

MATH-GA 2210.001 : Introduction to Number  
Theory I

ALENA PIRUTKA

# Contents

<b>1</b>	<b>Analytical tools</b>	<b>4</b>
1.1	Primes in arithmetic progressions . . . . .	4
1.1.1	Euler's identity and existence of infinitely many primes . . .	4
1.1.2	Zeta function . . . . .	5
1.1.3	Characters . . . . .	8
1.1.4	L-functions . . . . .	9
1.1.5	Proof of Dirichlet theorem . . . . .	13
1.2	Zeta function . . . . .	14
1.2.1	Fourier analysis . . . . .	15
1.2.2	Entire functions . . . . .	17
1.2.3	$\Gamma$ -function. . . . .	19
1.2.4	Zeros of the zeta function . . . . .	21
1.3	Distribution of primes . . . . .	26
<b>2</b>	<b>Local Fields</b>	<b>30</b>
2.1	Absolute values . . . . .	30
2.2	Completions . . . . .	36
2.3	Locally compact fields . . . . .	39
2.4	Extensions . . . . .	40
2.4.1	Basic facts . . . . .	40
2.4.2	Unramified extensions . . . . .	41
2.4.3	Totally ramified extensions . . . . .	42
2.4.4	Ramification groups and Krasner's lemma . . . . .	43
<b>3</b>	<b>Dedekind rings</b>	<b>45</b>
3.1	Dedekind rings . . . . .	45
3.1.1	Fractional ideals . . . . .	45
3.1.2	Discrete valuation rings . . . . .	46
3.1.3	Dedekind rings . . . . .	47
3.2	Extensions . . . . .	50
3.2.1	Extensions of Dedekind rings . . . . .	50
3.3	Decomposition of ideals . . . . .	52
3.3.1	Galois case . . . . .	53
3.3.2	Explicit factorisation . . . . .	53
3.3.3	Complement: extensions of valued fields . . . . .	54

<b>4</b>	<b>Number fields</b>	<b>56</b>
4.1	Geometry of numbers . . . . .	56
	4.1.1 Lattices in $\mathbb{R}^n$ . . . . .	56
	4.1.2 Applications in arithmetics . . . . .	60
4.2	Rings of integers of number fields . . . . .	64
	4.2.1 Number fields . . . . .	64
4.3	Ideal classes . . . . .	66
	4.3.1 Canonical embedding . . . . .	66
	4.3.2 Finiteness . . . . .	67
4.4	Applications . . . . .	70
	4.4.1 Quadratic fields . . . . .	70
	4.4.2 $Cl(D)$ versus $Cl(A_D)$ . . . . .	71
	4.4.3 Equation $y^2 = x^3 + k$ . . . . .	73
	4.4.4 An example of computation of $Cl(\mathcal{O}_K)$ . . . . .	74
4.5	The Dirichlet formula . . . . .	75

The course MATH-GA 2210.001 provides an introduction to the Number Theory, with analytic, algebraic and diophantine aspects. The analytic techniques allow to provide a proof of some classical results on prime numbers. In the algebraic part we will study the fundamental properties of local and global fields.

Here is the list of references used in the preparation of these notes:

1. Z. I Borevich and I.R. Shafarevich, *Number theory*.
2. K. Ireland and M.Rosen, *A classical introduction to modern number theory*.
3. J.S. Milne, *Algebraic Number theory*.
4. S. J. Miller and R. Takloo-Bighash, *An Invitation to Modern Number Theory*.
5. A.Karatsuba, *Basic Analytic Number Theory*.
6. Lecture notes on Algebraic number theory by Loïc Merel.
7. Lecture notes on Algebraic number theory by Gaëtan Chenevier.
8. P. Samuel, *Théorie algébrique des nombres*.
9. S. Lang, *Algebraic number theory*.
10. I. Stewart ad D. Tall, *Algebraic number theory*.

# Chapter 1

## Analytical tools

### 1.1 Primes in arithmetic progressions

The goal of this section is to prove the Dirichlet theorem:

**Theorem 1.1.1** (Dirichlet). *Every arithmetic progression*

$$a, a + q, a + 2q, \dots$$

*in which  $a$  and  $q$  have no common factor, includes infinitely many primes.*

#### 1.1.1 Euler's identity and existence of infinitely many primes

The series  $\sum_{n \geq 1} n^{-s}$  converges uniformly for  $s$  in a compact in the half-plane  $\operatorname{Re} s > 1$ , so that it defines an analytic function

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

(introduced by Riemann in 1859.)

**Proposition 1.1.2.** *The infinite product*

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

*converges uniformly on any compact in the half-plane  $\operatorname{Re} s > 1$  and defines an analytic function verifying*

$$\zeta(s) = \prod_p \frac{1}{1 - p^{-s}}.$$

*Proof.* We express  $\frac{1}{1 - p^{-s}}$  as a sum of a geometric series

$$\frac{1}{1 - p^{-s}} = \sum_{m \geq 0} \frac{1}{p^{ms}}.$$

Let  $X$  be a sufficiently big integer. Multiplying the identities above for primes  $\leq X$  we obtain:

$$\zeta(s) = \prod_{p \leq X} \frac{1}{1 - p^{-s}} = \prod_{p \leq X} \sum_{m \geq 0} \frac{1}{p^{ms}} = \sum_{n \in N(X)} \frac{1}{n^s},$$

where  $N(X)$  is the set of positive integers having all prime factors  $\leq X$ . Then for  $\operatorname{Re} s = t > 1$  we have

$$|\zeta(s) - \prod_{p \leq X} \frac{1}{1 - p^{-s}}| \leq \sum_{n \notin N(X)} \frac{1}{n^s} \leq \sum_{n > X} \frac{1}{n^t}.$$

To verify that the Euler product converges in remains to show that it is nonzero. Let us show that  $\zeta(s) \neq 0$  for  $\operatorname{Re} s > 1$ . We use the Talyor series expansion for the principal definition of the complex logarithm:  $\log(1 - p^{-s}) = -\sum_{m \geq 1} \frac{p^{-ms}}{m}$ , so that for  $\operatorname{Re} s > 1$  we obtain

$$\zeta(s) = \exp\left(\sum_p \sum_{m \geq 1} \frac{p^{-ms}}{m}\right)$$

is nonzero. □

The expression above provides a method to show the infinity of prime numbers. Write

$$\log \zeta(s) = \sum_p \sum_{m=1}^{\infty} m^{-1} p^{-ms}. \quad (1.1)$$

Since  $\zeta(s) \rightarrow \infty$  as  $s \rightarrow 1$  from the right, and since

$$\sum_p \sum_{m=2}^{\infty} m^{-1} p^{-ms} < \sum_p \sum_{m=2}^{\infty} p^{-m} = \sum_p \frac{1}{p(p-1)} < 1.$$

it follows that  $\sum_p p^{-s} \rightarrow \infty$  as  $s \rightarrow 1$  from the right. This proves the existence of an infinity of primes and, moreover, that the series  $\sum p^{-1}$  diverges.

The proof Dirichlet theorem is inspired by the same idea, but with more involved techniques. We first investigate some additional properties of the zeta function.

### 1.1.2 Zeta function

**Proposition 1.1.3.** *Assume  $s > 1$ . Then  $\lim_{s \rightarrow 1} (s-1)\zeta(s) = 1$ .*

*Proof.* We have

$$(n+1)^{-s} < \int_n^{n+1} t^{-s} dt < n^{-s}.$$

Taking the sum from 1 to  $\infty$ , one obtains

$$\zeta(s) - 1 < \int_1^{\infty} t^{-s} dt = \frac{1}{s-1} < \zeta(s).$$

Hence  $1 < (s-1)\zeta(s) < s$ . We obtain the result taking limit as  $s \rightarrow 1$ .  $\square$

**Corollary 1.1.4.**

$$\frac{\ln \zeta(s)}{\ln(s-1)^{-1}} \xrightarrow{s \rightarrow 1} 1.$$

*Proof.* Denote  $r(s) = (s-1)\zeta(s)$ . Then  $\ln(s-1) + \ln \zeta(s) = \ln r(s)$ , so that

$$\frac{\ln \zeta(s)}{\ln(s-1)^{-1}} = 1 + \frac{\ln r(s)}{\ln(s-1)^{-1}}.$$

By the proposition above,  $r(s) \rightarrow 1$  as  $s \rightarrow 1$ . Hence  $\ln r(s) \rightarrow 0$  and we deduce the result.  $\square$

**Proposition 1.1.5.**

$$\ln \zeta(s) = \sum_p p^{-s} + R(s)$$

where  $R(s)$  is bounded as  $s \rightarrow 1$ .

*Proof.* By proposition 1.1.2, we have  $\zeta(s) = \prod_{p \leq N} (1 - p^{-1})^{-1} a_N(s)$ , with  $a_N(s) \rightarrow 1, N \rightarrow \infty$ .

We then have

$$\ln \zeta(s) = \sum_{p \leq N} \sum_{m=1}^N m^{-1} p^{-ms} + \ln a_N(s)$$

and, taking the limit for  $N \rightarrow \infty$ ,

$$\ln \zeta(s) = \sum_p p^{-s} + \sum_p \sum_{m=2}^{\infty} m^{-1} p^{-ms},$$

where the second sum is less than  $\sum_p \sum_{m=2}^{\infty} p^{-ms} = \sum_p p^{-2s} (1 - p^{-s})^{-1} \leq (1 - 2^{-s})^{-1} \sum_p p^{-2s} \leq 2\zeta(2)$ .  $\square$

If  $s \in \mathbb{C}$ , from the definition we see that  $\zeta(s)$  is convergent for  $\operatorname{Re} s > 1$ .

**Proposition 1.1.6.** *The function  $\zeta(s) - (s-1)^{-1}$  can be continued to an analytic function on  $\{s \in \mathbb{C}, \operatorname{Re} s > 0\}$ .*

*Proof.* Assume  $\operatorname{Re} s > 1$ . Then, using the lemma below, one can write

$$\begin{aligned} \zeta(s) &= \sum_{n=1}^{\infty} n^{-s} = \sum_{n=1}^{\infty} n(n^{-s} - (n+1)^{-s}) = s \sum_{n=1}^{\infty} n \int_n^{n+1} x^{-s-1} dx = \\ &= s \sum_{n=1}^{\infty} \int_n^{n+1} [x] x^{-s-1} dx = s \int_1^{\infty} [x] x^{-s-1} dx = s \int_1^{\infty} x^{-s} dx - s \int_1^{\infty} \{x\} x^{-s-1} dx = \end{aligned}$$

$= \frac{s}{s-1} - s \int_1^\infty \{x\} x^{-s-1} dx$ , where  $[x]$  is the integral part of a real number  $x$  and  $\{x\} = x - [x]$  is its fractional part. Since  $0 \leq \{x\} \leq 1$  the last integral converges and defines an analytic function for  $Re s > 0$  and the result follows.  $\square$

**Lemma 1.1.7.** *Let  $(a_n), (b_n)$  be two sequences of complex numbers such that  $\sum a_n b_n$  converges. Let  $A_n = \sum_1^n a_i$  and suppose  $A_n b_n \rightarrow 0, n \rightarrow \infty$ . Then*

$$\sum_{n=1}^{\infty} a_n b_n = \sum_{n=1}^{\infty} A_n (b_n - b_{n+1}).$$

*Proof.* Let  $S_N = \sum_{n=1}^N a_n b_n$  and  $A_0 = 0$ . Then

$$S_N = \sum_{n=1}^N (A_n - A_{n-1}) b_n = \sum_{n=1}^N A_n b_n - \sum_{n=1}^N A_{n-1} b_n = \sum_{n=1}^N A_n b_n - \sum_{n=1}^{N-1} A_n b_{n+1} = A_N b_N + \sum_{n=1}^{N-1} A_n (b_n - b_{n+1}).$$

The result follows taking the limit as  $N \rightarrow \infty$ .  $\square$

The following formula will be useful:

**Corollary 1.1.8.** *For  $Re s > 0, N \geq 1$*

$$\zeta(s) = \sum_{n=1}^N \frac{1}{n^s} + \frac{N^{1-s}}{1-s} - \frac{1}{2} N^{-s} + s \int_N^\infty \rho(x) x^{-s-1} dx,$$

with  $\rho(x) = \frac{1}{2} - \{x\}$ .

*Proof.* Write

$$\begin{aligned} \zeta(s) - \left( \sum_{n=1}^N \frac{1}{n^s} + \frac{N^{1-s}}{1-s} - \frac{1}{2} N^{-s} + s \int_N^\infty \rho(x) x^{-s-1} dx \right) &= \\ &= \frac{s}{s-1} - \sum_{n=1}^N \frac{1}{n^s} - \frac{N^{1-s}}{1-s} + \frac{1}{2} N^{-s} - s \int_1^N (x - [x]) x^{-s-1} dx + \int_N^\infty \frac{(-s)x^{-s-1}}{2} dx = \\ &= \frac{s}{s-1} - \sum_{n=1}^N \frac{1}{n^s} - \frac{N^{1-s}}{1-s} + \frac{1}{2} N^{-s} - \int_1^N s x^{-s} dx + \sum_{n=1}^{N-1} n \left( \frac{1}{n^s} - \frac{1}{(n+1)^s} \right) - \frac{1}{2} N^s = \\ &= \frac{s}{s-1} - \sum_{n=1}^N \frac{1}{n^s} - \frac{N^{1-s}}{1-s} + \frac{s N^{-s+1}}{s-1} - \frac{s}{s-1} + \sum_{n=1}^{N-1} \frac{1}{n^s} - \frac{N-1}{N^s} = 0. \end{aligned}$$

$\square$



### 1.1.3 Characters

Let  $A$  be an abelian group.

**Definition 1.1.9.** A **character** on  $A$  is a group homomorphism  $A \rightarrow \mathbb{C}^*$ . The set of characters is denoted by  $\hat{A}$ .

Note that  $\hat{A}$  is an abelian group: if  $\chi, \psi \in \hat{A}$  we define  $\chi\psi$  by  $a \mapsto \chi(a)\psi(a)$ . The trivial character  $\chi_0$ , defined by  $\chi_0(a) = 1$  for all  $a \in A$ , is the neutral element of the group. Finally, for  $\chi \in \hat{A}$  we define  $\chi^{-1}$  as the character given by  $a \mapsto \chi(a)^{-1}$ .

If  $A$  is a finite group of order  $n$ , we have  $a^n = e$  for any  $a \in A$  hence the values of  $\chi$  are the roots of unity and  $\overline{\chi(a)} = \chi(a)^{-1} = \chi^{-1}(a)$ .

**Proposition 1.1.10.** *Let  $A$  be a finite abelian group. Then  $A \simeq \hat{\hat{A}}$ .*

*Proof.* Suppose first that  $A$  is cyclic, generated by an element  $g$  of order  $n$ . Then any character  $\chi$  is uniquely defined by its value  $\chi(g)$ . Since  $\chi(g)$  is a root of unity, there are at most  $n$  characters. Now, if  $\xi_n = e^{2\pi i/n}$  and  $\lambda$  is a character such that  $\lambda(g) = \xi_n$ , we obtain that the powers  $\lambda^k$ ,  $k = 1, \dots, n$  are distinct characters, hence  $\hat{A}$  is a cyclic group generated by  $\lambda$ . In the general case, since any finite abelian group is a direct product of cyclic groups, it is enough to check that if  $A \simeq A_1 \times A_2$ , then  $\hat{A} \simeq \hat{A}_1 \times \hat{A}_2$ , that we leave as an exercise. □

**Proposition 1.1.11.** *Let  $A$  be a finite abelian group and  $\chi, \psi \in \hat{A}$ ,  $a, b \in A$ . Then*

$$(i) \sum_{a \in A} \chi(a) \overline{\psi(a)} = n\delta(\chi, \psi)$$

$$(ii) \sum_{\chi \in \hat{A}} \chi(a) \overline{\chi(b)} = n\delta(a, b).$$

*Proof.* (i) We have  $\sum_{a \in A} \chi(a) \overline{\psi(a)} = \sum_a \chi\psi^{-1}(a)$ . It is enough to show that  $\sum_a \chi_0(a) = n$  and  $\sum_a \chi(a) = 0$  if  $\chi \neq \chi_0$ . The first assertion follows from the definition of  $\chi_0$ . For the second, we have that there is  $b \in A$ ,  $\chi(b) \neq \chi_0(b) = 1$ . Then  $\sum_a \chi(a) = \sum_a \chi(ba) = \chi(b) \sum_a \chi(a)$  and the result follows.

(ii) The proof is similar to (i), using that if  $a$  is nonzero in  $A$ , there is a character  $\psi$  such that  $\psi(a) \neq 0$ . We leave it as an exercise. □

**Definition 1.1.12.** A **Dirichlet character mod  $m$**  is a character for  $A = (\mathbb{Z}/m\mathbb{Z})^*$  the group of units in the ring  $\mathbb{Z}/m\mathbb{Z}$ .

Note that Dirichlet characters mod  $m$  induce, and are induced from the characters  $\chi : \mathbb{Z} \rightarrow \mathbb{C}^*$  such that

$$(i) \chi(n + m) = \chi(n) \text{ for all } n \in \mathbb{Z};$$

$$(ii) \chi(kn) = \chi(k)\chi(n) \text{ for all } k, n \in \mathbb{Z};$$

(iii)  $\chi(n) \neq 0$  if and only if  $(n, m) = 1$ .

Since the order of the group  $(\mathbb{Z}/m\mathbb{Z})^*$  is the value of Euler's function  $\phi(m)$ , there are  $\phi(m)$  Dirichlet characters mod  $m$ . The proposition above gives in this case:

**Proposition 1.1.13.** *Let  $\chi$  and  $\psi$  be Dirichlet characters modulo  $m$  and  $a, b \in \mathbb{Z}$ . Then*

$$(i) \sum_{a=0}^{m-1} \chi(a)\overline{\psi(a)} = \phi(m)\delta(\chi, \psi)$$

$$(ii) \sum_{\chi} \chi(a)\overline{\chi(b)} = \phi(m)\delta(a, b).$$

### 1.1.4 L-functions

Let  $\chi$  be a Dirichlet character modulo  $m$ .

**Definition 1.1.14.** The **Dirichlet  $L$ -function associated to  $\chi$**  is

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

Note that since  $|\chi(n)n^{-s}| \leq n^{-s}$ , the function  $L(s, \chi)$  converges and is continuous for  $s > 1$ .

**Proposition 1.1.15.** (i)  $L(s, \chi) = \prod_p (1 - \chi(p)p^{-s})^{-1}$ ;

$$(ii) L(s, \chi_0) = \prod_{p|m} (1 - p^{-s})\zeta(s).$$

(iii)  $\lim_{s \rightarrow 1} (s - 1)L(s, \chi_0) = \phi(m)/m$ . In particular,  $L(s, \chi_0) \rightarrow \infty$  as  $s \rightarrow 1$ .

*Proof.* The statement (i) follows as in proposition 1.1.2. For(ii) write

$$L(s, \chi_0) = \prod_{(p,m)=1} (1 - \chi_0(p)p^{-s})^{-1} = \prod_{p|m} (1 - p^{-s})^{-1} \prod_p (1 - p^{-s})^{-1} = \prod_{p|m} (1 - p^{-s})\zeta(s)$$

using proposition 1.1.2 again. To establish (iii) we use proposition 1.1.3 and we obtain

$$\lim_{s \rightarrow 1} (s - 1)L(s, \chi_0) = \prod_{p|m} (1 - p^{-1}) = \phi(m)/m.$$

□

**Proposition 1.1.16.** *Let  $\chi$  be a nontrivial Dirichlet character modulo  $m$ . Then  $L(s, \chi)$  can be continued to an analytic function for  $\operatorname{Re} s > 0$ .*

*Proof.* Let  $S(x) = \sum_{n \leq x} \chi(n)$ . By lemma 1.1.7, we have

$$L(s, \chi) = \sum_{n=1}^{\infty} S(n)(n^{-s} - (n+1)^{-s}) = s \sum_{n=1}^{\infty} S(x) \int_n^{n+1} x^{-s-1} dx =$$

$= s \int_1^{\infty} S(x)x^{-s-1} dx$ . By lemma below,  $|S(x)| \leq \phi(m)$  for all  $x$ . Hence the above integral converges and defines an analytic function for  $Re s > 0$ .  $\square$

**Lemma 1.1.17.** *Let  $\chi$  be a nontrivial character modulo  $m$ . For any  $N > 0$*

$$\left| \sum_{n=0}^N \chi(n) \right| \leq \phi(m).$$

*Proof.* Let  $N = qm + r, 0 \leq r < m$ . Since  $\chi(n+m) = \chi(n)$  and  $\sum_{n=0}^{m-1} \chi(n) = 0$  by the orthogonality relations, we obtain

$$\left| \sum_{n=0}^N \chi(n) \right| = \left| q \sum_{n=0}^{m-1} \chi(n) + \sum_{n=0}^r \chi(n) \right| \leq \left| \sum_{n=0}^r \chi(n) \right| \leq \sum_{n=0}^{m-1} |\chi(n)| = \phi(m).$$

$\square$

We now study Gauss sums associated to Dirichlet characters.

**Definition 1.1.18.** For  $\chi$  a Dirichlet character we define  $G(s, \chi) = \sum_p \sum_{k \geq 1} \frac{\chi(p^k)p^{-ks}}{k}$ .

Note that since  $|\frac{\chi(p^k)p^{-ks}}{k}| \leq p^{-ks}$  and  $\zeta(s)$  converges for  $s > 1$ , the same holds for  $G(s, \chi)$ .

**Proposition 1.1.19.** (i) For  $s > 1$ ,  $exp G(s, \chi) = L(s, \chi)$ ;

(ii)  $G(s, \chi) = \sum_{(p,m)=1} \chi(p)p^{-s} + R_{\chi}(s)$ , where  $R_{\chi}(s)$  is bounded as  $s \rightarrow 1$ ;

(iii)

$$\sum_{\chi} \overline{\chi(a)} G(s, \chi) = \phi(m) \sum_{p \equiv a(m)} p^{-s} + R_{\chi,a}(s), \quad (1.2)$$

where  $R_{\chi,a}(s)$  is bounded as  $s \rightarrow 1$ .

(iv)  $\lim_{s \rightarrow 1} G(s, \chi_0) / \ln(s-1)^{-1} = 1$ .

*Proof.* Note that for  $z \in \mathbb{C}, |z| < 1$  one has

$$exp\left(\sum_{k=1}^{\infty} \frac{z^k}{k}\right) = (1-z)^{-1}.$$

So that, for  $z = \chi(p)p^{-s}$  we obtain  $exp(\sum_{k=1}^{\infty} \frac{\chi(p^k)p^{-ks}}{k}) = (1 - \chi(p)p^{-1})^{-1}$  and we deduce (i).

The proof of (ii) is similar to proposition 1.1.5. To get (iii), we multiply the both sides of (ii) by  $\overline{\chi(a)}$  and sum over all Dirichlet characters modulo  $m$ :

$$\sum_{\chi} \overline{\chi(a)} G(s, \chi) = \sum_{(p,m)=1} p^{-s} \sum_{\chi} \overline{\chi(a)} \chi(p) + \sum_{\chi} \overline{\chi(a)} R_{\chi}(s).$$

By proposition 1.1.13, we obtain

$$\sum_{\chi} \overline{\chi(a)} G(s, \chi) = \phi(m) \sum_{p \equiv a(m)} p^{-s} + R_{\chi,a}(s),$$

where  $R_{\chi,a}(s)$  is bounded as  $s \rightarrow 1$ , as required.

For (iv), we use that  $L(s, \chi_0) = \prod_{p|m} (1 - p^{-s}) \zeta(s)$ . Hence  $G(s, \chi_0) = \sum_{p|m} \ln(1 - p^{-s}) + \ln \zeta(s)$ , so that the statement follows from Proposition 1.1.3.  $\square$

In particular, from (i) we obtain that the series  $G(s, \chi)$  provides a definition for  $\ln L(s, \chi)$ , with no choice of branch involved. Understanding the behaviour of  $G(s, \chi)$  for  $\chi$  non trivial is the crucial technical step in the proof of Dirichlet theorem. We present here a proof due to de la Vallée Poissin (1896).

**Proposition 1.1.20.** *Let  $F(s) = \prod_{\chi} L(s, \chi)$  where the product is over all Dirichlet characters modulo  $m$ . Then, for  $s$  real and  $s > 1$  we have  $F(s) \geq 1$ .*

*Proof.* By definition,  $G(s, \chi) = \sum_p \sum_{k \geq 1} \frac{\chi(p^k) p^{-ks}}{k}$ . Summing over  $\chi$  and using Proposition 1.1.13, we obtain

$$\sum_{\chi} G(s, \chi) = \phi(m) \sum_{p^k \equiv 1(m)} \frac{1}{k} p^{-ks}.$$

The right-hand side of this equation is positive, taking the exponential, we obtain  $\prod_{\chi} L(s, \chi) \geq 1$ .  $\square$

**Theorem 1.1.21.** *Let  $\chi$  be a nontrivial Dirichlet character modulo  $m$ . Then  $L(1, \chi) \neq 0$ .*

*Proof.* We first consider the case when  $\chi$  is a complex character. By definition, for  $s$  real, we have  $\overline{L(s, \chi)} = L(s, \bar{\chi})$ . Letting  $s \rightarrow 1$  we see that  $L(1, \chi) = 0$  implies  $L(1, \bar{\chi}) = 0$ . Assume  $L(1, \chi) = 0$ . Since  $L(s, \chi)$  and  $L(s, \bar{\chi})$  have zero at  $s = 1$ ,  $L(s, \chi_0)$  has a simple pole at  $s = 1$  by Proposition 1.1.15(iii) and the other factors are analytic at around  $s = 1$  we obtain  $F(1) = 0$ . But from Proposition 1.1.20, for  $s$  real and  $s > 1$  we have  $F(s) \geq 1$ , contradiction.

The case when  $\chi$  is nontrivial real character (i.e.  $\chi(n) = 0, 1$  or  $-1$ ) is more difficult. For such character assume  $L(1, \chi) = 0$  and consider

$$\psi(s) = \frac{L(s, \chi) L(s, \chi_0)}{L(2s, \chi_0)}.$$

Note that  $\psi(s)$  is analytic for  $\text{Re } s > 1/2$ : in fact, the zero of  $L(s, \chi)$  at  $s = 1$  cancels the simple pole of  $L(s, \chi_0)$  and the denominator is analytic for  $\text{Re } s > 1/2$ . Moreover, since  $L(2s, \chi_0)$  has a simple pole at  $s = 1$  we have that  $\psi(s) \rightarrow 0, s \rightarrow 1/2$ .

**Lemma 1.1.22.** *For  $s$  real and  $s > 1$  we have  $\psi(s) = \sum_{n=1}^{\infty} a_n n^{-s}$  where  $a_1 = 1, a_n \geq 0$  and the series is convergent for  $s > 1$ .*

*Proof.* We have

$$\psi(s) = \prod_p (1 - \chi(p)p^{-s})^{-1} (1 - \chi_0(p)p^{-s})^{-1} (1 - \chi_0(p)p^{-2s}) = \prod_{p|m} \frac{1 - p^{-2s}}{(1 - p^{-s})(1 - \chi(p)p^{-s})}.$$

If  $\chi(p) = -1$ , the  $p$ -factor is equal to 1. Hence

$$\psi(s) = \prod_{\chi(p)=1} \frac{1 + p^{-s}}{1 - p^{-s}}.$$

We have  $\frac{1+p^{-s}}{1-p^{-s}} = (1+p^{-s}) \sum_{k=0}^{\infty} p^{-ks} = 1 + 2p^{-s} + 2p^{-2s} + \dots$ . Applying lemma 1.1.23 below yields the result.  $\square$

Expanding  $\psi(s)$  (as a function of a complex variable) as a power series around  $s = 2$ , we obtain

$$\psi(s) = \sum_{m=0}^{\infty} b_m (s - 2)^m.$$

Since  $\phi(s)$  is analytic, the radius of convergence of this power series is at least  $3/2$ . We have

$$b_m = \psi^{(m)}(2)/m! = \sum_{n=1}^{\infty} a_n (-\ln n)^m n^{-2} = (-1)^m c_m, c_m \geq 0.$$

Hence  $\phi(s) = \sum_{m=0}^{\infty} c_m (2 - s)^m$  and  $c_0 = \psi(2) = \sum_{n=1}^{\infty} a_n n^{-2} \geq a_1 = 1$ . Hence for  $s$  real in  $(\frac{1}{2}, 2)$  we have  $\psi(s) \geq 1$ , contradiction with  $\psi(s) \rightarrow 0$  as  $s \rightarrow 1/2$ . This finishes the proof of the theorem.  $\square$

**Lemma 1.1.23.** *Let  $f$  be a nonnegative function on  $\mathbb{Z}$  such that  $f(mn) = f(m)f(n)$  for all  $(m, n) = 1$ . Assume that there is a constant  $c$  such that  $f(p^k) < c$  for all prime powers  $p^k$ . Then*

(i)  $\sum_{n=1}^{\infty} f(n)n^{-s}$  converges for all real  $s > 1$ ;

(ii)  $\sum_{n=1}^{\infty} f(n)n^{-s} = \prod_p (1 + \sum_{k=1}^{\infty} f(p^k)p^{-ks})$ .

