique maximal ideal of $\mathbb{Z}_p$ and $\mathbb{Z}_p/p\mathbb{Z}_p$ is isomorphic to $\mathbb{F}_p$.

Proof. Observe first that $\mathbb{Z}_p$ is a domain: indeed, if we take m, n two non-zero p-adic integers and we write $m = up^r$ and $n = vp^s$ with $u, v \in \mathbb{Z}_p^\times$ and $r, s \geq 0$, then $mn = (uv)p^{r+s}$. Since $uv \in \mathbb{Z}_p^\times$ and $p^{r+s} \neq 0 \in \mathbb{Z} \subset \mathbb{Z}_p$, we deduce that $mn \neq 0 \in \mathbb{Z}_p$.

Let now I be a non-zero ideal in $\mathbb{Z}_p$ and take an element $m \in I \setminus \{0\}$ such that $v_p(m)$ is minimal. Take a non-zero element $n \in I$ and write $m = up^r$ and $n = vp^s$ with $r = v_p(m)$, $s = v_p(n)$ and $u, v \in \mathbb{Z}_p^\times$ thanks to Corollary 3.2. We immediately see that $m \mid n$ since $s \geq r$, and hence $n \in m\mathbb{Z}_p$. This being true for every non-zero $n \in I$, we get $I = m\mathbb{Z}_p = p^r\mathbb{Z}_p$. In particular, $\mathbb{Z}_p$ is a principal ideal domain and its ideals are the $p^r\mathbb{Z}_p$ for $r \in \mathbb{N}_{\geq 0}$

It only remains to compute the quotient $\mathbb{Z}_p/p^r\mathbb{Z}_p$ for $r \geq 0$. To do so, we consider the natural morphism:

$$\varphi : \mathbb{Z} \to \mathbb{Z}_p/p^r\mathbb{Z}_p$$

induced by the inclusion $\mathbb{Z} \subset \mathbb{Z}_p$. Its kernel is:

$$\ker(\varphi) = \mathbb{Z} \cap p\mathbb{Z}_p = \{a \in \mathbb{Z} : v_p(a) \geq 1\} = p\mathbb{Z}.$$

Moreover, given $n = (n_r)_{r\geq 0} \in \mathbb{Z}_p$, we can choose $\tilde{n}_r \in \mathbb{Z}$ a lifting of $n_r \in \mathbb{Z}/p^r\mathbb{Z}$, and then $v_p(n_r - \tilde{n}_r) \geq r$. By Proposition 3.2.2 we deduce that $p^r \mid n_r - \tilde{n}_r$, so that $\varphi(n_r)$ is equal to the class of n modulo p^r. In other words, φ is surjective, and hence, by the universal property of the quotient, it induces an isomorphism between $\mathbb{Z}/p^r\mathbb{Z}$ and $\mathbb{Z}_p/p^r\mathbb{Z}_p$. □

Definition 3.2.7

The field of fractions of $\mathbb{Z}_p$ is called the **field of p-adic numbers** and is denoted $\mathbb{Q}_p$. The p-adic valuation on $\mathbb{Z}_p$ can be extended to $\mathbb{Q}_p$ by the formula $v_p(m/n) = v_p(m) - v_p(n)$ with $m, n \in \mathbb{Q}_p$. One immediately checks that its properties given by Proposition 3.2.1 extend to $\mathbb{Q}_p$.

Proposition 3.2.8

Inside $\mathbb{Q}_p$, the ring $\mathbb{Z}_p$ can be described as follows:

$$\mathbb{Z}_p = \{x \in \mathbb{Q}_p : v_p(x) \geq 0\}.$$

Proof. Of course, $\mathbb{Z}_p \subset \{x \in \mathbb{Q}_p : v_p(x) \geq 0\}$. Conversely, take $x \in \mathbb{Q}_p$ with $v_p(x) \geq 0$ and write $x = m/n$ with $m, n \in \mathbb{Z}_p$. Then $v_p(m) \geq v_p(n)$, and hence $n \mid m$ in $\mathbb{Z}_p$. We deduce that $x \in \mathbb{Z}_p$. □

3.3 The p-adic absolute value

In the previous sections, we have proved that $\mathbb{Z}_p$ is a principal ideal domain endowed with a p-adic valuation v_p. It is often also useful to endow $\mathbb{Z}_p$ with a topological structure. To do so, we introduce the p-adic absolute value and the corresponding distance:

Definition 3.3.1

Given $x \in \mathbb{Q}_p$, the **p-adic absolute value** of x is:

$$|x|_p := p^{-v_p(x)}.$$

Proposition 3.3.2

The map:

$$\begin{aligned} d_p : \mathbb{Q}_p \times \mathbb{Q}_p &\to \mathbb{R}_{\geq 0} \\ (x, y) &\mapsto |x - y|_p \end{aligned}$$

is a distance on $\mathbb{Q}_p$, called the p-**adic distance** on $\mathbb{Q}_p$.

Proof. We check the axioms of a distance.

- For $x, y \in \mathbb{Q}_p$:

$$d_p(x, y) = 0 \Leftrightarrow v_p(x - y) = +\infty \Leftrightarrow x - y = 0 \Leftrightarrow x = y.$$

- For $x, y \in \mathbb{Q}_p$, we obviously have $d_p(x, y) = d_p(y, x)$.
- For $x, y, z \in \mathbb{Q}_p$:

$$\begin{aligned} d_p(x, z) = p^{-v_p(x-z)} = p^{-v_p((x-y)+(y-z))} &\leq p^{-\min\{v_p(x-y), v_p(y-z)\}} \\ &\leq \max\{d_p(x, y), d_p(y, z)\} \leq d_p(x, y) + d_p(y, z). \end{aligned}$$

We deduce that d_p is a distance on $\mathbb{Q}_p$. □

Proposition 3.3.3

The three maps:

$$\begin{aligned} + : \mathbb{Q}_p \times \mathbb{Q}_p &\to \mathbb{Q}_p & \cdot : \mathbb{Q}_p \times \mathbb{Q}_p &\to \mathbb{Q}_p & \iota : \mathbb{Q}_p^\times &\to \mathbb{Q}_p^\times \\ (x, y) &\mapsto x + y, & (x, y) &\mapsto xy, & x &\mapsto x^{-1} \end{aligned}$$

are all continuous. We say that $\mathbb{Q}_p$ is a **topological field**.

Proof.

- Take $x, y, x', y' \in \mathbb{Q}_p$. Choose $r \geq 0$ such that $d_p(x, x') \leq p^{-r}$ and $d_p(y, y') \leq p^{-r}$. Then $v_p(x - x') \geq r$ and $v_p(y - y') \geq r$. Hence $v_p(x + y - x' - y') \geq r$ and $v_p(xy - x'y') \geq r$. We deduce that $d_p(x + y, x' + y') \leq p^{-r}$ and $d_p(xy, x'y') \leq p^{-r}$. We conclude that $+$ and $\cdot$ are continuous.
- Choose an integer $s \in \mathbb{Z}$ and take $x, x' \in \mathbb{Q}_p^\times$ with $v_p(x) \leq s$ and $v_p(x') \leq s$. Choose $r \geq 0$ such that $d_p(x, x') \leq p^{-r}$. Then $v_p(x - x') \geq r$, and hence:

$$v_p(x^{-1} - x'^{-1}) = v_p(x' - x) - v_p(x) - v_p(x') \geq r - 2s.$$

In other words, $d_p(\iota(x),\iota(x')) \leq p^{2s-r}$. We deduce that ι is continuous on $\{x \in \mathbb{Q}_p^\times : v_p(x) \leq s\}$. This being true for all $s \in \mathbb{Z}$, the map ι is continuous on $\mathbb{Q}_p^\times$.

□

Theorem 3.3.4

The metric space $\mathbb{Z}_p$ is compact.

Proof. Consider a sequence $(n^{(i)})_{i\geq 0}$ of p-adic integers. For each $i \geq 0$, write $n^{(i)} = (n_r^{(i)})_{r\geq 0}$. We construct an extraction of the sequence $(n^{(i)})_{i\geq 0}$ in the following way.

- Since the set $\mathbb{Z}/p\mathbb{Z}$ is finite, we can find $i_1 \geq 0$ such that the set S_1 of integers $i > i_1$ with $n_1^{(i)} = n_1^{(i_1)}$ is infinite.
- Similarly, since the set $\mathbb{Z}/p^2\mathbb{Z}$ is finite, we can find $i_2 \in S_1$ such that the set S_2 of integers $i \in S_1 \cap \mathbb{N}_{>i_2}$ with $n_2^{(i)} = n_2^{(i_2)}$ is infinite.
- Similarly, since the set $\mathbb{Z}/p^3\mathbb{Z}$ is finite, we can find $i_3 \in S_2$ such that the set S_3 of integers $i \in S_2 \cap \mathbb{N}_{>i_3}$ with $n_3^{(i)} = n_3^{(i_3)}$ is infinite.
- And so on.

Consider the sequence $(n^{(i_k)})_{k\geq 1}$. It is obviously an extraction of the sequence $(n^{(i)})_{i\geq 0}$. Let us prove that it converges. To do so, we introduce:

$$n^{(\infty)} := (n_r^{(\infty)}) \in \prod_{r\geq 0} \mathbb{Z}/p^r\mathbb{Z}$$

with $n_r^{(\infty)} := n_r^{(r)}$. We observe that, for $r \geq 1$:

$$\pi_r(n_r^{(\infty)}) = \pi_r(n_r^{(r)}) = n_{r-1}^{(r)} = n_{r-1}^{(r-1)} = n_{r-1}^{(\infty)}.$$

Hence $n^{(\infty)} \in \mathbb{Z}_p$. Moreover, given $i \geq 0$, we have $n_r^{(\infty)} = n_r^{(r)} = n_r^{(i)}$ for each $r \leq i$, and hence $v_p(n^{(\infty)} - n^{(i)}) \geq i$. We deduce that:

$$d_p(n^{(i)}, n^{(\infty)}) = p^{-v_p(n^{(\infty)}-n^{(i)})} \leq p^{-i} \xrightarrow[i\to\infty]{} 0,$$

and hence $n^{(i)} \xrightarrow[i\to\infty]{} n^{(\infty)}$. This shows that $\mathbb{Z}_p$ is compact. □

Corollary 3.3.5

The metric space $\mathbb{Q}_p$ is complete.

Proof. Take a Cauchy sequence $(x^{(i)})_{i\geq 0}$ of p-adic numbers. It is bounded, hence it is contained in a subset of $\mathbb{Q}_p$ of the form $p^r\mathbb{Z}_p$ for some $r \in \mathbb{Z}$. By Theorem 3.3.4, the sequence $(p^{-r}x^{(i)})_{i\geq 0}$ has a convergent subsequence, hence so does the sequence $(x^{(i)})_{i\geq 0}$ by Proposition 3.3.3. Now $(x^{(i)})_{i\geq 0}$ is a Cauchy sequence that has a convergent subsequence: it is therefore convergent. We deduce that $\mathbb{Q}_p$ is complete. □

3.4 The Weak Approximation Theorem

For each prime number p, the ring $\mathbb{Z}$ can be seen as a subring of $\mathbb{Z}_p$ and the field $\mathbb{Q}$ can be seen as a subfield of $\mathbb{Q}_p$. In order to better relate $\mathbb{Z}$ and $\mathbb{Q}$ to $\mathbb{Z}_p$ and $\mathbb{Q}_p$, it is crucial to understand how large $\mathbb{Z}$ and $\mathbb{Q}$ are in $\mathbb{Z}_p$ and $\mathbb{Q}_p$. For that purpose, we start with the following easy proposition:

Proposition 3.4.1

The ring $\mathbb{Z}$ is dense in $\mathbb{Z}_p$. The field $\mathbb{Q}$ is dense in $\mathbb{Q}_p$.

Proof. Take $n \in \mathbb{Z}_p$. For each $r \geq 0$, since $\mathbb{Z}/p^r\mathbb{Z} \cong \mathbb{Z}_p/p^r\mathbb{Z}_p$, there exists an integer n_r such that $n \equiv n_r \mod p^r$. We then have $n_r \xrightarrow[r\to\infty]{} n$. Hence $\mathbb{Z}$ is dense in $\mathbb{Z}_p$.

Now take $x \in \mathbb{Q}_p$. Then $p^{-v_p(x)}x \in \mathbb{Z}_p$. In particular, $p^{-v_p(x)}x$ belongs to the closure of $\mathbb{Z}$ in $\mathbb{Z}_p$. Hence $x = p^{v_p(x)}p^{-v_p(x)}x$ belongs to the closure of $p^{v_p(x)}\mathbb{Z}$ in $\mathbb{Q}_p$. Since $p^{v_p(x)}\mathbb{Z}$ is contained in $\mathbb{Q}$, we deduce that $\mathbb{Q}$ is dense in $\mathbb{Q}_p$. □

In order to do arithmetic in the field of rational numbers, it is often useful to work with several $\mathbb{Z}_p$'s or several $\mathbb{Q}_p$'s at the same time. This can be achieved by so-called Weak Approximation Theorem:

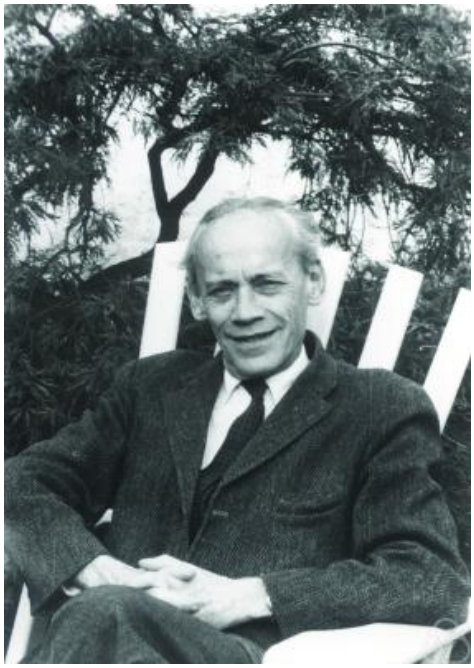
Theorem 3.4.2 (Artin-Whaples Weak Approximation)

Consider the diagonal embedding:

$$\iota : \mathbb{Q} \to \mathbb{R} \times \prod_{p \text{ prime}} \mathbb{Q}_p.$$

Its image is dense in $\mathbb{R} \times \prod_{p \text{ prime}} \mathbb{Q}_p$, where the product is endowed with the product topology. In other words, for any finite set of prime numbers

S, for any $(x_p)_{p\in S} \in \prod_{p\in S}\mathbb{Q}_p$ and any $x_\infty \in \mathbb{R}$, there exists a sequence of rational numbers $(y_n)_{n\geq 0}$ that converges to x_p in $\mathbb{Q}_p$ for each $p \in S$ and that converges to x_∞ in $\mathbb{R}$.



Emil Artin (Vienna 1898 - Hamburg 1962)
No picture seems to be available for George William Whaples
(Neponset 1914 - Amherst 1981)

Proof. Given $p \in S$, one can choose an integer $a(p) \geq 1$ such that $p^{a(p)} > \prod_{q\in S\setminus\{p\}} q$, and then:

$$\frac{1}{1+\left(\frac{p^{a(p)}}{\prod_{q\in S\setminus\{p\}} q}\right)^k} \xrightarrow[k\to\infty]{} \begin{cases} 1 & \text{in } \mathbb{Q}_p, \\ 0 & \text{in } \mathbb{Q}_{p'} \text{ for } p' \in S\setminus\{p\}, \\ 0 & \text{in } \mathbb{R}. \end{cases}$$

Similarly, one checks that:

$$\frac{1}{1+\left(\frac{1}{\prod_{q\in S} q}\right)^k} \xrightarrow[k\to\infty]{} \begin{cases} 0 & \text{in } \mathbb{Q}_p \text{ for } p \in S, \\ 1 & \text{in } \mathbb{R}. \end{cases}$$

Now take a sequence $(x_\infty^{(j)})_{k\geq 0}$ of rational numbers converging to x_∞ in $\mathbb{R}$ and, for each $p \in S$, a sequence $(x_p^{(j)})_{k\geq 0}$ of rational numbers converging to x_p in $\mathbb{Q}_p$.

These sequences exist because $\mathbb{Q}$ is dense in $\mathbb{R}$ and in each p-adic field. For each j, one then has:

$$\frac{x_\infty^{(j)}}{1+\left(\frac{1}{\prod_{q\in S} q}\right)^k} + \sum_{p\in S} \frac{x_p^{(j)}}{1+\left(\frac{p^{a(p)}}{\prod_{q\in S\setminus\{p\}} q}\right)^k} \xrightarrow[k\to\infty]{} \begin{cases} x_p^{(j)} & \text{in } \mathbb{Q}_p \text{ for } p\in S, \\ x_\infty^{(j)} & \text{in } \mathbb{R}. \end{cases}$$

Hence one can find a map $\varphi : \mathbb{N}_{>0} \to \mathbb{N}_{>0}$ such that:

$$\frac{x_\infty^{(j)}}{1+\left(\frac{1}{\prod_{q\in S} q}\right)^{\varphi(j)}} + \sum_{p\in S} \frac{x_p^{(j)}}{1+\left(\frac{p^{a(p)}}{\prod_{q\in S\setminus\{p\}} q}\right)^{\varphi(j)}} \xrightarrow[j\to\infty]{} \begin{cases} x_p & \text{in } \mathbb{Q}_p \text{ for } p\in S, \\ x_\infty & \text{in } \mathbb{R}. \end{cases}$$

□

3.5 p-adic equations and Hensel's Lemma

Proposition 3.5.1

Consider a finite number of polynomials $f_1,\dots,f_m \in \mathbb{Z}_p[X_1,\dots,X_n]$. The following assertions are equivalent:

(i) The system $f_1 = \cdots = f_m = 0$ has solutions in $\mathbb{Z}_p$.

(ii) The system $f_1 = \cdots = f_m = 0$ has solutions in $\mathbb{Z}/p^r\mathbb{Z}$ for each $r \geq 1$.

Proof. The direct implication is obvious. Conversely, assume that, for each $r \geq 1$, there exists $\mathbf{x}^{(r)} \in \mathbb{Z}_p^n$ such that:

$$f_1(\mathbf{x}^{(r)}) \equiv \cdots \equiv f_m(\mathbf{x}^{(r)}) \equiv 0 \mod p^r.$$

Since $\mathbb{Z}_p$ is compact, one can extract a subsequence of $(\mathbf{x}^{(r)})_{r\geq 1}$ that converges to some $\mathbf{x}^{(\infty)} \in \mathbb{Z}_p^n$. This limit is a solution to the system $f_1 = \cdots = f_m = 0$. □

Proposition 3.5.2

Consider a finite number of homogeneous polynomials $f_1,\dots,f_m \in \mathbb{Z}_p[X_1,\dots,X_n]$. The following assertions are equivalent:

(i) The system $f_1 = \cdots = f_m = 0$ has non-zero solutions in $\mathbb{Z}_p$.

(ii) The system $f_1 = \cdots = f_m = 0$ has non-zero solutions in $\mathbb{Q}_p$.

(iii) The system $f_1 = \cdots = f_m = 0$ has primitive solutions in $\mathbb{Z}/p^r\mathbb{Z}$ for each $r \geq 1$, i.e. solutions $(x_1, \ldots, x_n)$ such that not all the x_i's are divisible by p.

Proof. The implication $(i) \Rightarrow (ii)$ is obvious. Let us prove $(ii) \Rightarrow (iii)$. Take $\mathbf{x} = (x_1, \ldots, x_n)$ a non-zero solution to the system $f_1 = \cdots = f_m = 0$ in $\mathbb{Q}_p$. Then $\mathbf{y} := p^{-\min\{v_p(x_1), \ldots, v_p(x_n)\}}\mathbf{x}$ is an n-tuple of p-adic integers and at least one of its entries is a p-adic unit. Hence, for each $r \geq 0$, the reduction of this n-tuple modulo p^r is a primitive solution to the system $f_1 = \cdots = f_m = 0 \mod p^r$ in $\mathbb{Z}/p^r\mathbb{Z}$.

Let us finally prove the implication $(iii) \Rightarrow (i)$. For each $r \geq 1$, take $\mathbf{x}^{(r)} = (x_1^{(r)}, \ldots, x_n^{(r)}) \in \mathbb{Z}_p^n$ such that at least one of its entries is a p-adic unit and:

$$f_1(\mathbf{x}^{(r)}) \equiv \cdots \equiv f_m(\mathbf{x}^{(r)}) \equiv 0 \mod p^r.$$

Since $\mathbb{Z}_p$ is compact, one can extract a subsequence of $(\mathbf{x}^{(r)})_{r\geq 1}$ that converges to some $\mathbf{x}^{(\infty)} = (x_1^{(\infty)}, \ldots, x_n^{(\infty)}) \in \mathbb{Z}_p^n$. This limit is a solution to the system $f_1 = \cdots = f_m = 0$, and it is non-zero since:

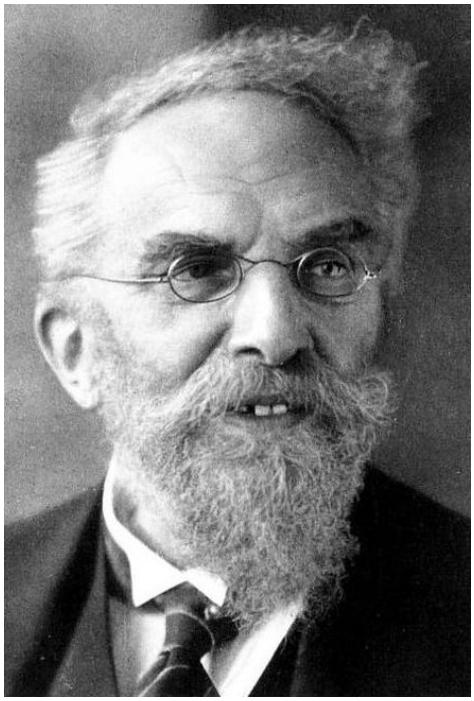
$$\min\{v_p(x_1^{(\infty)}), \ldots, v_p(x_n^{(\infty)})\} = \lim_{r\to+\infty} \min\{v_p(x_1^{(r)}), \ldots, v_p(x_n^{(r)})\} = 0.$$

□

Theorem 3.5.3 (Hensel)

Consider a polynomial $f \in \mathbb{Z}_p[X]$ and an element $a \in \mathbb{Z}_p$ such that $v_p(f(a)) > 2v_p(f'(a))$. Then there exists $b \in \mathbb{Z}_p$ such that:

$$\begin{gathered} f(b) = 0, \\ v_p(b-a) = v_p(f(a)) - v_p(f'(a)) > v_p(f'(a)), \\ v_p(f'(b)) = v_p(f'(a)). \end{gathered}$$



Kurt Hensel (Königsberg 1861 - Marburg 1941)

Proof. Set $n := v_p(f(a))$ and $m := v_p(f'(a))$. We are going to construct b by approximating it by an inductively constructed sequence $(b_i)_{i\geq 0}$. To do so, start with $b_0 = a$. By assumption, we have:

$$f(b_0) \equiv 0 \mod p^n.$$

Now we would like to construct $b_1 \in \mathbb{Z}_p$ such that $f(b_1) \equiv 0 \mod p^{n+1}$. To do so, we look for b_1 in the form $b_1 = b_0 + p^{r_0}c_0$ for some $r_0 \geq 1$ and $c_0 \in \mathbb{Z}_p$. We therefore want:

$$f(b_0 + p^{r_0}c_0) \equiv 0 \mod p^{n+1}.$$

By Taylor's formula, we have:

$$f(b_0 + p^{r_0}c_0) \equiv f(b_0) + p^{r_0}c_0 f'(b_0) \mod p^{2r_0}.$$

Recall that $v_p(f(b_0)) = n$ and $v_p(f'(b_0)) = m$, so we may write $f(b_0) = p^n u_0$ and $f'(b_0) = p^m u_0'$ for some units $u_0, u_0' \in \mathbb{Z}_p^\times$. We therefore get:

$$f(b_0 + p^{r_0}c_0) \equiv p^n u_0 + p^{r_0+m}c_0 u_0' \mod p^{2r_0},$$

and so we want $2r_0 \geq n+1$ and:

$$p^n u_0 + p^{r_0+m}c_0 u_0' \equiv 0 \mod p^{n+1}. \tag{3.1}$$

This requires to choose $r_0 = n - m$. The inequality $2r_0 \geq n+1$ then immediately follows from the assumption that $v_p(f(a)) > 2v_p(f'(a))$, and congruence 3.1 is equivalent to:

$$p^n(u_0 + c_0 u_0') \equiv 0 \mod p^{n+1},$$

and hence to:

$$u_0 + c_0 u_0' \equiv 0 \mod p.$$

Since $u_0' \in \mathbb{Z}_p^\times$, we may therefore choose $c_0 = -u_0 u_0'^{-1}$. Note also that $v_p(f'(b_1)) = v_p(f'(b_0)) = m$ since $b_1 \equiv b_0 \mod p^{n-m}$ and $n - m > m$.

Summarizing, by setting $b_1 = b_0 - p^{n-m} u_0 u_0'^{-1}$, we have:

$$f(b_1) \equiv 0 \mod p^{n+1}, \quad v_p(b_1 - b_0) = n - m, \quad v_p(f'(b_1)) = m.$$

By repeating the previous construction, we can inductively find $b_2, b_3, \cdots \in \mathbb{Z}_p$ such that:

$$\begin{aligned}
f(b_2) \equiv 0 \mod p^{n+2}, \quad b_2 \equiv b_1 \mod p^{n+1-m}, \quad v_p(f'(b_2)) = m, \\
f(b_3) \equiv 0 \mod p^{n+3}, \quad b_3 \equiv b_2 \mod p^{n+2-m}, \quad v_p(f'(b_3)) = m, \\
\cdots, \\
f(b_i) \equiv 0 \mod p^{n+i}, \quad b_i \equiv b_{i-1} \mod p^{n+i-1-m}, \quad v_p(f'(b_i)) = m,
\end{aligned}$$

for all i. This implies that the sequence $(b_i)_{i \geq 0}$ converges to some $b \in \mathbb{Z}_p$ such that $f(b) = 0$, $v_p(b - a) = n - m > m$ and $v_p(f'(b)) = m$. □

Example 3.5.4. Consider $f = X^3 - 19 \in \mathbb{Z}_3[X]$. We have $v_3(f(-2)) = 3$ and $v_3(f'(-2)) = 1$. By Hensel's lemma, we deduce that 19 has a cubic root in $\mathbb{Z}_3$.

Chapter 4

Riemann zeta function, L-functions and Dirichlet's Theorem on arithmetic progressions

In this chapter, we aim at proving Dirichlet's Theorem on arithmetic progressions, which asserts that a sequence of the form $(a+bn)_{n\geq 0}$ with $a, b \in \mathbb{N}$ coprime contains infinitely many prime numbers. For that purpose, we will need to develop techniques from analytic number theory.

4.1 Dirichlet series and Euler products

Definition 4.1.1

A **Dirichlet series** is a series of the form:

$$L(s) = \sum_{n\geq 1} \frac{a_n}{n^s}$$

where $(a_n)_{n\geq 1}$ is a sequence of complex numbers.

Lemma 4.1.2. *Let $L(s) = \sum_{n\geq 1} \frac{a_n}{n^s}$ be a Dirichlet series such that the sequence $(a_n)_{n\geq 1}$ is bounded. Then the series defining $L(s)$ converges absolutely and normally on any half-plane of the form $\mathrm{Re}(s) > s_0$ with $s_0 > 1$. In particular, L defines a holomorphic function on the half-plane $\mathrm{Re}(s) > 1$.*

Proof. Choose $C > 0$ an upper bound for the sequence $(a_n)_{n\geq 1}$. For $n \geq 1$ and $s \in \mathbb{C}$, we have:

$$\left|\frac{a_n}{n^s}\right| \leq \frac{C}{n^{\mathrm{Re}(s)}}.$$

The lemma follows from the basic fact that the series $\sum \frac{1}{n^{\mathrm{Re}(s)}}$ converges uniformly on any half-plane of the form $\mathrm{Re}(s) > s_0$ with $s_0 > 1$. □

Peter Gustav Lejeune Dirichlet (Düren 1805 - Göttingen 1859)

Definition 4.1.3

Let $f : \mathbb{N} \to \mathbb{C}$. We say that f is **multiplicative** if $f(1) = 1$ and $f(mn) = f(n)f(m)$ whenever m and n are coprime. We say that f is **completely multiplicative** if $f(1) = 1$ and $f(mn) = f(m)f(n)$ for all $m, n \in \mathbb{N}$.