*Proof.* Let  $s > 1$  and  $a(p) = \sum_{k=1}^{\infty} f(p^k)p^{-ks}$ . Then

$$a(p) < cp^{-s} \sum_{k=0}^{\infty} p^{-ks} = cp^{-s}(1 - p^{-s})^{-1},$$

so that  $a(p) < 2cp^{-s}$ . Since for  $x > 0$  we have  $1 + x < \exp x$  we deduce

$$\prod_{p \leq N} (1 + a(p)) < \prod_{p \leq N} \exp a(p) = \exp \sum_{p \leq N} a(p) < \exp(2c \sum_p p^{-s}) := M.$$

By the definition of  $a(p)$  and the multiplicativity of  $f$  we deduce

$$\sum_{n=1}^{\infty} f(n)n^{-s} < \prod_{p \leq N} (1 + a(p)) < M.$$

Since  $f$  is nonnegative, we obtain (i). We deduce (ii) similarly to Proposition 1.1.2.  $\square$

We now deduce as a corollary:

**Proposition 1.1.24.** *If  $\chi$  is a nontrivial character modulo  $m$ , then  $G(s, \chi)$  remains bounded as  $s \rightarrow 1$  through real values  $s > 1$ .*

*Proof.* Since  $L(1, \chi) \neq 0$  by theorem 1.1.21, there is a disk  $D$  around  $L(1, \chi)$ , not containing 0. Let  $\ln z$  be a single-valued branch of the logarithm, defined on  $D$ . Let  $\delta > 0$  be such that  $L(s, \chi) \in D$  for  $s \in (1, 1 + \delta)$ . Then for  $s$  in this interval the exponential of both functions  $\ln L(s, \chi)$  and  $G(s, \chi)$  is  $L(s, \chi)$ . Hence, there is an integer  $N$  such that for  $s \in (1, 1 + \delta)$  one has

$$G(s, \chi) = 2\pi i N + \ln L(s, \chi),$$

so that  $\lim_{s \rightarrow 1} G(s, \chi)$  exists and is equal to  $2\pi i N + \ln L(1, \chi)$ , in particular  $G(s, \chi)$  is bounded.  $\square$

### 1.1.5 Proof of Dirichlet theorem

**Definition 1.1.25.** Let  $S$  and  $T$  be two sets of positive integers, with  $T$  infinite. The **upper natural density** and **lower natural density** of  $S$  in  $T$  are defined as

$$\limsup_{N \rightarrow \infty} \frac{\#\{n \in S, n \leq N\}}{\#\{n \in T, n \leq N\}} \text{ and } \liminf_{N \rightarrow \infty} \frac{\#\{n \in S, n \leq N\}}{\#\{n \in T, n \leq N\}}.$$

If the upper and lower densities coincide, the common value is called **the natural density** of  $S$  in  $T$ .

**Definition 1.1.26.** Let  $S$  and  $T$  be two sets of positive integers, with  $\sum_{n \in T} n^{-1}$  divergent. The **upper Dirichlet density** and **lower Dirichlet density** of  $S$  in  $T$  are defined as

$$\limsup_{s \rightarrow +1} \frac{\sum_{n \in S} n^{-s}}{\sum_{n \in T} n^{-s}} \text{ and } \liminf_{s \rightarrow +1} \frac{\sum_{n \in S} n^{-s}}{\sum_{n \in T} n^{-s}}.$$

If the upper and lower densities coincide, the common value is called **the Dirichlet density**  $d(S)$  of  $S$  in  $T$ .

Note that Proposition 1.1.5 implies that a subset  $S$  of the set of all primes  $\mathcal{P}$  has Dirichlet density if

$$\lim_{s \rightarrow 1} \frac{\sum_{p \in S} p^{-s}}{\ln(s-1)^{-1}}$$

exists.

The following properties are straightforward:

**Proposition 1.1.27.** *Let  $S \subset \mathcal{P}$ .*

- (i) *If  $S$  is finite, then  $d(S) = 0$ ;*
- (ii) *If  $S$  consists of all but finitely many primes, then  $d(S) = 1$ ;*
- (iii) *If  $S = S_1 \cup S_2$ , where  $S_1$  and  $S_2$  are disjoint and  $d(S_1)$  and  $d(S_2)$  both exist, then  $d(S) = d(S_1) + d(S_2)$ .*

We will prove a more precise statement of the Dirichlet theorem:

**Theorem 1.1.28** (Dirichlet). *Let  $a, m \in \mathbb{Z}$ ,  $(a, m) = 1$ . Let*

$$\mathcal{P}(a, m) = \{p \text{ prime}, p \equiv a \pmod{m}\}.$$

*Then  $d(\mathcal{P}(a, m)) = \frac{1}{\phi(m)}$ , in particular, this set is infinite.*

*Proof.* Recall the identity (1.2):

$$\sum_{\chi} \overline{\chi(a)} G(s, \chi) = \phi(m) \sum_{p \equiv a(m)} p^{-s} + R_{\chi, a}(s),$$

where  $R_{\chi, a}(s)$  is bounded as  $s \rightarrow 1$ . We divide this identity by  $\ln(s-1)^{-1}$  and take the limit as  $s \rightarrow 1$ . By Proposition 1.1.24, the limit of the left-hand side is 1, and the limit of the right-hand-side is  $\phi(m)d(\mathcal{P}(a, m))$ . We obtain  $d(\mathcal{P}(a, m)) = \frac{1}{\phi(m)}$  as claimed.  $\square$

## 1.2 Zeta function

Deeper properties concerning the distribution of primes are related to the properties of the zeta function. We continue investigating these properties using tools from real and complex analysis.

### 1.2.1 Fourier analysis

**Definition 1.2.1.** If  $f \in L^1(\mathbb{R})$  we denote

$$\hat{f}(y) = \int_{\mathbb{R}} f(x)e^{-2\pi ixy} dx.$$

Examples:

- (Fourier inversion formula) If  $f, \hat{f} \in L^1(\mathbb{R})$ , then  $f(x) = \int_{\mathbb{R}} \hat{f}(y)e^{2\pi ixy} dy$ .
- for  $f(x) = e^{-\pi x^2}$  one has  $\hat{f}(y) = e^{-\pi y^2}$ , i.e. one could think about this function as being 'self-dual'.

Let  $\mathcal{L} \subset L^1(\mathbb{R})$  be the vector space of twice continuously differentiable functions, such that the functions  $f, f', f''$  are rapidly decreasing (i.e. as  $x^{-(1+\eta)}$  for some  $\eta > 0$ .)

**Theorem 1.2.2.** (*Poisson summation formula*) For  $f \in \mathcal{L}$ , we have

$$\sum_{m \in \mathbb{Z}} f(m) = \sum_{n \in \mathbb{Z}} \hat{f}(n).$$

For the proof, see for example section 11.4.2 in [S. J. Miller and R. Takloo-Bighash, *An Invitation to Modern Number Theory*]. The formula holds under weaker assumptions, but the version above is enough for applications here.

Recall that the  $\Gamma$ -function is defined for  $Re(s) > 0$  by

$$\Gamma(s) = \int_0^{\infty} t^{s-1} e^{-t} dt.$$

One has the following properties (see the next section for some of proofs):

- $\Gamma(n+1) = n!$  and  $\Gamma(1) = 1$ .
- $\Gamma(s)$  has a meromorphic continuation to the entire complex plane with simple poles at  $s = 0, -1, -2, \dots$  and the residue at  $s = -k$  is  $\frac{(-1)^k}{k!}$ .
- (reflexion formula)  $\Gamma(s)\Gamma(1-s) = \frac{\pi}{\sin(\pi s)}$ .
- (functional equation)  $\Gamma(s+1) = s\Gamma(s)$ .
- (duplication formula)  $\Gamma(s)\Gamma(s + \frac{1}{2}) = 2^{1-2s} \pi^{\frac{1}{2}} \Gamma(2s)$ .

**Definition 1.2.3.** For  $Re(s) > 1$  define  $\xi(s) = \frac{1}{2}s(s-1)\Gamma(\frac{s}{2})\pi^{-s/2}\zeta(s)$ .

The following analytic continuation theorem is of high importance.



**Theorem 1.2.4.** (*Analytic continuation of the zeta function*) The function  $\xi(s)$  has an analytic continuation to an entire function and satisfies the functional equation

$$\xi(s) = \xi(1 - s).$$

*Proof.* By change of variables in the definition of the Gamma function we get

$$\int_0^\infty x^{\frac{1}{2}s-1} e^{-n^2\pi x} dx = \frac{\Gamma(\frac{s}{2})}{n^s \pi^{\frac{s}{2}}}.$$

Summing over  $n \in \mathbb{N}$ , for  $\operatorname{Re}(s) > 1$ , we obtain

$$\pi^{-\frac{s}{2}} \Gamma(\frac{s}{2}) \zeta(s) = \int_0^\infty x^{\frac{1}{2}s-1} \left( \sum_{n=1}^\infty e^{-n^2\pi x} \right) dx = \int_0^\infty x^{\frac{1}{2}s-1} w(x) dx,$$

where  $w(x) = \sum_{n=1}^\infty e^{-n^2\pi x}$ . Note that the absolute convergence of the sum justifies that one could exchange the order sum-integral in the first equality.

Dividing the last integral into two pieces for  $x > 1$  (resp.  $x < 1$ ) and changing variables by  $x \mapsto x^{-1}$  in the second we obtain:

$$\pi^{-\frac{s}{2}} \Gamma(\frac{s}{2}) \zeta(s) = \int_1^\infty x^{\frac{1}{2}s-1} w(x) dx + \int_1^\infty x^{-\frac{1}{2}s-1} w(\frac{1}{x}) dx.$$

By lemma below, one deduces from the functional equation for  $w(x)$  that

$$\pi^{-\frac{s}{2}} \Gamma(\frac{s}{2}) \zeta(s) = \frac{1}{s(s-1)} + \int_1^\infty (x^{\frac{1}{2}s-1} + x^{-\frac{1}{2}s-\frac{1}{2}}) w(x) dx.$$

Since  $w(x)$  is rapidly decreasing, the integral on the right converges absolutely for any  $s$  and defines an entire function of  $s$ . The remaining assertions follow from the location of poles of  $\frac{1}{s(s-1)}$  and the invariance of the right hand side of the last equality under the change  $s \mapsto 1 - s$ .  $\square$

**Lemma 1.2.5.** The function  $w(x) = \sum_{n=1}^\infty e^{-n^2\pi x}$  satisfies the functional equation

$$w(\frac{1}{x}) = -\frac{1}{2} - \frac{1}{2}x^{\frac{1}{2}} + x^{\frac{1}{2}}w(x).$$

*Proof.* Write  $w(x) = \frac{\theta(x)-1}{2}$  with  $\theta(x) = \sum_{n=-\infty}^{+\infty} e^{-\pi n^2 x}$ . Note that this series is converging rapidly for  $x > 0$ . By the Poisson summation formula, we have

$$\theta(x^{-1}) = \sum_{n=-\infty}^{+\infty} e^{-\pi n^2 x^{-1}} = \sum_{m=-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-\pi t^2 x^{-1} + 2\pi i m t} dt = x^{\frac{1}{2}} \theta(x)$$

and the functional equation for  $w(x)$  easily follows.  $\square$

**Remark 1.2.6.** Using the duplication and the reflexion formulas for the  $\Gamma$ -function, one could obtain the functional equation for the zeta function in the following form:

$$\zeta(s) = \frac{1}{\pi}(2\pi)^s \sin \frac{\pi s}{2} \Gamma(1-s) \zeta(1-s).$$

**Corollary 1.2.7.**  $\zeta(-2m) = 0$  for all  $m \in \mathbb{N}$ .

*Proof.* The result follows from the analytic continuation and the fact that the  $\Gamma$ -function has poles at  $-m$ ,  $m \in \mathbb{N}$ .  $\square$

The zeros  $-2m$  of the zeta function given in the corollary above are called the **trivial zeros**. For  $0 \leq \operatorname{Re}(s) \leq 1$  the functional equation implies that zeros must lie symmetrically around the critical line  $\operatorname{Re}(s) = \frac{1}{2}$ . The **Riemann Hypothesis** asserts that all zeros  $s$  of the zeta function with  $0 \leq \operatorname{Re}(s) \leq 1$  lie on the critical line.

## 1.2.2 Entire functions

In this section we discuss properties of entire functions with prescribed set of zeros. More details could be found in [A.Karatsuba, *Basic Analytic Number Theory*.]

**Theorem 1.2.8.** Let  $a_1, \dots, a_n, \dots$  be an infinite sequence of complex numbers with

$$0 < |a_1| \leq |a_2| \leq \dots \leq |a_n| \leq \dots$$

and  $\lim_{n \rightarrow \infty} \frac{1}{|a_n|} = 0$ . Then there exists an entire function  $g : \mathbb{C} \rightarrow \mathbb{C}$  whose set of zeros coincide with set  $\{a_n\}$  (with multiplicities).

*Proof.* For  $n = 1, 2, \dots$  we set

$$u_n = u_n(s) = \left(1 - \frac{s}{a_n}\right) \exp\left(\frac{s}{a_n} + \frac{1}{2}\left(\frac{s}{a_n}\right)^2 + \dots + \frac{1}{n-1}\left(\frac{s}{a_n}\right)^{n-1}\right).$$

Consider the infinite product  $\prod_{n=1}^{\infty} u_n(s)$ . Let us show that the product converges for any  $s \neq a_n$ , and defines an entire function  $g(s)$  with zeros  $a_1, \dots, a_n, \dots$ . Consider a disk of radius  $|a_n|$  and the product  $\prod_{r=n}^{\infty} u_r(s)$ . It is enough to establish that this product converges to an analytic function inside the disc  $|s| < |a_n|$ : in fact then the product  $\prod_{n=1}^{\infty} u_n(s)$  is an analytic function in this disk, having only zeros  $a_i$  with  $|a_i| < |a_n|$  and since  $|a_n| \rightarrow \infty$ , we deduce the theorem.

For  $|s| < |a_n|$ ,  $r \geq n$  we have

$$\ln u_r(s) = \ln\left(1 - \frac{s}{a_r}\right) + \frac{s}{a_r} + \frac{1}{2}\left(\frac{s}{a_r}\right)^2 + \dots + \frac{1}{r-1}\left(\frac{s}{a_r}\right)^{r-1}.$$

Hence for  $r = n, n+1, \dots$  and  $|s| < |a_n|$ ,

$$\ln u_r(s) = -\frac{1}{r}\left(\frac{s}{a_r}\right)^r - \frac{1}{r+1}\left(\frac{s}{a_r}\right)^{r+1} - \dots$$

and

$$u_r(s) = \exp\left(-\frac{1}{r}\left(\frac{s}{a_r}\right)^r - \frac{1}{r+1}\left(\frac{s}{a_r}\right)^{r+1} - \dots\right).$$

Hence it is enough to establish that the series

$$\sum_{r=n}^{\infty} \left[ \frac{1}{r} \left(\frac{s}{a_r}\right)^r + \frac{1}{r+1} \left(\frac{s}{a_r}\right)^{r+1} + \dots \right] \quad (1.3)$$

is absolutely convergent for  $|s| < |a_n|$ . But for any  $0 < \epsilon < \frac{1}{2}$  and  $|s| \leq (1 - \epsilon)|a_n|$  we have

$$\left| \frac{1}{r} \left(\frac{s}{a_r}\right)^r + \frac{1}{r+1} \left(\frac{s}{a_r}\right)^{r+1} + \dots \right| \leq \frac{1}{r} (1 - \epsilon)^r + \frac{1}{r+1} (1 - \epsilon)^{r+1} + \dots < \frac{(1 - \epsilon)^r}{\epsilon r}.$$

hence, using proposition 1.2.9 below, the series (1.3) is absolutely convergent for  $|s| \leq (1 - \epsilon)|a_n|$ , so that we obtain that  $\prod_{n=1}^{\infty} u_n(s)$  is analytic on  $\mathbb{C}$  and we finish the proof of the theorem.  $\square$

**Proposition 1.2.9.** *Let  $u_n(s)$ ,  $n \geq 1$  be an infinite sequence of analytic functions on the domain  $\Omega$ , such that*

- $u_n(s) \neq -1$  for all  $n$  and  $s \in \Omega$ ;
- $|u_n(s)| \leq a_n$  for all  $n$  and  $s \in \Omega$  and the series  $\sum_{n=1}^{\infty} a_n$  converges.

*The infinite product*

$$\prod_{n=1}^{\infty} (1 + u_n(s))$$

*converges for any  $s \in \Omega$  and defines an analytic function  $v(s)$  on  $\Omega$ , such that  $v(s) \neq 0$  for  $s \in \Omega$ .*

**Remark 1.2.10.** If  $\sum_{n=1}^{\infty} \frac{1}{|a_n|^{1+s}} < \infty$ , then the function

$$g(s) = \prod_{n=1}^{\infty} \left(1 - \frac{s}{a_n}\right) \exp\left(\sum_{j=1}^{\infty} \frac{1}{j} \left(\frac{s}{a_n}\right)^j\right)$$

satisfies the conditions of the theorem above.

One could also show that any entire function has the form

$$g(s) = e^{h(s)} s^m \prod_{n=1}^{\infty} \left(1 - \frac{s}{a_n}\right) \exp\left(\sum_{j=1}^{\infty} \frac{1}{j} \left(\frac{s}{a_n}\right)^j\right)$$

with  $h$  entire. This expression is more precise for functions of finite order.

**Definition 1.2.11.** Let  $g(s)$  be an entire function and let  $M(r) = M_g(r) = \max_{|s|=r} |g(s)|$ . We say that  $g$  is an entire function of **finite order** if there exists  $a > 0$  such that  $M(r) < \exp(r^a)$  for  $r > r_0(a)$  for some constant  $r_0(a)$ . We then call  $\alpha = \inf a$  the **order** of  $g(s)$ . If such  $a$  does not exist, we say that  $g$  is of infinite order.

**Definition 1.2.12.** Let  $s_1, \dots, s_n$  be a sequence of complex numbers, such that

$$0 < |s_1| \leq |s_2| \leq \dots \leq |s_n| \leq \dots$$

If there exists  $b > 0$  such that  $\sum_{n=1}^{\infty} |s_n|^{-b} < \infty$  then we say that  $(s_n)$  has a finite order of convergence, and we call  $\beta = \inf b$  the **order** of convergence. If such  $b$  does not exist, we say that the order of convergence of  $(s_n)$  is  $\infty$ .

We have the following properties:

**Theorem 1.2.13.** Let  $g(s)$  be an entire function of finite order  $\alpha$ , such that  $g(0) \neq 0$  and let  $s_1, \dots, s_n$  be zeros of  $g$  with  $0 < |s_1| \leq |s_2| \leq \dots \leq |s_n| \leq \dots$ . Then

- (i) the sequence  $s_n$  has a finite convergence order  $\beta \leq \alpha$ ;
- (ii)  $g(s) = e^{h(s)} s^m \prod_{n=1}^{\infty} (1 - \frac{s}{s_n}) \exp(\sum_{j=1}^p \frac{1}{j} (\frac{s}{s_n})^j)$ , where  $p \geq 0$  is the smallest integer such that  $\sum_{n=1}^{\infty} |s_n|^{-(p+1)} < \infty$  and  $h(s)$  is a polynomial of degree  $d \leq \alpha$  and  $\alpha = \max(d, \beta)$ .
- (iii) If, in addition, for any  $c > 0$  there is an infinite sequence  $r_1, \dots, r_n, \dots$  with  $r_n \rightarrow \infty$  such that

$$\max |g(s)| > \exp(cr_n^\alpha), |s| = r_n, n = 1, 2, \dots$$

then  $\alpha = \beta$  and the series  $\sum_{n=1}^{\infty} |s_n|^{-\beta}$  diverges.

### 1.2.3 $\Gamma$ -function.

In this section we use the properties of entire functions to study the  $\Gamma$ -function. We start with the following definition:

$$\frac{1}{\Gamma(s)} = s \cdot e^{\gamma s} \prod_{n \geq 1} (1 + \frac{s}{n}) e^{-\frac{s}{n}},$$

where  $\gamma$  is the Euler constant

$$\gamma = \lim_{N \rightarrow \infty} 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{N} - \log N.$$

Proposition 1.2.17 below says that this definition coincides with the one in the previous sections.

**Proposition 1.2.14.**

$$\Gamma(s) = \frac{1}{s} \prod_{n \geq 1} \left(1 + \frac{1}{n}\right)^s \left(1 + \frac{s}{n}\right)^{-1}.$$

*Proof.* From the definition of the infinite product and the  $\Gamma$ -function we have

$$\begin{aligned} \frac{1}{\Gamma(s)} &= \text{slim}_{m \rightarrow \infty} \exp\left(s\left(1 + \frac{1}{2} + \dots + \frac{1}{m} - \log(m)\right)\right) \cdot \text{slim}_{m \rightarrow \infty} \prod_{n=1}^m \left(1 + \frac{s}{n}\right) e^{-s/n} = \\ &= \text{slim}_{m \rightarrow \infty} m^{-s} \prod_{n=1}^m \left(1 + \frac{s}{n}\right) = \text{slim}_{m \rightarrow \infty} \prod_{n=1}^{m-1} \left(1 + \frac{1}{n}\right)^{-s} \prod_{n=1}^m \left(1 + \frac{s}{n}\right) = \\ &= \text{slim}_{m \rightarrow \infty} \prod_{n=1}^m \left(1 + \frac{1}{n}\right)^{-s} \left(1 + \frac{s}{n}\right) \left(1 + \frac{1}{m}\right)^s = s \prod_{n=1}^{\infty} \left(1 + \frac{1}{n}\right)^{-s} \left(1 + \frac{s}{n}\right). \end{aligned}$$

□

**Proposition 1.2.15.** (i)  $\Gamma(s) = \lim_{N \rightarrow \infty} \frac{1 \cdot 2 \cdot \dots \cdot (n-1)n^s}{s(s+1)\dots(s+n-1)}$ .

(ii)  $\Gamma(s+1) = s\Gamma(s)$ . In particular,  $\Gamma(n+1) = n!$ .

*Proof.* (i) is straightforward. For (ii), using the previous proposition, we obtain:

$$\begin{aligned} \frac{\Gamma(s+1)}{\Gamma(s)} &= \frac{s}{s+1} \lim_{m \rightarrow \infty} \prod_{n=1}^m \frac{\left(1 + \frac{1}{n}\right)^{s+1} \left(1 + \frac{s+1}{n}\right)^{-1}}{\left(1 + \frac{1}{n}\right)^s \left(1 + \frac{s}{n}\right)^{-1}} = \\ &= \frac{s}{s+1} \lim_{m \rightarrow \infty} \prod_{n=1}^m \frac{n+1}{n} \frac{n+s}{n+s+1} = \frac{s}{s+1} \lim_{m \rightarrow \infty} \frac{(m+1)(s+1)}{m+1+s} = s. \end{aligned}$$

□

**Proposition 1.2.16.** For  $s \in \mathbb{C} \setminus \mathbb{Z}$  one has  $\Gamma(s)\Gamma(1-s) = \frac{\pi}{\sin(\pi s)}$ .

*Proof.* The function  $\sin(\pi s)$  is an entire function of the first order, with zeros  $s = 0, \pm 1, \pm 2, \dots$ . By theorem 1.2.13, we can write

$$\sin(\pi s) = s e^{h(s)} \prod_{n=1}^{\infty} \left(1 - \frac{s^2}{n^2}\right),$$

with  $h(s) = as + b$ . We take a logarithmic differential:

$$\pi \frac{\cos(\pi s)}{\sin(\pi s)} = \frac{1}{s} + h'(s) - \sum_{n=1}^{\infty} \frac{2s}{n^2 - s^2}.$$

Taking limit for  $s \rightarrow 0$  we obtain  $a = 0$ , i.e.  $h(s) = b$ .

Next,  $\frac{\sin(\pi s)}{s} = c \prod_{n=1}^{\infty} (1 - \frac{s^2}{n^2})$ . Taking limit for  $s \rightarrow 0$  we obtain  $c = \pi$ , i.e.

$$\sin(\pi s) = \pi s \prod_{n=1}^{\infty} (1 - \frac{s^2}{n^2}).$$

By definition of the  $\Gamma$  function we have

$$\Gamma(s)\Gamma(-s) = -\frac{1}{s^2} \prod_{n=1}^{\infty} (1 - \frac{s^2}{n^2})^{-1} = -\frac{\pi}{s \sin(\pi s)}$$

and from previous proposition,  $\Gamma(1-s) = -s\Gamma(-s)$ . The result follows.  $\square$

**Proposition 1.2.17.**  $\Gamma(s) = \int_0^{\infty} e^{-t} t^{s-1} dt$ .

*Proof.* See Thm.4 p.53 in [A.Karatsuba, *Basic Analytic Number Theory*].  $\square$

We list also additional properties of the  $\Gamma$ -function:

1.  $\Gamma(\frac{1}{2}) = \sqrt{\pi}$  (exercise);
2. (Stirling's formula)  $\log \Gamma(s) = (s - \frac{1}{2}) \log(s) - s + \log \sqrt{2\pi} + \mathcal{O}(\frac{1}{|s|})$ .
3.  $-\frac{\Gamma'(s)}{\Gamma(s)} = \frac{1}{s} + \gamma + \sum_{n=1}^{\infty} [\frac{1}{n+s} - \frac{1}{n}]$

As a consequence,

$$\frac{\Gamma'(n)}{\Gamma(n)} = -\gamma + \sum_{k=1}^{n-1} \frac{1}{k}.$$

4.  $\Gamma(s)^{-1}$  is entire of order  $\alpha = 1$  and

$$\frac{\Gamma'(s)}{\Gamma(s)} = \log(s) + \mathcal{O}(\frac{1}{|s|}).$$

## 1.2.4 Zeros of the zeta function

Let  $\xi(s)$  be defined as in theorem 1.2.4.

**Theorem 1.2.18.** • *The function  $\xi(s)$  is an entire function of order 1 with infinitely many zeros  $\rho_n$  such that  $0 \leq \text{Re} \rho_n \leq 1$ ;*

- *the series  $\sum |\rho_n|^{-1}$  diverges;*
- *the series  $\sum |\rho_n|^{-1-\epsilon}$  converges for any  $\epsilon > 0$ ;*
- *the zeros of  $\xi(s)$  are nontrivial zeros of  $\zeta(s)$ .*

*Proof.* For  $Re(s) > 1$  the zeta function, and, hence, the function  $\xi(s)$  has no zeros. Theorem 1.2.4 implies that  $\xi(s) \neq 0$  for  $Re(s) > 0$  as well. Since  $\xi(0) = \xi(1) \neq 0$ , zeros of  $\xi(s)$  coincide with nontrivial zeros of  $\zeta(s)$ .