Theorem 4.1.4

Let $f : \mathbb{N} \to \mathbb{C}$ be a multiplicative bounded function and consider the Dirichlet series:

$$L(s) = \sum_{n \geq 1} \frac{f(n)}{n^s}.$$

We then have, for $s \in \mathbb{C}$ with $\mathrm{Re}(s) > 1$:

$$L(s) = \prod_{p \text{ prime}} L_p(s),$$

where L_p stands for the *local Dirichlet series*:

$$L_p(s) = \sum_{r \geq 0} \frac{f(p^r)}{p^{rs}}.$$

Proof. Fix a positive integer N and let $p_1, \dots, p_m$ be the prime numbers smaller or equal to N. Since all the local Dirichlet series are absolutely convergent, we can compute:

$$\prod_{p \leq N \text{ prime}} L_p(s) = \sum_{r_1,\dots,r_m \geq 0} \frac{f(p_1^{r_1} \dots p_m^{r_m})}{(p_1^{r_1} \dots p_m^{r_m})^s} = \sum_{n \in X_N} \frac{f(n)}{n^s}$$

where X_N is the set of positive integers whose prime factors are all $\leq N$. Since $\mathbb{N} = \bigcup_{N \in \mathbb{N}} X_N$ and the series $\sum \frac{f(n)}{n^s}$ is absolutely convergent, we deduce that:

$$\prod_{p \leq N \text{ prime}} L_p(s) \xrightarrow[N \to \infty]{} L(s),$$

as wished. □

Remark 4.1.5. In the previous theorem, if f is completely multiplicative, then for each prime number p and each $s \in \mathbb{C}$ with $\mathrm{Re}(s) > 1$:

$$L_p(s) = \sum_{r \geq 0} \frac{f(p^r)}{p^{rs}} = \sum_{r \geq 0} \frac{f(p)^r}{p^{rs}} = \frac{1}{1 - f(p)p^{-s}}.$$

Hence we get:

$$L(s) = \prod_{p \text{ prime}} \frac{1}{1 - f(p)p^{-s}}.$$

4.2 Riemann's zeta function

Definition 4.2.1

The **Riemann zeta function** is the Dirichlet series:

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

According to the results we settled in the previous section, the zeta function defines a holomorphic function on the half-plane $\mathrm{Re}(s) > 1$, and we have the equality:

$$\zeta(s) = \prod_{p \text{ prime}} \frac{1}{1 - p^{-s}}$$

on that half-plane.

Bernhard Riemann (Breselenz 1826 - Selasca 1866)

Theorem 4.2.2

The zeta function can be extended to a meromorphic function on the half-plane $\mathrm{Re}(s) > 0$ whose unique pole is $s = 1$ and is simple.

Proof. For $s \in \mathbb{C}$ with $\mathrm{Re}(s) > 1$, we have:

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

For each $n \geq 1$, consider the function:

$$f_n : s \mapsto \int_n^{n+1} \left(\frac{1}{n^s} - \frac{1}{x^s}\right) dx.$$

It is a holomorphic function on $\mathrm{Re}(s) > 0$. Moreover, for each $s \in \mathbb{C}$ such that $\mathrm{Re}(s) > 0$, we have the bound:

$$\left|f_n(s)\right| \leq \int_n^{n+1} \left|\frac{1}{n^s} - \frac{1}{x^s}\right| \leq \sup_{n\leq x\leq n+1} \left|\frac{s}{x^{s+1}}\right| \leq \frac{|s|}{n^{\mathrm{Re}(s)+1}}.$$

Since the series $\sum \frac{|s|}{n^{\mathrm{Re}(s)+1}}$ converges normally on compact subsets of the half plane $\mathrm{Re}(s) > 0$, so does the series $\sum f_n$, and hence $\sum f_n$ defines a holomorphic function on that half-plane. We can therefore extend the zeta function into a

meromorphic function on the half-plane $\mathrm{Re}(s) > 0$ whose unique pole is $s = 1$ and is simple by the formula:

$$\zeta(s) := \sum_{n\geq 1} \int_n^{n+1} \left(\frac{1}{n^s} - \frac{1}{x^s}\right) dx + \frac{1}{s-1}.$$

□

4.3 Dirichlet L-functions

Definition 4.3.1

Let G be a finite abelian group. A **character** of G is a group homomorphism $\chi : G \to \mathbb{C}^\times$. The set of characters of G, denoted $\hat{G}$, is an abelian group called the **dual** of G. Its neutral element is called the **trivial character**.

Remark 4.3.2. If χ is a character of a finite abelian group G, then, for any $g \in G$, the complex number $\chi(g)$ is a root of unity. Indeed, if n is the order of g, then, $\chi(g)^n = \chi(g^n) = \chi(1) = 1$.

Proposition 4.3.3

Let G be a finite abelian group. Then G and $\hat{G}$ have the same order.

Proof. Take $g_1, \ldots, g_r$ some generators of G. We proceed by induction on r. For $r = 1$, the group G is isomorphic to a cyclic group $\mathbb{Z}/N\mathbb{Z}$. Characters of $\mathbb{Z}/N\mathbb{Z}$ can be described as:

$$\chi_b : a \mod N \mapsto e^{2i\pi ab/N}, \quad b \in \mathbb{Z}/N\mathbb{Z}.$$

Hence $\hat{G}$ has N elements.

Now assume that $r \geq 2$ and set $H := \langle g_1, \ldots, g_{r-1} \rangle$, so that G/H is a cyclic group generated by g_r. The inclusion $H \subset G$ induces a restriction morphism $\varphi : \hat{G} \to \hat{H}$ whose kernel is immediately seen to be $\widehat{G/H}$. Let us prove that φ is surjective. To do so, take $\chi \in \hat{H}$, let ρ be the order of g_r in G/H, let $z \in \mathbb{C}^\times$ be a ρ-th root of $\chi(\rho g_r)$ and define:

$$\begin{aligned} \tilde{\chi} : G &\to \mathbb{C}^\times \\ h + a g_r &\mapsto \chi(h) \cdot z^a \end{aligned}$$

for $h \in H$ and $a \in \mathbb{Z}$. One easily checks that $\tilde{\chi}$ is a well-defined group homomorphism such that $\tilde{\chi}|_H = \chi$. Hence φ is surjective and:

$$|\hat{G}| = |\ker\varphi| \cdot |\mathrm{im}\varphi| = |\widehat{G/H}| \cdot |\hat{H}|.$$

By induction, we have $|\widehat{G/H}| = |G/H|$ and $|\hat{H}| = |H|$, and hence:

$$|\hat{G}| = |G/H| \cdot |H| = |G|.$$

□

Corollary 4.3.4

Let G be a finite abelian group. The map:

$$\begin{aligned} G &\to \hat{\hat{G}} \\ g &\mapsto \big(\chi \mapsto \chi(g)\big) \end{aligned}$$

is an isomorphism.

Proof. The map is obviously an injective group homomorphism. Since G and $\hat{\hat{G}}$ have the same order, it is an isomorphism. □

Proposition 4.3.5

Let G be a finite abelian group.

(i) For any $\chi \in \hat{G}$, we have:

$$\sum_{g \in G} \chi(g) = \begin{cases} |G| & \text{if } \chi \text{ is trivial,} \\ 0 & \text{otherwise.} \end{cases}.$$

(ii) For any $g \in G$:

$$\sum_{\chi \in \hat{G}} \chi(g) = \begin{cases} |\hat{G}| & \text{if } g = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. (i) If χ is the trivial character:

$$\sum_{g \in G} \chi(g) = \sum_{g \in G} 1 = |G|.$$

Otherwise, there exists $g_0 \in G$ such that $\chi(g_0) \neq 1$. Then:

$$\sum_{g\in G} \chi(g) = \sum_{g\in G} \chi(g_0 g) = \chi(g_0) \sum_{g\in G} \chi(g).$$

Hence:

$$\sum_{g\in G} \chi(g) = 0.$$

(ii) By Proposition 4.3, it suffices to apply statement (i) to the finite abelian group $\hat{G}$.

□

Definition 4.3.6

Fix an integer $N \geq 1$. A **Dirichlet character modulo** N is a character $\chi_0 : (\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{C}^\times$. It can also be seen as a map $\chi : \mathbb{N}_{>0} \to \mathbb{C}^\times$ by setting:

$$\chi(a) := \begin{cases} \chi_0(a \mod N) & \text{if } a \wedge N = 1, \\ 0 & \text{otherwise.} \end{cases}$$

When χ_0 is the trivial character, the corresponding function χ is called the **principal** Dirichlet character modulo N and is denoted $\mathbf{1}$.

Seen as a function on $\mathbb{N}$, a Dirichlet character modulo N is a completely multiplicative function.

Definition 4.3.7

Let N be a positive integer and let χ be a Dirichlet character modulo N. The **Dirichlet** L**-function** associated to χ is defined as:

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

According to the results we settled in Section 4.1, the Dirichlet L-function $L(\chi, -)$ defines a holomorphic function on the half-plane $\mathrm{Re}(s) > 1$, and we have the equality:

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

on that half-plane. In particular, when $\chi = \mathbf{1}$ is the principal character, we get:

$$L(\mathbf{1}, s) = \prod_{\substack{p \text{ prime} \\ p \nmid N}} \frac{1}{1 - p^{-s}} = \zeta(s) \cdot \prod_{p \mid N} (1 - p^{-s}).$$

According to Theorem 4.2.2, this function can be extended to a meromorphic function on the half-plane $\mathrm{Re}(s) > 0$ whose unique pole is $s = 1$ and is simple. Let us now see what happens for non-principal characters.

Theorem 4.3.8

Let χ be a non-principal Dirichlet character modulo N. Then the Dirichlet series defining the Dirichlet L-function $L(\chi, -)$ converges uniformly on compact subsets of the half-plane $\mathrm{Re}(s) > 0$, and hence $L(\chi, -)$ is a holomorphic function on that half-plane.

Proof. The proof is based on the so-called Abel's summation formula. Let us briefly recall it. Take two sequences of complex numbers $(u_n)_{n \geq 1}$ and $(v_n)_{n \geq 1}$ and set $U_n := u_1 + \cdots + u_n$ for $n \geq 1$. Then for all $N \geq M \geq 1$, we have:

$$\begin{aligned} \sum_{n=M}^{N} u_n v_n &= \sum_{n=M}^{N} (U_n - U_{n-1}) v_n = \sum_{n=M}^{N} U_n v_n - \sum_{n=M}^{N} U_{n-1} v_n \\ &= \sum_{n=M}^{N} U_n v_n - \sum_{n=M-1}^{N-1} U_n v_{n+1} \\ &= U_N v_N - U_{M-1} v_M + \sum_{n=M}^{N-1} U_n (v_n - v_{n+1}). \end{aligned}$$

Coming back to the proof of Theorem 4.3.8, let H_0 be the half-plane $\mathrm{Re}(s) > 0$. For $s \in H_0$, we can write:

$$f(s) = \sum_{n \geq 1} u_n v_n(s)$$

where $u_n = \chi(n)$ and $v_n(s) = \frac{1}{n^s}$. By setting $U_n := u_1 + \cdots + u_n$ for $n \geq 1$, Abel's summation formula, tells us that, for $N \geq M \geq 1$:

$$\sum_{n=M}^{N} u_n v_n(s) = U_N v_N(s) - U_{M-1} v_M(s) + \sum_{n=M}^{N-1} U_n (v_n(s) - v_{n+1}(s)).$$

Now observe that the sequence $(U_n)_{n \geq 1}$ is bounded by some constant C. We

deduce that:

$$\begin{aligned}\left|\sum_{n=M}^{N} u_n v_n(s)\right| &\leq |U_N v_N(s)| + |U_{M-1} v_M(s)| + \sum_{n=M}^{N-1} |U_n| \cdot |v_n(s) - v_{n+1}(s))| \\ &\leq C \cdot \left(\left|\frac{1}{N^s}\right| + \left|\frac{1}{M^s}\right| + \sum_{n=M}^{N-1} \left|\frac{1}{n^s} - \frac{1}{(n+1)^s}\right| \right) \\ &\leq C \cdot \left(\frac{1}{N^{\mathrm{Re}(s)}} + \frac{1}{M^{\mathrm{Re}(s)}} + \sum_{n=M}^{N-1} \sup_{n\leq x\leq n+1} \left|\frac{s}{x^{s+1}}\right| \right) \\ &\leq C \cdot \left(\frac{1}{N^{\mathrm{Re}(s)}} + \frac{1}{M^{\mathrm{Re}(s)}} + |s| \sum_{n=M}^{N-1} \frac{1}{n^{\mathrm{Re}(s)+1}} \right).\end{aligned}$$

Note that $\mathrm{Re}(s) > 0$ and $\mathrm{Re}(s+1) > 1$, so that the sequence $\left(\frac{1}{n^{\mathrm{Re}(s)}}\right)_{n\geq 1}$ converges uniformly on compact subsets of H_0 to 0 and the series $\sum \frac{1}{n^{\mathrm{Re}(s+1)}}$ converges uniformly on compact subsets of H_0. We deduce that the Dirichlet L-function $L(\chi, s) = \sum_{n\geq 1} u_n v_n(s)$ converges uniformly on compact subsets of H_0. Since each term of the series defining $L(\chi, -)$ is holomorphic on H_0, we deduce that $L(\chi, -)$ is holomorphic on H_0. □

4.4 Behaviour of L-functions at $s = 1$

Theorem 4.4.1

Let χ be a non-principal Dirichlet character modulo N. Then $L(\chi, 1) \neq 0$.

In order to prove this result, we introduce the function:

$$\zeta_N = \prod_{\chi} L(\chi, -),$$

where χ describes all Dirichlet characters modulo N. It is a meromorphic function on the half-plane $\{\mathrm{Re}(s) > 0\}$, and it has no poles outside $s = 1$. At $s = 1$, the L-function $L(\mathbf{1}, -)$ has a simple pole while for each non-principal character χ the L-function $L(\chi, -)$ has no pole. Hence two possibilities arise:

- either $L(\chi, -)$ vanishes at $s = 1$ for some non-principal Dirichlet character χ, in which case ζ_N is has no pole at $s = 1$,
- or $L(\chi, 1) \neq 0$ for every non-principal Dirichlet character χ, in which case ζ_N has a pole of order 1 at $s = 1$.

Therefore, in order to prove Theorem 4.4.1, it suffices to check that ζ_N has no pole at $s = 1$. For that purpose, we proceed by contradiction, assuming that ζ_N does have a pole at $s = 1$, and we start by writing the function ζ_N as a Dirichlet series:

Lemma 4.4.2. *Let H_1 be the half-plane $\mathrm{Re}(s) > 1$.*

1. *For each prime number p, there exist integers $\alpha(p), \beta(p) \leq \varphi(N)$ such that, for $s \in H_1$:*
$$\zeta_N(s) = \prod_{\substack{p \text{ prime} \\ p \nmid N}} \left(\frac{1}{1 - p^{-\alpha(p)s}} \right)^{\beta(p)}.$$

2. *On the half-plane H_1, the function ζ_N can be written as a Dirichlet series:*
$$\zeta_N(s) = \sum_{n \geq 1} \frac{f(n)}{n^s}$$
for some multiplicative function $f : \mathbb{N}_{>0} \to \mathbb{N}_{\geq 0}$.

Proof. 1. For $s \in H_1$:
$$\zeta_N(s) = \prod_{\chi} L(\chi, s) = \prod_{\chi} \prod_{\substack{p \text{ prime} \\ p \nmid N}} \frac{1}{1 - \chi(p)p^{-s}} = \prod_{\substack{p \text{ prime} \\ p \nmid N}} \prod_{\chi} \frac{1}{1 - \chi(p)p^{-s}}.$$
For each prime number p with $p \nmid N$, denote by $\alpha(p)$ the order of p in $(\mathbb{Z}/N\mathbb{Z})^\times$ and make the following observations:

- for any Dirichlet character modulo N, the number $\chi(p)$ is an $\alpha(p)$-th root unity;
- for any $\alpha(p)$-th root of unity w, there are $\beta(p) := \frac{\varphi(N)}{\alpha(p)}$ Dirichlet characters χ modulo N such that $\chi(p) = w$.