To determine the order of  $\xi(s)$ , we consider  $|s| \rightarrow \infty$ . By corollary 1.1.8,  $\zeta(s) = \mathcal{O}(|s|)$  for  $Re(s) \geq \frac{1}{2}$ . Since  $|\Gamma(s)| \leq e^{c|s|} |ln|s||$ , the order of  $\xi$  is at most one. But for  $s \rightarrow +\infty$ ,  $ln\Gamma(s) \equiv s ln(s)$ , so that the order of  $\xi(s)$  is 1. Theorem 1.2.13 imply that  $\sum |\rho_n|^{-1}$ , where  $\rho_n$  are zeros of  $\xi(s)$  is divergent. In particular,  $\xi(s)$  has infinitely many zeros, and the series  $\sum |\rho_n|^{-1-\epsilon}$  is convergent for any  $\epsilon > 0$ .  $\square$

**Corollary 1.2.19.** (i)  $\xi(s) = e^{A+B_s} \prod_{n=1}^{\infty} (1 - \frac{s}{\rho_n}) e^{\frac{s}{\rho_n}}$ ;

(ii) nontrivial zeros of zeta-function are symmetric with respect to the lines  $Re(s) = \frac{1}{2}$  and  $Im(s) = 0$ .

In what follows we enumerate zeros of zeta function in an increasing order (with respect to the absolute value).

**Proposition 1.2.20.**

$$\frac{\zeta'(s)}{\zeta(s)} = -\frac{1}{s-1} + \sum_{n=1}^{\infty} \left( \frac{1}{s-\rho_n} + \frac{1}{\rho_n} \right) + \sum_{n=1}^{\infty} \left( \frac{1}{s+2n} - \frac{1}{2n} \right) + B_0,$$

where  $\rho_n$  are all nontrivial zeros of  $\zeta(s)$  and  $B_0$  is a constant.

*Proof.* It is enough to take the logarithmic derivative in corollary 1.2.19(i).  $\square$

**Theorem 1.2.21.** Let  $\rho_n = \beta_n + i\gamma_n$ ,  $n = 1, 2, \dots$  are all nontrivial zeros of  $\zeta(s)$ ,  $T \geq 2$ . Then

$$\sum_{n=1}^{\infty} \frac{1}{1 + (T - \gamma_n)^2} \leq c \log T. \quad (1.4)$$

*Proof.* For  $s = 2 + iT$ , one has

$$\left| \sum_{n=1}^{\infty} \left( \frac{1}{s+2n} - \frac{1}{2n} \right) \right| \leq \sum_{n \leq T} \left( \frac{1}{2n} + \frac{1}{2n} \right) + \sum_{n > T} \frac{|s|}{4n^2} \leq c_0 \log(T), \quad (1.5)$$

so that by proposition 1.2.20

$$\begin{aligned} -Re\left(\frac{\zeta'(s)}{\zeta(s)}\right) &= Re\left(\frac{1}{s-1} - B_0 - \sum_{n=1}^{\infty} \left(\frac{1}{s+2n} - \frac{1}{2n}\right)\right) - \\ &\quad - Re \sum_{n=1}^{\infty} \left(\frac{1}{s-\rho_n} + \frac{1}{\rho_n}\right) \leq c_1 \log(T) - Re \sum_{n=1}^{\infty} \left(\frac{1}{s-\rho_n} + \frac{1}{\rho_n}\right). \end{aligned}$$

From the Euler product expression (proposition 1.1.2), we have

$$-\frac{\zeta'(s)}{\zeta(s)} = \sum_{n \geq 1} \frac{\Lambda(n)}{n^s}, \quad (1.6)$$

where  $\Lambda(n) = \log(p)$ ,  $n = p^k$  and 0 otherwise. Hence

$$\left| \frac{\zeta'(s)}{\zeta(s)} \right| = \left| \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^{2+iT}} \right| < c_2,$$

so that

$$\operatorname{Re} \sum_{n=1}^{\infty} \left( \frac{1}{s - \rho_n} + \frac{1}{\rho_n} \right) \leq c_3 \log(T).$$

We deduce the theorem from the following inequalities

$$\begin{aligned} \operatorname{Re} \frac{1}{s - \rho_n} &= \operatorname{Re} \frac{1}{(2 - \beta_n) + i(T - \gamma_n)} = \frac{2 - \beta_n}{(2 - \beta_n)^2 + (T - \gamma_n)^2} \geq \\ &\geq \frac{0.5}{1 + (T - \gamma_n)^2} \end{aligned}$$

and  $\operatorname{Re} \frac{1}{\rho} = \frac{\beta_n}{\beta_n^2 + \gamma_n^2} \geq 0$ . □

**Corollary 1.2.22.** *The number of zeros  $\rho_n$  of the zeta function, such that*

$$T \leq |\operatorname{Im}(\rho_n)| \leq T + 1$$

*is at most  $c \log(T)$ .*

**Corollary 1.2.23.** *For  $T \geq 2$ , one has  $\sum_{|T - \gamma_n| > 1} \frac{1}{|T - \gamma_n|^2} = \mathcal{O}(\log(T))$ .*

**Corollary 1.2.24.** *For  $-1 \leq \sigma \leq 2$ ,  $s = \sigma + it$ ,  $|t| \geq 2$ , one has*

$$\frac{\zeta'(s)}{\zeta(s)} = \sum_{|t - \gamma_n| \leq 1} \frac{1}{s - \rho_n} + \mathcal{O}(\log|t|).$$

*Proof.* The inequality 1.5 is valid for  $s = \sigma + it$ ,  $|t| \geq 2$ ,  $-1 \leq \sigma \leq 2$ , so that

$$\frac{\zeta'(s)}{\zeta(s)} = \sum_{n=1}^{\infty} \left( \frac{1}{s - \rho_n} + \frac{1}{\rho_n} \right) + \mathcal{O}(\log|t|).$$

We subtract the same inequality for  $s = 2 + it$ :

$$\frac{\zeta'(s)}{\zeta(s)} = \sum_{n=1}^{\infty} \left( \frac{1}{s - \rho_n} - \frac{1}{2 + it - \rho_n} \right) + \mathcal{O}(\log(|t|)).$$



If  $|\gamma_n - t| > 1$ , then

$$\left| \frac{1}{\sigma + it - \rho_n} - \frac{1}{2 + it - \rho_n} \right| \leq \frac{2 - \sigma}{(\gamma_n - t)^2} \leq \frac{3}{(\gamma_n - t)^2}.$$

Now the statement follows from the previous corollaries 1.2.22 and 1.2.23.  $\square$

**Theorem 1.2.25.** (*de la Vallée Poussin*) *There exists a constant  $c > 0$  such that the zeta function has no zeros for*

$$\operatorname{Re}(s) = \sigma \geq 1 - \frac{c}{\log(|t| + 2)}.$$

*Proof.* The function  $\zeta(s)$  has a pole at  $s = 1$ , hence for some  $\gamma_0$  there is no zeros  $s$  with  $|s - 1| \leq \gamma_0$ . Let  $\rho_n = \beta_n + i\gamma_n$  be a zero of  $\zeta$  with  $|\gamma_n| > |\gamma_0|$ . For  $\operatorname{Re}(s) = \sigma > 1$  we have as in (1.6)

$$-\frac{\zeta'(s)}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s} = \sum_{n=1}^{\infty} \Lambda(n) n^{-\sigma} e^{-it \log(n)},$$

so that

$$-\operatorname{Re} \frac{\zeta'(s)}{\zeta(s)} = \sum_{n=1}^{\infty} \Lambda(n) n^{-\sigma} \cos(t \log(n)).$$

Since for all real  $\phi$  we have

$$3 + 4\cos \phi + \cos 2\phi = 2(1 + \cos \phi)^2 \geq 0,$$

we deduce

$$3 \frac{\zeta'(\sigma)}{\zeta(\sigma)} + 4 \operatorname{Re} \frac{\zeta'(\sigma + it)}{\zeta(\sigma + it)} + (-\operatorname{Re} \frac{\zeta'(\sigma + i2t)}{\zeta(\sigma + i2t)}) \geq 0. \quad (1.7)$$

We will provide a majoration for each summand in the formula (1.7). By proposition 1.2.20 and corollary 1.2.22 for  $s = \sigma$  and  $1 < \sigma \leq 2$  we obtain

$$-\frac{\zeta'(s)}{\zeta(s)} < \frac{1}{\sigma - 1} + B_1,$$

where  $B_1$  is a constant. For  $s = \sigma + it, 1 < \sigma \leq 2, |t| > \gamma_0$  we find, again by proposition 1.2.20:

$$-\operatorname{Re} \frac{\zeta'(s)}{\zeta(s)} < A \log(|t| + 2) - \sum_{k=1}^{\infty} \operatorname{Re} \left( \frac{1}{s - \rho_k} + \frac{1}{\rho_k} \right),$$

where  $A > 0$  is an absolute constant. Since  $0 \leq \beta_k \leq 1$ , we have  $\rho_k = \beta_k + i\gamma_k$  we deduce

$$\operatorname{Re} \frac{1}{s - \rho_k} = \operatorname{Re} \frac{1}{\sigma - \beta_k + i(t - \gamma_k)} = \frac{\sigma - \beta_k}{(\sigma - \beta_k)^2 + (t - \gamma_k)^2},$$

in addition  $\operatorname{Re} \frac{1}{\rho_k} = \frac{\beta_k}{\beta_k^2 + \gamma_k^2} \geq 0$ . We deduce

$$-\operatorname{Re} \frac{\zeta'(\sigma + it)}{\zeta(\sigma + it)} < A \log(|t| + 2) - \frac{\sigma - \beta_n}{(\sigma - \beta_n)^2 + (t - \gamma_n)^2}$$

and

$$-\operatorname{Re} \frac{\zeta'(\sigma + it)}{\zeta(\sigma + it)} < A \log(2|t| + 2).$$

We now substitute these estimations in (1.7):

$$\frac{3}{\sigma - 1} - 4 \frac{\sigma - \beta_n}{(\sigma - \beta_n)^2 + (t - \gamma_n)^2} + A_1 \log(|t| + 2) \geq 0$$

for  $A_1 > 0$  a constant. This inequality works for any  $t$ ,  $|t| > \gamma_0$  and any  $\sigma$ ,  $1 < \sigma \leq 2$ . For instance, for  $t = \gamma_n$ ,  $\sigma = 1 + \frac{1}{2A_1 \log(|\gamma_n| + 2)}$ , so that

$$\frac{4}{\sigma - \beta_n} \leq \frac{3}{\sigma - 1} + A_1 \log(|\gamma_n| + 2),$$

$$\beta_n \leq 1 - \frac{1}{14A_1 \log(|\gamma_n| + 2)},$$

and we finish the proof of the theorem.  $\square$

**Corollary 1.2.26.** *Let  $T \geq 2$  and  $c > 0$  a constant. Then for*

$$\sigma \geq 1 - \frac{c}{2 \log(T + 2)}, \quad 2 \leq |t| \leq T$$

one has an estimation  $|\frac{\zeta'(s)}{\zeta(s)}| = \mathcal{O}(\log^2 T)$ , where  $s = \sigma + it$ .

*Proof.* Using corollary 1.2.24, we have

$$\left| \frac{\zeta'(s)}{\zeta(s)} \right| = \sum_{|t - \gamma_n| \leq 1} \frac{1}{s - \rho_n} + \mathcal{O}(\log(T)).$$

Hence

$$\left| \frac{\zeta'(s)}{\zeta(s)} \right| \leq \sum_{|t - \gamma_n| \leq 1} \frac{1}{|\sigma - \beta_n| + i(t - \gamma_n)} + \mathcal{O}(\log(T)).$$

Since  $\beta_n \leq 1 - \frac{c}{\log(T+2)}$  and  $\sigma \geq 1 - \frac{c}{2 \log(T+2)}$ , we have

$$\left| \frac{\zeta'(s)}{\zeta(s)} \right| \leq \frac{2}{c} \log(T + 2) \sum_{|t - \gamma_n| \leq 1} 1 + \mathcal{O}(\log(T)) = \mathcal{O}(\log^2 T).$$

.

$\square$

### 1.3 Distribution of primes

Let  $\pi(x) = \sum_{\substack{p \text{ prime} \\ p \leq x}} 1$ . We will be interested in the asymptotic description of this function. First we need some facts on the Dirichlet sums.

**Definition 1.3.1.** The Dirichlet series is a series of the form

$$f(s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s}, \quad (1.8)$$

where the coefficients  $a_n$  are complex numbers and  $s = \sigma + it$ .

To the Dirichlet series one associates the function

$$\Phi(x) = \sum_{n \leq x} a_n.$$

**Theorem 1.3.2.** [Tauberian theorem] Assume that the series (1.8) converges for  $\sigma > 1$ ,  $|a_n| \leq A(n)$  where  $A(n) > 0$  is a monotonic, increasing function and for  $\sigma \rightarrow 1 + 0$  one has

$$\sum_{n=1}^{\infty} |a_n| n^{-\sigma} = \mathcal{O}((\sigma - 1)^{-\alpha}), \alpha > 0.$$

Then for any  $b_0 \geq b > 1$ ,  $T \geq 1$ ,  $x = N + \frac{1}{2}$  the following formula holds

$$\Phi(x) = \sum_{n \leq x} a_n = \frac{1}{2\pi i} \int_{b-iT}^{b+iT} f(s) \frac{x^s}{s} ds + \mathcal{O}\left(\frac{x^b}{T(b-1)^\alpha}\right) + \mathcal{O}\left(\frac{x A(2x) \log(x)}{T}\right).$$

In addition, the constant in the  $\mathcal{O}$ -sign depends only on  $b_0$ .

*Proof.* First we prove that

$$\frac{1}{2\pi i} \int_{b-iT}^{b+iT} \frac{a^s}{s} ds = \epsilon + \mathcal{O}\left(\frac{a^b}{T|\log(a)|}\right) \quad (1.9)$$

where  $\epsilon = 1$  if  $a > 1$  and  $\epsilon = 0$  if  $0 < a < 1$ . Let us consider the case  $a > 1$  (we left the case  $0 < a < 1$  as an exercise). Consider  $U > b$  and the rectangular path  $\Gamma$  with sides  $[-U + iT, -U - iT]$ ,  $[-U - iT, b - iT]$ ,  $[b - iT, b + iT]$ ,  $[b + iT, -U + iT]$ .

By Cauchy theorem,  $\frac{1}{2\pi i} \int_{\Gamma} \frac{a^s ds}{s} = 1$ , so that

$$\frac{1}{2\pi i} \int_{b-iT}^{b+iT} \frac{a^s ds}{s} = 1 + R \quad (1.10)$$

where  $R$  is the the integral on the left, upper and bottom sides. The integrals on the upper and bottom sides have the same absolute value, so that on each of these sides we have

$$\frac{1}{2\pi i} \left| \int \frac{a^s ds}{s} \right| \leq \frac{1}{2\pi} \int_{-U}^b \frac{a^\sigma d\sigma}{\sqrt{T^2 + \sigma^2}} \leq \frac{a^b}{T \log(a)}.$$

Also we have for the left side

$$\frac{1}{2\pi} \int \frac{a^s ds}{s} \leq \frac{1}{2\pi} \int_{-T}^{+T} \frac{a^{-U} dt}{\sqrt{U^2 + t^2}} = \mathcal{O}(a^U) \rightarrow 0$$

for  $U \rightarrow \infty$ . Passing to the limit in (1.10) when  $U \rightarrow \infty$ , we obtain the formula (1.9).

The series (1.8) is absolutely convergent for  $s = b + it$ . We obtain, exchanging the integral-sum:

$$\frac{1}{2\pi i} \int_{b-iT}^{b+iT} f(s) \frac{x^s}{s} ds = \sum_{n=1}^{\infty} a_n \left( \frac{1}{2\pi i} \int_{b-iT}^{b+iT} \left(\frac{x}{n}\right)^s \frac{ds}{s} \right) = \sum_{n \leq x} a_n + R,$$

where

$$R = \mathcal{O}\left(\sum_{n=1}^{\infty} |a_n| \left(\frac{x}{n}\right)^b T^{-1} \left|\log \frac{x}{n}\right|^{-1}\right).$$

Note that, since  $x = N + 1/2$ , we have  $x/n \neq 1$  for an integer  $n$ . We divide the sum under the  $\mathcal{O}$ -sign into two parts. For the first part, take  $\frac{x}{n} \leq \frac{1}{2}$  or  $\frac{x}{n} \geq 2$ , so that  $|\log \frac{x}{n}| \geq 2$ . From the assumptions  $\sum_{n=1}^{\infty} \frac{|a_n|}{n^b} = \mathcal{O}\left(\frac{1}{(b-1)^\alpha}\right)$ , the first sum is  $\mathcal{O}\left(\frac{x^b}{T(b-1)^\alpha}\right)$ . The remaining part is

$$\sum_{\frac{1}{2}x < n < 2x} |a_n| \left(\frac{x}{n}\right)^b T^{-1} \left|\log \frac{x}{n}\right|^{-1} \leq T^{-1} A(2x) 2^b \sum_{\frac{1}{2}x < n < 2x} \left|\log \frac{N+0.5}{n}\right|^{-1}.$$

The summands with  $n = N - 1, N, N + 1$  in the last sum are of order  $\mathcal{O}(x)$  and for the remaining part  $r$  we obtain

$$r \leq \int_{x/2}^{N-1} \left(\log \frac{N+0.5}{u}\right)^{-1} du + \int_{N+1}^{2x} \left(\log \frac{u}{N+0.5}\right)^{-1} du = \mathcal{O}(x \log x).$$

and the theorem follows.  $\square$

Now we are ready to prove the prime number theorem. Recall that we defined  $\Lambda(n) = \log(p)$  if  $n = p^k$  and  $\Lambda(n) = 0$  otherwise.

**Theorem 1.3.3.** *There exists a constant  $c > 0$  such that*

$$\psi(x) = \sum_{n \leq x} \Lambda(x) = x + \mathcal{O}(xe^{-c\sqrt{\ln(x)}});$$

$$\pi(x) = \int_2^x \frac{du}{\ln(u)} + \mathcal{O}(xe^{-\frac{c}{2}\sqrt{\ln(x)}}).$$

*Proof.* For  $Re(s) > 1$ , using the Euler product argument (1.1) we write

$$-\frac{\zeta'(s)}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s}.$$

Using theorem 1.2.4, in the previous theorem we could take  $\alpha = 1$ ,  $A(n) = \log(n)$ .

Consider  $b = 1 + \frac{1}{\log(x)}$ ,  $T = e^{\sqrt{\log(x)}}$ . Then

$$\psi(x) = \frac{1}{2\pi i} \int_{b-iT}^{b+iT} \left(-\frac{\zeta'(s)}{\zeta(s)}\right) \frac{x^s}{s} ds + \mathcal{O}\left(\frac{x \log^2 x}{T}\right).$$

By theorem 1.2.25 and its corollary, for some constant  $c_1 > 0$ , the zeta function has no zeros with  $Re(s) = \sigma \geq \sigma_1 = 1 - \frac{c_1}{2\log(T+2)}$ ,  $|t| \leq T$ , and  $\frac{\zeta'(s)}{\zeta(s)} = \mathcal{O}(\log^2 T)$  for  $s = \sigma_1 + iT$ ,  $s = \sigma \pm iT$ ,  $\sigma_1 \leq \sigma \leq b$ ,  $s = b + it$ . Consider the integral

$$J = \frac{1}{2\pi i} \int_{\Gamma} \left(-\frac{\zeta'(s)}{\zeta(s)}\right) \frac{x^s}{s} ds$$

along the rectangle  $\Gamma$  with sides  $[\sigma_1 + iT, \sigma_1 - iT]$ ,  $[\sigma_1 - iT, b - iT]$ ,  $[b - iT, b + iT]$ ,  $[b + iT, \sigma_1 + iT]$ .

Since the only nontrivial pole inside  $\Gamma$  of the function  $\left(-\frac{\zeta'(s)}{\zeta(s)}\right) \frac{x^s}{s}$  is  $s = 1$  with residue  $x$ , we have

$$\frac{1}{2\pi i} \int_{b-iT}^{b+iT} \left(-\frac{\zeta'(s)}{\zeta(s)}\right) \frac{x^s}{s} ds = x + R$$

with  $R$  the sum of integrals along the left, upper and bottom sides. We will estimate these integrals. For the upper and bottom sides we have

$$\left| \frac{1}{2\pi i} \int_{\sigma_1+iT}^{b+iT} -\frac{\zeta'(s)}{\zeta(s)} \frac{x^s}{s} ds \right| \leq \int_{\sigma_1}^b \left| \frac{\zeta'(\sigma + iT)}{\zeta(\sigma + iT)} \right| \frac{x^\sigma}{T} d\sigma = \mathcal{O}\left(\frac{x \log^2 T}{T}\right),$$

and the integral by the left side is

$$\begin{aligned} \left| \frac{1}{2\pi i} \int_{\sigma_1-iT}^{\sigma_1+iT} -\frac{\zeta'(s)}{\zeta(s)} \frac{x^s}{s} ds \right| &= \left| \frac{1}{2\pi i} \int_{-T}^T \frac{\zeta'(\sigma_1 + it)}{\zeta(\sigma_1 + it)} \frac{x^{\sigma_1 + it}}{\sigma_1 + it} dt \right| = \\ &= \mathcal{O}(x^{\sigma_1} \log^2 T \left( \int_0^1 \frac{dt}{\sigma_1} + \int_1^T \frac{dt}{t} \right)) = \mathcal{O}(x^{\sigma_1} \log^3 T). \end{aligned}$$

From the inequalities above, the definition of  $T$  and  $\sigma_1$  we deduce the first assertion of the theorem.

Consider

$$S = \sum_{n \leq x} \frac{\Lambda(n)}{\log(n)} = \sum_{p \leq x} 1 + \sum_{n=p^k, k \geq 2} \frac{\Lambda(n)}{\log(n)}.$$

In the second sum  $k \leq \log(x)$  and for a fixed  $k$  we have at most  $\sqrt{x}$  summands,  $\leq 1$ . We deduce

$$S = \pi(x) + \mathcal{O}(\sqrt{x} \log(x)). \quad (1.11)$$

In the lemma 1.3.4 below we put  $c_n = \Lambda(n)$ ,  $f(x) = \frac{1}{\log(x)}$ , i.e.  $C(x) = \sum_{n \leq x} c_n = \psi(x) = x + \mathcal{O}(xe^{-c\sqrt{\log(x)}})$ ,  $f'(x) = -\frac{1}{x \log^2 x}$ , so that we obtain

$$S = \int_2^x \frac{\psi(u)}{u \log^2 u} du + \frac{\psi(x)}{\log(x)} = \int_2^x \frac{du}{\log^2 u} + \frac{x}{\log(x)} + R$$

with

$$\begin{aligned} R &= \mathcal{O}\left(\int_2^x e^{-c\sqrt{\log u}} \frac{du}{\log^2 u} + xe^{-c\sqrt{\log x}}\right) = \\ &= \mathcal{O}\left(\int_2^{\sqrt{x}} du + \int_{\sqrt{x}}^x e^{-c\sqrt{\log u}} du + xe^{-c\sqrt{\log x}}\right) = \mathcal{O}(xe^{-\frac{c}{2}\sqrt{\ln(x)}}) \end{aligned}$$

and

$$\int_2^x \frac{du}{\log^2 u} + \frac{x}{\log(x)} = -\frac{u}{\log(u)} \Big|_2^x + \int_2^x \frac{du}{\log(u)} + \frac{x}{\log(x)} = \int_2^x \frac{du}{\log(u)} + \frac{2}{\log 2}.$$

The theorem follows from this equality and (1.11). □

**Lemma 1.3.4.** (*Abel transform*) Let  $f(x)$  be a continuously differentiable function on the interval  $[a, b]$ ,  $c_n$  be complex numbers and

$$C(x) = \sum_{a < n \leq x} c_n.$$

Then

$$\sum_{a < n \leq b} c_n f(n) = - \int_a^b C(x) f'(x) dx + C(b) f(b).$$

*Proof.* We have

$$\begin{aligned} C(b)f(b) - \sum_{a < n \leq b} c_n f(n) &= \sum_{a < n \leq b} c_n (f(b) - f(n)) = \\ &= \sum_{a < n \leq b} \int_n^b c_n f'(x) dx = \sum_{a < n \leq b} \int_a^b c_n g(n, x) f'(x) dx, \end{aligned}$$

where  $g(n, x) = 1$  for  $n \leq x \leq b$  and  $g(n, x) = 0$  for  $x < n$ . To finish the proof of the lemma we exchange the order integral-sum in the last sum and notice that

$$\sum_{a < n \leq b} c_n g(n, x) = \sum_{a < n \leq x} c_n = C(x).$$

□

# Chapter 2

## Local Fields

### 2.1 Absolute values

Consider the following congruence:

$$x^2 \equiv 2 \pmod{7^n}.$$

If  $n = 1$ , there are two solutions:  $x_0 \equiv \pm 3 \pmod{7}$ . Consider now the case  $n = 2$ .  
If

$$x^2 \equiv 2 \pmod{7^2},$$

we deduce in particular that  $x^2 \equiv 2 \pmod{7}$ . Hence we are looking for the solutions of the form  $x_0 + 7t_1$ . For example, for  $x_0 \equiv 3 \pmod{7}$ , we obtain:

$$(3 + 7t_1)^2 \equiv 2 \pmod{7^2},$$

so that

$$t_1 \equiv 1 \pmod{7} \text{ and } x_1 \equiv 3 + 7 \cdot 1 \pmod{7^2}.$$

Similarly, for  $n = 3$ , we obtain  $x_2 \equiv 3 + 7 \cdot 1 + 7^2 \cdot 2 \pmod{7^3}$ . Proceeding in the same way, we obtain an infinite sequence

$$x_0, x_1, \dots, x_n, \dots$$

satisfying  $x_0 \equiv 3 \pmod{7}$ ,  $x_n \equiv x_{n-1} \pmod{7^n}$ ,  $x_n^2 \equiv 2 \pmod{7^{n+1}}$ . This construction is similar to the approximate computation of  $\sqrt{2}$  in  $\mathbb{R}$ : one then is interested in a sequence  $r_1, r_2, \dots, r_n, \dots$  of rational numbers such that  $|r_n^2 - 2| < \frac{1}{10^n}$ . In our case, we are implicitly using a different «metric»:  $x_n^2 - 2$  is divisible by  $7^{n+1}$ . This example provides a motivation to study various absolute values (metrics) on a given field  $K$ .

**Definition 2.1.1.** Let  $K$  be a field. An **absolute value** on  $K$  is a function  $|\cdot| : K \rightarrow \mathbb{R}$ , satisfying the following properties:

- $|0| = 0$  and  $|x| > 0$  for  $x \neq 0$ ;

- $|xy| = |x||y|$ ;
- (triangle inequality)  $|x + y| \leq |x| + |y|$ .

We call  $|\cdot|$  : **nonarchimedean** if a stronger condition  $|x + y| \leq \max\{|x|, |y|\}$  holds. Otherwise, we call the absolute value **archimedean**.

**Remark 2.1.2.** A classical 'archimedean' property (introduced by Archimedes) is that for any two positive real numbers  $x, y$ , there exists an integer  $n$  such that  $x < ny$ . We will see that an absolute value is nonarchimedean precisely if this property does not hold.

**Definition 2.1.3.** Let  $K$  be a field. A **valuation** on  $K$  is a function  $v : K \rightarrow \Gamma \cup \{\infty\}$ , where  $\Gamma$  is a commutative group (written additively), satisfying the following properties:

- $v(x) = \infty \Leftrightarrow x = 0$ ;
- $v(xy) = v(x) + v(y)$ ;
- (triangle inequality)  $v(x + y) \geq \min\{v(x), v(y)\}$ .

**Remark 2.1.4.** We also use the notation  $v : K^* \rightarrow \mathbb{Z}$  for a valuation, where  $K^* = K \setminus \{0\}$  and where we exclude the value  $\infty$  at 0.

Note that for  $v : K \rightarrow \mathbb{Z} \cup \{\infty\}$  a valuation and  $0 < \alpha < 1$  a real number one can associate a nonarchimedean absolute value  $|\cdot|_v$  defined by

$$|x|_v = \alpha^{v(x)}.$$

**Example 2.1.5.** 1. If  $K = \mathbb{R}$ , then we have the usual absolute value  $|\cdot|$  on  $K$ . This absolute value is archimedean.

2. If  $K = \mathbb{C}$  then  $|z| = \sqrt{z \cdot \bar{z}}$  defines an archimedean absolute value. More generally, if  $K$  is a number field and  $\sigma : K \hookrightarrow \mathbb{C}$  is a complex embedding of  $K$ , then we have an archimedean absolute value  $|x|_\sigma = |\sigma(x)|$  on  $K$ .
3. On any field  $K$  one can define a **trivial** absolute value by  $|x| = 1$  for  $x \neq 0$ .
4. If  $k = \mathbb{Q}$  and  $p$  is a prime, we have a  $p$ -adic valuation  $v_p(x) = r$  if  $x = p^r \frac{m}{n}$  with  $m$  and  $n$  prime to  $p$  and  $r > 0$ . We then define a nonarchimedean absolute value  $|x|_p = (1/p)^{v_p(x)}$ .
5. if  $k = \mathbb{C}(t)$  and  $a \in \mathbb{C}$ , we define the absolute value  $v_{t-a}$  as the order of vanishing at  $a$ : if  $x \in \mathbb{C}(t)$  we can write  $x$  as  $x = (t - a)^r \frac{P(t)}{Q(t)}$  with  $P(t)$  and  $Q(t)$  prime to  $t - a$  and  $r \in \mathbb{Z}$ . We then define  $v_{t-a}(x) = r$ .

In the other direction, if  $|\cdot|$  is an absolute value, then one can define  $v(x) = -\log|x|$ .



**Proposition 2.1.6.** *Let  $|\cdot|$  be a nontrivial nonarchimedean absolute value on a field  $K$  and let  $v(x) = -\log|x|$ . Then:*

(i)  $v(xy) = v(x) + v(y)$

(ii)  $v(x + y) \geq \min\{v(x), v(y)\}$ .

If  $v(K^*)$  is discrete in  $\mathbb{R}$ , then  $v$  is a multiple of a discrete valuation  $w : K^* \rightarrow \mathbb{Z}$ .

*Proof.* The statements (i) and (ii) are straightforward. For the last statement, since  $v(K^*)$  is discrete additive subgroup of  $\mathbb{R}$ , one has  $v(K^*) = \mathbb{Z} \cdot c$  for some  $c > 0$  (take  $c$  the minimal positive element in  $v(K^*)$  that exists since  $v(K^*)$  is discrete). Then  $w = c^{-1}v$  is a discrete valuation as claimed.  $\square$

If the last property holds, we call  $|\cdot|$  a discrete absolute value.

**Example 2.1.7.** There exists nondiscrete nonarchimedean absolute values. For example, if  $\bar{\mathbb{Q}}$  is an algebraic closure of  $\mathbb{Q}$ , then we will see that one can extend the  $p$ -adic absolute value to  $\bar{\mathbb{Q}}$ . Then one should necessarily have  $|p^{1/n}| = 1/\sqrt[n]{p} \rightarrow 1, n \rightarrow \infty$ .

**Proposition 2.1.8.** *An absolute value  $|\cdot|$  is nonarchimedean if and only if it takes bounded values on  $\{m \cdot 1, m \in \mathbb{Z}\}$ .*

*Proof.* If  $|\cdot|$  is nonarchimedean, then, for  $m > 0$  one has  $|m \cdot 1| = |1 + 1 + \dots + 1| \leq |1| = 1$ . Also  $|-1| = |1|$ .

Conversely, let  $N$  be an integer such that  $|m \cdot 1| \leq N$ . Then for  $x, y \in K$  we have

$$|x + y|^n \leq \sum \binom{n}{r} |x|^r |y|^{n-r} \leq N(n + 1) \max\{|x|, |y|\}^n.$$

Hence  $|x + y| \leq N^{1/n}(n + 1)^{1/n} \max\{|x|, |y|\}$  and  $N^{1/n}(n + 1)^{1/n} \rightarrow 1, n \rightarrow \infty$  (since  $\frac{1}{n} \log \frac{n+1}{N} \rightarrow 0$ ).  $\square$

**Corollary 2.1.9.** *If  $\text{char} K > 0$ , then  $K$  has only nonarchimedean absolute values.*

*Proof.* The set  $\{m \cdot 1, m \in \mathbb{Z}\}$  is finite if  $\text{char} K > 0$ .  $\square$

**Proposition 2.1.10.** *Let  $|\cdot|$  be a nonarchimedean absolute value on a field  $K$ . Then:*

- $A = \{x \in K, |x| \leq 1\}$  is a subring of  $K$ ;
- $U = \{x \in K, |x| = 1\}$  is a group of units of  $A$ ;
- $\mathfrak{m} = \{x \in K, |x| < 1\}$  is a unique maximal ideal of  $A$ .

The absolute value  $|\cdot|$  is discrete if and only if the ideal  $\mathfrak{m}$  is principal.

*Proof.* The first three properties are straightforward. Assume  $|\cdot|$  is discrete and let  $v, w$  be valuations as in Proposition 2.1.6. Then  $\mathfrak{m}$  is generated by any element  $\pi$  with  $w(\pi) = 1$ . Conversely, if  $\mathfrak{m} = (\pi)$  is principal, then  $v(K^*) = \mathbb{Z}v(\pi)$ .  $\square$

An absolute value  $|\cdot|$  defines a metric on  $K$ , where we put

$$d(x, y) = |x - y|$$

for the distance function. Then we have a topology on  $K$  with

$$B(x, r) = \{y \in K, |y - x| < r\}, r > 0$$

the fundamental system of open neighborhoods of  $x$ .

**Example 2.1.11.** If  $K = \mathbb{Q}$ , the  $p$ -adic valuation on  $K$  defines the  **$p$ -adic topology**: two rational numbers are «close» if their difference is divisible by a large power of  $p$ .

**Proposition 2.1.12.** *Let  $|\cdot|_1, |\cdot|_2$  be two absolute values on a field  $K$ , with  $|\cdot|_1$  nontrivial. The following conditions are equivalent:*

- (i)  $|\cdot|_1$  and  $|\cdot|_2$  define the same topology;
- (ii)  $|x|_1 < 1 \Rightarrow |x|_2 < 1$ ;
- (iii)  $|x|_2 = |x|_1^a$  for some  $a > 0$ .

*Proof.* (iii) $\Rightarrow$ (i) is obvious, for (i) $\Rightarrow$ (ii) note that  $|x|^n = |x|_1^n$  hence  $x^n \rightarrow 0 \Leftrightarrow |x| < 1$ , hence (i) implies  $|x|_1 < 1 \Rightarrow |x|_2 < 1$ . Let us show that (ii) $\Rightarrow$ (iii). Since  $|\cdot|_1$  is nontrivial, there exists  $y \in K$  such that  $|y|_1 > 1$ . Let  $a$  be such that  $|y|_2 = |y|_1^a$ . Note that  $a > 0$  by (ii).

Let  $x \in K^*$  and let  $b \in \mathbb{R}$  be such that  $|x|_1 = |y|_1^b$ . It is enough to establish that  $|x|_2 = |y|_2^b$ . Let  $m/n > b$  be a rational number. Then  $|x|_1 = |y|_1^b < |y|_1^{\frac{m}{n}}$ , so that  $|x^n/y^m|_1 < 1$ . By (ii) we deduce  $|x^n/y^m|_2 < 1$ , so that  $|x|_2 < |y|_2^{\frac{m}{n}}$ . Since this is true for all rational numbers  $m/n > b$ , we deduce that  $|x|_2 \leq |y|_2^b$ . By a similar argument for rationals  $m/n < b$  we deduce  $|x|_2 \geq |y|_2^b$ , this finishes the proof of (iii).  $\square$

**Definition 2.1.13.** The absolute values  $|\cdot|_1$  and  $|\cdot|_2$  are equivalent if they satisfy the equivalent conditions of the proposition above. A **place** of  $K$  is an equivalence class of absolute values on  $K$ .

In the case  $K = \mathbb{Q}$  we have a complete list of absolute values:

**Theorem 2.1.14** (Ostrowski). *Let  $|\cdot|$  be a nontrivial absolute value on  $\mathbb{Q}$ . Then*

- (i) *If  $|\cdot|$  is archimedean, then  $|\cdot|$  is equivalent to  $|\cdot|_\infty : x \mapsto |x|$  the usual absolute value in  $\mathbb{R}$ .*

(ii) If  $|\cdot|$  is nonarchimedean, then  $|\cdot|$  is equivalent to  $|\cdot|_p$  for exactly one prime  $p$ .

*Proof.* Let  $m, n > 1$  be integers. Write

$$m = a_0 + a_1n + \dots + a_rn^r,$$

where the  $a_i$  are integers,  $0 \leq a_i < n$ ,  $n^r \leq m$  and  $r \leq \log(m)/\log(n)$ . In particular,

$$|a_i| \leq |1| + \dots + |1| = a_i|1| = a_i \leq n.$$

Let  $N = \max\{1, |n|\}$ . Then:

$$|m| \leq \sum |a_i||n|^i \leq \sum |a_i|N^r \leq (1+r)nN^r \leq \left(1 + \frac{\log(m)}{\log(n)}\right)nN^{\frac{\log(m)}{\log(n)}}.$$

Replacing  $m$  with  $m^t$  for  $t$  an integer and letting  $t \rightarrow \infty$ , we obtain:

$$|m| \leq \left(1 + \frac{t\log(m)}{\log(n)}\right)^{1/t} n^{\frac{1}{t}} N^{\frac{\log(m)}{\log(n)}}$$

and  $|m| \leq N^{\frac{\log(m)}{\log(n)}}$ .

Assume that for some  $n$  one has  $|n| < 1$ . Then  $N = 1$  and the inequality above implies that  $|m| \leq 1$  for all integers  $m$ . In particular, the absolute value is nonarchimedean by Proposition 2.1.8. Let  $A$  and  $\mathfrak{m}$  be as in Proposition 2.1.10. By definition of  $A$ , we have  $\mathbb{Z} \subset A$ . Then  $\mathfrak{m} \cap \mathbb{Z}$  is a prime ideal in  $\mathbb{Z}$ , hence  $\mathfrak{m} \cap \mathbb{Z} = (p)$  for some prime  $p$ . Then  $|m| = 1$  if  $m$  is not divisible by  $p$ , so that  $|\frac{\alpha}{\beta}p^r| = |p|^r$  if  $\alpha, \beta$  are integers prime to  $p$ . If  $a$  is such that  $|p| = (1/p)^a$ , then  $|x| = |x|_p^a$ .

Assume now that for all integers  $n > 1$  we have  $|n| > 1$ . Then  $N = |n|$  and we obtain  $|m|^{\frac{1}{\log(m)}} \leq |n|^{\frac{1}{\log(n)}}$ . Inverting the roles of  $m$  and  $n$ , we get  $|n|^{\frac{1}{\log(n)}} \leq |m|^{\frac{1}{\log(m)}}$ , so that

$$c := |m|^{\frac{1}{\log(m)}} = |n|^{\frac{1}{\log(n)}}.$$

If  $a = \log(c)$ , then  $|n| = |n|_\infty^a$  for all integers  $n > 1$ , hence, by multiplicativity of  $|\cdot|$ , for all rationals.  $\square$

**Remark 2.1.15.** Similarly, if  $K$  is a number field, we have a complete list of places of  $K$ . Assume that  $K$  has  $r_1$  real embeddings and  $r_2$  pairs of conjugated complex embeddings. Then the places of  $K$  are in one-to-one correspondence with, on the one hand, the  $r_1$  real embeddings and  $r_2$  pairs of complex embeddings, corresponding to archimedean absolute values (infinite places of  $K$ ), and on the other hand, the nonzero prime ideals of  $\mathcal{O}_K$ , corresponding to the nonarchimedean absolute values (finite places of  $K$ ).

Again in the case  $K = \mathbb{Q}$  recall **the Chinese remainder theorem**: if  $n_1, \dots, n_k$  are integers that are pairwise coprime, then the system of congruences

$$x \equiv a_i \pmod{n_i}, i = 1, \dots, k$$

has a solution, unique modulo  $N = n_1 \cdot \dots \cdot n_k$ .

The following statement of the **weak approximation theorem** could be seen as a generalization of the Chinese remainder theorem in the context of valued fields.

**Theorem 2.1.16.** *Let  $| \cdot |_1, | \cdot |_2, \dots, | \cdot |_k$  be nontrivial inequivalent absolute values on a field  $K$ , and let  $a_1, \dots, a_k$  be elements of  $K$ . For every  $\epsilon > 0$ , there is an element  $x \in K$  such that*

$$|x - a_i|_i < \epsilon \text{ for all } i.$$

*Proof.* We start with the following two lemmas:

**Lemma 2.1.17.** *In the notation of the theorem, there is an element  $a \in K$  such that  $|a|_1 > 1$  and  $|a_i| < 1$  for  $i \neq 1$ .*

*Proof.* If  $k = 2$ , since  $| \cdot |_1$ , and  $| \cdot |_2$  are inequivalent, there are two elements  $b$  and  $c$  such that

$$|b|_1 < 1, |b|_2 \geq 1, |c|_1 \geq 1, |c|_2 < 1.$$

Then  $x = \frac{c}{b}$  works.

By induction, assume the lemma holds for  $n - 1$  absolute values. Then there are two elements  $b$  and  $c$  such that

$$|b|_1 > 1, |b|_i < 1, i = 1, \dots, n - 1, |c|_1 > 1, |c|_n < 1.$$

If  $|b|_n < 1$ , we can take  $x = b$ . If  $|b|_n = 1$ , then  $x = cb^r$  works for sufficiently large  $r$ . If  $|b|_n > 1$ , then  $x = \frac{cb^r}{1+b^r}$  works for sufficiently large  $r$ .  $\square$

**Lemma 2.1.18.** *In the notation of the theorem, there is an element  $a \in K$  that is close to 1 for the value  $| \cdot |_1$  and close to 0 for the values  $| \cdot |_i$ ,  $i \neq 1$ .*

*Proof.* Consider  $x$  as in the previous lemma and put  $x_r = \frac{x^r}{1+x^r}$ . Then

$$|x_r - 1|_1 = \frac{1}{|1 + x^r|_1} \leq \frac{1}{|x|_1^r - 1} \rightarrow 0, r \rightarrow \infty.$$

For  $i \geq 2$  we have

$$|x_r|_i = \frac{|x|_i^r}{|1 + x|_i^r} \leq \frac{|x|_i^r}{1 - |x|_i^r} \rightarrow 0, r \rightarrow \infty.$$

$\square$

To prove the theorem, consider  $b_i$ ,  $i = 1, \dots, n$  close to 1 for  $| \cdot |_i$  and close to 0 for  $| \cdot |_j, j \neq i$ . Then  $x = a_1 b_1 + \dots + a_n b_n$  works.

$\square$

**Corollary 2.1.19.** *Let  $|\cdot|_1, |\cdot|_2, \dots, |\cdot|_k$  be nontrivial inequivalent absolute values on a field  $K$ . If  $|a|^{r_1} \dots |a|^{r_k} = 1$  with  $r_i \in \mathbb{R}$  for all  $a \in K^*$ , then  $r_i = 0$  for all  $i$ .*

*Proof.* If there is  $r_i \neq 0$ , consider  $x$  such that  $|x|_i$  is sufficiently large and  $|x|_j$  is sufficiently small for  $j \neq i$ , on then cannot have  $|x|^{r_1} \dots |x|^{r_k} = 1$ , contradiction.  $\square$

## 2.2 Completions

Let  $K$  be a field and  $|\cdot|$  a nontrivial absolute value on  $K$ .

**Definition 2.2.1.** A sequence  $(a_n)$  of elements of  $K$  is a **Cauchy sequence** if for every  $\epsilon > 0$  there is an integer  $N > 0$  such that  $|a_m - a_n| < \epsilon \forall m, n > N$ .

**Definition 2.2.2.** A field  $K$  is **complete** if every Cauchy sequence converges in  $K$ .

**Theorem 2.2.3.** *Let  $K$  be a field and  $|\cdot|$  a nontrivial absolute value on  $K$ . Then there is a complete valued field  $(\hat{K}, |\cdot|)$  and a homomorphism  $K \rightarrow \hat{K}$  preserving the absolute value, satisfying the following universal property: every homomorphism  $K \rightarrow L$  with  $(L, |\cdot|)$  complete valued field, preserving the absolute value, extends uniquely to a homomorphism  $\hat{K} \rightarrow L$ .*

*Proof.* We construct  $\hat{K}$  as the set of limits of Cauchy sequences in  $K$ . More precisely, let  $(a_n)$  and  $(b_n)$  be two Cauchy sequences in  $K$ . We say that these sequences are equivalent if  $\lim |a_n - b_n| = 0$ . Define  $\hat{K}$  to be the set of equivalence classes of Cauchy sequences in  $K$ . One has the addition and the multiplication on  $\hat{K}$  induced by the addition and the multiplication on  $K$  and one verifies that  $\hat{K}$  is a field. The canonical map  $K \rightarrow \hat{K}$  sends  $a \in K$  to the Cauchy sequence  $(a, a, \dots, a)$ . A homomorphism  $K \rightarrow L$  extends to  $\hat{K} \rightarrow L$  by mapping the Cauchy sequence in  $K$  to its limit in  $L$ .  $\square$

If  $|\cdot|$  corresponds to a valuation  $v$ , we will write  $K_v$  for the completion  $\hat{K}$  and  $\hat{\mathcal{O}}_v$  for the ring of integers in  $K_v$ . For  $p$ -adic valuation on  $\mathbb{Q}$  we will write  $\mathbb{Q}_p$  for the completion and  $\mathbb{Z}_p$  for the ring of integers in  $\mathbb{Q}_p$ .

In the case of a discrete nonarchimedean value, we have the following description:

- Let  $A$  and  $\mathfrak{m}$  be as in Proposition 2.1.10. Let  $\pi$  be a generator of the maximal ideal  $\mathfrak{m}$  (equivalently,  $\pi$  is an element of  $K$  with largest value  $< 1$ ). We call  $\pi$  a **uniformizing parameter** and  $k = A/\mathfrak{m}$  the **residue field** of  $K$ . Up to changing  $|\cdot|$  by an equivalent one we may assume that  $|\pi| = 1$ . Then the set of values  $|K| = \{|\pi|^m, m \in \mathbb{Z}\} \cup \{0\} = \mathbb{Z}$ .
- $|\hat{K}| = |K| = \mathbb{Z}$ : in fact, if  $a \in \hat{K}$  and if  $(a_n)$  is a Cauchy sequence converging to  $a$ , then  $|a_n| \rightarrow |a|$ . But  $|K^*|$  is discrete, so that  $|a| \in |K^*|$ .

- If  $\hat{A} = \{a \in \hat{K}, |a| \leq 1\}$  and  $\hat{\mathfrak{m}} = \{a \in \hat{K}, |a| < 1\}$ , then  $\hat{A}$  (resp.  $\hat{\mathfrak{m}}$ ) is the set of limits of Cauchy sequences in  $A$  (resp. in  $\mathfrak{m}$ ), so that  $\hat{A}$  (resp.  $\hat{\mathfrak{m}}$ ) is the closure of  $A$  (resp.  $\mathfrak{m}$ ) in  $\hat{K}$ . Also  $\pi$  generates  $\hat{A}$ .
- For every integer  $n$ , the map  $A/\mathfrak{m}^n \rightarrow \hat{A}/\hat{\mathfrak{m}}^n$  is an isomorphism : in fact

$$\mathfrak{m}^n = \{a \in A, |a| \leq |\pi|^n\} = \{a \in A, |a| < |\pi^{n-1}|\}$$

is closed and open in  $A$ . The injectivity follows from the fact that it is closed and the surjectivity follows from the fact that  $\hat{\mathfrak{m}}^n$  open.

**Proposition 2.2.4.** *Let  $S$  be a set of representatives for  $A/\mathfrak{m}$ . Then the series*

$$a_{-n}\pi^{-n} + \dots + a_0 + a_1\pi + \dots + a_m\pi^m + \dots, a_i \in S$$

*is a Cauchy series. Every Cauchy series is of this form, in particular, any element of  $\hat{K}$  has a unique representative in the form above.*

*Proof.* Let  $x_N = \sum_{i=-n}^N a_i\pi^i$ . Then  $x_N$  is a Cauchy sequence.

Conversely, let  $\alpha \in \hat{K}$ . Then  $\alpha = \pi^n\alpha_0$  with  $\alpha_0$  a unit in  $\hat{A}$ , since  $|\hat{K}| = |K|$ . Then, by the definition of  $S$ , there exists  $a_0 \in S$  such that  $\alpha_0 - a_0 \in \hat{\mathfrak{m}}$ , so that  $\frac{\alpha_0 - a_0}{\pi} \in \hat{A}$ , so there exists  $a_1 \in S$  with  $\frac{\alpha_0 - a_0}{\pi} - a_1 \in \hat{\mathfrak{m}}$ . One then constructs inductively  $a_2, a_3, \dots$  such that

$$\alpha_0 = a_0 + a_1\pi + a_2\pi^2 + \dots$$

and we obtain the existence. Note that  $|\sum a_i\pi^i| = |\pi^m|$ , where  $a_m$  is the first nonzero coefficients. Hence  $\sum a_i\pi^i = 0$  if and only if  $a_i = 0$  for all  $i$  and we obtain the uniqueness.  $\square$

**Example 2.2.5.** We can think about elements of  $\mathbb{Q}_p$  as a series

$$a_{-n}p^{-n} + \dots + a_0 + a_1p + a_1p^2 + \dots, 0 \leq a_i < p.$$

In terms of limits,  $\mathbb{Z}_p = \varprojlim \mathbb{Z}/p^n\mathbb{Z}$ .

**Example 2.2.6.** The completion of the field  $k(t)$  with respect to the absolute value induced by the valuation  $v_t$  (the order of vanishing at 0) is the field  $k((t))$  of formal power series over  $k$ .

In the remaining part of this paragraph we study a the fundamental results on solving polynomial equations in complete valued fields. In what follows  $K$  is a complete discretely valued field and  $A$  is the corresponding ring of integers.

**Lemma 2.2.7.** *Let  $f \in A[x]$  and let  $a_0$  is a simple root of  $f(x) \bmod \pi$ . Then there is a unique root  $a$  of  $f(x)$  congruent to  $a_0 \bmod \pi$ .*

*Proof.* We will construct inductively a sequence  $(a_n)$  such that  $f(a_n) \equiv 0 \pmod{\pi^{n+1}}$ . By assumption, we have  $a_0$ . Put

$$a_{n+1} = a_n + h\pi^{n+1}.$$

Then

$$f(a_n + h\pi^{n+1}) = f(a_n) + h\pi^{n+1}f'(a_n) + \dots$$

We take

$$h = -\frac{f(a_n)}{\pi^{n+1}}f'(a_n)^{-1} \pmod{\pi}.$$

This is possible since  $f(a_n) \equiv 0 \pmod{\pi^{n+1}}$  and  $f'(a_n) = f'(a_0) \pmod{\pi}$ , which is nonzero.

The sequence  $(a_n)$  is a Cauchy sequence and its limit  $a$  in  $K$  is the root of  $f$ , as required. This root is unique, since from the construction  $a \pmod{\pi^n}$  is uniquely determined for each  $n$ .  $\square$

**Theorem 2.2.8** (Newton's lemma). *Let  $f(x) \in A[x]$  and let  $a_0 \in A$  such that  $|f(a_0)| < |f'(a_0)|^2$ . Then there is a root  $a \in A$  of  $f(x)$  such that  $|a - a_0| \leq |f(a_0)|/|f'(a_0)|^2$ .*

*Proof.* The sequence  $(a_n)$  defined inductively by  $a_{n+1} = a_n - \frac{f(a_n)}{f'(a_n)}$  is a Cauchy sequence converging to a root of  $f(x)$ .  $\square$

**Theorem 2.2.9** (Hensel's lemma). *Let  $f(x) \in A[x]$  and  $\bar{f}(x)$  be the image of  $f$  in  $k[x]$ . Assume that  $f$  is monic and that  $\bar{f}$  factors as  $\bar{f} = g_0h_0$  with  $g_0$  and  $h_0$  monic relatively prime in  $k[X]$ . Then  $f$  factors as  $f = gh$  with  $g$  and  $h$  monic such that  $\bar{g} = g_0$  and  $\bar{h} = h_0$ . The polynomials  $g$  and  $h$  are uniquely determined and  $(g, h) = A[x]$ .*

*Proof.* We first prove the uniqueness of  $g$  and  $h$ . Assume we have  $g'$  and  $h'$  monic such that  $\bar{g}' = g_0$  and  $\bar{h}' = h_0$  and  $f = g'h'$ . From the lemma below,  $(g, h') = A[x]$ , so there exist  $r, s \in A[x]$  with  $gr + h's = 1$ . We deduce

$$g' = g'gr + g'h's = g'gr + ghs,$$

and so  $g$  divides  $g'$ . Since these polynomials are both monic and of the same degree, they must be equal. Then  $h' = h$ , since  $f = gh = gh'$ .

Let us now prove the existence of  $f$  and  $g$ . We know that  $f - g_0h_0 \in \pi A[x]$ . We construct by induction polynomials  $g_n, h_n$  such that  $f - g_nh_n \equiv 0 \pmod{\pi^{n+1}A[x]}$  and  $g_n \equiv g_0, h_n \equiv h_0 \pmod{\pi A[x]}$ . Then we claim that there are two polynomials  $u, v \in A[x]$  with  $\deg u < \deg g_0$  and  $\deg v < \deg h_0$  and such that  $f - (g_n + \pi^{n+1}u)(h_n + \pi^{n+1}v) \equiv 0 \pmod{\pi^{n+1}A[x]}$ , i.e.

$$uh_n + g_nv \equiv (f - g_nh_n)/\pi^{n+1} \pmod{\pi A[x]}.$$

In fact, the lemma below provides such polynomials since  $g_0$  and  $h_0$  are monic and relatively prime. Passing to the limit, we obtain the factorisation of  $f$  as required.  $\square$

**Lemma 2.2.10.** *If  $f, g \in A[x]$  are such that  $\bar{f}$  and  $\bar{g}$  are relatively prime and  $f$  is monic, then  $(f, g) = A[x]$ : there exist  $u, v \in A[x]$  with  $\deg u < \deg g$  and  $\deg v < \deg f$  such that  $uf + vg = 1$ .*

*Proof.* Put  $M = A[x]/(f, g)$ . Note that  $M$  is a finitely generated  $A$ -module since  $f$  is monic. Since  $\bar{f}$  and  $\bar{g}$  are relatively prime, we have  $(\bar{f}, \bar{g}) = k[x]$ , hence

$$(f, g) + \mathfrak{m}A[x] = A[x].$$

We deduce that  $\mathfrak{m}M = M$ . By Nakayama's lemma,  $M = 0$ . We deduce that there exist  $u, v \in A[x]$  with  $uf + vg = 1$ . It remains to insure that one can assume that  $\deg u < \deg g$  and  $\deg v < \deg f$ . In fact, if  $\deg v \geq \deg f$  write  $v = fq + r$  with  $\deg r < \deg f$ . Then  $(u + qg)f + rg = 1$ , so that we also must have  $\deg u + qg < \deg g$ .  $\square$

**Remark 2.2.11.** By induction, a factorisation of  $f$  into a product of relatively prime polynomials in  $k[x]$  lifts to  $A[x]$ . In particular,  $x^p - x$  splits into  $p$  distinct factors in  $\mathbb{Z}_p[x]$ . We deduce that  $\mathbb{Z}_p$  contains  $(p - 1)$  roots of unity. More generally, if the residue field  $k = A/\mathfrak{m}$  is finite with  $q$  elements, then  $K$  contains  $q$  roots of the polynomial  $x^q - x$ . If  $S$  is the set of this roots, we have a bijection  $S \rightarrow k, a \mapsto \bar{a}$ , preserving the multiplication. We call the set  $S$  **Teichmüller representatives** of the elements of  $k$ .

## 2.3 Locally compact fields

A valued field  $K$  carries a natural topology, as explained in the previous section. We now study the compactness properties of  $K$ .

**Proposition 2.3.1.** *Let  $K$  be complete with respect to a nonarchimedean discrete absolute value. Let  $A$  be the ring of integers in  $K$  and let  $\mathfrak{m}$  be maximal ideal in  $A$ . Then  $A$  is compact if and only if  $A/\mathfrak{m}$  is finite.*

*Proof.* Put  $S$  the set of representatives of  $A/\mathfrak{m}$ .