We deduce that, for $s \in H_1$:
$$\zeta_N(s) = \prod_{\substack{p \text{ prime} \\ p \nmid N}} \left(\prod_{w^{\alpha(p)}=1} \frac{1}{1 - wp^{-s}} \right)^{\beta(p)} = \prod_{\substack{p \text{ prime} \\ p \nmid N}} \left(\frac{1}{1 - p^{-\alpha(p)s}} \right)^{\beta(p)}.$$

2. For each prime number p with $p \nmid N$ and each $s \in H_1$:
$$\frac{1}{1 - p^{-\alpha(p)s}} = \sum_{n \geq 0} p^{-n\alpha(p)s}.$$

Hence there exist non-negative integers $a_n^{(p)}$ for $n \geq 0$ such that the equality:

$$\left(\frac{1}{1-p^{-\alpha(p)s}}\right)^{\beta(p)} = \sum_{n\geq 0} a_n^{(p)} p^{-ns}$$

holds on the half-plane H_1. We deduce the formula:

$$\zeta_N(s) = \prod_{\substack{p \text{ prime} \\ p \nmid N}} \left(\sum_{n\geq 0} a_n^{(p)} p^{-ns}\right),$$

in which each factor converges absolutely.

We now want to expand this product by proceeding similarly to Theorem 4.1.4. For that pupose, let $f : \mathbb{N}_{>0} \to \mathbb{N}_{\geq 0}$ be the unique multiplicative function such that:

$$f(p^n) = \begin{cases} a_n^{(p)} & \text{for each prime number } p \nmid N \text{ and each } n \geq 0, \\ 0 & \text{for each prime number } p \mid N \text{ and each } n \geq 0, \end{cases}$$

fix a positive integer M and let $p_1, \ldots, p_m$ be the prime numbers smaller or equal to M. Since all the series $\sum a_n^{(p)} p^{-ns}$ are absolutely convergent, we can compute:

$$\prod_{p \leq M \text{ prime}} \left(\sum_{n\geq 0} a_n^{(p)} p^{-ns}\right) = \sum_{r_1,\ldots,r_m \geq 0} \frac{f(p_1^{r_1} \ldots p_m^{r_m})}{(p_1^{r_1} \ldots p_m^{r_m})^s} = \sum_{n \in X_M} \frac{f(n)}{n^s} \tag{4.1}$$

where X_M is the set of positive integers whose prime factors are all $\leq M$. When s is real, all terms in the series $\sum \frac{f(n)}{n^s}$ are non-negative, and since $\mathbb{N} = \bigcup_{M \in \mathbb{N}} X_M$, equation (4.1) implies that:

$$\zeta_N(s) = \sum_{n\geq 1} \frac{f(n)}{n^s} < +\infty.$$

We deduce that, when s is any complex number in H_1, the series $\sum \frac{f(n)}{n^s}$ is absolutely convergent, and hence by using equation (4.1) once again:

$$\zeta_N(s) = \sum_{n\geq 1} \frac{f(n)}{n^s}.$$

□

Lemma 4.4.3. *Consider a Dirichlet series $L(s) = \sum_{n\geq 1} \frac{a_n}{n^s}$ with $a_n \geq 0$ for all n. Assume that the series converges on each half-plane of the form $\mathrm{Re}(s) > s_0$ with $s_0 > 1$, and assume that it extends to a holomorphic function on the half-plane $\mathrm{Re}(s) > 0$. Then the Dirichlet series converges absolutely on the latter.*

Remark 4.4.4. Observe that, with the assumptions of the lemma, if the Dirichlet series defining L converges at some $s_0 > 0$, then it converges normally on the half-plane $\mathrm{Re}(s) > s_0$ (since the inequality $\left|\frac{a_n}{n^s}\right| \leq \frac{a_n}{n^{s_0}}$ holds on that half-plane). This simple observation will be used in the proof.

Proof. Let D be the open disk of center 2 and radius 2. By assumption, L defines a holomorphic function on the half-plane $\mathrm{Re}(s) > 0$. We deduce that its Taylor series at $s = 2$ converges on D. By remark 4.4.4, the Dirichlet series defining L converges normally on the half-plane $\mathrm{Re}(s) > 3/2$, and hence we can compute the successive derivatives of L at $s = 2$ term by term:

$$L^{(r)}(2) = \sum_{n\geq 1} (-1)^r \frac{a_n (\log n)^r}{n^2}.$$

We deduce that, for $s \in D$:

$$L(s) = \sum_{r\geq 0} \frac{1}{r!} \left(\sum_{n\geq 1} (-1)^r \frac{a_n (\log n)^r}{n^2} \right) (s-2)^r.$$

In particular, for $s \in (0,1)$:

$$L(s) = \sum_{r\geq 0} \sum_{n\geq 1} \frac{1}{r!} \frac{a_n (\log n)^r}{n^2} (2-s)^r.$$

This is a double sum in which all terms are non-negative, so we may compute:

$$\begin{aligned} L(s) &= \sum_{r\geq 0} \sum_{n\geq 1} \frac{1}{r!} \frac{a_n (\log n)^r}{n^2} (2-s)^r = \sum_{n\geq 1} \sum_{r\geq 0} \frac{1}{r!} \frac{a_n (\log n)^r}{n^2} (2-s)^r \\ &= \sum_{n\geq 1} a_n \sum_{r\geq 0} \frac{1}{r!} \frac{(\log n)^r}{n^2} (2-s)^r = \sum_{n\geq 1} \frac{a_n}{n^2} \exp((2-s)\log n) \\ &= \sum_{n\geq 1} \frac{a_n}{n^2} n^{2-s} = \sum_{n\geq 1} \frac{a_n}{n^s}. \end{aligned}$$

This shows that the Dirichlet series defining L converges at every $s \in (0,1)$, and hence, by remark 4.4.4, it converges normally on each half-plane of the form $\mathrm{Re}(s) > s_0$ with $s_0 > 0$. □

From the previous lemmas, we deduce that the equality:

$$\zeta_N(s) = \prod_{\substack{p \text{ prime} \\ p \nmid N}} \left(\frac{1}{1 - p^{-\alpha(p)s}} \right)^{\beta(p)}$$

holds for every s with $\mathrm{Re}(s) > 0$. However, for $s > 0$:

$$\prod_{\substack{p \text{ prime} \\ p \nmid N}} \left(\frac{1}{1-p^{-\alpha(p)s}}\right)^{\beta(p)} \geq \prod_{\substack{p \text{ prime} \\ p \nmid N}} \left(\frac{1}{1-p^{-\varphi(N)s}}\right)$$
$$\geq \zeta(\varphi(N)s) \cdot \prod_{\substack{p \text{ prime} \\ p \mid N}} \left(1-p^{-\varphi(N)s}\right) \xrightarrow[s \to \left(\frac{1}{\varphi(N)}\right)^+]{} +\infty,$$

since the Riemann zeta function has a pole at $s = 1$.

4.5 Dirichlet's Theorem on arithmetic progressions

Theorem 4.5.1 (Dirichlet)

Consider two positive coprime integers a and b. There exist infinitely many prime numbers of the form $an + b$ with $n \in \mathbb{N}$.

Proof. Take $s > 1$. Recall that according to Theorem 4.1.4, we have:

$$L(\chi, s) = \prod_{p \text{ prime}} L_p(\chi, s),$$

with

$$L_p(\chi, s) = \frac{1}{1-\chi(p)p^{-s}}.$$

Now for each prime p, we have $\left|\chi(p)p^{-s}\right| \leq \frac{1}{2}$. Hence:

$$\frac{L_p'(\chi, s)}{L_p(\chi, s)} = \frac{\chi(p) \cdot \log p \cdot p^{-s}}{1-\chi(p)p^{-s}} = \log p \cdot \sum_{n \geq 1} \frac{\chi(p)^n}{p^{ns}} = \log p \cdot \left(\frac{\chi(p)}{p^s} + \sum_{n \geq 2} \frac{\chi(p)^n}{p^{ns}}\right),$$

with:

$$\left|\sum_{n \geq 2} \frac{\chi(p)^n}{p^{ns}}\right| \leq \frac{1}{p^{2s}} \sum_{n \geq 2} \frac{1}{2^{n-2}} \leq \frac{2}{p^2}.$$

We deduce that:

$$\frac{L'(\chi, s)}{L(\chi, s)} = \sum_{p \text{ prime}} \frac{L_p'(\chi, s)}{L_p(\chi, s)} = \sum_{p \text{ prime}} \log p \cdot \frac{\chi(p)}{p^s} + O(1).$$

Multiplying by $\overline{\chi(b)}$ and summing up over all Dirichlet characters modulo a:

$$\begin{aligned}\sum_{\chi}\overline{\chi(b)}\frac{L_p'(\chi,s)}{L_p(\chi,s)} &= \sum_{\chi}\sum_{p\text{ prime}}\frac{\overline{\chi(b)}\cdot\log p\cdot\chi(p)}{p^s}+O(1)\\ &= \sum_{p\text{ prime}}\frac{\log p}{p^s}\sum_{\chi}\overline{\chi(b)}\chi(p)+O(1)\\ &= \sum_{p\text{ prime}}\frac{\log p}{p^s}\sum_{\chi}\chi(b^{-1}p)+O(1)\\ &= \varphi(a)\sum_{\substack{p\text{ prime}\\ p\equiv b \bmod a}}\frac{\log p}{p^s}+O(1),\end{aligned}$$

where b^{-1} stands for an inverse of b modulo a and the last line follows from lemma 4.3.5. Now observe that, according to Theorems 4.2.2, 4.3.8 and 4.4.1, the L-function $L(\mathbf{1},-)$ has a pole at $s=1$, while the L-functions $L(\chi,-)$ with $\chi\neq\mathbb{1}$ have neither a zero nor a pole at $s=1$. We deduce that the function $\sum_{\chi}\overline{\chi(b)}\log L(\chi,-)$ is unbounded as $s\to 1^+$. Hence so is the function:

$$s\mapsto\sum_{\substack{p\text{ prime}\\ p\equiv b \bmod a}}\frac{\log p}{p^s}.$$

We deduce that there are infinitely many prime numbers congruent to b modulo a. □

Chapter 5

Quadratic forms over $\mathbb{Q}$

5.1 Short reminders on quadratic forms

In this chapter, a **quadratic form** over a field K will be a polynomial of the form:

$$q(X_1, \ldots, X_n) = \sum_{1 \leq i \leq j \leq n} a_{ij} X_i X_j.$$

When K has characteristic different from 2, the matrix of q is the matrix $M_q := (b_{ij})_{i,j}$ with $b_{ii} = a_{ii}$ for each i and $b_{ij} = a_{ij}/2$ for $i \neq j$. We then have:

$$q(X_1, \ldots, X_n) = (X_1, \ldots, X_n) \cdot M_q \cdot \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}.$$

Two quadratic forms q and q' in n variables over K are said to be **isometric** if there exists a matrix $A \in \mathrm{GL}_n(K)$ such that:

$$q'(X_1, \ldots, X_n) = q((X_1, \ldots, X_n) \cdot A).$$

This equality can be rewritten in matrix terms as:

$$M_{q'} = A M_q {}^t A.$$

In particular, the **discriminant** of a quadratic form q, that is the square class of determinant of the matrix M_q, is invariant under isometry.

As the following proposition shows, up to isometry, quadratic forms can be written in a very simple way:

Proposition 5.1.1

Let K be a field with characteristic different from 2. Then any quadratic form q in n variables over K is isometric to a quadratic form of the form:

$$q'(X_1,\dots,X_n) = a_1X_1^2 + \cdots + a_nX_n^2$$

with $a_1,\dots,a_n \in K$.

Remark 5.1.2. To simplify notations, the quadratic form $a_1X_1^2 + \cdots + a_nX_n^2$ is usually denoted $\langle a_1,\dots,a_n\rangle$.

Proof. We proceed by induction on n. For $n = 1$, there is nothing to prove. Assume now that $n \geq 2$ and write:

$$q(X_1,\dots,X_n) = \sum_{1\leq i\leq j\leq n} a_{ij}X_iX_j.$$

If all the a_{ij}'s are zero, there is nothing to prove. We henceforth assume that there exist i_0, j_0 such that $a_{i_0j_0} \neq 0$. The following two cases arise:

- *First case:* $a_{i_0i_0} \neq 0$ or $a_{j_0j_0} \neq 0$. In this case, up to reordering the variables, we may assume that $a_{11} \neq 0$.
- *Second case:* $a_{i_0i_0} = a_{j_0j_0} = 0$. In this case, we may perform the variable change:
$$X_1 \mapsto X_{i_0} + X_{j_0}, X_{j_0} \mapsto X_{i_0} - X_{j_0}$$
to get a quadratic form q_1 that is isometric to q and whose coefficient in front of $X_{i_0}^2$ is $a_{i_0j_0} \neq 0$. Hence, as before, up to replacing q by q_1 and up to reordering the variables, we may assume that $a_{11} \neq 0$.

We henceforth assume that $a_{11} \neq 0$. We may then write q as:

$$q(X_1,\dots,X_n) = a_{11}\left(X_1 + \frac{a_{12}}{2a_{11}}X_2 + \cdots + \frac{a_{1n}}{2a_{11}}X_n\right)^2 + q_2(X_2,\dots,X_n)$$

for some quadratic form q_2 in $n-1$ variables. By the inductive assumption, q_2 is isometric to $\langle a_2,\dots,a_n\rangle$ for some $a_2,\dots,a_n \in K$. Hence q is isometric to $\langle a_{11}, a_2,\dots,a_n\rangle$. □

A quadratic form q is said to be **isotropic** if there exist $x_1,\dots,x_n \in K$ not all zero such that:

$$q(x_1,\dots,x_n) = 0.$$

Otherwise q is said to be **anisotropic**.

Proposition 5.1.3

Let K be a field with characteristic different from 2 and consider the quadratic form $q = \langle a_1, \dots, a_n \rangle$ for some $a_1, \dots, a_n \in K^\times$. If q is isotropic, then q induces a surjective map $K^n \to K$.

Proof. Let q be an isotropic quadratic form over K in n variables. Choose some $\mathbf{x} := (x_1, \dots, x_n) \in K^n \setminus \{0\}$ such that $q(\mathbf{x}) = 0$. Now, for each $\lambda_1, \dots, \lambda_n \in K$, we have:

$$\begin{aligned} q(x_1 + \lambda_1 T, \dots, x_n + \lambda_n T) &= q(\mathbf{x}) + 2(a_1\lambda_1 x_1 + \cdots + a_n\lambda_n x_n)T + (a_1\lambda_1^2 + \cdots + a_n\lambda_n^2) \\ &= 2(a_1\lambda_1 x_1 + \cdots + a_n\lambda_n x_n)T + (a_1\lambda_1^2 + \cdots + a_n\lambda_n^2). \end{aligned}$$

Since the a_i's all are all non-zero and at least one of the x_i's is non-zero, we may choose $\lambda_1, \dots, \lambda \in K$ so that:

$$a_1\lambda_1 x_1 + \cdots + a_n\lambda_n x_n \neq 0.$$

The map:

$$\begin{aligned} K &\mapsto K \\ t &\mapsto 2(a_1\lambda_1 x_1 + \cdots + a_n\lambda_n x_n)t + (a_1\lambda_1^2 + \cdots + a_n\lambda_n^2) \end{aligned}$$

is then surjective, and hence so is the map $K^n \to K$ induced by q. □

5.2 Quadratic forms over finite fields

In this section, we aim at studying the isotropy of quadratic forms over finite fields. We start by focusing on quadratic forms in two variables over $\mathbb{F}_p$ for some prime number p. The goal then consists in understanding which integers are squares modulo p. That is why we introduce the Legendre symbol:

Definition 5.2.1

Let p be a prime number and a an integer. We define the **Legendre symbol**:

$$\left(\frac{a}{p}\right) = \begin{cases} 1 & \text{if } p \nmid a \text{ and the class of } a \text{ modulo } p \text{ is a square} \\ -1 & \text{if } p \nmid a \text{ and the class of } a \text{ modulo } p \text{ is not a square} \\ 0 & \text{if } p \mid a. \end{cases}$$

Adrien-Marie Legendre (Paris 1752 - Paris 1833)

Example 5.2.2. We have $\left(\frac{2}{7}\right) = 1$ and $\left(\frac{3}{7}\right) = -1$.

Proposition 5.2.3 (Euler's criterion)

Let p be an odd prime number and a an integer not divisible by p. Then:

$$\left(\frac{a}{p}\right) \equiv a^{\frac{p-1}{2}} \mod p.$$

In particular:

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

and the map $\left(\frac{\cdot}{p}\right) : \mathbb{F}_p^\times \to \{-1, 1\}$ is a group homomorphism.

Proof. Consider the group homomorphisms:

$$\begin{aligned} f : \mathbb{F}_p^\times &\to \mathbb{F}_p^\times \\ x &\mapsto x^2 \end{aligned} \qquad \begin{aligned} g : \mathbb{F}_p^\times &\to \mathbb{F}_p^\times \\ x &\mapsto x^{\frac{p-1}{2}}. \end{aligned}$$

The kernel of f contains $\{\pm 1\}$ and has at most 2 elements, hence $\ker(f) = \{\pm 1\}$. We deduce that $\mathrm{im}(f)$ is a subgroup of $\mathbb{F}_p^\times$ with order $\frac{p-1}{2}$. But $\mathrm{im}(f)$ is contained in $\ker(g)$. Since $\ker(g)$ has at most $\frac{p-1}{2}$ elements, we conclude that $\mathrm{im}(f) =$

$\ker(g)$. In other words, an element $\overline{a} \in \mathbb{F}_p^\times$ is a square if, and only if, $\overline{a}^{\frac{p-1}{2}} = 1$. This is exactly what we wanted to prove since $\overline{a}^{\frac{p-1}{2}} \in \{\pm 1\}$ for all $\overline{a} \in \mathbb{F}_p^\times$ □

Proposition 5.2.3 implies in particular that there are exactly $\frac{p-1}{2}$ squares in $\mathbb{F}_p^\times$, and hence that:

$$\sum_{c \in \mathbb{F}_p^\times} \left(\frac{c}{p}\right) = 0.$$

Leonhard Euler (Bâle 1707 - Saint-Petersburg 1783)

Theorem 5.2.4 (Quadratic Reciprocity Law)

(i) For any odd prime number:

$$\left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}}.$$

(ii) Given two distinct odd prime numbers p and ℓ:

$$\left(\frac{p}{\ell}\right)\left(\frac{\ell}{p}\right) = (-1)^{\frac{p-1}{2}\cdot\frac{\ell-1}{2}}.$$

Proof.

(i) We want to prove that $\left(\frac{2}{p}\right) = 1$ if, and only if, $p \equiv \pm 1 \mod 8$. To do so, choose a primitive 8-th root of unity α in $\overline{\mathbb{F}_p}$. Since $\alpha^4 + 1 = 0$, we have:

$$(\alpha + \alpha^{-1})^2 = 2.$$

Hence the square roots of 2 in $\overline{\mathbb{F}_p}$ are $\alpha + \alpha^{-1}$ and $-(\alpha + \alpha^{-1})$. In particular, $\left(\frac{2}{p}\right) = 1$ if, and only if, $\alpha + \alpha^{-1} \in \mathbb{F}_p$. Since $\mathbb{F}_p$ is the subfield of $\overline{\mathbb{F}_p}$ fixed by the Frobenius morphism, that is also equivalent to the equality $(\alpha + \alpha^{-1})^p = \alpha + \alpha^{-1}$.

If $p \equiv \pm 1 \mod 8$, then:

$$(\alpha + \alpha^{-1})^p = \alpha^p + \alpha^{-p} = \alpha^{\pm 1} + \alpha^{\mp 1} = \alpha + \alpha^{-1}$$

and hence $\left(\frac{2}{p}\right) = 1$. If $p \equiv \pm 3 \mod 8$, then:

$$(\alpha + \alpha^{-1})^p = \alpha^p + \alpha^{-p} = \alpha^3 + \alpha^{-3} = -(\alpha + \alpha^{-1})$$

and hence $\left(\frac{2}{p}\right) = -1$.

(ii) Let $\xi \in \overline{\mathbb{F}_p}$ be a primitive ℓ-th root of unity. Consider the Gauss sum:

$$S := \sum_{x \in \mathbb{F}_\ell^\times} \left(\frac{x}{\ell}\right) \xi^x.$$

We compute:

$$\begin{aligned} S^2 &= \sum_{x \in \mathbb{F}_\ell^\times} \sum_{y \in \mathbb{F}_\ell^\times} \left(\frac{xy}{\ell}\right) \xi^{x+y} \\ &= \sum_{c \in \mathbb{F}_\ell^\times} \sum_{x \in \mathbb{F}_\ell^\times} \left(\frac{cx^2}{\ell}\right) \xi^{(c+1)x} \\ &= \sum_{c \in \mathbb{F}_\ell^\times} \left(\frac{c}{\ell}\right) \sum_{x \in \mathbb{F}_\ell^\times} \xi^{(c+1)x} \\ &= (\ell - 1)\left(\frac{-1}{\ell}\right) - \sum_{c \in \mathbb{F}_\ell^\times \setminus \{-1\}} \left(\frac{c}{\ell}\right) = \ell(-1)^{\frac{\ell-1}{2}}, \end{aligned}$$

where the last equality follows from $\sum_{c \in \mathbb{F}_\ell^\times} \left(\frac{c}{\ell}\right) = 0$. Hence the square roots of $(-1)^{\frac{\ell-1}{2}} \ell$ in $\overline{\mathbb{F}_p}$ are $\pm S$.

Now, since $\mathbb{F}_p$ is the subfield of $\overline{\mathbb{F}_p}$ fixed by the Frobenius morphism, S belongs to $\mathbb{F}_p$ if, and only if, $S^p = S$. But:

$$S^p = \sum_{x \in \mathbb{F}_\ell^\times} \left(\frac{x}{\ell}\right) \xi^{xp} = \sum_{y \in \mathbb{F}_\ell^\times} \left(\frac{p}{\ell}\right)^{-1} \left(\frac{y}{\ell}\right) \xi^y = \left(\frac{p}{\ell}\right) S.$$

Hence $(-1)^{\frac{\ell-1}{2}}\ell$ is a square in $\mathbb{F}_p$ if, and only if, $\left(\frac{p}{\ell}\right) = 1$. In other words, $\left(\frac{(-1)^{\frac{\ell-1}{2}}\ell}{p}\right) = \left(\frac{p}{\ell}\right)$, and hence, by using Proposition 5.2.3:

$$\left(\frac{p}{\ell}\right)\left(\frac{\ell}{p}\right) = \left(\frac{(-1)^{\frac{\ell-1}{2}}}{p}\right) = (-1)^{\frac{\ell-1}{2}\frac{p-1}{2}}.$$

□

The quadratic reciprocity law allows to determine in a practical way if some integer is a square or not modulo some prime number p. For instance, in order to determine whether 186 is a square modulo 2027, we compute:

$$\begin{aligned}\left(\frac{186}{2027}\right) &= \left(\frac{2}{2027}\right)\left(\frac{3}{2027}\right)\left(\frac{31}{2027}\right) = (-1)\cdot(-1)\left(\frac{2027}{3}\right)\cdot(-1)\left(\frac{2027}{31}\right) \\ &= -\left(\frac{2}{3}\right)\left(\frac{12}{31}\right) = -(-1)\cdot\left(\frac{3}{31}\right) = -\left(\frac{31}{3}\right) = -\left(\frac{1}{3}\right) = -1.\end{aligned}$$

Hence 186 is not a square modulo 2027.

Now that we get a quite good understanding of quuadratic forms in two variables over the field $\mathbb{F}_p$ for p a prime number, let us move on to quadratic forms in at least 3 variables. In order to study them, we prove the following general result about equations in a large number of variables and of low degree over finite fields:

Theorem 5.2.5 (Chevalley-Warning)

Let p be a prime number and let q be a power of p. Let $\mathbb{F}_q$ be a finite field with q elements. Let $f \in \mathbb{F}_q[X_0, ..., X_n]$ be a homogeneous polynomial of degree $d \leq n$. Then the equation $f = 0$ has non-zero solutions in $\mathbb{F}_q^{n+1}$.

Claude Chevalley (Johannesburg 1909 - Paris 1884)
No picture seems to be available for Ewald Warning, a student of Emil Artin

Before proving Theorem 5.2.5, we need the settle the followinf easy lemma:

Lemma 5.2.6. *Let $i \in \mathbb{Z}$. Then:*

$$S_i := \sum_{x \in \mathbb{F}_q} x^i = \begin{cases} 0 \text{ if } i = 0 \text{ or } q-1 \nmid i, \\ -1 \text{ otherwise.} \end{cases}$$

Proof. We distinguish three cases.

- If $i = 0$, then $S_0 = \sum_{x \in \mathbb{F}_q} 1 = 0$.
- If $i \neq 0$ and $q-1 \mid i$, Lagrange's Theorem shows that:

$$S_i = \sum_{x \in \mathbb{F}_q} x^i = 0 + \sum_{x \in \mathbb{F}_q^\times} 1 = -1.$$

- For $i \neq 0$ and $q-1 \nmid i$, there exists y with $y^i \neq 1$. Moreover:

$$y^i S_i = \sum_{x \in \mathbb{F}_q} (xy)^i = \sum_{x \in \mathbb{F}_q} x^i = S_i.$$

  Hence $S_i = 0$.

□

Proof of Theorem 5.2.5. Set

$$N := \sharp\{(x_0, ..., x_n) \in \mathbb{F}_q^{n+1} | f(x_0, ..., x_n) = 0\},$$

and consider the polynomial $F = 1 - f^{q-1}$. By Lagrange's Theorem, we have:

$$\forall \mathbf{x} \in \mathbb{F}_q^{n+1}, F(\mathbf{x}) = \begin{cases} 0 \text{ if } f(\mathbf{x}) \neq 0 \\ 1 \text{ if } f(\mathbf{x}) = 0. \end{cases}$$

Hence the following equality holds in $\mathbb{F}_q$:

$$N = \sum_{\mathbf{x} \in \mathbb{F}_q^{n+1}} F(\mathbf{x}).$$

Writing:

$$F(\mathbf{x}) = \sum_{\mathbf{i} \in \mathbb{N}_{\geq 0}^{n+1}} a_{\mathbf{i}} \mathbf{x}^{\mathbf{i}} = \sum_{i_0 + \cdots + i_n \leq d(q-1)} a_{\mathbf{i}} \mathbf{x}^{\mathbf{i}},$$

applying lemma 5.2.6 and using the convention $0^0 = 1$, we get:

$$N = \sum_{\mathbf{x} \in \mathbb{F}_q^{n+1}} F(\mathbf{x}) = \sum_{\mathbf{x} \in \mathbb{F}_q^{n+1}} \sum_{i_0 + \cdots + i_n \leq d(q-1)} a_{\mathbf{i}} \mathbf{x}^{\mathbf{i}} = \sum_{i_0 + \cdots + i_n \leq d(q-1)} a_{\mathbf{i}} \underbrace{\sum_{\mathbf{x} \in \mathbb{F}_q^{n+1}} \mathbf{x}^{\mathbf{i}}}_{\substack{= \left(\sum_{x \in \mathbb{F}_q} x^{i_0}\right) \cdots \left(\sum_{x \in \mathbb{F}_q} x^{i_n}\right) \\ = 0 \text{ provided at least one of} \\ \text{the } i_k\text{'s is} \leq q-1}}.$$

But since $d \leq n$, for every $(i_0, \ldots, i_n) \in \mathbb{N}_{\geq 0}^{n+1}$ with $i_0 + \cdots + i_n \leq d(q-1)$, at least one of the i_k's is $\leq q-1$. Hence $N = 0$ in $\mathbb{F}_q$. In other words, p divides N, and N has to be at least 2. In particular, the equation $f = 0$ has a non-zero solution. □

As a corollary, we get:

Corollary 5.2.7

Any quadratic form in at least 3 variables over a finite field is isotropic.

5.3 Quadratic forms over p-adic fields

In this section, we aim at studying the isotropy of quadratic forms over p-adic. We start by focusing on quadratic forms in two variables over $\mathbb{Q}_p$ for some prime number p. The goal then consists in understanding which elements in $\mathbb{Q}_p^\times$ are squares. Our main tool to do so is Hensel's lemma.

Proposition 5.3.1

Let p be a prime number and let $x \in \mathbb{Q}_p^\times$. Write $x = up^\alpha$ with $u \in \mathbb{Z}_p^\times$ and $\alpha \in \mathbb{Z}$.

(i) If p is odd, then x is a square in $\mathbb{Q}_p$ if, and only if, α is even and the reduction $[u]$ of u modulo p is a square in $\mathbb{F}_p$.

(ii) If $p = 2$, then x is a square in $\mathbb{Q}_2$ if, and only if, α is even and $u \equiv 1 \mod 8$.

Proof. In both cases, it is obvious that, if x is a square in $\mathbb{Q}_p$, then $\alpha = v_p(x)$ is even and u is a square in $\mathbb{Q}_p$. We are therefore reduced to studying in each case when u is a square in $\mathbb{Q}_p$.

(i) We assume p is odd. Of course, if u is a square in $\mathbb{Z}_p^\times$, by reducing modulo p, we see that $[u]$ is a square in $\mathbb{F}_p$. Conversely, if $[u]$ is a square in $\mathbb{F}_p$, Hensel's lemma (Theorem 3.5.3) applied to the equation $x^2 = u$ shows that u is a square in $\mathbb{Q}_p$.