Assume first that  $A$  is compact. Since  $\mathfrak{m} = \{x \in K, |x| < 1\}$  is open in  $K$  and  $A$  is disjoint union of sets  $s + \mathfrak{m}$ ,  $s \in S$ , we deduce that  $S$  must be finite since  $A$  is compact.

Assume now that  $S$  is finite. Recall that every element of  $A$  could be written as  $s_0 + s_1\pi + s_2\pi^2 + \dots$ . The number of finite sums  $s_0 + s_1\pi + s_2\pi^2 + \dots + s_n\pi^n$  is finite and any element of  $A$  is at distance  $\pi^{n+1}$  of such element. We deduce that for any  $r > 0$  there is a finite covering of  $A$  by open balls of radius  $r$  ( $A$  is totally bounded). Thus  $A$  is complete and totally bounded, hence compact.  $\square$



**Definition 2.3.2.** A **local field** is a field  $K$  with a nontrivial absolute value  $|\cdot|$  such that  $K$  is locally compact.

Note that a local field, being locally compact, is complete. We have a full classification of local fields:

**Proposition 2.3.3.** *Let  $K$  be a local field. Then*

- (i) *If the absolute value  $|\cdot|$  of  $K$  is archimedean, then  $K$  is isomorphic to  $\mathbb{R}$  or  $\mathbb{C}$  and  $|\cdot|$  is equivalent to the usual absolute value.*
- (ii) *If  $K$  is a nonarchimedean local field of zero characteristic, then  $K$  is isomorphic to  $\mathbb{Q}_p$ .*
- (iii) *A nonarchimedean local field of characteristic  $p > 0$  is isomorphic to the field  $k((t))$  of the formal power series over a finite field  $k$ , and the absolute value is equivalent to the absolute value induced by the order of vanishing  $v_t$  at 0.*

*Proof.* For (i), note that  $\mathbb{Q} \subset K$  and the restriction of  $|\cdot|$  to  $\mathbb{Q}$  is archimedean, hence is the usual absolute value. Hence  $K$  contains  $\mathbb{R}$ , and we may even assume that  $K$  contains  $\mathbb{C}$ , adjoining a root of  $-1$ , if necessary. Assume there is  $x \in K \setminus \mathbb{C}$  and let  $c \in \mathbb{C}$  be the closest element to  $x$ . Put  $x' = x - c$ , then  $|x' - z| \geq |x'| \forall z \in \mathbb{C}$ . We deduce, for  $\zeta$  a primitive  $n^{\text{th}}$  root of unity, that  $|x'^n - z^n| = |x' - z||x' - \zeta| \dots \geq |x' - z||x'|^{n-1}$ . If  $|z| < 1$ , letting  $n \rightarrow \infty$ , we obtain  $|x'| \geq |x' - z|$ . Hence  $|x' - z| = |x'|$ . Taking  $x' - z$  in place of  $x'$ , we obtain  $|x' - 2z| = |x'|$ . By induction,  $|x' - nz| = |x'|$ , and we obtain a contradiction with the archimedean property.

For (ii), note that  $\mathbb{Q} \subset K$  and the restriction of the absolute value of  $K$  to  $\mathbb{Q}$  is nonarchimedean, hence should be equivalent to a  $p$ -adic absolute value. Hence  $\mathbb{Q}_p \subset K$ . Since the residue field of  $K$  is finite by the previous proposition, we deduce that  $K$  is a finite extension of  $\mathbb{Q}_p$ .

For (iii), since the residue field  $k = S$  of  $K$  is finite, one verifies that the map  $K \rightarrow k((T)), \sum s_i \pi^i \mapsto \sum s_i T^i$  is an isomorphism of valued fields.  $\square$

**Remark 2.3.4.** In the case (ii) the absolute value is equivalent to a (unique) extension of the  $p$ -adic absolute value to  $K$ .

## 2.4 Extensions

### 2.4.1 Basic facts

Let  $K$  be a field, complete with respect to a discrete absolute value  $|\cdot|_K$  and let  $L/K$  be a finite separable extension. One then wonders if one could extend the absolute value to  $L$ . Later in this course we will prove the following:

**Theorem 2.4.1.** *The absolute value  $|\cdot|_K$  extends uniquely to a discrete absolute value  $|\cdot|_L$  on  $L$ . In addition,  $L$  is complete and the absolute value is characterized by  $|\beta|_L = |Nm_{L/K}(\beta)|_K^{\frac{1}{n}}$ .*

Let  $A$  (resp.  $B$ ) be a valuation ring for  $K$  (resp.  $L$ ), i.e.  $A = \{x, |x| \leq 1\}$ ,  $\mathfrak{m}_A$  (resp.  $\mathfrak{m}_B$ ) is the maximal ideal. Sometimes we will also write  $\mathfrak{m}_K$  (resp.  $\mathfrak{m}_L$ ). Then  $A \subset B$ ,  $\mathfrak{m}_A \subset \mathfrak{m}_B$  and we have an extension of residue fields  $k_A \subset k_B$ . Let  $\pi_A$  (resp.  $\pi_B$ ) be the uniformizing parameter for  $A$  (resp. for  $B$ ). Then, viewing  $\pi_A$  as an element of  $B$ , one could write

$$\pi_A = u\pi_B^e$$

where  $e > 0$  is the **ramification index** of the extension  $L/K$  and  $u$  is a unit in  $B$ .

**Definition 2.4.2.** We say that the extension  $L/K$  is **unramified** if  $e = 1$  and that it is **totally ramified** if  $e = [L : k]$ .

We will also establish the following:

**Theorem 2.4.3.** *In the notations above,*

$$[L : K] = e[k_B : k_A].$$

**Corollary 2.4.4.** *If  $\Omega$  is infinite (separable) extension (for example, a separable closure), then  $|\cdot|_K$  extends uniquely on  $\Omega$ .*

*Proof.* The corollary follows from Theorem 2.4.1, since  $|\cdot|_K$  extends uniquely to each finite extension.

## 2.4.2 Unramified extensions

**Theorem 2.4.5.** *Let  $L/K$  be an algebraic extension and let  $l/k$  be the corresponding extension of residue fields. Assume  $K$  and  $k$  are perfect.*

- (i) *There is a one-to-one correspondence between the finite unramified extensions  $K' \subset L$  of  $K$  and the finite extensions  $k' \subset l$  of  $k$ .*
- (ii) *Moreover,  $K'$  is Galois over  $K$  iff  $k'$  is Galois over  $k$ . In this case, there is a canonical isomorphism  $\text{Gal}(K'/K) \simeq \text{Gal}(k'/k)$ .*

*Proof.* (i) We first establish the surjectivity: for  $k'/k$  of degree  $n$  write  $k' = k[\alpha]$  and  $f$  minimal polynomial of  $\alpha$ . We may assume that  $f$  is the reduction modulo the maximal ideal of  $A$ , of a polynomial  $P \in A[x]$ . By Newton's lemma, there exists  $\beta \in L$  with  $P(\beta) = 0$  and  $\beta \equiv \alpha \pmod{\mathfrak{m}_L}$  (in particular,  $\beta \notin \mathfrak{m}_L$ ). Then  $K' = K[\beta]$  is of degree  $n$  and has residue field extension  $k'/k$ , so that by theorem 2.4.3, we have  $e = 1$ . We left the injectivity property as an exercise: in fact, if  $K', K''$  have the same residual extension  $k'/k$ , then one easily shows that  $K \cdot K''/K$  also has the residual extension  $k'/k$  and this leads to a contradiction for the degree.

- (ii) If  $K'/K$  is Galois, then the Galois group preserves  $A'$  and the maximal ideal  $\mathfrak{m}_{A'}$ , so that there is a map

$$\text{Gal}(K'/K) \rightarrow \text{Aut}(k'/k).$$

If  $k' = k[\alpha]$  with minimal polynomial  $\bar{g}$  that lifts to  $g \in A[X]$ , we obtain, by Newton lemma again, that there is a root  $\beta$  of  $g$  with  $\beta \equiv \alpha \pmod{\mathfrak{m}_L}$ . Then  $g$  splits in  $K'$  since  $K'/K$  is Galois, so that  $\bar{g}$  splits in  $k'$ . If  $\beta_1, \dots, \beta_n$  are the roots of  $g$  and  $\bar{\beta}_1, \dots, \bar{\beta}_n$  are the roots of  $\bar{g}$ , then the action of  $\text{Gal}(k'/k)$  is induced by the action of  $\text{Gal}(K'/K)$ , this gives an isomorphism  $\text{Gal}(K'/K) \rightarrow \text{Gal}(k'/k)$ . Conversely, if  $K'/K$  is Galois, then for  $k' = k[\alpha]$ ,  $\beta$  a lift of  $\alpha$ , Hensel lemma shows that all the conjugates of  $\beta$  are in  $K' = K[\beta]$ , so that  $K'/K$  is Galois.  $\square$

**Corollary 2.4.6.** *There is  $K_0 \subset L$  an extension of  $K$  such that  $K_0$  contains all the unramified extensions of  $K$ . If  $k$  is finite, then  $K_0$  is obtained by adjoining roots of unity.*

*If  $L = \bar{K}$ , we obtain maximal unramified extension  $K^{nr}$  of  $K$  and, a finite extension  $K'/K$  is unramified iff  $K' \subset K^{nr}$ . The residue field of  $K^{nr}$  is  $\bar{k}$ .*

### 2.4.3 Totally ramified extensions

**Definition 2.4.7.** A polynomial  $f \in K[x]$  is Eisenstein, if

$$f(x) = a_0x^n + \dots + a_{n-1}x + a_n$$

with  $|a_0| = 1$ ,  $|a_i| < 1$ ,  $|a_n| = |\pi|$ .

Equivalently, in the definition above we ask that  $v(a_0) = 0$ ,  $v(a_i) > 0$  and  $v(a_n) = 1$  for the normalized valuation on  $K$  corresponding to  $|\cdot|_K$  (i.e.  $v(\pi) = 1$ ).

**Proposition 2.4.8.** *A finite extension  $L/K$  is totally ramified if  $L = K[\alpha]$  with  $\alpha$  a root of an Eisenstein polynomial.*

*Proof.* Assume that  $\alpha$  is a root of an Eisenstein polynomial. Let  $n$  be the degree of the extension  $L/K$ . Then we deduce that  $v(\alpha^n) = 1$ , for the valuation extending the valuation  $v$  on  $K$ , i.e.  $v(\alpha) = \frac{1}{n}$ , so that  $e \geq n$ , using Theorem 2.4.3, we deduce  $e = n$ .

Conversally, if  $L/K$  is totally ramified, then if  $\alpha$  generates  $\mathfrak{m}_L$ , we have that  $1, \alpha, \dots, \alpha^{n-1}$  all have different valuations, so that the relation  $a_n + a_{n-1}\alpha + \dots + a_1\alpha^{n-1} = 0$  is impossible. Since the degree of  $L/K$  is  $n$ , we have a relation  $a_n + a_{n-1}\alpha + \dots + a_1\alpha^{n-1} + \alpha^n = 0$  and comparing the absolute values of the summands we deduce that the polynomial  $f(x) = x^n + \dots + a_{n-1}x + a_n$  is Eisenstein.  $\square$

## 2.4.4 Ramification groups and Krasner's lemma

Let  $L/K$  be a finite Galois extension. Assume that the residue field  $k$  of  $K$  is perfect. Put  $G = \text{Gal}(L/K)$ .

**Definition 2.4.9.**  $G_0 = \{\sigma \in G, |\sigma\alpha - \alpha| < 1 \forall \alpha \in B\}$  is called **the inertia group**.

**Theorem 2.4.10.** *Let  $L/K$  be a Galois extension and assume that the residue field extension  $l/k$  is separable. Then the fixed field  $K_0 = L^{G_0}$  is the largest unramified extension of  $K$  in  $L$ ,  $G_0$  is normal and*

$$G/G_0 = \text{Gal}(K_0/K) = \text{Gal}(l/k).$$

*Proof.* Note that  $G_0$  is indeed normal: if  $\sigma, \tau \in G$ , then

$$|\tau^{-1}\sigma\tau\alpha - \alpha| = |\sigma(\tau\alpha) - \tau\alpha|,$$

so that if  $\sigma \in G_0$ , then  $\tau^{-1}\sigma\tau \in G_0$ . If  $K'$  is the largest unramified extension in  $L$ , then  $\sigma K'$  is also unramified, for any  $\sigma : K' \rightarrow \bar{K}$  preserving  $K$ , so that by maximality,  $\sigma K' = K'$ . We deduce that  $K'$  is Galois and that we have an isomorphism  $\text{Gal}(K'/K) \simeq \text{Gal}(l/k)$  and  $G_0$  is the kernel of the map  $G \rightarrow \text{Gal}(l/k)$ , so that  $K_0$  is the fixed field.  $\square$

**Proposition 2.4.11.** *[Krasner's lemma] Let  $K$  be complete with respect to a nonarchimedean absolute value. Let  $\alpha, \beta \in \bar{K}$  with  $\alpha$  separable over  $K[\beta]$ . If  $\alpha$  is closer to  $\beta$  than any conjugate of  $\alpha$  (over  $K$ ), then  $K[\alpha] \subset K[\beta]$ .*

*Proof.* It is enough to show that for  $\sigma : K[\alpha, \beta] \subset \bar{K}$  fixing  $K[\beta]$ , one has  $\sigma\alpha = \alpha$ . Note that  $\sigma\beta = \beta$  (by assumption) and  $\sigma$  preserves  $|\cdot|$ , so that

$$|\sigma\alpha - \beta| = |\sigma\alpha - \sigma\beta| = |\alpha - \beta|.$$

We deduce

$$|\sigma\alpha - \alpha| = |\sigma\alpha - \beta + \beta - \alpha| \leq |\alpha - \beta|.$$

But  $\alpha$  is closer to  $\beta$  than any of its conjugate, so that we deduce  $\sigma\alpha = \alpha$ .

If now  $K$  is of characteristic zero and if  $h(x) \in K[x]$ , with  $h(x) = \sum c_i x^i$ , we define

$$\|h(x)\| = \max\{|c_i|\}.$$

Next, if  $f$  is monic irreducible, with

$$f(x) = \prod (x - \alpha_i), \alpha_i \in \bar{K}$$

and  $g$  is a polynomial such that  $\|f - g\|$  is small enough, we have for  $\beta$  a root of  $g$  that

$$|(f - g)(\beta)| = |f(\beta)| = \prod |\beta - \alpha_i|.$$

In particular, if  $\|f - g\|$  is small enough, there should exist  $\alpha_i$  with  $|\beta - \alpha_i|$  small enough, so that we may arrange that  $\beta$  is closer to  $\alpha_i$  than any conjugate of  $\alpha_i$ :

$$|\beta - \alpha_i| < |\alpha_i - \alpha_j|, i \neq j.$$

Then Krasner's Lemma says that  $K[\alpha] \subset K[\beta]$  and by a degree argument we obtain  $K[\alpha] = K[\beta]$ .

We obtain:

**Proposition 2.4.12.** *Let  $F$  be a monic irreducible polynomial in  $K[x]$ . Then for any  $g$  monic, sufficiently close to  $f$ , we have that  $g$  is also irreducible and for each root  $\beta$  of  $g$  there exists  $\alpha_i$  a root of  $f$  such that*

$$|\beta - \alpha_i| < |\alpha_i - \alpha_j|$$

for any  $i \neq j$  and  $\alpha_i$ 's are the roots of  $f$ . For such  $\alpha$ , we have  $K[\alpha] = K[\beta]$ .

**Corollary 2.4.13.** *Let  $K/\mathbb{Q}_p$  be a finite extension. Then there is a finite extension  $L/\mathbb{Q}$ , contained in  $K$  and such that*

$$[L : \mathbb{Q}] = [K : \mathbb{Q}_p] \text{ and } L \cdot \mathbb{Q}_p = K.$$

# Chapter 3

## Dedekind rings

### 3.1 Dedekind rings

This section is devoted to some general properties of rings and ideals.

#### 3.1.1 Fractional ideals

Fix an integral ring  $A$  and let  $K$  be its field of fractions. Recall the following notions:

- $A$  is **noetherian** if any increasing sequence of ideals stabilizes: for  $I_1 \subseteq \dots \subseteq I_n \subseteq \dots$  ideals of  $A$  there is  $N > 0$  such that  $I_N = I_m$  for any  $m \geq N$ . This property is equivalent to the following: any ideal of  $A$  is generated by a finite number of elements.
- $A$  is integrally closed if any element  $x$  of  $K$ , that is integral over  $A$  (i.e.  $x$  is a root of a monic polynomial with coefficients in  $A$ ) is in  $A$ . For example,  $\mathbb{Z}$  is integrally closed,  $A = \mathbb{Z} + \mathbb{Z}[\sqrt{-3}]$  is not,  $x = \frac{1+\sqrt{-3}}{2}$  is not in  $A$  and verifies the equation  $x^2 - x + 1 = 0$ . If  $L$  is a field containing  $A$ , the set of elements of  $L$  that are integral over  $A$  is a ring that we call **integral closure** of  $A$  in  $L$ .

Let  $I \subset K$  be an  $A$ -module (we do not assume that  $I$  is an ideal of  $A$ , i.e. that  $I \subset A$ ).

**Definition 3.1.1.**  $I^{-1} = \{x \in K \mid xI \subset A\}$  and  $R(I) = \{x \in K \mid xI \subset I\}$ . We call  $I$  a **fractional ideal** if  $I \neq 0$  and if there exists  $a \in A$  with  $aI \subset A$ . Note that  $I \neq 0$  is fractional iff there exists  $a \in K$  with  $aI \subset A$ .

**Definition 3.1.2.** A fractional ideal  $I$  is **invertible** if  $II^{-1} = A$ .

Let  $I_1$  and  $I_2$  be two fractional ideals with  $I_1 \subset a_1^{-1}A, I_2 \subset a_2^{-1}A, a_1, a_2 \in A$ . The following properties are straightforward:

1.  $I_1 + I_2, I_1I_2, I_1 \cap I_2$  are fractional ideals contained in  $(a_1a_2)^{-1}A$ .

2.  $I = \{x \in K, | xI_2 \subset I_1\}$  is a fractional ideal: in fact,  $I \neq 0$  and for  $v \neq 0 \in I_2$  we have  $a_1vI \subset A$ .
3. Applying the previous property to  $I_1 = I$  and  $I_2 = A$ , we obtain that  $I_1^{-1}$  is fractional; applying it to  $I_1 = I$  and  $I_2 = I$  we obtain that  $R(I_1)$  is fractional.
4. if  $A$  is noetherian, then any fractional ideal is of finite type.

Another important notion is the localization.

**Definition 3.1.3.** Let  $\mathfrak{p}$  be a prime ideal of  $A$ . The **localization**  $A_{\mathfrak{p}}$  of  $A$  is the subring of  $K$  of elements  $\frac{u}{v}$  with  $u \in A$  and  $v \in A \setminus \mathfrak{p}$ .

We have the following properties:

1.  $A_{\mathfrak{p}}$  is local ring (i.e. it has a unique maximal ideal);
2.  $\mathfrak{p} = \mathfrak{p}A_{\mathfrak{p}} \cap A$ .
3. More generally, if  $I \subset A_{\mathfrak{p}}$  is an ideal, then  $I = (I \cap A)A_{\mathfrak{p}}$ : in fact, we have an inclusion  $(I \cap A)A_{\mathfrak{p}} \subset I$ , and for another direction for  $x \in I$  we could write  $x = \frac{u}{v}$  with  $u \in A$  and  $v \in A \setminus \mathfrak{p}$ , since  $v$  is invertible in  $A$  we deduce that  $u \in I \cap A$ . Hence  $x \in (I \cap A)A_{\mathfrak{p}}$ .
4. If  $A$  is integrally closed, then  $A_{\mathfrak{p}}$  is integrally closed as well: let  $x \in K$  be integral over  $A_{\mathfrak{p}}$ . Then

$$x^n + \frac{a_{n-1}}{b}x^{n-1} + \dots + \frac{a_0}{b}$$

with  $b \in A \setminus \mathfrak{p}$  and  $a_i \in A$ . Hence  $bx$  is integral over  $A$ , hence  $bx \in A$  and  $x \in A_{\mathfrak{p}}$ .

### 3.1.2 Discrete valuation rings

Recall that if  $A$  is a local ring with the (unique) maximal ideal generated by one element  $\mathfrak{m} = (\pi)$ , then we can define a discrete valuation  $v : K^* \rightarrow \mathbb{Z}$ ,  $v(x) = r$  if  $x = u\pi^r$  with  $r \in \mathbb{Z}$  and  $u \in A^* \setminus \mathfrak{m}$ .

**Proposition 3.1.4.** *Let  $A$  be an integral noetherian, integrally closed ring, such that  $A$  has a unique nonzero prime ideal. Then  $A$  is a discrete valuation ring.*

*Proof.* It is enough to show that  $A$  is principal. Let  $\mathfrak{m}$  be the maximal ideal of  $A$ . We divide the proof into several steps:

1. If  $x \in \mathfrak{m}$ , then  $A[\frac{1}{x}] = K$ . In fact, it is enough to prove that any prime ideal of  $A[\frac{1}{x}]$  is zero, this will imply that it is a field, hence equals to  $K$ . Let  $\mathfrak{p}$  be a prime ideal of  $A[\frac{1}{x}]$ . Note that  $x \notin \mathfrak{p}$  since  $x$  is invertible. The ideal  $\mathfrak{p} \cap A$  is a prime ideal in  $A$ , and it is different from  $\mathfrak{m}$  since  $x \in \mathfrak{m}$ . By assumption,  $\mathfrak{p} \cap A = 0$ . But if  $\frac{y}{x^n} \in \mathfrak{p}$ , we may assume  $y \in A$ , so that  $y \in \mathfrak{p} \cap A = \{0\}$ . We then have that  $\mathfrak{p} = 0$ .

2. Let  $z$  be a nonzero element of  $A$ . If  $x \in \mathfrak{m}$ , there exists  $n \geq 0$  such that  $x^n \in zA$ . By the previous step  $K = A[\frac{1}{x}]$ , hence for  $z \in K$  there exists  $n$  such that  $x^n z \in A$ .
3.  $\mathfrak{m}^m \subset zA$  for some  $m$ . In fact, since  $A$  is noetherian,  $\mathfrak{m}$  is generated by a finite number elements, call them  $x_1, \dots, x_k$ . Then  $m = kn$  with  $n$  such that  $x_i^n \in zA$  (from the previous step) works.
4.  $\mathfrak{m}^{-1} \neq A$ . In fact, assume  $z \in \mathfrak{m}$  and  $m$  minimal for the previous step. Let  $y \in \mathfrak{m}^{m-1} \setminus zA$  (exists by minimality). Then  $\mathfrak{m}y \subset zA$ , so that  $y/z \in \mathfrak{m}^{-1} \setminus A$ .
5.  $\mathfrak{m}\mathfrak{m}^{-1} = A$ . Note that  $\mathfrak{m}\mathfrak{m}^{-1}$  is a sub  $A$  module containing  $\mathfrak{m}$ , hence it equals  $\mathfrak{m}$  or  $A$ . Let  $t \in \mathfrak{m}^{-1} \setminus A$ . Since  $A$  is integrally closed,  $t$  is not integral over  $A$ . Hence the sequence of  $A$ -modules

$$A \subset A + At \subset A + At + At^2 \subset \dots$$

is strictly increasing. Since  $A$  is noetherian, it is not included in any  $A$ -module of finite type, in particular, in  $\mathfrak{m}^{-1}$  (since it is a fractional ideal, it is of finite type). Then there is a minimal  $n$  such that  $t^n \notin \mathfrak{m}^{-1}$ . Hence  $t^n \mathfrak{m}$  is not in  $A$ , hence not in  $\mathfrak{m}\mathfrak{m}^{-1}$  and  $t\mathfrak{m}$ . Hence  $t^{n-1}\mathfrak{m}$  is not included in  $\mathfrak{m}$ . We deduce that  $\mathfrak{m}\mathfrak{m}^{-1}$  is not contained in  $\mathfrak{m}$ , hence  $\mathfrak{m}\mathfrak{m}^{-1} = A$  by maximality of  $\mathfrak{m}$ .

6.  $\mathfrak{m}$  is principal. Since  $\mathfrak{m}\mathfrak{m}^{-1} = A$ , there exists an element  $u \in A \setminus \mathfrak{m}$  with  $u = vw$ ,  $v \in \mathfrak{m}$ ,  $w \in \mathfrak{m}^{-1}$ . Then  $u$  is invertible (since  $\mathfrak{m}$  is maximal). Let  $t \in \mathfrak{m}$ . Then  $t = (tw/u)v$ . Since  $w \in \mathfrak{m}^{-1}$ , we have that  $tw \in A$ . We then have  $t \in vA$ . Hence  $\mathfrak{m}$  is generated by  $v$ , hence principal.
7.  $\bigcap_n \mathfrak{m}^n = 0$ . Since  $\mathfrak{m}$  is principal, we could write  $\mathfrak{m} = (\pi)$ . The ideal  $I = \bigcap_n \pi^n A$  is prime, in fact, if  $xy \in I$  with  $x \in \pi^n A$  and  $y \in \pi^m A$  we have  $(\frac{x}{\pi^n})(\frac{y}{\pi^m}) \in \bigcap_n \pi^n A$ , so that either  $x \in \pi^{n+1} A$  or  $y \in \pi^{m+1} A$ , by induction  $x \in I$  or  $y \in I$ . But  $I$  is different from  $\mathfrak{m}$ : if not, we would have  $\pi A = \pi^2 A$  and  $A = \pi A$ , that is not possible. Hence  $I=0$  as claimed.
8.  $A$  is principal. Let  $I \subset A$  be an ideal. Let  $N$  be maximal integer such that  $I \subset \mathfrak{m}^N$ . Then  $I = \pi^N A$ .

This finishes the proof of the proposition. □

Let  $A$  be a noetherian and integrally closed ring. Assume  $\mathfrak{p}$  is a minimal and maximal prime ideal. Then  $A_{\mathfrak{p}}$  is a discrete valuation ring with maximal ideal  $\mathfrak{p}A_{\mathfrak{p}}$ . We denote  $v_{\mathfrak{p}}$  the corresponding valuation. Note that the fractional ideals of  $A_{\mathfrak{p}}$  are  $\mathfrak{p}^n, n \in \mathbb{Z}$ .

### 3.1.3 Dedekind rings

Let  $A$  be an integral noetherian ring.



**Proposition 3.1.5.** *The following properties are equivalent:*

- (i) *A is integrally closed and any prime ideal of A is maximal;*
- (ii) *For any prime nonzero ideal  $\mathfrak{p}$  of A the localization  $A_{\mathfrak{p}}$  is a discrete valuation ring.*
- (iii) *Any fractional ideal of A is invertible.*

*Proof.* (i)  $\Rightarrow$  (ii). We've proved that  $A_{\mathfrak{p}}$  is noetherian and integrally closed. Any prime ideal of  $A_{\mathfrak{p}}$  comes from an ideal (hence prime) of  $A$ , and hence is maximal. By previous section,  $A_{\mathfrak{p}}$  is a discrete valuation ring.

(ii)  $\Rightarrow$  (iii). Let  $I$  be a fractional ideal of  $A$ . Consider  $II^{-1}$ : assume it is strictly contained in  $A$ . Then it is contained in some prime ideal  $\mathfrak{p}$  of  $A$ . Write  $I = (a_1, \dots, a_r)$ . Let  $x \in A_{\mathfrak{p}}$  with  $IA_{\mathfrak{p}} = xA_{\mathfrak{p}}$ . Write  $a_i = x \frac{u_i}{v_i}$  with  $u_i \in A$  and  $v_i \in A \setminus \mathfrak{p}$ . Let  $v = \prod v_i \in A \setminus \mathfrak{p}$ . Since  $va_i/x \in A$  we have  $v/x \in I^{-1}$ . Hence  $v \in II^{-1}A_{\mathfrak{p}} \subset \mathfrak{p}A_{\mathfrak{p}}$ . But  $\mathfrak{p}A_{\mathfrak{p}} \cap A = \mathfrak{p}$ , so that  $v \in \mathfrak{p}$ , contradiction.