(ii) We assume that $p = 2$. Of course, if u is a square in $\mathbb{Z}_2^\times$, by reducing modulo 8, we see that the reduction of u modulo 8 is a square in $\mathbb{Z}/8\mathbb{Z}$. The squares in $\mathbb{Z}/8\mathbb{Z}$ being 0, 1 and 4, we deduce that $u \equiv 1 \mod 8$. Conversely, if $u \equiv 1 \mod 8$, then Hensel's lemma (Theorem 3.5.3) applied to the equation $x^2 = u$ shows that u is a square in $\mathbb{Q}_2$.

□

Remark 5.3.2. Proposition 5.3.1 implies that, for each prime number p, the set of squares in $\mathbb{Q}_p^\times$ is open in $\mathbb{Q}_p^\times$.

Now that we get a good understanding of quuadratic forms in two variables over the field $\mathbb{Q}_p$ for p a prime number, let us move on to quadratic forms in 3 variables. In order to study them, we introduce the so-called Hilbert symbol, that can be defined over any field with characteristic different from 2:

Definition 5.3.3

Let K be a field of characteristic other than 2 and let $a, b \in K^\times$. We define the **Hilbert symbol** of a and b as:

$$(a,b)_K := \begin{cases} 1 & \text{if the quadratic form } \langle 1, -a, -b\rangle \text{ is isotropic} \\ -1 & \text{otherwise.} \end{cases}$$

Example 5.3.4.

(i) One easily checks that:

$$\forall a,b,c,d \in K^\times, (ac^2, bd^2)_K = (a,b)_K,$$
$$\forall a \in K^\times, (a,-a)_K = 1,$$
$$\forall a \in K \setminus \{0,1\}, (a,1-a)_K = 1.$$

(ii) For any $a,b \in \mathbb{C}^\times$, we have $(a,b)_{\mathbb{C}} = 1$.

(iii) By Corollary 5.2, if K is a finite field, then $(a,b)_K = 1$ for any $a,b \in K^\times$.

(iv) For $a,b \in \mathbb{R}^\times$, we have $(a,b)_{\mathbb{R}} = -1$ if, and only if, $a < 0$ and $b < 0$.

David Hilbert (Königsberg 1862 - Göttingen 1943)

The next proposition provides a useful characterization of Hilbert symbols:

Proposition 5.3.5

Let K be a field of characteristic other than 2 and let $a,b \in K^\times$. The following conditions are equivalent:

(i) $(a,b)_K = 1$;

(ii) a belongs to the image of $N_{K(\sqrt{b})/K} : K(\sqrt{b})^\times \to K^\times$.

Proof. If b is a square in K, then the norm map $N_{K(\sqrt{b})/K}$ is surjective and $\left(\sqrt{b}, 0, 1\right)$ is a solution of $x^2 - ay^2 - bz^2 = 0$: hence (i) and (ii) both hold. We henceforth assume that b is not a square in K.

If assertion (ii) holds, then we can find $s, t \in K$ such that:

$$a = N_{K(\sqrt{b})/K}(s + t\sqrt{b}) = s^2 - bt^2,$$

and hence $(s, 1, t)$ is a solution of $x^2 - ay^2 - bz^2 = 0$. Conversely, assume (i) holds, and let (s, t, u) be a non-zero triple in K^3 such that $s^2 - at^2 - bu^2 = 0$. Since b is not a square in K, we have $t \neq 0$, and hence:

$$a = \frac{s^2 - bu^2}{t^2} = N_{K(\sqrt{b})/K}\left(\frac{s + u\sqrt{b}}{t}\right) \in \mathrm{Im}(N_{K(\sqrt{b})/K}).$$

□

We now focus on the case where K is a p-adic field. We are going to provide a general formula that allows to compute all Hilbert symbols in that context.

Proposition 5.3.6

Let p be a prime number. Let $u, v \in \mathbb{Z}_p^\times$. Then:

(i)
$$(pu, v)_{\mathbb{Q}_p} = \begin{cases} \left(\frac{[v]}{p}\right) & \text{if } p \text{ is odd,} \\ (-1)^{\frac{u-1}{2}\frac{v-1}{2} + \frac{v^2-1}{8}} & \text{if } p = 2, \end{cases}$$

(ii)
$$(u, v)_{\mathbb{Q}_p} = \begin{cases} 1 & \text{if } p \text{ is odd,} \\ (-1)^{\frac{u-1}{2}\frac{v-1}{2}} & \text{if } p = 2. \end{cases}$$

Proof.

(i) We start with the case p odd. Assume that $(pu, v)_{\mathbb{Q}_p} = 1$. Then the equation $x^2 - vy^2 - puz^2 = 0$ has a non-zero solution (x_0, y_0, z_0) in $\mathbb{Z}_p$ such that x_0, y_0 and z_0 are not all divisible by p. One easily checks that x_0 and y_0 are not divisible by p. Reducing modulo p, we get $[x_0]^2 = [v]\left[y_0\right]^2$, and hence $\left(\frac{[v]}{p}\right) = 1$. Conversely, if $\left(\frac{[v]}{p}\right) = 1$, then by Proposition 5.3.1, there exists $x_0 \in \mathbb{Q}_p^\times$ such that $x_0^2 = v$, and hence $(x_0, 1, 0)$ is a solution of the equation $x^2 - vy^2 - puz^2 = 0$. We deduce that $(pu, v)_{\mathbb{Q}_p} = 1$.

We now move on to the case $p = 2$. The following equivalence is elementary and will be very useful in the proof:

$$(-1)^{\frac{u-1}{2}\frac{v-1}{2}+\frac{v^2-1}{8}} = 1 \Leftrightarrow \begin{cases} & u \equiv 1 \mod 4 \text{ and } v \equiv \pm 1 \mod 8 \\ \text{OR} & u \equiv -1 \mod 4 \text{ and } v \equiv 1 \mod 8 \\ \text{OR} & u \equiv -1 \mod 4 \text{ and } v \equiv 3 \mod 8. \end{cases} \tag{5.1}$$

Assume first that $(2u, v)_{\mathbb{Q}_2} = 1$. Then the equation $x^2 - vy^2 - 2uz^2 = 0$ has a non-zero solution (x_0, y_0, z_0) in $\mathbb{Z}_2$ such that x_0, y_0 and z_0 are not all divisible by 2. One easily checks that x_0 and y_0 are not divisible by 2. Reducing modulo 8, we get

$$1 - v \equiv x_0^2 - vy_0^2 \equiv 2uz_0^2 \mod 8.$$

By inspection, we see that:

- if $u \equiv 1 \mod 4$, then $2uz_0^2$ is 0 or 2 modulo 8, and hence $v \equiv \pm 1$ mod 8;
- if $u \equiv 3 \mod 4$, then $2uz_0^2$ is either 0 or -2 modulo 8, and hence v is either 1 or 3 modulo 8.

By equivalence (5.1), in both cases the integer $(-1)^{\frac{u-1}{2}\frac{v-1}{2}+\frac{v^2-1}{8}} = 1$. Conversely, assume that $(-1)^{\frac{u-1}{2}\frac{v-1}{2}+\frac{v^2-1}{8}} = 1$. By equivalence (5.1), we have the following cases to consider:

- if $v \equiv 1 \mod 8$, Proposition 5.3.1 shows that there exists $x_0 \in \mathbb{Q}_p^\times$ such that $x_0^2 = v$. In particular, $(x_0, 1, 0)$ is a solution of the equation $x^2 - vy^2 - 2uz^2 = 0$ and $(pu, v)_{\mathbb{Q}_p} = 1$.
- if $u \equiv 1 \mod 4$ and $v \equiv -1 \mod 8$ or if $u \equiv -1 \mod 4$ and $v \equiv 3$ mod 8, then $v+2u \equiv 1 \mod 8$ and Proposition 5.3.1 shows that there exists $x_0 \in \mathbb{Q}_p^\times$ such that $x_0^2 = v + 2u$. In particular, $(x_0, 1, 1)$ is a solution of the equation $x^2 - vy^2 - 2uz^2 = 0$ and $(pu, v)_{\mathbb{Q}_p} = 1$.

(ii) We start with the case p odd. By Proposition 5.2, the equation $x^2 - [u]\, y^2 - [v]\, z^2 = 0$ has a non-zero solution (x_0, y_0, z_0) in $\mathbb{F}_p$. If $x_0 \neq 0$, we may choose some liftings $\tilde{y}_0$ and $\tilde{z}_0$ of y_0 and z_0 in $\mathbb{Z}_p$ and by Hensel's lemma (Theorem 3.5.3), we can find $\tilde{x}_0 \in \mathbb{Z}_p$ a lifting of x_0 such that $\tilde{x}_0^2 - u\tilde{y}_0^2 - v\tilde{z}_0^2 = 0$. We deduce that $(u, v)_{\mathbb{Q}_p} = 1$.

We now move on to the case $p = 2$. The following equivalence is elementary and will be very useful in the proof:

$$(-1)^{\frac{u-1}{2}\frac{v-1}{2}} = 1 \Leftrightarrow (u \equiv 1 \mod 4 \text{ or } v \equiv 1 \mod 4). \tag{5.2}$$

Assume first that $(u,v)_{\mathbb{Q}_2} = 1$. Then the equation $x^2 - uy^2 - vz^2 = 0$ has a non-zero solution (x_0, y_0, z_0) in $\mathbb{Z}_2$ such that x_0, y_0 and z_0 are not all divisible by 2. Reducing modulo 4, we get a primitive solution in $\mathbb{Z}/4\mathbb{Z}$. By inspection, this shows that $u \equiv 1 \mod 4$ or $v \equiv 1 \mod 4$, and hence that $(-1)^{\frac{u-1}{2}\frac{v-1}{2}} = 1$. Conversely, assume that $(-1)^{\frac{u-1}{2}\frac{v-1}{2}} = 1$. This means that $u \equiv 1 \mod 4$ or $v \equiv 1 \mod 4$. Without loss of generality, we may and do assume that $u \equiv 1 \mod 4$. If $u \equiv 1 \mod 8$, Proposition 5.3.1 shows that there exists $x_0 \in \mathbb{Q}_2^\times$ such that $x_0^2 = u$. In particular, $(x_0, 1, 0)$ is a solution of the equation $x^2 - uy^2 - vz^2 = 0$ and $(u,v)_{\mathbb{Q}_2} = 1$. If $u \equiv 5 \mod 8$, Proposition 5.3.1 shows that there exists $x_0 \in \mathbb{Q}_2^\times$ such that $x_0^2 = u + 4v$. In particular, $(x_0, 1, 2)$ is a solution of the equation $x^2 - uy^2 - vz^2 = 0$ and $(u,v)_{\mathbb{Q}_2} = 1$.

□

Theorem 5.3.7

Fix a prime number p and take $a, b \in \mathbb{Q}_p^\times$. Writing $a = p^\alpha u$ and $b = p^\beta v$ with $\alpha, \beta \in \mathbb{Z}$ and $u, v \in \mathbb{Z}_p^\times$ and denoting by $[\cdot]$ the reduction modulo p:

$$(a,b)_{\mathbb{Q}_p} = \begin{cases} (-1)^{\alpha\beta\frac{p-1}{2}} \left(\frac{[u]}{p}\right)^\beta \left(\frac{[v]}{p}\right)^\alpha & \text{if } p \text{ is odd} \\ (-1)^{\left[\frac{u-1}{2}\frac{v-1}{2} + \alpha\frac{v^2-1}{8} + \beta\frac{u^2-1}{8}\right]} & \text{if } p = 2. \end{cases}$$

Proof. Up to exchanging a and b and by using the first formula in Example 5.3.4(i), we may and do assume that $(\alpha, \beta) \in \{(0,0), (0,1), (1,1)\}$. If $(\alpha, \beta) = (0,0)$, the statement boils down to Proposition 5.3.6(ii). If $(\alpha, \beta) = (0,1)$, the statement boils down to Proposition 5.3.6(i). It therefore remains to check the case $(\alpha, \beta) = (1,1)$. To do so, observe that the equation $x^2 - puy^2 - pvz^2 = 0$ has a non-zero solution in $\mathbb{Q}_p$ if, and only if, so does the equation $p^2y^2 - pu^{-1}x^2 + p^2u^{-1}vz^2 = 0$. Equivalently, this means that the equation $y^2 - pux^2 + uvz^2 = 0$ has a non-zero solution in $\mathbb{Q}_p$. Hence:

$$(a,b)_{\mathbb{Q}_p} = (pu, -uv)_{\mathbb{Q}_p}$$

If p is odd, we get thanks to Proposition 5.2.3:

$$(a,b)_{\mathbb{Q}_p} = \left(\frac{[-uv]}{p}\right) = (-1)^{\alpha\beta\frac{p-1}{2}} \left(\frac{[u]}{p}\right)\left(\frac{[v]}{p}\right).$$

If $p = 2$, we get:

$$(a,b)_{\mathbb{Q}_p} = (-1)^{\left[\frac{u-1}{2}\frac{-uv-1}{2}+\frac{u^2v^2-1}{8}\right]} = (-1)^{\left[\frac{u-1}{2}\frac{v-1}{2}+\frac{u^2-1}{8}+\frac{v^2-1}{8}\right]}$$

where the last equality is easily checked by inspection. □

Corollary 5.3.8

Assume that $K = \mathbb{R}$ or $K = \mathbb{Q}_p$ for some prime number p. The map:

$$\begin{aligned} K^\times \times K^\times &\to \{1,-1\} \\ (a,b) &\mapsto (a,b)_K \end{aligned}$$

is bilinear.

Proof. This is a straightforward computation thanks to the previous Theorem and to Proposition 5.2.3. □

Theorem 5.3.9 (Hilbert's Reciprocity Law)

Fix $a, b \in \mathbb{Q}^\times$. We have $(a,b)_{\mathbb{Q}_p} = 1$ for almost every prime number p and:

$$(a,b)_{\mathbb{R}} \cdot \prod_p (a,b)_{\mathbb{Q}_p} = 1$$

where p runs over all prime numbers.

Proof. Since Hilbert symbols over $\mathbb{R}$ and over p-adic fields are bilinear, we may and do assume that a and b are each equal to -1 or to a prime number. Hence, up to exchanging a and b, we have the following six cases, in each of which we use Theorem 5.3.7:

- $a = b = -1$. In this case:

$$\begin{cases} (-1,-1)_{\mathbb{R}} = (-1,-1)_{\mathbb{Q}_2} = -1 \\ (-1,-1)_{\mathbb{Q}_p} = 1 \quad \text{for } p \text{ odd.} \end{cases}$$

- $a = -1$ and $b = 2$. In this case,

$$\begin{cases} (-1,2)_{\mathbb{R}} = 1 \\ (-1,2)_{\mathbb{Q}_p} = 1 \quad \text{for every prime number } p. \end{cases}$$

- $a = -1$ and b is an odd prime number. In this case:

$$\begin{cases} (-1,b)_{\mathbb{R}} = 1 \\ (-1,b)_{\mathbb{Q}_p} = 1 \quad \text{for } p \notin \{2,b\} \\ (-1,b)_{\mathbb{Q}_p} = (-1)^{\frac{b-1}{2}} \quad \text{for } p \in \{2,b\}. \end{cases}$$

- a and b are the same prime number. In this case, we know that, for $K = \mathbb{R}$ and $K = \mathbb{Q}_p$:

$$(a,b)_K = (-1,b)_K \cdot (-b,b)_K = (-1,b)_K,$$

  and hence we are reduced to the previous case.