(iii)  $\Rightarrow$  (i). Let us show that  $A$  is integrally closed: if  $x \in K$  is integral over  $A$ , the ring  $A[x]$  is of finite type over  $A$  and is contained in  $K$ . Hence it's a fractional ideal. We have  $A[x]^2 = A[x]$ , hence

$$A[x] = A[x](A[x]A[x]^{-1}) = A[x]A[x]^{-1} = A,$$

since  $A[x]$  is invertible. Let  $\mathfrak{p}$  be a prime ideal in  $A$ , let  $\mathfrak{m}$  be a maximal ideal containing  $\mathfrak{p}$ . The fractional ideal  $\mathfrak{m}^{-1}\mathfrak{p}$  is contained in  $A$ , hence  $(\mathfrak{m}^{-1}\mathfrak{p})\mathfrak{m} = \mathfrak{p}$ . Since  $\mathfrak{p}$  is prime,  $\mathfrak{p} = \mathfrak{m}$  or  $\mathfrak{p}\mathfrak{m}^{-1} \subset \mathfrak{p}$ . Assume the later case. Since  $A$  is strictly included in  $\mathfrak{m}^{-1}$ , we deduce

$$\mathfrak{m}^{-1} \subset \mathfrak{p}\mathfrak{p}^{-1}\mathfrak{m}^{-1} \subset \mathfrak{p}\mathfrak{p}^{-1} = A,$$

since  $\mathfrak{p}$  is invertible, contradiction. □

**Definition 3.1.6.** An integral noetherian ring satisfying the equivalent conditions of the proposition above is called **Dedekind ring**.

**Examples.**  $\mathbb{Z}$  is a Dedekind ring,  $k[t]$  is Dedekind,  $k[t_1, t_2]$  is not a Dedekind ring.

Let  $A$  be a Dedekind ring and  $\mathfrak{p}$  a maximal ideal of  $A$ . Let  $I$  be a fractional ideal of  $K$  (or, more generally, a nonempty subset). We set

$$v_{\mathfrak{p}}(I) = \inf_{x \in I} v_{\mathfrak{p}}(x).$$

One easily verifies that  $v_{\mathfrak{p}}(I)$  is well defined and is in  $\mathbb{Z}$ .

Since any fractional ideal of a Dedekind ring  $A$  is invertible, the set of fractional ideals of  $A$  is the multiplicative group, that we denote by  $I(A)$ .

**Proposition 3.1.7.** *The group  $I(A)$  is isomorphic to a free abelian group generated by nonzero prime ideals. More precisely, for  $I$  be a fractional ideal in  $K$  we have*

$$I = \prod_{\mathfrak{p}} \mathfrak{p}^{v_{\mathfrak{p}}(I)}.$$

*Proof.* We first show that any ideal  $I$  of  $A$  is a product of prime ideals in  $A$ . In fact, let  $\mathfrak{p}_1$  be a prime ideal in  $A$ , containing  $I$ . Then  $I \subset I\mathfrak{p}_1^{-1} \subset A$ . If  $I$  is not a product of maximal ideals, we construct an increasing sequence

$$I \subset I\mathfrak{p}_1^{-1} \subset I\mathfrak{p}_1^{-1}\mathfrak{p}_2^{-1} \subset \dots,$$

but this sequence should be finite since  $A$  is noetherian, contradiction.

Let now  $\mathfrak{p}$  be a maximal ideal of  $\mathbb{Z}$  and let

$$i_{\mathfrak{p}} : I(A) \rightarrow I(A_{\mathfrak{p}})$$

be the map  $I \mapsto IA_{\mathfrak{p}}$ . It is a group homomorphism. Let  $\mathfrak{p}' \neq \mathfrak{p}$  be a prime ideal of  $A$ . Then  $i_{\mathfrak{p}}(\mathfrak{p}') = A_{\mathfrak{p}}$ . In fact,  $\mathfrak{p}'$  is not contained in the maximal ideal of  $A_{\mathfrak{p}}$ , the ideal of  $A_{\mathfrak{p}}$  generated by  $\mathfrak{p}'$  is then equals to  $A_{\mathfrak{p}}$ .

Let us show that  $I(A)$  is generated by maximal ideals of  $A$ . Let  $I$  be a fractional ideal of  $A$ . Let  $t \in A$  with  $tI \subset A$ . From above, the ideals  $tI$  and  $tA$  are products of maximal ideals. Since  $I$  is invertible, it is a quotient of these products.

Now we show that the subgroup of  $I(A)$  generated by prime ideals is free: if not,  $A = \prod \mathfrak{p}^{r_{\mathfrak{p}}}$  with  $r_{\mathfrak{p}} \in \mathbb{Z}$  almost all zeros. Let  $\mathfrak{p}_0$  be a prime ideal in  $A$  such that  $r_{\mathfrak{p}_0}$  is nonzero. Then

$$A_{\mathfrak{p}_0} = i_{\mathfrak{p}_0}(A) = i_{\mathfrak{p}_0}(\mathfrak{p}_0^{r_{\mathfrak{p}_0}}) = \mathfrak{p}_0^{r_{\mathfrak{p}_0}} A_{\mathfrak{p}_0} \neq A_{\mathfrak{p}_0}.$$

Let  $I$  be a fractional ideal of  $K$ , then  $I = \prod_{\mathfrak{p}} \mathfrak{p}^{r_{\mathfrak{p}}}$ . We have

$$i_{\mathfrak{p}}(I) = \mathfrak{p}^{r_{\mathfrak{p}}} A_{\mathfrak{p}} = (\mathfrak{p}A_{\mathfrak{p}})^{r_{\mathfrak{p}}}.$$

Hence  $r_{\mathfrak{p}} = v_{\mathfrak{p}}(IA_{\mathfrak{p}}) = v_{\mathfrak{p}}(I)$ , and we finish the proof of the proposition.  $\square$

Note that a Dedekind ring  $A$  is not necessarily factorial, hence we do not have a unique factorization for elements in  $A$ , but we have this decomposition property for ideals in  $A$ .

**Definition 3.1.8.** A **principal** fractional ideal of  $A$  is an ideal of type  $aA$  with  $a \in K$ . The **ideal classes** group is the quotient of  $I(A)$  by the subgroup of principal ideals.

**Corollary 3.1.9.** *Let  $A$  be a Dedekind ring with  $K$  the field of fractions. Let  $x \in K$ . We then have  $v_{\mathfrak{p}}(x) = 0$  for all but finitely many maximal ideals  $\mathfrak{p}$  in  $A$ .*

**Corollary 3.1.10.** *Let  $A$  be a discrete valuation ring, then  $I(A) = \mathbb{Z}$ .*

**Corollary 3.1.11.** *Let  $A$  be a Dedekind ring, then the maps  $i_{\mathfrak{p}}$  define an isomorphism between  $I(A)$  and the direct sum  $I(A_{\mathfrak{p}})$  over maximal ideals of  $A$ .*

Note that for  $I_1, I_2$  fractional ideals, and  $\mathfrak{p}$  a prime ideal in a Dedekind ring  $A$ , one has the following properties:

1.  $v_{\mathfrak{p}}(I_1 I_2) = v_{\mathfrak{p}}(I_1) + v_{\mathfrak{p}}(I_2)$ ;
2.  $v_{\mathfrak{p}}(I_1 + I_2) = \min(v_{\mathfrak{p}}(I_1), v_{\mathfrak{p}}(I_2))$ ;

$$3. v_{\mathfrak{p}}(I_1 \cap I_2) = \max(v_{\mathfrak{p}}(I_1), v_{\mathfrak{p}}(I_2));$$

$$4. v_{\mathfrak{p}}(I_1 \cap I_2^{-1}) = v_{\mathfrak{p}}(I_1) - v_{\mathfrak{p}}(I_2).$$

**Proposition 3.1.12.** *Let  $A$  be a Dedekind ring with  $K$  field of fractions. Let  $I$  be a finite set and let  $(x_i)_{i \in I}$  (resp.  $(n_i)$ , resp.  $(\mathfrak{p}_i)$ ) be a family of elements of  $K$  (resp. of integers, resp. of maximal ideals of  $A$ ). Then there is  $y \in K$  such that  $v_{\mathfrak{p}_i}(y - x_i) \geq n_i$  for  $i \in I$  and  $v_{\mathfrak{p}}(y) \geq 0$  for any  $\mathfrak{p} \neq \mathfrak{p}_i$ ,  $i \in I$ .*

*Proof.* Assume first  $x_i \in A$ . We may assume  $n_i \geq 0$  (up to replacing by zero the negative ones). We could also assume that only one element  $x_{i_0}$  is nonzero. In fact, let  $y_{i_0}$  be the solution in this case, then the general solution is  $y = \sum y_i$ .

We have  $A = \mathfrak{p}_{i_0}^{n_{i_0}} + \prod_{i \neq i_0} \mathfrak{p}^{n_i}$ . Hence  $x_{i_0} = y + z$  with  $y$  in the first term and  $z$  in the second. Then  $y$  works.

In general case (without assumption  $x_i \in A$ ) we set  $x_i = a_i/s$  with  $a_i \in A$  and  $s \in A$ . We will find  $y = a/s$  with  $a$  determined by conditions  $v_{\mathfrak{p}_i}(a - a_i) \geq n_i + v_{\mathfrak{p}_i}(s)$  and  $v_{\mathfrak{p}}(a) \geq v_{\mathfrak{p}}(s)$  for  $\mathfrak{p}$  distinct from  $\mathfrak{p}_i$ . This is the same problem as in the previous case with a biggest family of prime ideals.  $\square$

## 3.2 Extensions

### 3.2.1 Extensions of Dedekind rings

Let  $L/K$  be a finite extension of fields. Let  $x \in L$ . Recall that the trace (resp. the determinant) of the linear map

$$L \rightarrow L, y \mapsto xy$$

is denoted by  $Tr_{L/K}(x)$  (resp.  $N_{L/K}(x)$ ) and is called the trace (resp. the norm) of  $x$ . If  $L' = K(x)$  is a Galois extension of  $K$ , then

$$Tr_{L'/K}(x) = \sum_{\sigma \in Gal(L'/K)} \sigma(x)$$

and

$$Nm_{L'/K} = \prod_{\sigma \in Gal(L'/K)} \sigma(x).$$

Assume that  $A$  is an integral noetherian and integrally closed ring with field of fractions  $K$ . Let  $L$  be a finite separable extension of degree  $n$ . Let  $B$  be the integral closure of  $A$  in  $L$ . Note that  $B \cap K = A$  since  $A$  is integrally closed.

**Proposition 3.2.1.**  *$B$  is an  $A$ -module of finite type.*

*Proof.* We will show that  $B$  is included in an  $A$ -module of finite type. Since  $A$  is noetherian, we then deduce that  $B$  is of finite type over  $A$ . Let  $M$  be a sub  $A$ -module of  $L$ . Let  $M^*$  be the set of elements  $x$  of  $M$ , such that  $Tr_{L/K}(xy) \in A$  for any  $y \in M$ . If  $M$  is free, then  $M^*$  is free (since the bilinear form  $(x, y) \rightarrow Tr(xy)$  is

nondegenerate for a separable extension). Let  $X$  be a family of elements of  $B$ , forming a basis of  $L$  as a  $K$ -vector space. Let  $V$  be a free  $A$ -module generated by this base. Then the image of  $B$  by the map  $Tr$  is contained in  $A$  by definition, and we have an inclusion:

$$V \subset B \subset B^* \subset V^*$$

and we deduce that  $B$  is of finite type.  $\square$

**Corollary 3.2.2.**  *$B$  is integrally closed.*

*Proof.* It follows from the fact that  $A$  is integrally closed and  $B$  is of finite type over  $A$ .

**Proposition 3.2.3.** *The field of fractions of  $B$  is  $L$ .*

*Proof.* If  $x \in L$  a root of a polynomial  $P(x) = a_n X^n + \dots + a_1 X + a_0$ , then  $a_n x$  is integral over  $A$ , so that  $x$  is a quotient of two elements in  $B$ .  $\square$

**Theorem 3.2.4.** *If  $A$  is a Dedekind ring, then  $B$  is a Dedekind ring.*

*Proof.* From propositions above, it only remains to prove that the localization of  $B$  at any nonzero prime ideal is a discrete valuation ring, so that it is enough to show that any prime nonzero ideal of  $B$  is maximal. If  $\mathfrak{p}$  is a prime ideal in  $B$ , that is not maximal, there is a maximal ideal  $\mathfrak{m}$  of  $B$  containing  $\mathfrak{p}$ . Then  $\mathfrak{p} \cap A$  and  $\mathfrak{m} \cap A$  are prime ideals in  $A$ , so that they should coincide since  $A$  is Dedekind. We obtain a contradiction by the lemma below.  $\square$

**Lemma 3.2.5.** *Let  $A \subset B$  be two rings with  $A \subset B$  and  $B$  integral over  $A$ . Let  $\mathfrak{p} \subset \mathfrak{q}$  be prime ideals of  $B$ . If  $\mathfrak{p} \cap A = \mathfrak{q} \cap A$ , then  $\mathfrak{p} = \mathfrak{q}$ .*

*Proof.* Assume that the inclusion  $\mathfrak{p} \subset \mathfrak{q}$  is strict, let  $x \in \mathfrak{q} \setminus \mathfrak{p}$ . Since  $x$  is integral, there is a monic polynomial  $P(X) = X^n + a_{n-1} X^{n-1} + \dots + a_0$  with coefficients in  $A$  such that  $P(x) \in \mathfrak{p}$ , we choose  $P$  of minimal degree  $> 1$ . Since  $\mathfrak{p}$  is prime,  $a_0 \in \mathfrak{q} \cap A = \mathfrak{p} \cap A$ . Hence  $(P(x) - a_0)/x \in \mathfrak{p}$  since  $\mathfrak{p}$  is prime. Contradiction, since  $P$  is of minimal degree.  $\square$

Recall that if  $L/\mathbb{Q}$  is a finite extension, then the ring of integers of  $L$  is the integral closure of  $\mathbb{Z}$  in  $L$ .

**Corollary 3.2.6.** *The ring of integers of a number field is a Dedekind ring.*

*Proof.* We apply the proposition above to  $A = \mathbb{Z}$ .

### 3.3 Decomposition of ideals

Let  $A, B, K, L$  be as before and  $\mathfrak{p} \subset B$  be a nonzero prime ideal. Let

$$\mathfrak{q} = \mathfrak{p} \cap A.$$

**Definition 3.3.1.** We say that  $\mathfrak{p} \mid \mathfrak{q}$  ( $\mathfrak{p}$  divides  $\mathfrak{q}$ ). For the ideal  $\mathfrak{q}B$  of  $B$ , we define

$$e_{\mathfrak{p}} = v_{\mathfrak{p}}(\mathfrak{q}B)$$

the ramification index in the extension  $L/K$ .

By definition,

$$\mathfrak{q}B = \prod_{\mathfrak{p} \mid \mathfrak{q}} \mathfrak{p}^{e_{\mathfrak{p}}},$$

we say that  $\mathfrak{q}$  **ramifies** in  $B$  if there exists  $\mathfrak{p}$  with  $e_{\mathfrak{p}} \neq 1$ .

**Definition 3.3.2.** The **residual degree** of  $\mathfrak{p}$  in  $L/K$  is

$$f_{\mathfrak{p}} = [B/\mathfrak{p} : A/\mathfrak{q}].$$

**Definition 3.3.3.** The extension  $L/K$  is **totally ramified** if there is a unique prime ideal  $\mathfrak{p}$  dividing  $\mathfrak{q}$  and  $f_{\mathfrak{p}} = 1$ . If  $B/\mathfrak{p}$  is separable extension  $A/\mathfrak{q}$  and  $e_{\mathfrak{p}} = 1$ , we say that  $L/K$  is **unramified**.

**Proposition 3.3.4.** Let  $\mathfrak{q} \subset A$  maximal ideal, then

$$[L : K] = [B/\mathfrak{q}B : A/\mathfrak{q}] = \sum_{\mathfrak{p} \mid \mathfrak{q}} e_{\mathfrak{p}} f_{\mathfrak{p}}.$$

*Proof.* For the second equality, we write

$$\mathfrak{q}B = \prod \mathfrak{p}_i^{e_{\mathfrak{p}_i}} \subset \dots \subset \mathfrak{p}_2 \mathfrak{p}_1^{e_{\mathfrak{p}_1}} \subset \mathfrak{p}_1^{e_{\mathfrak{p}_1}} \subset \dots \subset \mathfrak{p}_1^2 \subset \mathfrak{p}_1 \subset B$$

with all intermediate quotients of dimension 1 over  $B/\mathfrak{p}_i$ , hence of dimension  $f_{\mathfrak{p}_i}$  over  $A/\mathfrak{q}$ . Each quotient of dimension  $f_{\mathfrak{p}_i}$  appears  $e_{\mathfrak{p}_i}$  times, hence  $[B/\mathfrak{q}B : A/\mathfrak{q}] = \sum_{\mathfrak{p} \mid \mathfrak{q}} e_{\mathfrak{p}} f_{\mathfrak{p}}$ .

The first equality holds if  $A$  is principal since a basis of the  $A$ -module  $B$  is also the basus of  $B/\mathfrak{p}$  as  $A/\mathfrak{q}$ -module (any nontorsion finite type module over a principal ring is free). For the general case denote  $A_0 = A_{\mathfrak{q}}$ , it is a discrete valuation ring, with field of fractions  $L$ . Denote  $B_0$  the integral closure of  $A_0$  in  $L$ , then it is easy to check that  $B_0 = A_{\mathfrak{q}}B$ . By the case when  $A$  principal we deduce

$$[L : K] = n = [B_0/\mathfrak{q}B_0 : A_0/\mathfrak{q}A_0].$$

Write

$$\mathfrak{q}B_0 = \prod_i (B_0\mathfrak{p}_i)^{e_{\mathfrak{p}_i}}.$$

By construction,  $B_0\mathfrak{p}_i$  are prime ideals of  $B_0$ . Hence

$$[B_0/\mathfrak{q}B_0 : A_0/\mathfrak{q}] = \sum e_{\mathfrak{p}_i} [B_0/B_0\mathfrak{p}_i : A_0/\mathfrak{q}].$$

Here  $B_0/B_0\mathfrak{p}_i \simeq B/\mathfrak{p}_i$  and  $A_0/\mathfrak{q} \simeq A/\mathfrak{q}$  and we deduce the result.  $\square$

### 3.3.1 Galois case

Assume  $L/K$  is Galois. Then

**Proposition 3.3.5.** *Let  $\mathfrak{q} \subset A$  be maximal ideal. The group  $\text{Gal}(L/K)$  acts transitively on the set of prime ideals of  $B$  dividing  $\mathfrak{q}$ .*

*Proof.* Let  $\mathfrak{p} \subset B$  be such that  $\mathfrak{p} | \mathfrak{q}$ . Then the image  $\mathfrak{p}'$  of  $\mathfrak{p}$  under the  $\text{Gal}(L/K)$ -action is contained in  $B$  and is a sub- $B$ -module of  $L$ . Hence  $\mathfrak{p}'$  is a prime ideal dividing  $\mathfrak{q}$ .

For the transitivity, let  $x \in \mathfrak{p}$  and  $\mathfrak{p}' \subset B$  with  $\mathfrak{p}' | \mathfrak{q}$ . By the approximation lemma we can choose such  $x$  such that  $x$  does not belong to all conjugates of  $\mathfrak{p}'$  distinct from  $\mathfrak{p}$ . Then  $N_{L/K}(x) = \prod \sigma(x) \in \mathfrak{p} \cap A$ , hence  $N_{L/K}(x) \in \mathfrak{q} \cap A = \mathfrak{p}' \cap A$ , hence  $N_{L/K}(x) \in \mathfrak{p}'$  and there exists  $\sigma \in \text{Gal}(L/K)$  such that  $x \in \sigma(\mathfrak{p}')$ , so that  $\sigma(\mathfrak{p}')$  should be equal to  $\mathfrak{p}'$ , by the choice of  $x$ .  $\square$

**Corollary 3.3.6.** *The invariants  $e_{\mathfrak{p}}$  and  $f_{\mathfrak{p}}$  depend only on  $\mathfrak{q}$ . If  $g_{\mathfrak{q}}$  is the number of prime ideals of  $A$  dividing  $\mathfrak{q}$ , then*

$$[L : K] = g_{\mathfrak{q}} e_{\mathfrak{q}} f_{\mathfrak{q}}.$$

### 3.3.2 Explicit factorisation

Let  $A, B, K, L$  be as above. Assume  $B = A[\alpha]$ .

**Proposition 3.3.7.** *Let  $\mathfrak{p} \subset A$  be a prime ideal. If  $f$  is minimal irreducible polynomial of  $\alpha$  and  $\bar{f}(\alpha) = \prod \bar{P}_i(x)^{e_i}$  factorisation in  $A/\mathfrak{p}$  with  $\bar{P}_i$  irreducible monic, then*

$$\mathfrak{p}B = \prod \beta_i^{e_i}$$

*is the factorisation of  $\mathfrak{p}$  and  $\beta_i = \mathfrak{p}B + P_i(\alpha)B$ , where  $P_i \in A[X]$  irreducible, that lifts  $\bar{P}_i$ .*

*Proof.* Let  $\bar{\alpha}$  be a root of  $\bar{P}_i$ . Then

$$\beta := \ker(A[\alpha] \rightarrow \bar{A}[\bar{\alpha}])$$

satisfies

$$\mathfrak{p}B + P(\alpha)B \subset \beta.$$

If  $b = g(\alpha) \in \beta$  with  $g \in A[x]$  (i.e.  $b \in B$ ), then

$$\bar{g} = \overline{Ph}$$

with  $\bar{h} \in \bar{A}[x]$ , so that  $g - Ph$  has coefficients in  $\mathfrak{p}$  and  $g(\alpha) \in \mathfrak{p}B + P(\alpha)B$ .

Now, if  $e'_i$  is the ramification index of  $\beta_i$ , then  $f_{\beta_i} = \deg \bar{P}_i$ . But  $f(\alpha) = 0$ , so that  $f(x) - \prod P_i^{e_i} \in \mathfrak{p}A[x]$ , i.e.  $\prod P_i(\alpha)^{e_i} \in \mathfrak{p}B$  and

$$\beta_i^{e_i} \subset \mathfrak{p}B + P_i(\alpha)^{e_i}B,$$

hence

$$\prod \beta_i^{e_i} \subset \mathfrak{p}B + \prod P_i(\alpha)^{e_i}B \subset \mathfrak{p}B = \prod \beta_i^{e'_i}.$$

We deduce that  $e_i \geq e'_i$ , but

$$\sum e_i f_i = \deg f = [L : F] = \sum e'_i f_i,$$

so that  $e_i = e'_i$ . □

The following lemma on ideals containing an integer will be also useful:

**Lemma 3.3.8.** *Let  $A$  be a ring and let  $I$  be an ideal of  $A$ . Let  $M, N \in \mathbb{Z}$  be two relatively prime integers, such that  $MN \in I$ . Then  $I = JJ'$  where  $J = I + MA$  and  $J' = I + NA$ .*

*Proof.* We observe that  $JJ' = I^2 + NI + MI + MNA$ . But  $I \subset NI + MI$  (since  $M$  and  $N$  are relatively prime, by Bezout theorem), hence  $I = NI + MI$ . Since  $MN \in I$  and  $I^2 \subset I$ , so that  $JJ' = I$ , as claimed. □

### 3.3.3 Complement: extensions of valued fields

Using properties of the Dedekind rings, we can now establish theorem 2.4.1. Let  $K$  be a field, complete with respect to a discrete absolute value  $|\cdot|_K$  and let  $L/K$  be a finite separable extension.

**Theorem 3.3.9.** *The absolute value  $|\cdot|_K$  extends uniquely to a discrete absolute value  $|\cdot|_L$  on  $L$ . In addition,  $L$  is complete and the absolute value is characterized by  $|\beta|_L = |Nm_{L/K}(\beta)|_K^{\frac{1}{n}}$ .*

*Proof.* Let  $A$  be the discrete valuation ring in  $K$ , and let  $B$  be its integral closure in  $L$ . Let  $\mathfrak{p}$  be the maximal ideal of  $A$ . Then  $B$  is a Dedekind ring, and the absolute values of  $L$  extending  $|\cdot|_{\mathfrak{p}}$  correspond to the ideals of  $B$  lying over  $\mathfrak{p}$ . Assume that there are distinct prime ideals  $\mathfrak{p}_1$  and  $\mathfrak{p}_2$  in  $B$  dividing  $\mathfrak{p}$ . Then there is an element  $\beta \in B$  such that  $\mathfrak{p}_1 \cap A[\beta] \neq \mathfrak{p}_2 \cap A[\beta]$  (for instance for  $\beta \in \mathfrak{p}_1 \setminus \mathfrak{p}_2$ ). Let  $f(X)$  be the minimal polynomial of  $\beta$  over  $K$ . Since  $f$  is irreducible and  $A$  is complete, Hensel's lemma shows that the image  $\bar{f}(X)$  in  $k[X]$ ,  $k = A/\mathfrak{p}$  should be a power of an irreducible polynomial. Then  $A[\beta]/\mathfrak{p}A[\beta] = k[X]/(\bar{f})$  is a local ring, and we obtain a contradiction since  $A[\beta]$  contains two prime ideals containing  $\mathfrak{p}$ .

Hence  $|\cdot|_{\mathfrak{p}}$  extends uniquely to an absolute value  $|\cdot|_L$  on  $L$ . Similarly,  $|\cdot|_{\mathfrak{p}}$  also extends uniquely to an absolute value  $|\cdot|_{L'}$  on a Galois closure  $L'$  of  $L$ . For each  $\sigma \in \text{Gal}(L'/K)$ , consider the map  $L \rightarrow \mathbb{C}$ ,  $\beta \mapsto |\sigma\beta|_{L'}$ . This is again an absolute value on  $L$ , and so the uniqueness implies that  $|\beta|_L = |\sigma\beta|_{L'}$ , so that

$$|\beta|_L = |Nm_{L/K}(\beta)|_K^{\frac{1}{n}}.$$

Let us show that  $L$  is complete. If  $e_1, \dots, e_n$  is a basis of  $B$  as an  $A$ -module, and let  $a_m$  be a Cauchy sequence in  $L$ . Then  $a_m = \sum a_m^i e_i$  and each sequence  $a_m^i$  is also a Cauchy sequence (for a fixed  $i$ , hence converges in  $K$ , so that  $a_m$  converges in  $L$ .  $\square$



# Chapter 4

## Number fields

In this section we will investigate the properties of finite extensions of  $\mathbb{Q}$  and their rings of integers. We will apply and extend the general results from the previous chapter.

### 4.1 Geometry of numbers

In this section we review classical results on lattices in  $\mathbb{R}^n$ , this theory was initiated by Minkowski in 1896.

#### 4.1.1 Lattices in $\mathbb{R}^n$

Let  $V$  be the vector space  $\mathbb{R}^n$ , we do not specify the norm on  $V$  (since they are all equivalent). Recall the following notions:

1.  $D \subset V$  is **discrete** if for any real number  $r > 0$ , the set  $\{v \in D, |v| \leq r\}$  is finite;
2. a **lattice** in  $V$  is a discrete subgroup  $L$  of  $V$  that generates  $V$  as an  $\mathbb{R}$ -vector space.

For example,  $L_1 = \mathbb{Z}^2$  in  $\mathbb{R}^2$ , or  $L_2 = \{(a, b) \in \mathbb{Z}^2, a \equiv 2b \pmod{3}\}$ .

**Exercise:** Find a  $\mathbb{Z}$ -basis of  $L_2$ .

If  $e = e_1, \dots, e_n$  is a basis of  $V$ , we denote:

$$L(e) = \{m_1 e_1 + \dots + m_n e_n, \mid m_1, m_2, \dots, m_n \in \mathbb{Z}\}$$

and we denote

$$\Pi(e) = \left\{ \sum v_i e_i, \mid v_i \in [0, 1] \right\} \subset V.$$

**Theorem 4.1.1.** *Let  $L \subset V$  be a subgroup. The following conditions are equivalent:*

- (i)  $L$  is a lattice;

(ii)  $L$  is generated over  $\mathbb{Z}$  by a finite number of elements, forming an  $\mathbb{R}$ -basis of  $V$ . In particular, any lattice in  $\mathbb{R}^n$  admits a  $\mathbb{Z}$ -basis with  $n$  elements.