- $a = 2$ and b is an odd prime number. In this case:

$$\begin{cases} (2,b)_{\mathbb{R}} = 1 \\ (2,b)_{\mathbb{Q}_p} = 1 \quad \text{for } p \notin \{2,b\} \\ (2,b)_{\mathbb{Q}_2} = (-1)^{\frac{b^2-1}{8}} \\ (2,b)_{\mathbb{Q}_b} = \left(\frac{2}{b}\right) = (-1)^{\frac{b^2-1}{8}} \quad \text{by the quadratic reciprocity law.} \end{cases}$$

- a and b are distinct odd prime numbers. In this case:

$$\begin{cases} (a,b)_{\mathbb{R}} = 1 \\ (a,b)_{\mathbb{Q}_p} = 1 \quad \text{for } p \notin \{a,b\} \\ (a,b)_{\mathbb{Q}_2} = (-1)^{\frac{a-1}{2}\frac{b-1}{2}} = \left(\frac{a}{b}\right)\left(\frac{b}{a}\right) \quad \text{by the quadratic reciprocity law} \\ (a,b)_{\mathbb{Q}_a} = \left(\frac{b}{a}\right) \\ (a,b)_{\mathbb{Q}_b} = \left(\frac{a}{b}\right). \end{cases}$$

In each case, we see that $(a,b)_{\mathbb{Q}_p} = 1$ for almost every prime number p and:

$$(a,b)_{\mathbb{R}} \cdot \prod_p (a,b)_{\mathbb{Q}_p} = 1$$

. □

Theorem 5.3.10

Let p be a prime number. If p is odd, we let $u \in \mathbb{Z}_p^\times$ be any element that is not a square. If $p = 2$ we take $u \in \mathbb{Z}_p^\times$ any element that is congruent to ± 3 modulo 8 (in particular it is not a square). A quadratic form in

four variables over $\mathbb{Q}_p$ is anisotropic if, and only if, it is isometric to the quadratic form:

$$q_0 = \langle 1, -u, -p, pu \rangle.$$

Before proving this Theorem, we settle the following easy lemma that is very convenient to prove that some quadratic forms are isometric:

Lemma 5.3.11. *Let K be a field with characteristic not 2. Set $q := \langle a_1, a_2 \rangle$ and $q' := \langle a_1', a_2' \rangle$ with $a_1, a_2, a_1', a_2' \in K^\times$. The quadratic forms q and q' are isometric if, and only if, there exists $(x_1, x_2) \in K^2$ such that $q'(x_1, x_2) = a_1$ and q and q' have the same discriminant.*

Proof. The direct implication is obvious. Assume that there exist $x_1, x_2 \in K$ such that $q'(x_1, x_2) = a_1$ and q and q' have the same discriminant. Since x_1, x_2 are not both zero, we can find $x_1', x_2' \in K$ such that $a_1' x_1 x_1' + a_2' x_2 x_2' = 0$. We then have:

$$M_q = \begin{pmatrix} x_1 & x_2 \\ x_1' & x_2' \end{pmatrix} M_{q'} \begin{pmatrix} x_1 & x_1' \\ x_2 & x_2' \end{pmatrix},$$

and hence q and q' are isometric. □

Proof of Theorem 5.3.10. Let us first check that q_0 is anisotropic. To do so, observe that the equation $q_0(x, y, z, t) = 0$ can be rewritten as:

$$x^2 - uy^2 = p(z^2 - ut^2). \tag{5.3}$$

Moreover, since u is not a square in $\mathbb{Z}_p$, the quadratic form $\langle 1, -u \rangle$ is anisotropic, and hence equation (5.3) can be rewritten as:

$$N_{\mathbb{Q}_p(\sqrt{u})/\mathbb{Q}_p}\left(\frac{x + y\sqrt{u}}{z + t\sqrt{u}}\right) = p. \tag{5.4}$$

In other words, the equation $q_0(x, y, z, t) = 0$ has non-zero solutions in $\mathbb{Q}_p$ if, and only if, $p \in \mathrm{im}(N_{\mathbb{Q}_p(\sqrt{u})/\mathbb{Q}_p})$. By Proposition 5.3.5, this is equivalent to the equality $(u, p)_{\mathbb{Q}_p} = 1$. But by Theorem 5.3.7, we have:

$$(u, p)_{\mathbb{Q}_p} = \begin{cases} \left(\frac{[u]}{p}\right) = -1 & \text{if } p \text{ is odd} \\ (-1)^{\left[\frac{u^2-1}{8}\right]} = -1 & \text{if } p = 2. \end{cases}$$

We deduce that the equation $q_0(x, y, z, t) = 0$ has no non-zero solutions in $\mathbb{Q}_p$.

Let now q be any anisotropic quadratic form in 4 variables over $\mathbb{Q}_p$. By Theorem 5.1.1, we may and do assume that q is of the form:

$$q = a_1 \cdot \langle 1, a_2, a_3, \ldots, a_n \rangle = a_1(X_1^2 + a_2 X_2^2 + \ldots, a_n X_n^2),$$

with $0 \leq v_p(a_2) \leq v_p(a_3) \leq v_p(a_4) \leq 1$.

Let us first assume that p is odd. In that case, we make the following considerations.

- Since the form $\langle a_2, a_3, a_4 \rangle$ is anisotropic, so is the form $\langle 1, a_2^{-1}a_3, a_2^{-1}a_4 \rangle$ and Proposition 5.3.6 implies that $v_p(a_1) = v_p(a_2) = 0$.
- Since the quadratic form $\langle 1, a_2, a_3 \rangle$ is anisotropic, Proposition 5.3.6 implies that $v_p(a_3) = v_p(a_4) = 1$ and that $[-a_2]$ is not a square in $\mathbb{F}_p^\times$. In particular, since $\mathbb{F}_p^\times$ has exactly two square classes, $[-a_2]$ lies in the same square class as $[u]$. By Proposition 5.3.1, we deduce that $-a_2$ lies in the same square class as u.
- Write now $a_3 = pu_3$ and $a_4 = pu_4$. Since the quadratic form $\langle 1, a_3, a_4 \rangle$ is anisotropic, so is the quadratic form $\langle pu_3, 1, u_3^{-1}u_4 \rangle$, and hence, by Proposition 5.3.6, we get that $[-u_3^{-1}u_4]$ is not a square in $\mathbb{F}_p^\times$. In particular, since $\mathbb{F}_p^\times$ has exactly two square classes, one among $[-u_3]$ and $[u_4]$ is a square and the other lies in the same square class as $[u]$. By Proposition 5.3.1, we deduce that one among $-u_3$ and u_4 is a square and the other lies in the same square class as u.

The previous considerations show that the quadratic form q is isometric to $a_1 \cdot \langle 1, -u, -p, pu \rangle$ or to $a_1 \cdot \langle 1, -u, -pu, p \rangle$. Moreover, by Proposition 5.3.1, the element a_1 is in the same square class as 1, u, p or pu. Hence q is isometric to one of the following quadratic forms:

$$\langle 1, -u, -p, pu \rangle, \langle -1, u, p, -pu \rangle, \langle 1, -u, p, -pu \rangle, \langle -1, u, -p, pu \rangle. \tag{5.5}$$

Now, according to Proposition 5.3.6, there exist $x, y \in \mathbb{Q}_p$ such that $x^2 - uy^2 = -1$. By Lemma 5.3.11, we deduce that:

$$\langle 1, -u \rangle \cong \langle -1, u \rangle$$

are isometric, and hence the four quadratic forms in equation (5.5) are isometric. We deduce that q is isometric to q_0.

Let us now move on to the case $p = 2$. According to Proposition 5.3.1:

$$\mathbb{Q}_2^\times / \mathbb{Q}_2^{\times 2} = \{\pm 1, \pm 2, \pm 3, \pm 6\}.$$

We may henceforth assume that $a_1, a_2, a_3, a_4 \in \{\pm 1, \pm 2, \pm 3, \pm 6\}$.

Since q is anisotropic, so are the quadratic forms $\langle 1, a_i, a_j\rangle$ with $2 \leq i < j \leq 4$ and the quadratic form $\langle a_2, a_3, a_4\rangle$. But by Theorem 5.3.7:

- given $a, b \in \{\pm 1, \pm 2, \pm 3, \pm 6\}$ with $v_2(a) \leq v_2(b)$, the quadratic form $\langle 1, a, b\rangle$ is anisotropic if, and only if:

$$(a,b) \in \{1,-3\}^2 \cup \{(1,2),(1,-6),(3,2),(3,-6),(\pm 3,-2),(\pm 3,6),\\ (2,2),(2,6),(-2,\pm 6),(-6,-6),(6,2),(\pm 6,-2)\}.$$

- given $a, b, c \in \{\pm 1, \pm 3\}$, the quadratic form $\langle 2a, 2b, 2c\rangle$ is anisotropic if, and only if:

$$(a,b,c) \in \{1,-3\}^3 \cup \{-1,3\}^3.$$

We deduce that the triple (a_2, a_3, a_4) is in:

$$\{1,-3\}^3 \cup \{(1,1,2),(1,1,-6),(-3,-3,-2),(-3,-3,6),(1,2,2),(1,-6,-6),\\ (3,2,2),(3,2,6),(3,-2,\pm 6),(3,-6,-6),(3,6,\pm 2),(3,-6,-2),(-3,-2,6),(-3,6,-2),\\ (2,2,2),(-6,-6,-6)\}.$$

Now, by Proposition 5.3.1, since q is anisotropic, the congruence $X_1^2 + a_2X_2^2 + a_3X_3^2 + a_4X_4^2 \equiv 0 \mod 8$ cannot have a solution (x_1, x_2, x_3, x_4) with $x_1 \equiv 1 \mod 2$ or with $a_i \equiv x_i \equiv 1 \mod 2$ for some $i \in \{2,3,4\}$. Hence the triple (a_2, a_3, a_4) is in:

$$\{(1,1,1),(1,-3,-3),(1,2,2),(1,-6,-6),\\ (3,2,6),(3,-6,-2),(3,-2,-6),(-3,-2,6),(3,6,2),(-3,6,-2),(2,2,2),(-6,-6,-6)\}.$$

By using Lemma 5.3.11 and Theorem 5.3.7, we get the following isometries:

$$\langle 1,1,1\rangle \cong \langle 1,-3,-3\rangle \cong \langle 1,-6,-6\rangle,$$
$$\langle 1,1,1\rangle \cong \langle 1,2,2\rangle,$$
$$\langle 1,2,2\rangle \cong \langle 3,2,6\rangle \cong \langle 3,6,2\rangle,$$
$$\langle 1,-6,-6\rangle \cong \langle 3,-6,-2\rangle \cong \langle 3,-2,-6\rangle,$$
$$\langle 1,-3,-3\rangle \cong \langle -3,-2,6\rangle \cong \langle -3,-6,-2\rangle,$$
$$\langle 2,2,2\rangle \cong \langle 1,2,4\rangle,$$
$$\langle -6,-6,-6\rangle \cong \langle -3,-6,-12\rangle.$$

We deduce that $\langle 1, a_2, a_3, a_4\rangle$ is isometric to one of the following quadratic forms:

$$\langle 1,-2,-3,6\rangle, \quad \langle 1,1,2,4\rangle, \quad \langle 1,-3,-6,-12\rangle.$$

But the quadratic form $\langle 1,-3,-6,-12\rangle$ is isotropic since $3^2-3-6=0$ and the quadratic form $\langle 1,1,2,4\rangle$ is isotropic since, by Proposition 5.3.1, the field $\mathbb{Q}_2$ contains a square root $\sqrt{-7}$ of -7 and $\sqrt{-7}^2+1+2+4=0$. We deduce that $\langle 1,a_2,a_3,a_4\rangle$ is isometric to $\langle 1,-2,-3,6\rangle$ and hence that q is isometric to:

$$a_1\cdot\langle 1,-2,-3,6\rangle.$$

In other words, q is isometric to one of the following two quadratic forms:

$$\langle 1,-2,-3,6\rangle, \langle -1,2,3,-6\rangle.$$

By using Lemma 5.3.11, we have:

$$\langle 1,-2,-3,6\rangle \cong \langle -1,2,-3,6\rangle \cong \langle -1,2,3,-6\rangle.$$

Hence:

$$q\cong\langle 1,-2,-3,6\rangle.$$

□

Theorem 5.3.12

Let p be a prime number. Every quadratic form in five variables over $\mathbb{Q}_p$ is isotropic.

Proof. Let q be an anisotropic quadratic form in 5 variables over $\mathbb{Q}_p$. One can find $a_1,\ldots,a_5\in\mathbb{Q}_p^\times$ such that:

$$q\cong\langle a_1,\ldots,a_5\rangle.$$

The quadratic form $\langle a_1,\ldots,a_4\rangle$ being anisotropic, Theorem 5.3.10 implies that it is isometric to $\langle 1,-u,-p,up\rangle$ where u is an element of $\mathbb{Z}_p$ that is not a square. Hence:

$$q\cong\langle 1,-u,-p,up,a_5\rangle.$$

Similarly, the quadratic forms $\langle 1,-u,-p,a_5\rangle$ and $\langle 1,-p,up,a_5\rangle$ should be isometric to $\langle 1,-u,-p,up,a_5\rangle$. By comparing the discriminants, we see that a_5 should be in the square class as pu and as $-u$. But pu and $-u$ are not in the same square class: contradiction! We deduce that every quadratic form in 5 variables over $\mathbb{Q}_p$ is isotropic. □

5.4 Quadratic forms over $\mathbb{Q}$

Theorem 5.4.1 (Hasse-Minkowski)

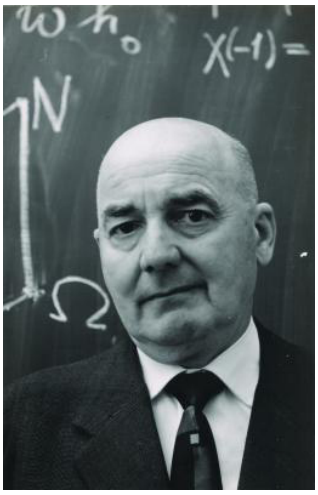
Let q be a quadratic form over $\mathbb{Q}$. Assume that the equation $q = 0$ has non-zero solutions in $\mathbb{R}$ and in $\mathbb{Q}_p$ for all p. Then it has a non-zero solution in $\mathbb{Q}$.

Proof. By a change of variables, we can always assume that q is of the form:

$$q = a_1 X_1^2 + \cdots + a_n X_n^2.$$

In the case one of the a_i's is zero, there is nothing to prove. We henceforth assume that $a_1, \ldots, a_n$ are all non-zero. Replacing q by $a_1^{-1} q$, we may and do assume that $a_1 = 1$. We now proceed by induction on n.

Case $n = 2$. Assume first that $n = 2$, so that $q = X_1^2 + a_2 X_2^2$. Since the equation $q = 0$ has non-zero real solutions, $a_2 < 0$. Moreover, for each prime number p, since the equation $q = 0$ has non-zero p-adic solutions, $v_p(a_2)$ is even. We deduce that $-a_2$ is the square of a rational number, and hence the equation $q = 0$ has a non-zero rational solution.



Helmut Hasse (Kassel 1898 - Ahrensburg 1979)
and Hermann Minkowski (Aleksotas 1864 - Göttingen 1909)

Case $n = 3$. Assume that $n = 3$, so that $q = X_1^2 + a_2 X_2^2 + a_3 X_3^2$. By performing a change of variables of the form $X_2 \mapsto \lambda X_2, X_3 \mapsto \mu X_3$, we may and do assume that a_2 and a_3 are square-free integers. Also, by exchanging the variables X_2 and X_3, we assume that $|a_2| \leq |a_3|$. Let us first deal with two easy cases:

(i) If $|a_3| = 1$, then q is of the form:

$$X_1^2 \pm X_2^2 \pm X_3.$$

At least one of the signs has to be negative, since the equation $q = 0$ has non-zero solutions in $\mathbb{R}$. But then it is obvious that $q = 0$ has a non-zero rational solution.

(ii) If $a_2 = -1$, the triple $(1,1,0)$ is a rational solution to the equation $q = 0$.

We now move on to the harder case where $|a_3| > 1$ and $a_2 \neq -1$. By contradiction, assume also that the equation $q = 0$ has no non-zero rational solutions and that $|a_2| + |a_3|$ is as small as possible. Since $|a_3| > 1$, we can write:

$$|a_3| = p_1 \dots p_r$$

for some pairwise distinct prime numbers $p_1, \dots, p_r$. Fix some $i \in \{1, 2, \dots, r\}$. Since the equation $q = 0$ has non-zero solutions in $\mathbb{Q}_{p_i}$, we can find three p_i-adic integers $x_{1,i}, x_{2,i}, x_{3,i} \in \mathbb{Z}_{p_i}$, not all divisible by p_i, such that $x_{1,i}^2 + a_2 x_{2,i}^2 + a_3 x_{3,i}^2 = 0$. One then easily checks that $x_{2,i}$ is not divisible by p_i, and hence the class of $-a_2$ modulo p_i is a square in $\mathbb{Z}/p_i\mathbb{Z}$. This is true for all i, and hence the Chinese remainder Theorem implies that the class of $-a_2$ modulo a_3 is a square in $\mathbb{Z}/a_3\mathbb{Z}$. In other words, we can find two integers t and a_3' such that:

$$t^2 = -a_2 + a_3 a_3' \quad \text{and} \quad 0 \leq t \leq \frac{|a_3|}{2}.$$

In particular:

$$a_3 a_3' = t^2 + a_2 \in N_{\mathbb{Q}(\sqrt{-a_2})}(\mathbb{Q}(\sqrt{-a_2})^\times).$$

But since the equation $q = 0$ has no non-zero rational solutions, Proposition 5.3.5 implies that:

$$-a_3 \notin N_{\mathbb{Q}(\sqrt{-a_2})}(\mathbb{Q}(\sqrt{-a_2})^\times).$$

We deduce that:

$$-a_3' \notin N_{\mathbb{Q}(\sqrt{-a_2})}(\mathbb{Q}(\sqrt{-a_2})^\times)$$

and hence, by using once again Proposition 5.3.5, the equation $x_1^2 + a_2 x_2^2 + a_3' x_3^2 = 0$ has no non-zero rational solutions. Of course, if we write $a_3' = u^2 a_3''$ with $u, a_3'' \in \mathbb{Z}$ and a_3'' square-free, the same holds for the equation $x_1^2 + a_2 x_2^2 + a_3'' x_3^2 = 0$. But:

$$\begin{aligned} |a_2| + |a_3''| \leq |a_2| + |a_3'| \leq |a_2| + \left| \frac{t^2 + a_2}{a_3} \right| &\leq |a_2| + \frac{a_3^2 + 4|a_3|}{4|a_3|} \\ &\leq |a_2| + \frac{|a_3|}{4} + 1 < |a_2| + |a_3|. \end{aligned}$$

This contradicts the minimality of $|a_2| + |a_3|$.

Case $n = 4$. Assume that $n = 4$, so that $q = X_1^2 + a_2X_2^2 + a_3X_3^2 + a_4X_4^2 = q_0 + q_1$ with $q_0 = X_1^2 + a_2X_2^2$ and $q_1 = a_3X_3^2 + a_4X_4^2$. Set:

$$S := \{2, \infty\} \cup \{p \text{ prime: } v_p(a_2),\ v_p(a_3) \text{ and } v_p(a_4) \text{ are not all } 0\}$$

and:

$$\mathbb{Q}_\infty := \mathbb{R}.$$

Given an element $p \in S$, we have three cases:

(i) the quadratic form q_0 is isotropic over $\mathbb{Q}_p$, in which case the map $q_0 : \mathbb{Q}_p^2 \to \mathbb{Q}_p$ is surjective by Proposition 5.1.3;

(ii) the quadratic form q_1 is isotropic over $\mathbb{Q}_p$, in which case the map $q_1 : \mathbb{Q}_p^2 \to \mathbb{Q}_p$ is surjective by Proposition 5.1.3;

(iii) the quadratic forms q_0 and q_1 are both anisotropic over $\mathbb{Q}_p$.

Since q is isotropic over $\mathbb{Q}_p$, we deduce in the three previous cases that there exist $y_p \in \mathbb{Q}_p^\times$ and $x_{1,p}, x_{2,p}, x_{3,p}, x_{4,p} \in \mathbb{Q}_p$ such that:

$$y_p = q_0(x_{1,p}, x_{2,p}) = -q_1(x_{3,p}, x_{4,p}). \tag{5.6}$$

Moreover, when $p \neq \infty$, we may and do assume that $y_p \in \mathbb{Z}_p \setminus p\mathbb{Z}_p$. According to the Chinese Remainder Theorem, the set of integers $y \in \mathbb{Z}$ such that:

$$\begin{cases} y \equiv y_2 \mod 2^{v_2(y_2)+3}, \\ \forall\, p \in S \setminus \{2\}, y \equiv y_p \mod p^{v_p(y_p)+1}, \\ y y_\infty > 0 \end{cases} \tag{5.7}$$

is of the form $\{a + nb : n \in \mathbb{N}\}$ with $b := \pm 2^{v_2(y_2)+3} \cdot \prod_{p \in S\setminus\{2,\infty\}} p^{v_p(y_p)+1}$ and a some integer coprime to b. According to Dirichlet's Theorem on Arithmetic Progressions, we may choose an integer $n_0 \in \mathbb{N}$ such that $y_0 := a + n_0 b$ has a unique prime divisor outside S, that we denote p_0.

We now consider the equation:

$$y_0 = q_0(x_1, x_2) = -q_1(x_3, x_4). \tag{5.8}$$

Let us prove that it has solutions in $\mathbb{R}$ and in every $\mathbb{Q}_p$:

- Given an element $p \in S$ (it could happen that $p = \infty$), the congruences and the inequality given by (5.7) together with Proposition 5.3.1 show that $y_p y_0^{-1}$ is a square in $\mathbb{Q}_p^\times$, and hence equality (5.6) shows that equation (5.8) has solutions in $\mathbb{Q}_p$.

- Given a prime number p not in S and different from p_0, we see that y_0, a_2, a_3 and a_4 are all units in $\mathbb{Z}_p$, and hence, by Proposition 5.3.6(ii), we deduce that equation (5.8) has solutions in $\mathbb{Q}_p$.
- Finally, we study equation (5.8) over $\mathbb{Q}_{p_0}$. According to the previous arguments, we know that equation (5.8) has solutions in $\mathbb{R}$ and in every $\mathbb{Q}_p$ for $p \neq p_0$. By definition of Hilbert symbols, this means that:

$$\forall\, p \neq p_0, \quad (y_0, -a_2)_{\mathbb{Q}_p} = (-a_3 y_0, -a_4 y_0)_{\mathbb{Q}_p} = 1,$$
$$(y_0, -a_2)_{\mathbb{R}} = (-a_3 y_0, -a_4 y_0)_{\mathbb{R}} = 1.$$

  But by Hilbert's Reciprocity Law (Theorem 5.3.9), we have:

$$(y_0, -a_2)_{\mathbb{R}} \cdot \prod_p (y_0, -a_2)_{\mathbb{Q}_p} = 1,$$
$$(-a_3 y_0, -a_4 y_0)_{\mathbb{R}} \cdot \prod_p (-a_3 y_0, -a_4 y_0)_{\mathbb{Q}_p} = 1.$$

  Hence $(y_0, -a_2)_{\mathbb{Q}_{p_0}} = (-a_3 y_0, -a_4 y_0)_{\mathbb{Q}_{p_0}} = 1$ and equation (5.8) has solutions in $\mathbb{Q}_{p_0}$.

Now that we know that equation (5.8) has solutions in $\mathbb{R}$ and in every $\mathbb{Q}_p$, we can apply Hasse-Minkowski's Theorem in the case $n = 3$ to deduce that equation (5.8) has rational solutions.

Case $n \geq 5$. Assume that $n \geq 5$ and write $q = q_0 + q_1$ with $q_0 = X_1^2 + a_2 X_2^2$ and $q_1 = a_3 X_3^2 + \cdots + a_n X_n^2$. As in the case $n = 4$, we set:

$$S := \{2, \infty\} \cup \{p \text{ prime: } v_p(a_2),\ v_p(a_3) \text{ and } v_p(a_4) \text{ are not all } 0\}$$

and:

$$\mathbb{Q}_\infty := \mathbb{R}.$$

Given an element $p \in S$, we may argue as in the case $n = 4$ and check that there exist $y_p \in \mathbb{Q}_p^\times$ and $x_{1,p}, \ldots, x_{n,p} \in \mathbb{Q}_p$ such that:

$$y_p = q_0(x_{1,p}, x_{2,p}) = -q_1(x_{3,p}, \ldots, x_{n,p}). \tag{5.9}$$

By the Weak Approximation Theorem (Theorem 3.4.2) and by Remark 5.3.2, we can find $x_{1,0}, x_{2,0} \in \mathbb{Q}$ and $y_0 \in \mathbb{Q}^\times$ such that $y_0 = q_0(x_{1,0}, x_{2,0})$ and $y_p y_0^{-1}$ is a square in $\mathbb{Q}_p^\times$ for each $p \in S$. In particular, for $p \in S$, the equation:

$$y_0 = -q_1(x_3, \ldots, x_n) \tag{5.10}$$

has solutions in $\mathbb{Q}_p$. But this statement also holds for p a prime number outside S according to Propositions 5.3.6(ii) and 5.1.3 applied to the quadratic form $q_1(X_3, X_4, X_5, 0, \ldots, 0)$. Hence, by induction, we deduce that the equation (5.10) has solutions in $\mathbb{Q}$. □

Together with Hensel's lemma, the Hasse-Minkowski Theorem provides an algorithmic way to check whether degree 2 equations with rational coefficients have rational solutions:

Corollary 5.4.2

Let $n \geq 3$ and let $q = a_1 X_1^2 + \cdots + a_n X_n^2$ be a quadratic form with $a_1, \ldots, a_n \in \mathbb{Z} \setminus \{0\}$. Let $p_1, \ldots, p_r$ be the odd prime divisors of the product $a_1 \ldots a_n$ and set:

$$N := 8 p_1 \ldots p_r \cdot (a_1 \vee \cdots \vee a_n).$$

Then the following assertions are equivalent:

(i) the equation $q = 0$ has a non-zero solution in $\mathbb{Q}$;

(ii) not all the a_i's have the same sign and the equation $q = 0$ has primitive solutions in $\mathbb{Z}/N\mathbb{Z}$.

Proof. The direct implication is obvious. Let us prove the converse. Up to dividing q by $a_1 \wedge \cdots \wedge a_n$, we may assume that $a_1 \wedge \cdots \wedge a_n = 1$. By Theorem 5.4.1, it suffices to prove that q is isotropic over $\mathbb{R}$ and over every $\mathbb{Q}_p$.

- Since the a_i's do not all have the same sign, q is isotropic over $\mathbb{R}$.
- For p a prime number outside $\{2, p_1, \ldots, p_r\}$, the quadratic form q is isotropic over $\mathbb{Q}_p$ according to Proposition 5.3.6.
- Take $p \in \{2, p_1, \ldots, p_r\}$ and choose $x_1, \ldots, x_n \in \mathbb{Z}$ such that $x_1 \wedge \cdots \wedge x_n = 1$ and:
$$a_1 x_1^2 + \cdots + a_n x_n^2 \equiv 0 \mod N.$$
Choose i such that x_i is not divisible by p. If $p \neq 2$, since $p a_i$ divides N, we have:
$$x_i^2 \equiv \frac{a_1 x_1^2 + \cdots + a_{i-1} x_{i-1}^2 + \cdots + a_{i+1} x_{i+1}^2 + a_n x_n^2}{a_i} \mod p.$$
If $p = 2$, since $8 a_i$ divides N, we have:
$$x_i^2 \equiv \frac{a_1 x_1^2 + \cdots + a_{i-1} x_{i-1}^2 + \cdots + a_{i+1} x_{i+1}^2 + a_n x_n^2}{a_i} \mod 8.$$
In both cases, by Proposition 5.3.1, we can find $\tilde{x}_i \in \mathbb{Z}_2$ such that:
$$\tilde{x}_i^2 = \frac{a_1 x_1^2 + \cdots + a_{i-1} x_{i-1}^2 + \cdots + a_{i+1} x_{i+1}^2 + a_n x_n^2}{a_i}.$$
We then have $q(x_1, \ldots, \tilde{x}_i, \ldots, x_n) = 0$ and q is isotropic over $\mathbb{Q}_p$.

□

Corollary 5.4.3

Let q be a quadratic form in more than 4 variables over $\mathbb{Q}$. If the equation $q = 0$ has non-zero solutions in $\mathbb{R}$, then it has a non-zero solution in $\mathbb{Q}$.

Proof. This corollary follows immediately from Theorems 5.3.12 and 5.4.1. □

Bibliography

[IR72] **K. Ireland & M. Rosen.** A classical introduction to modern number theory, Springer GTM 84 (1972).

[L93] **S. Lang.** Algebra, 3rd edition, Addison-Wesley (1993).

[N99] **J. Neukirch.** Algebraic Number Theory, Grundlehren der mathematischen Wissenschaften, Springer (1999).

[S77] **J.-P. Serre.** Cours d'arithmétique, P.U.F. (1977).