*Proof.* The proof is divided into following steps:

- (i)  $L(e)$  is a lattice and any element of  $V$  could be written in a form  $v = \lambda(v) + x(v)$  with  $\lambda(v) \in L(e)$  and  $x(v) \in \Pi(e)$ .
- (ii) if for  $e = \{e_1, \dots, e_n\}$  a basis of  $V$  we have that  $e_i \in L$ , then  $L(e)$  is of finite index in  $L$  and there is an integer  $N \geq 1$  such that  $L \subset \frac{1}{N}L(e)$ .
- (iii)  $L = L(e)$  for  $e$  a basis in  $V$  such that  $|\det(e_1, \dots, e_n)|$  is minimal for the set  $B$  of all the bases in  $V$  with  $e_i \in L$ .

The part (i) is standard. We just note that to see that  $L(e)$  is discrete one can use the norm  $|\sum_i v_i e_i| = \sup_i |v_i|$  on  $V$ .

To show (ii) first observe that  $L(e) \subset L$ . Let  $v \in L$  and write  $v = \lambda(v) + x(v)$  with  $\lambda(v) \in L(e)$  and  $x(v) \in \Pi(e)$ , hence  $x(v) = v - \lambda(v) \in L \cap \Pi(e)$  which is a finite set since  $L$  is a lattice. Hence

$$N = |L/L(e)| \leq |L \cap \Pi(e)| < \infty.$$

By Lagrange theorem, this  $N$  works.

To show (iii), we first observe that  $B$  is nonempty. We fix  $e = \{e_1, \dots, e_n\} \in B$ , then  $L(e) \subset L$  and let  $N$  be as in (ii). Then any element in  $L$  could be written as  $\sum_i \frac{m_i}{N} e_i$  with  $m_i \in \mathbb{Z}$ . In particular, for  $f \in B$ , we have  $\det(f) \in N^{-n} \det(e) \mathbb{Z}$ . This set is discrete, hence there exists  $e \in B$  with  $|\det(e)|$  minimal. To show that  $L = L(e)$  by the part (i) it is enough to show that  $\Pi(e) \cap L = \emptyset$ . Let  $v = \sum v_i e_i \in \Pi(e) \cap L$ ,  $0 < v_i < 1$ . Let  $1 \leq i \leq n$ . If  $v_i \neq 0$ , then  $v, e_1, \dots, \hat{e}_i, \dots, e_n$  is in  $B$ . We compute the determinant:

$$|\det(v, e_1, \dots, \hat{e}_i, \dots, e_n)| = |v_i \det(e)| < \det(e).$$

contradiction with the choice of  $e$ . □

We deduce as a corollary:

**Corollary 4.1.2.** *Let  $L \subset \mathbb{R}^n$  be a lattice generated over  $\mathbb{Z}$  by  $a = \{a_1, \dots, a_m\}$ . The following conditions are equivalent:*

- (i)  $a$  is a  $\mathbb{Z}$ -basis of  $L$ ;
- (ii)  $a$  is a basis of  $V$ ;
- (iii)  $m = n$ . In particular, all the  $\mathbb{Z}$ -bases of  $L$  have the same cardinality  $n$ .

We now investigate how to change the basis of a lattice  $L$ .

Let

$$GL_n(\mathbb{Z}) = \{A \in M_n(\mathbb{Z}), \exists B \in M_n(\mathbb{Z}), AB = I_n\}.$$

The following properties are easy to verify:

**Proposition 4.1.3.** (i)  $GL_n(\mathbb{Z}) = \{A \in M_n(\mathbb{Z}), \det(A) = \pm 1\}$ .

(ii) Let  $e$  and  $f$  are two bases of  $V$  and let  $P$  be the change of basis matrix from  $e$  to  $f$ . Then  $L(e) = L(f)$  iff  $P \in GL_n(\mathbb{Z})$ .

In particular,  $(a, b), (c, d) \in \mathbb{Z}^2$  generate  $\mathbb{Z}^2$  iff  $ad - bc = \pm 1$ .

The decomposition in the part (i) of theorem 4.1.1 leads to the following more general notion:

**Definition 4.1.4.** Let  $L$  be a lattice in  $\mathbb{R}^n$  and let  $X \subset \mathbb{R}^n$  is a (Lebesgue) measurable set. We say that  $X$  is a **fundamental domain** for the action of  $L$  if any  $v \in \mathbb{R}^n$  could be written, in a unique way, as  $v = \lambda + x$ .

There exists many fundamental domains for  $L$ , for example,  $\Pi(e)$  is a fundamental domain of  $L(e)$  with measure  $|\det(e_1, \dots, e_n)|$ , for  $e = \{e_1, \dots, e_n\}$ .

**Lemma 4.1.5.** Let  $L$  be a lattice of  $V$ ,  $X, Y \subset V$  two measurable subsets. Assume that  $X$  is a fundamental domain of  $V$  and that for  $x, y \in Y$  we have

$$x - y \in L \Rightarrow x = y.$$

Then  $\mu(Y) \leq \mu(X)$ .

*Proof.* Since  $X$  is a fundamental domain, by definition we have a decomposition:

$$V = \bigsqcup_{\lambda \in L} (X + \lambda).$$

Then

$$Y = \bigsqcup_{\lambda \in L} (Y \cap (X + \lambda)),$$

hence

$$\mu(Y) = \sum_{\lambda \in L} (\mu(X \cap (Y - \lambda))).$$

By hypothesis,  $Y$  and  $Y - \lambda$  are disjoint, hence the right hand side is at most  $\mu(X)$ .  $\square$

In particular, all the fundamental domains of  $L$  have the same (nonzero) measure, that equals to the measure of  $\Pi(e)$  for  $e$  a basis of  $V$  such that  $L = L(e)$ . We call this measure **covolume**  $\text{covol}(L)$  of the lattice  $L$ .

**Proposition 4.1.6.** Let  $L \subset \mathbb{R}^n$  be a lattice and let  $L' \subset L$  be a subgroup. The following are equivalent:

- $L'$  is a lattice;
- $L'$  is of finite index in  $L$ .

If these assumptions are satisfied, then  $\text{covol}(L') = |L/L'| \text{covol}(L)$ .

*Proof.* The equivalence of (i) and (ii) follows easily from the proof of theorem 4.1.1. We show that  $\text{covol}(L') = |L/L'| \text{covol}(L)$ . Let  $h = |L/L'|$ . Then

$$L = \bigsqcup_{i=1}^h (L' + \lambda_i).$$

for  $\lambda_i \in L$ . Let  $X$  be a fundamental domain for  $L$  and let

$$X' = \bigsqcup_{i=1}^h (X + \lambda_i),$$

then  $\mu(X') = h\mu(X)$  (since the Lebesgue measure is invariant under translations), and  $X'$  is a fundamental domain for  $L'$ :

$$\mathbb{R}^n = \bigsqcup_{\lambda \in L} (X + \lambda) = \bigsqcup_{i=1}^h \bigsqcup_{\lambda' \in L'} (X + x_i + \lambda') = \bigsqcup_{\lambda' \in L'} (X' + \lambda').$$

□

For example, for  $L = \{(a, b) \in \mathbb{Z}^2, a \equiv 2b \pmod{3}\}$ , we could use that we have a map

$$\mathbb{Z}^2 \rightarrow \mathbb{Z}/3\mathbb{Z}, (a, b) \mapsto a - 2b,$$

that is surjective and has kernel  $L$ , hence  $\mathbb{Z}^2/L \simeq \mathbb{Z}/3\mathbb{Z}$  and by the proposition above we obtain  $\text{covol}(L) = 3$ .

The following property will be very important for the applications.

**Theorem 4.1.7.** (*Minkowski*) *Let  $C \subset V$  be a measurable subset, convex and symmetric (i.e. for  $x \in C$  we have  $-x \in C$ ). Let  $L$  be a lattice in  $V$ . If  $\text{covol}(L) < \mu(C)/2^n$ , or if  $C$  is compact and  $\text{covol}(L) \leq \mu(C)/2^n$ , then the intersection  $L \cap C$  contains a nonzero element.*

*Proof.* Let  $L' = 2L$ , then  $L'$  is a lattice and  $\text{covol}(L') = 2^n \text{covol}(L)$ .

In the first case, assume  $\text{covol}(L) \leq \mu(C)/2^n$ . Then lemma 4.1.5 says that there are distinct  $x, y \in C$  with  $x - y \leq 2L \cap C$ . Since  $C$  is convex and symmetric,  $\frac{x-y}{2}$  is in  $C \cap L$  nonzero, as claimed.

In the second case, if  $C$  is compact and  $\text{covol}(L) \leq \mu(C)/2^n$ , then we introduce for  $\epsilon > 0$  the set

$$C_\epsilon = \{v \in V, \exists c \in C, |v - c| < \epsilon\}.$$

It is an open convex symmetric domain, with measure  $> \mu(C)$ . By the previous case,  $(L \setminus \{0\}) \cap C_\epsilon \neq \emptyset$ . Since  $L$  is discrete, this set is also finite, and decreasing  $\epsilon$  it becomes smaller (with respect to the natural inclusion), hence equals to a (hence nonempty) set  $(L \setminus \{0\}) \cap C$ .

□

## 4.1.2 Applications in arithmetics

In this paragraph we will use the properties of lattices developed above, to establish various classical results in arithmetics.

**Theorem 4.1.8.** *(Fermat, Euler) Let  $p$  be a prime number,  $p \equiv 1 \pmod{4}$ . Then there exists  $a, b \in \mathbb{Z}$  with  $p = a^2 + b^2$ .*

*Proof.* First note that  $-1$  is a square modulo  $p$  by assumption  $p \equiv 1 \pmod{4}$ . We introduce the following lattice:

$$L = \{(a, b) \in \mathbb{Z}^2, a \equiv ub \pmod{p}\},$$

where we fix  $u \in \mathbb{Z}$  with  $u^2 \equiv -1 \pmod{p}$ . We consider the map

$$\psi : \mathbb{Z}^2 \rightarrow \mathbb{Z}/p\mathbb{Z}, (a, b) \mapsto a - bu.$$

It is a surjective map with kernel  $L$ . Hence  $L$  is a lattice in  $\mathbb{R}^2$  of covolume  $p$ , by proposition 4.1.6. Note that for any  $(a, b) \in L$ , we have

$$a^2 + b^2 \equiv (u^2 + 1)b^2 \equiv 0 \pmod{p}.$$

Let  $C(r) = \{(x, y) \in \mathbb{R}^2, x^2 + y^2 \leq r\}$ . Using that  $2p\pi > 4\text{covol}(L) = 4p$  and Minkowski's theorem 4.1.7, we deduce that  $L \cap C(2p) \neq \{0\}$ . But then  $0 < a^2 + b^2 < 2p$  and  $a^2 + b^2 \equiv 0 \pmod{p}$ , the only possibility is that  $p \equiv a^2 + b^2$ .  $\square$

A similar argument leads to the following:

**Theorem 4.1.9.** *Let  $d \geq 1$  be an integer and let  $p > 0$  be an integer such that  $-d$  is a square modulo  $p$ . Then at least one integer in the following list:*

$$p, 2p, 3p, \dots, hp,$$

where  $h$  is the integral part of  $4\frac{\sqrt{d}}{\pi}$ , is of the form  $a^2 + db^2$ .

*Proof.* We first note that the condition that  $-d$  is a square modulo  $p$  is necessary in order to have that  $p \mid a^2 + db^2$ . As in the previous theorem, let  $u \in \mathbb{Z}$  with  $u^2 \equiv -d \pmod{p}$  and let

$$L = \{(a, b) \in \mathbb{Z}^2, a \equiv ub \pmod{p}\}.$$

Similarly, we obtain that  $L$  is a lattice of covolume  $p$  with  $a^2 + db^2 \equiv 0 \pmod{p}$  for any  $(a, b) \in L$ . Denote

$$C(r) = \{(x, y) \in \mathbb{R}^2, x^2 + dy^2 \leq r\},$$

its volume is  $\frac{\pi r}{\sqrt{d}}$ . Using Minkowski's theorem, there is a nonzero element in  $L \cap C(\frac{4\sqrt{d}}{\pi}p)$ . If  $(a, b) \in \mathbb{Z}^2$  is such element, then  $a^2 + db^2 = kp$ , where  $k$  is an integer such that  $0 < k \leq \frac{4\sqrt{d}}{\pi}p$ .  $\square$

As a corollary, we obtain:

**Corollary 4.1.10.** (i) A prime  $p$  is of the form  $p = a^2 + 2b^2$  iff  $p \equiv 1, 3 \pmod{8}$ .

(ii) A prime  $p \neq 3$  is of the form  $p = a^2 + 3b^2$  iff  $p \equiv 1 \pmod{3}$ .

(iii) Let  $p$  be a prime. Then  $p$  (resp.  $2p$ ) is of the form  $a^2 + 5b^2$  iff  $p \equiv 1, 9 \pmod{20}$  (resp.  $p \equiv 3, 7 \pmod{20}$ ).

*Proof.* (i) It is enough to observe that in the theorem above  $h = 1$  for  $d = 2$ .

(ii) By the theorem above,  $p$  or  $2p$  is of the form  $a^2 + 3b^2$ , by an argument modulo 4, we see that  $2p$  is impossible.

(iii) By quadratic reciprocity,  $-5$  is a square modulo  $p$  iff

$$p \equiv 1, 3, 7, 9 \pmod{20}.$$

We obtain the result again by an argument modulo 4.

As another application we obtain Lagrange's theorem:

**Theorem 4.1.11.** (Lagrange) Any positive integer is a sum of four squares.

*Proof.* We use the classical fact that  $-1$  is a sum of two squares in  $\mathbb{Z}/n\mathbb{Z}$  if  $n$  has no square factors. Let  $u, v \in \mathbb{Z}$  with  $1 + u^2 + v^2 \equiv 0 \pmod{n}$ . Then one verifies that

$$L = \{(a, b, c, d) \in \mathbb{Z}^4, c \equiv au + bv \pmod{n}, d \equiv av - bu \pmod{n}\}$$

is a lattice in  $\mathbb{Z}^4$  of covolume  $n^2$ , for any  $(a, b, c, d) \in L$ , we have  $a^2 + b^2 + c^2 + d^2 \equiv 0 \pmod{n}$ . If

$$C(r) = \{x_1, x_2, x_3, x_4 \in \mathbb{R}^4, \sum x_i^2 < r, \}$$

then the measure of  $C(r)$  is  $\frac{\pi^2}{2}r^2$ . Then we can apply Minkowski's theorem to get that  $L \cap C(2n) \neq \{0\}$ , and  $n = a^2 + b^2 + c^2 + d^2$  for an element  $(a, b, c, d)$  in this intersection. □

More generally, we will be interested in values of integral quadratic forms.

**Definition 4.1.12.** A **binary integral quadratic form** is a function  $q : \mathbb{Z}^2 \rightarrow \mathbb{Z}$  such that

$$q(x, y) = ax^2 + bxy + cy^2$$

for  $a, b, c \in \mathbb{Z}$ .

- We say that  $q$  **represents**  $n \in \mathbb{Z}$  if there exists  $(x, y) \in \mathbb{Z}^2$  such that  $q(x, y) = n$ ; for  $n = 0$  the condition  $(x, y) \neq (0, 0)$  is required.
- The discriminant of the form  $q$  is

$$\text{disc}(q) = b^2 - 4ac.$$

- We say that two forms  $q, q'$  are **equivalent** (resp. **properly equivalent**) and we write  $q \equiv q'$  if there exists  $A \in GL_2(\mathbb{Z})$  (resp. and  $\det(A) = 1$ ) with  $q'(x, y) = q(A(x, y))$ .

We write  $q = (a, b, c)$  for the form  $q$  as above. Note that the integers  $a, b, c$  depend only on  $q$  since  $a = q(1, 0), b = q(0, 1)$  and  $c = q(1, 1) - a - b$ . If  $\text{disc}(q) < 0$  we will always assume that  $q$  is positive, i.e.  $a > 0$ . We will always assume that  $a, b, c$  are relatively prime, and we will say that  $n$  is represented **primitively** if  $x$  and  $y$  are relatively prime. It is easy to see that the set of integers represented by a form  $q$  depend only on the equivalence class of  $q$ .

**Theorem 4.1.13.** (Lagrange) *An integer  $n$  is primitively represented by a quadratic form of discriminant  $D$  iff  $D$  is a square modulo  $4n$ .*

*Proof.* If  $D$  is a square mod  $4n$ , then  $D = b^2 - 4nc$  for some  $b, c \in \mathbb{Z}$ . Then  $q = (n, b, c)$  represents primitively  $n$ .

For the converse, we use the lemma below:  $q$  is properly equivalent to  $(n, b, c)$ , but then  $\text{disc}(q) = \text{disc}(n, b, c) = b^2 - 4nc$ , so that  $\text{disc}(q)$  is a square mod  $4n$ .  $\square$

**Lemma 4.1.14.** *An integer  $n$  is primitively represented by a form  $q$  iff  $q$  is properly equivalent to a form  $(n, b, c)$  with  $-|n| < b \leq |n|$ , if  $n \neq 0$ .*

*Proof.* Let  $n = q(u)$  with  $u$  primitive (i.e. the coordinates of  $u$  are relatively prime). Then there is a vector  $v \in \mathbb{Z}^2$ , such that  $(u, v)$  is a basis of  $\mathbb{Z}^2$  and  $\det(u, v) = 1$ . Then the form  $q'(x, y) = q(ux + vy)$  is properly equivalent to  $q$  and satisfies  $q' = (n, *, *)$ . Using the following proper equivalences, we obtain a form  $(n, b, c)$  properly equivalent to  $q'$  with  $-|n| < b \leq |n|$ :

$$(a, b, c) \sim (a, -b, c) \sim (a, b + 2a, c + b + a) \sim (a, b - 2a, c - b + a).$$

$\square$

**Theorem 4.1.15.** *Let  $D \in \mathbb{Z}$  and let  $p$  be a prime such that  $D$  is a square modulo  $4p$ . Then there is a unique, up to equivalence, quadratic form  $q$  of discriminant  $D$  such that  $p$  is represented by  $q$ .*

*Proof.* As in the previous theorem,  $p$  is represented by a form of type  $(p, b, c)$  with  $-p < b \leq p$ . Note that we have

$$(p, b, c) \sim (p, \pm b, c),$$

and we may assume  $0 \leq b \leq p$ , so that  $D = b^2 - 4pc$  and  $c = \frac{b^2 - D}{4p}$ . Hence we only need to show that there is a unique integer  $b$  with  $0 \leq b \leq p$  and  $D \equiv b^2 \pmod{4p}$ . It is straightforward if  $p = 2$ , so that we assume that  $p$  is odd. If  $b, b'$  satisfy these conditions, then we have  $b^2 \equiv (b')^2 \pmod{4p}$ , hence  $b = \pm b' \pmod{p}$  since  $p$  is a prime,

hence  $b' = b$  or  $b' = p - b$ , but we also have  $b \equiv b' \pmod{2}$  and we deduce  $b = b'$ .  $\square$

The theorem above could be viewed as a motivation to investigate the forms of a given discriminant. Note first that if  $D$  is a discriminant of a binary integral quadratic form, then one should have

$$D \equiv 0, 1 \pmod{4}.$$

We then consider the following forms:

$$q(x, y) = x^2 - \frac{D}{4}y^2, D \equiv 0 \pmod{4}, \text{ and } q(x, y) = x^2 + xy + \frac{1-D}{4}y^2, D \equiv 1 \pmod{4}.$$

We call the forms above a **principal form** of discriminant  $D$ .

**Theorem 4.1.16.** (*Lagrange*) *Up to equivalence (resp. proper equivalence), there is a finite number of quadratic forms of a given discriminant.*

*Proof.* We leave the case  $D$  is a square as an exercise. It is then enough to show that a form  $q$  with discriminant  $D$  is properly equivalent to a form  $(a, b, c)$  with  $-|a| < b \leq |a| \leq |c|$ . The finiteness then follows, since there is a finite number of such triples.

Note that  $1 \leq |a| \leq \sqrt{\frac{|D|}{3}}$ , since

$$4|a|^2 \leq 4|ac| = |b^2 - D| \leq |a|^2 + |D|.$$

Let now  $q$  be a form of discriminant  $D$  and let  $\mathcal{V} = \{|x|\}$  with  $x$  represented primitively by  $q$ , this set is nonempty and does not contain 0 since  $D$  is not a square, hence we can choose  $|a| = \min \mathcal{V}$  and  $u \in \mathbb{Z}^2$  primitive with  $a = q(u)$ . By lemma 4.1.14,  $q$  is properly equivalent to  $(a, b, c)$  and  $-|a| < b \leq |a|$ . We also have  $|a| \leq |c|$  from the choice of  $u$ , since  $c$  is primitively represented by  $(a, b, c)$ , and hence by  $q$ .  $\square$

The theorem above allows to introduce the following notions:

**Definition 4.1.17.** We denote  $Cl(D)$  the set of proper equivalence classes of quadratic forms of discriminant  $D$ , we denote  $P(D)$  the set of proper equivalence classes of primitive quadratic forms of discriminant  $D$  and  $h(D) = |P(D)|$ .

**Corollary 4.1.18.** *If  $h(D) = 1$ , then any odd prime  $p$  such that  $D$  is a square mod  $p$  is represented by a principal form of discriminant  $D$ .*

We have also the following result, that we state here without proof, for forms of negative discriminant:

**Theorem 4.1.19.** (*Stark, Heegner, Baker*) *If  $D < 0$ , then  $h(D) = 1$  iff*

$$D \in \{-3, -4, -7, -8, -11, -12, -16, -19, -27, -28, -43, -67, -163.\}$$



## 4.2 Rings of integers of number fields

### 4.2.1 Number fields

Recall that a **number field** is a finite extension  $L$  of  $\mathbb{Q}$ . We also denote  $\bar{\mathbb{Q}} \subset \mathbb{C}$  the algebraic closure of  $\mathbb{Q}$  in  $\mathbb{C}$ :  $\bar{\mathbb{Q}} = \{x \in \mathbb{C}, \exists P \in \mathbb{Q}[X], | P(x) = 0\}$ .

For  $K$  a number field and  $x \in \bar{\mathbb{Q}}$  we denote  $P_{x,K}$  the minimal polynomial of  $x$  over  $K$ . As in section 3.2.1, for  $L/K$  an extension of number fields, we consider the linear map

$$L \rightarrow L, y \mapsto xy,$$

we define its trace  $Tr_{L/K}(x)$  (resp. norm  $N_{L/K}(x)$ ) and we call it the trace (resp. the norm) of  $x$ . We also denote  $\chi_{x,L/K}$  the characteristic polynomial of this map. The following properties are standard:

- $\chi_{x,L/K} = P_{x,K}^r$  where  $r = [L : K(x)]$ ;
- $Tr_{L/K}(x) = \sum_{\sigma \in Gal(L/K)} \sigma(x)$
- $Nm_{L/K} = \prod_{\sigma \in Gal(L/K)} \sigma(x)$
- $\chi_{x,L/K}(X) = \prod_{\sigma \in Gal(L/K)} (X - \sigma(x))$ .

**Definition 4.2.1.** Let  $L/K$  be a degree  $n$  extension of number fields. The **discriminant** of a family  $e_1, \dots, e_n$  of elements of  $L$  is the determinant of the  $n \times n$  matrix  $Tr_{L/K}(e_i e_j)$ , we denote it  $disc_{L/K}(e_1, \dots, e_n)$ , note that it is an element of  $K$ .

**Proposition 4.2.2.** Let  $L/K$  be a degree  $n$  extension and let  $e = \{e_1, \dots, e_n\} \in L$ . Then

- (i)  $disc_{L/K}(e_1, \dots, e_n) \neq 0$  iff  $e_1, \dots, e_n$  is a  $K$ -basis of  $L$ .
- (ii) If  $f = \{f_1, \dots, f_n\}$ ,  $f = Pe$  with  $P \in M_n(K)$  is another family, then

$$disc_{L/K}(f_1, \dots, f_n) = \det(P)^2 disc_{L/K}(e_1, \dots, e_n).$$

- (iii)  $disc_{L/K}(e_1, \dots, e_n) = disc_{L/K}(\sigma_i(e_j))^2$ .

- (iv) if  $L = K(x)$  and if  $x_1, \dots, x_n \in \mathbb{C}$  are the conjugates of  $x$ , then

$$disc_{K(x)/K}(1, x, \dots, x^{n-1}) = \prod_{i < j} (x_i - x_j)^2 = (-1)^{\frac{n(n-1)}{2}} N_{K(x)/K}(P'_{x,K}(x)).$$

*Proof.* The statements (i) and (ii) are standard linear algebra properties, for (ii) we use that  $Tr_{L/K} = \sum_{k=1}^n \sigma_k$ , so that  $Tr_{L/K}(e_i e_j) = (\sigma_i(e_j), \sigma_i(e_j))^t$ , that leads to the Vandermode determinant computation.  $\square$

Let  $K$  be a number field and let  $\mathcal{O}_K$  be the ring of integers of  $K$ : it is the integral closure of  $\mathbb{Z}$  in  $K$ . The following properties follow from section 3.2.1:

- (i) For  $x \in \mathcal{O}_K$  we have  $P_{x,K}, \chi_{x,K/\mathbb{Q}} \in \mathbb{Z}[X], \text{Tr}_{K/\mathbb{Q}}(x), N_{K/\mathbb{Q}}(x) \in \mathbb{Z}$ .
- (ii)  $\mathcal{O}_K$  is a Dedekind ring.

We also have the following structural property:

**Theorem 4.2.3.** (Dedekind) *Let  $K$  be a number field and let  $n = [K : \mathbb{Q}]$ . Then there exists  $e_1, \dots, e_n \in \mathcal{O}_K$  forming a basis of  $\mathcal{O}_K$  over  $\mathbb{Z}$ , i.e.  $\mathcal{O}_K = \sum_{i=1}^n \mathbb{Z}e_i$ .*

*Proof.* By a primitif element theorem, we can write  $K = \mathbb{Q}(x)$ , and, up to multiplying  $x$  by an integer, we may assume that  $x$  is integral over  $\mathbb{Z}$ , so that we have  $\mathbb{Z}[x] \subset \mathcal{O}_K$ . Consider now  $z \in \mathcal{O}_K$ . Then one checks that  $d = \text{disc}_{K/\mathbb{Q}}(1, x, \dots, x^{n-1})$  is a nonzero integer. From the inclusions

$$\mathbb{Z}[x] \subset \mathcal{O}_K \subset \frac{1}{d}\mathbb{Z}[x]$$

we deduce that  $\mathcal{O}_K$  is a sublattice of  $\frac{1}{d}\mathbb{Z}[x]$  and hence it could be generated by  $n$  elements.  $\square$

We call an **integral basis** of  $K$  a family as in the theorem above. Using proposition 4.2.2(ii) we easily see that absolute value  $|\text{disc}_{K/\mathbb{Q}}(e_1, \dots, e_n)|$  of the discriminant of an integral basis does not depend on a choice of the integral basis, we call it the **discriminant** of  $K$ .

We now determine the rings of integers in two particular cases of number fields.

**Proposition 4.2.4.** *Let  $d \in \mathbb{Z} \setminus \{0, 1\}$  with no square factors and let  $K = \mathbb{Q}(\sqrt{d})$ . Then*

$$\mathcal{O}_K = \begin{cases} \mathbb{Z} + \mathbb{Z}[\sqrt{d}], & d \equiv 2, 3 \pmod{4}; \\ \mathbb{Z} + \mathbb{Z}\left[\frac{1+\sqrt{d}}{2}\right], & d \equiv 1 \pmod{4}. \end{cases}$$

*Proof.* The inclusions  $\supset$  are straightforward: in the second case, for  $\alpha = \frac{1+\sqrt{d}}{2}$ , note that  $\alpha^2 - \alpha + \frac{1-d}{4} = 0$ . The inclusions  $\subset$  follow from the fact that for  $x = a + b\sqrt{d} \in \mathcal{O}_K$  we have that the norm and the trace of  $x$  are in  $\mathbb{Z}$ .  $\square$

We have that if  $K = \mathbb{Q}(\sqrt{d})$ , with  $d \in \mathbb{Z} \setminus \{0, 1\}$  without square factors, then  $\text{disc}(K) = d$  or  $4d$ , corresponding to  $d \equiv 1 \pmod{4}$  or not (exercise). Note that a polynomial  $x^2 + ax + b \in \mathbb{Z}[x]$  with discriminant  $D = a^2 - 4b$  is irreducible modulo  $p > 2$  iff  $D$  is not a square modulo  $4p$ , and has a double root iff  $p$  divides  $D$ . This observation, proposition above and proposition 3.3.7 then allow to deduce

**Corollary 4.2.5.** *Let  $d \in \mathbb{Z} \setminus \{0, 1\}$  be without square factors, let  $K = \mathbb{Q}(\sqrt{d})$  and let  $D = \text{disc}(K)$ . Then*

- if  $p|D$ , then there is a prime ideal  $P \subset \mathcal{O}_K$  containing  $p$  with  $(p) = P^2$ ;
- if  $D$  is not a square modulo  $4p$ , the ideal  $(p)$  is prime;
- if  $(p, D) = 1$  and  $D$  is a square modulo  $4p$ , then  $(p) = PP'$  where  $P$  and  $P'$  are two distinct prime ideals.

## 4.3 Ideal classes

In this section we will be interested in the group  $Cl(A)$  of ideal classes.

### 4.3.1 Canonical embedding

Let  $K$  be a number field and let  $n = [K : \mathbb{Q}]$ . Then we have  $n$  embeddings  $K \hookrightarrow \mathbb{C}$ . If  $\sigma$  is such embedding and if  $\sigma(K) \subset \mathbb{R}$  we call  $\sigma$  a **real** embedding. If not, then composition of  $\sigma$  and the complex conjugation provide another embedding, different from  $\sigma$ . We could then write  $n = r_1 + 2r_2$ , where  $r_1$  is the number of real embeddings  $\sigma_1, \dots, \sigma_{r_1}$ , and  $r_2$  is the number of complex conjugated embeddings  $\sigma_{r_1+1}, \bar{\sigma}_{r_1+1}, \dots, \sigma_{r_1+2r_2}, \bar{\sigma}_{r_1+2r_2}$ .

**Definition 4.3.1.** The **canonical embedding** of  $K$  is a map  $\iota : K \rightarrow \mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$  defined by

$$\iota(x) = (\sigma_i(x))_{i=1, \dots, r_1+2r_2}.$$

We identify  $\mathbb{C}$  with  $\mathbb{R}^2$  via the basis  $\{1, i\}$ .

**Lemma 4.3.2.** *Let  $e_1, \dots, e_n$  be the  $\mathbb{Q}$ -basis of  $K$ . Then  $\iota(e_1), \dots, \iota(e_n)$  is an  $\mathbb{R}$ -basis of  $\mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$  and generates a lattice of covolume*

$$2^{-r_2} |\text{disc}_{K/\mathbb{Q}}(e_1, \dots, e_n)|^{\frac{1}{2}}.$$

*Proof.* Consider an  $\mathbb{R}$ -basis  $f$  of  $\mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$ , consisting of elements  $(0, 0, \dots, 0, *, 0, \dots, 0)$ , where  $*$  = 1 (resp.  $1$  or  $i$ ) if its place is  $\leq r_1$  (resp.  $> r_1$ .)

Let  $P$  be a matrix of vectors  $\iota(e_j)$  in this basis. If  $r_2 = 0$ , then  $|\det(P)| = |\det(\sigma_j(e_i))| = |\text{disc}_{K/\mathbb{Q}}(e_1, \dots, e_n)|^{\frac{1}{2}} \neq 0$ . In general,

$$\begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} \begin{pmatrix} \text{Re}\sigma(e_1) & \dots & \text{Re}\sigma(e_n) \\ \text{Im}\sigma(e_1) & \dots & \text{Im}\sigma(e_n) \end{pmatrix} = \begin{pmatrix} \sigma(e_1) & \dots & \sigma(e_n) \\ \bar{\sigma}(e_1) & \dots & \bar{\sigma}(e_n) \end{pmatrix}$$

so that  $|\det(P)| = 2^{-r_2} |\det(\sigma_j(e_i))| = 2^{-r_2} |\text{disc}_{K/\mathbb{Q}}(e_1, \dots, e_n)|^{\frac{1}{2}}$ . □

**Theorem 4.3.3.**  $\iota(\mathcal{O}_K)$  is a lattice in  $\mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$ .

*Proof.* Using the lemma above, the ring  $\mathcal{O}_K$  contains a  $\mathbb{Q}$ -basis of  $K$ , the image of this basis, hence  $\iota(\mathcal{O}_K)$  generates  $\mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$ . It remains to show that  $\iota(\mathcal{O}_K)$  is discrete. Consider a norm on  $V$  that is a sup norm with respect to the basis  $f$ , as in the proof of lemma above. Let  $r > 0$ . If  $x \in \mathcal{O}_K$  such that  $\iota(x) < r$ , then  $|\sigma(x)| < r$  for  $\sigma$  real and  $|\text{Re}\sigma(x)|, |\text{Im}\sigma(x)| < r$  for  $\sigma$  complex, so that  $|\sigma(x)| < \sqrt{2}r$ . Since  $x$  is annihilated by its characteristic polynomial  $\prod (X - \sigma(x)) \in \mathbb{Z}[x]$ , the coefficients of this polynomial are bounded with respect to  $r$ , since these are integers, only a finite number of such polynomials is possible, hence  $\iota(\mathcal{O}_K)$  is discrete.

□

In particular, if  $A \subset \mathcal{O}_K$  is an additive subgroup containing a  $\mathbb{Q}$ -basis of  $K$ , then  $\iota(A)$  is a sublattice of  $\iota(\mathcal{O}_K)$ . In particular,  $A$  is of finite index  $\frac{\text{covol}(\iota(A))}{\text{covol}(\iota(\mathcal{O}_K))}$  in  $\mathcal{O}_K$ , and has a  $\mathbb{Z}$ -basis of  $n$  elements:

**Proposition 4.3.4.** *Let  $K$  be a number field and let  $A \subset \mathcal{O}_K$  be an additive subgroup containing a  $\mathbb{Q}$ -basis of  $K$ . Then*

(i)  $A$  has a  $\mathbb{Z}$ -basis of  $n = [K : \mathbb{Q}]$  elements.

(ii) If  $\text{disc}(A) \in \mathbb{Z}$  is a discriminant of a  $\mathbb{Z}$ -basis of  $A$ , then

$$|\text{disc}(A)|^{\frac{1}{2}} = \text{covol}(\iota(A))2^{r_2(K)}.$$

(iii) If  $B$  is a subgroup of  $\mathcal{O}_K$  containing  $A$ , then  $A$  is of finite index in  $B$ , that equals to

$$\frac{|\text{disc}(A)|^{\frac{1}{2}}}{|\text{disc}(B)|^{\frac{1}{2}}}.$$

### 4.3.2 Finiteness

**Definition 4.3.5.** An **order** in a number field  $K$  is a subring  $A \subset \mathcal{O}_K$  containing a  $\mathbb{Q}$ -basis of  $K$ .

For example, if  $\alpha$  is an algebraic integer, then  $\mathbb{Z}[\alpha]$  is an order in  $\mathbb{Q}(\alpha)$ .

Let  $A$  be an order in a number field  $K$  and let  $I$  be an ideal in  $A$ , then we define a norm of  $I$  by the formula

$$N(I) = |A/I|.$$

**Proposition 4.3.6.** *For any nonzero ideal  $I \subset A$  and  $z \in A$  nonzero,  $N(zI) = |N(z)|N(I)$ . In particular,  $N(zA) = |N(z)|$ .*

*Proof.* Let  $a = (a_1, \dots, a_{r_1+r_2}) \in \mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$  with  $a_i \neq 0$  for any  $i$  and let

$$d_a : \mathbb{R}^{r_1} \times \mathbb{C}^{r_2} \rightarrow \mathbb{R}^{r_1} \times \mathbb{C}^{r_2}, \quad d_a((x_i)) = (a_i x_i).$$

The map  $d_a$  multiplies the Lebesgue measure by

$$n(a) := \prod_{i=1}^{r_1} |a_i| \prod_{i=1}^{r_2} |a_{r_1+i}|^2.$$

We then deduce that  $\iota(zI) = d_{\iota(z)}(\iota(I))$ . Hence

$$\text{covol}(zI) = n(\iota(z))\text{covol}(\iota(I))$$

but  $n(\iota(z)) = |N(z)|$ . □

Now the results from the previous section allow to establish the following important theorem:

**Theorem 4.3.7.** (Minkowski) *Let  $A$  be an order in  $K$  and let  $I$  be a nonzero ideal in  $A$ , then there exists a nonzero  $x \in I$ , such that*

$$|N(x)| \leq C(r_2, n)N(I)|disc(A)|^{\frac{1}{2}},$$

where

$$C(r, n) = \left(\frac{4}{\pi}\right)^r \frac{n!}{n^n}$$

is the Minkowski constant and  $r_2$  is the number of pairs of conjugated complex embeddings of  $K$ .

*Proof.* We view  $\iota(I)$  as a lattice in  $\mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$ , where  $\iota$  is a canonical embedding. Consider the following norm on  $V$ :

$$|(x_i)| = \sum_{i=1}^{r_1} |x_i| + 2 \sum_{i=r_1+1}^{r_1+r_2} |x_i|$$

and for any real  $t > 0$  consider

$$C_t = \{v \in V, |v| \leq t\},$$

it is a compact, convex and symmetric domain. Note then

$$|N(x)| \leq n^{-n} |\iota(x)|^n$$

for any  $x \in K$ , since

$$\left(\prod_{i=1}^{r_1} |x_i| \prod_{i=r_1+1}^{r_1+r_2} |x_i|^2\right)^{\frac{1}{n}} \leq \frac{1}{n} \left(\sum_{i=1}^{r_1} |x_i| + 2 \sum_{i=r_1+1}^{r_1+r_2} |x_i|\right).$$

We also easily see (exercise) that we have for the Lebesgue measure of  $C_t$  that

$$\mu(C_t) = 2^{r_1} \left(\frac{\pi}{2}\right)^{r_2} \frac{t^n}{n!}.$$

Consider now  $t$  such that  $\mu(C_t) = 2^n covol(\iota(I))$ . By Minkowski's theorem 4.1.7, we have that there is a nonzero element, that we denote  $\iota(x)$ , in  $\iota(I)$ , such that  $|\iota(x)| \leq t$ , hence

$$N(x) \leq n^{-n} t^n = \frac{n!}{2^{r_1} \left(\frac{\pi}{2}\right)^{r_2} n^n} 2^n covol(\tau(I)) = \left(\frac{4}{\pi}\right)^{r_2} \frac{n!}{n^n} N(I) 2^{r_2} covol(\iota(A)).$$

□

As a corollary, we obtain:

**Proposition 4.3.8.** *Let  $A$  be an order in a number field  $K$ . There exists an integer  $C \geq 1$  such that for any nonzero ideal  $I$  of  $A$  there is a nonzero element  $x \in I$  with  $|I/Ax| \leq C$ .*

*Proof.* The result follows from the Minkowski theorem above, with  $C \leq C(r_2, n)|disc(A)|^{\frac{1}{2}}$ . In fact, the kernel of the canonical projection  $A/xA \rightarrow A/I$  is exactly  $I/xA$ , so that

$$|I/xA||A/I| = |A/xA|.$$

But  $|A/I| = N(I)$  and  $|A/xA| = N(xA) = |N_{K/\mathbb{Q}}(x)|$  by proposition 4.3.6, so that Minkowski theorem gives

$$|I/Ax| \leq C(r_2, n)|disc(A)|^{\frac{1}{2}}.$$

□

**Definition 4.3.9.** Let  $A$  be an order in a number field. The smallest integer  $C$  satisfying the conditions of proposition 4.3.8 is denoted  $C(A)$ .

These results allow to establish the central result of this section:

**Theorem 4.3.10.** (*finiteness of the class number*) *If  $A$  is an order in a number field  $K$ , then  $Cl(A)$  is finite.*

*Proof.* Let  $I$  be a nonzero ideal of  $A$  and let  $x \in I$  be as in proposition 4.3.8. If  $N$  is the cardinality of the finite group  $I/Ax$ , then  $NI \subset Ax$ . We then see

$$NA \subset \frac{N}{x}I \subset A.$$

In particular  $J = \frac{N}{x}I$  is an ideal of  $A$  equivalent to  $I$ , and  $J$  is in between  $NA$  and  $A$ . But the ring  $A/NA$  is finite and hence there is only a finite number of ideals containing  $N$ . □

**Corollary 4.3.11.** *If  $A$  is an order in the number field  $K$ , then  $C(A) \leq C(r_2, n)|disc(A)|^{\frac{1}{2}}$ , where  $n = [K : \mathbb{Q}]$ .*

**Corollary 4.3.12.** (*Minkowski*) *If  $K \neq \mathbb{Q}$ , then  $|disc(K)| \geq \frac{\pi}{3}(\frac{3\pi}{4})^{n-1}$ . In particular,  $disc(K) \neq \pm 1$ .*

*Proof.* We apply theorem theorem 4.3.7 to  $A = I = \mathcal{O}_K$  and we use that for any nonzero  $x \in \mathcal{O}_K$ , we have that  $N(x)$  is a nonzero integer, so that  $|N(x)| \geq 1$ . This implies that

$$|disc(K)| \geq C(r_2, n)^{-2} \geq a_n, \text{ with } a_n = C(\frac{n}{2}, n)^{-2},$$

since  $r_2 \leq \frac{n}{2}$  and  $\pi < 4$ . But then  $a_2 = \frac{\pi^2}{4}$  and  $\frac{a_{n+1}}{a_n} = \frac{\pi}{4}(1 + \frac{1}{n})^{2n} \geq \frac{3\pi}{4}$  for  $n > 1$ , hence the conclusion. □

## 4.4 Applications

### 4.4.1 Quadratic fields

In this section, we consider, with more elementary techniques, the case of

$$A_D = \mathbb{Z} + \mathbb{Z}\alpha,$$

where  $D \equiv 0, 1 \pmod{4}$  and  $D < 0$ , where we set  $\alpha = \sqrt{D/4}$  if  $D \equiv 0 \pmod{4}$ , or  $\alpha = \frac{1+\sqrt{D}}{2}$ , if  $D \equiv 1 \pmod{4}$ . Here, as a choice of  $\sqrt{D}$ , we consider the root with positive imaginary part. We then have a norm map:  $N(z) = z\bar{z}$ , i.e.,

$$N(x + y\alpha) = \begin{cases} x^2 - \frac{D}{4}y^2, & D \equiv 0 \pmod{4}; \\ x^2 + xy + \frac{1-D}{4}y^2, & D \equiv 1 \pmod{4}. \end{cases}$$

In particular, we see that the function  $(x, y) \rightarrow N(x + y\alpha)$  is a principal form of the discriminant  $D$ .

Using the norm map, one obtains the following standard properties:

- The units of  $A_D$  are the elements  $u \in A_D$  such that  $N(u) = 1$ .
- The element  $\pi \in A_D$  is irreducible iff no proper divisor of  $N(\pi)$  is of type  $N(z)$  for  $z \in A_D$ . In particular, this holds if  $N(\pi)$  is prime.
- $A_D$  has a factorization property (but the factorization is not necessarily unique!)

Recall the notion of an euclidean ring:

**Definition 4.4.1.** A ring  $A$  is **euclidean** if there is a map (an **euclidean function**)  $\phi : A \setminus \{0\} \rightarrow \mathbb{N}$ , such that for any nonzero  $a, b \in A$  there are  $q, r \in A$  such that  $a = bq + r$  and, if  $r \neq 0$ , then  $\phi(r) < \phi(b)$ .

In some cases the ring  $A_D$  is euclidean:

**Theorem 4.4.2.** *If  $D \in \{-3, -4, -7, -8, -11\}$ , then  $A_D$  is euclidean for the euclidean function  $N$ . In particular,  $A_D$  is principal and factorial. For other values of  $D$ , the ring  $A_D$  is not euclidean (for any euclidean function).*

*Proof.* Consider first the case when  $D \in \{-3, -4, -7, -8, -11\}$ . Observe that it is enough to show that for any  $z \in \mathbb{Q}(\alpha)$  there exists  $t \in A_D$  such that  $N(z - t) < 1$ . In fact, then for  $a, b \in A_D$  nonzero and if  $t \in A_D$ , such that  $N(\frac{a}{b} - t) < 1$ , then  $N(a - tb) < N(b)$  (by multiplicativity of the norm), so that  $N : A_D \rightarrow \mathbb{N}$  defines the euclidean function. We now verify the approximation property. Fix  $x, y \in \mathbb{R}$  and let  $q(x, y) = N(x + \alpha y)$ . We show more generally that there are  $u, v \in \mathbb{Z}$  such that  $q(x - u, y - v) < 1$ . Assume  $D \equiv 0 \pmod{4}$ . Then we can take  $u, v \in \mathbb{Z}$  such that  $|x - u|, |y - v| \leq \frac{1}{2}$ , so that

$$q(x - u, y - v) = (x - u)^2 - \frac{D}{4}(y - v)^2 \leq \frac{1 - \frac{D}{4}}{4}$$

and the last value is smaller than 1 if  $D = -4, -8$ . Assume  $d \equiv 1 \pmod{4}$ , then note that

$$4q(x, y) = (2x + y)^2 - Dy^2.$$

Take  $v \in \mathbb{Z}$  such that  $|y - v| \leq \frac{1}{2}$ , then  $u \in \mathbb{Z}$  such that  $|2u + \frac{v-2x-y}{2}| \leq 1$ . We then have  $4q(x - u, y - v) \leq 1 - \frac{D}{4}$  that is  $< 4$  as soon as  $-D < -12$ , that finishes the analysis for  $D \in \{-3, -4, -7, -8, -11\}$ .

Now assume that  $D$  is not in the set above. Assume that  $A_D$  has an euclidean function  $f$ . Let  $x \in A$  be nonzero and not a unit, such that  $f(x)$  is minimal. Then any  $a \in A_D$  is of the form  $bx + r$  with  $r = 0$  or  $r \in A_D^*$ . In particular,  $N(x) = N((x)) \leq 1 + |A_D^*|$ . But  $|D| > 4$ , so that we observe that  $A_D^* = \{\pm 1\}$ , so that  $N(x) \in \{2, 3\}$ . In particular, the principal form of the discriminant  $-D$  represents 2 or 3. Hence  $|D/4| \leq 3$  if  $D \equiv 0 \pmod{4}$  and  $|D| \leq 12$  if  $D \equiv 1 \pmod{4}$ . The only remaining case is  $A_{-12} = \mathbb{Z}[\sqrt{-3}]$ , but this ring cannot be euclidean, since it is not factorial. □

More generally, one can show that for  $d < 0$ , the ring  $\mathbb{Z}[\sqrt{d}]$  is principal iff  $d = -1$  or  $-2$ . And if  $d \equiv 1 \pmod{4}$ , then  $\mathbb{Z}[\frac{1+\sqrt{d}}{2}]$  is principal iff  $d \in \{-3, -7, -11, -19, -43, -67, -163\}$ .

In the case  $D > 0$  it is an open problem to determine if the ring  $A_D$  (defined similarly as for  $D < 0$ ), is euclidean.

#### 4.4.2 $Cl(D)$ versus $Cl(A_D)$

We keep the same notations  $(D, \alpha, A_D, N)$  as in the previous section. Recall that we defined:

$Cl(D) :=$  the set of proper equivalence classes of integral quadratic forms of discriminant  $D$ .

**Theorem 4.4.3.** (*Dedekind*) *There is bijection  $Cl(A_D) \rightarrow Cl(D)$ ,  $I \mapsto q_I$ .*

*Proof.* The proof will occupy the rest of this section. We proceed by several steps, defining the maps between  $Cl(A_D)$  and  $Cl(D)$ .

*Step 1.* For an ideal  $I$  with a fixed basis we will associate a quadratic form. Since  $I \subset A_D$  is an ideal, we can view  $I$  as a lattice in  $\mathbb{C}$ , hence  $I$  has a basis of two elements  $u, v$ . Up to changing  $v$  by  $-v$ , we may assume that the determinant of  $(u, v)$ , in the basis  $1, i$ , is positive, we call such a basis **direct**. We then define:

$$q_{u,v}(x, y) = \frac{1}{N(I)} N(xu + yv).$$

Then

1.  $q_{u,v}$  has integral values: for  $z = xu + yv$ , we have an inclusion of ideals  $(z) \subset I$ , hence  $N(z) | N(I)$  by definition.



2.  $q_{u,v}$  is an integral quadratic form:  $q_{u,v}(x, y) = ax^2 + bxy + cy^2$ , with

$$a = \frac{N(u)}{N(I)}, c = \frac{N(v)}{N(I)}, b = \frac{\text{Tr}_{\mathbb{Q}(\sqrt{D})/\mathbb{Q}}(u\bar{v})}{N(I)},$$

since  $N(z) = z\bar{z}$ .

*step 2.* We now investigate properties of  $q_{u,v}$ :

1. **Discriminant:** By definition of  $N$  and  $A_D$ , we have that  $N$  is a principle form with discriminant  $D$ , and  $I = \mathbb{Z}u + \mathbb{Z}v$  has index  $N(I)$ , by proposition 4.3.4,  $N|_I$  has discriminant  $DN(I)^2$ . But  $q_{u,v} = \frac{1}{N(I)}N|_I$ , so its discriminant is  $D$ .
2. **On what  $q$  depends:** we will show that  $q$  depends only on the class of  $I$  in  $Cl(A_D)$ . First, if  $u', v'$  is another direct basis of  $I$ , then by definition  $\frac{1}{N(I)}N(xu' + yv')$  is properly equivalent to  $q$ . Next, for an equivalent ideal  $zI$  with  $z \in A_D$  nonzero, we have that  $zI$  has a direct basis  $zu, zv$ , so that

$$\frac{1}{N(z)}N(xzu + yzv) = \frac{N(z)}{N(zI)}N(xu + yv) = \frac{1}{N(I)}N(xu + yv).$$

*step 3.* From quadratic forms to ideals:

$$q = ax^2 + bxy + cy^2 = a\left(x + \frac{b}{2a}y\right)^2 - \frac{D}{4a^2}y^2 = qN(x + \tau y),$$

where

$$\tau = \frac{b + \sqrt{D}}{2a}.$$

With this choice of  $\tau$  we see that  $1, \tau$  is a direct basis and

$$a\tau = \frac{b + \sqrt{D}}{2} \in \alpha + \mathbb{Z}$$

(since  $b \equiv D \pmod{2}$ ), hence  $a\tau \in A_D$  and we can define

$$I(q) = a\mathbb{Z} + a\tau\mathbb{Z}.$$

*step 4. Properties of  $I(q)$ : computing the norm.* First note that  $I(q)$  is an ideal: we need only to check that  $a\alpha$  and  $a\alpha\tau$  are in  $I(q)$ . Since  $a\tau \in \alpha + \mathbb{Z}$ , we have that

$$a\tau - \alpha \in \mathbb{Z},$$

hence  $a\alpha \in a^2\tau + a\mathbb{Z}$  is in  $I(q)$ . For  $a\alpha\tau$  it is enough to check that  $(a\tau)^2 \in I(q)$ , but  $a\tau^2 - b\tau + c = 0$ , so that  $(a\tau)^2 = ab\tau - ac$  is in  $I(q)$ .

Now we claim that  $N(I) = a$ . In fact,  $A_D$  has basis  $\{1, \alpha\}$ , but  $a\tau - \alpha \in \mathbb{Z}$ , hence  $A_D$  also has basis  $\{1, a\tau\}$ . By definition,  $I(q)$  has basis  $\{1, a\tau\}$ , so that  $A_D/I \simeq \mathbb{Z}/a\mathbb{Z}$  and  $N(I) = a$ .

*step 5.* Comparing  $I(q)$  and  $q_{u,v}$ : for the direct basis  $\{a, a\tau\}$  of  $I(q)$  we have that the associated form is

$$\frac{1}{a}N(ax + a\tau y) = aN(x + \tau y) = q(x, y).$$

Next, if we change  $q$  by a properly equivalent form, we check that  $I(q) \sim I(q')$ : in fact, let  $eh - fg = 1$ , then

$$q'(x, y) = q(ex + fy, gx + hy) = a'N(x + \tau'y),$$

where  $a = N(e + \tau g)$  and  $\tau' = \frac{f + \tau h}{e + \tau g}$ . A direct verification shows that  $\text{Im } \tau' > 0$ . By definition  $I(q') = a'\mathbb{Z} + a'\tau'\mathbb{Z}$ . Then we check that

$$(e + \tau g)I(q') = a'(\mathbb{Z}(f + \tau h) + \mathbb{Z}(e + \tau g)) = a'(\mathbb{Z} + \tau\mathbb{Z}) = I(q).$$

*step 6.* We now finish the proof that  $Cl(A_D)$  is bijective to  $Cl(D)$ . We have defined two maps  $I \rightarrow q_I$  and  $q \rightarrow I(q)$ . It is enough to verify that if  $q = q_I$ , then  $I(q) = I$ . Indeed, by definition, if  $I$  has a basis  $u, v$ , then

$$q_{u,v}(x, y) = \frac{1}{N(I)}N(ux + vy) = aN(x + \tau y),$$

where  $a = \frac{N(u)}{N(I)}$ ,  $\tau = \frac{v}{u}$ . Hence

$$I(q_{u,v}) = a\mathbb{Z} + a\tau\mathbb{Z},$$

so that  $N(I) \cdot I(q) = \bar{u}I$  and  $I(q_{u,v}) = I$ . □

### 4.4.3 Equation $y^2 = x^3 + k$

We consider an example of applications to diophantine equations.

**Theorem 4.4.4.** *Let  $k < 0$  be without square factors,  $k \equiv 2, 3 \pmod{4}$ . Assume that the order of the group  $Cl(\mathcal{O}_K)$  is not a multiple of 3. Then the equation  $y^2 = x^3 + k$  has at most two solutions.*

*Proof.* Let  $A = \mathbb{Z}[\sqrt{k}]$  and note that  $A$  is the ring of integers of  $\mathbb{Q}(\sqrt{k})$  since  $k$  is not  $1 \pmod{4}$ . Then in  $A$  the equation factorizes:

$$(y + \sqrt{k})(y - \sqrt{k}) = x^3.$$

We can also view this equation in terms of principal ideals

$$II' = (x)^3,$$

where  $I = (y + \sqrt{k})$  and  $I' = (y - \sqrt{k})$ .

We first check that  $I$  and  $I'$  are relatively prime ideals. Let  $D = I + I'$  be their greatest common divisor. Then  $2y, 2k \in D$ . Since  $k$  has no square factors,  $x, y, k$  are pairwise relatively prime. In particular,  $(2y, 2k) = 2$  and then  $2 \in D$  by Bezout theorem. The ideals containing 2 are  $2A, A$ , and a prime ideal  $P = 2A + \sqrt{k}A$  or  $2A + (\sqrt{k} + 1)A$ , depending on  $k$  is even or not, in addition  $P^2 = (2)$ . The case  $D = 2A$  is not possible since  $y + \sqrt{k} \notin 2A$ , so that  $D = P$ . Since  $P^2 = (2)$  and  $y \pm \sqrt{k} \notin 2A$ , we see that  $v_P(I) = v_P(I') = 1$ . But then

$$v_P(II') = 3v_P(x) = v_P(I) + v_P(I') = 2$$

a contradiction.

Hence  $I$  and  $I'$  are relatively prime. Since  $II'$  is a cube of an ideal in  $A$ , we deduce that  $I = J^3$  for some ideal  $J$ . But  $I = J^3$  is principal, hence the order of the class  $[J]$  in  $Cl(A)$  divides 3. Since the order of the group  $Cl(\mathcal{A})$  is not a multiple of 3, we deduce that  $J$  is principal.

Let  $a, b \in \mathbb{Z}$  such that  $J = (a + b\sqrt{k})$ . The identity  $I = J^3$  implies that  $y + \sqrt{k}$  is a cube of an element  $a + b\sqrt{k}$ , up to multiplication by a unit. But any unit in  $A$  is a cube, we may thus assume that we have the following identity:

$$y + \sqrt{k} = (a + b\sqrt{k})^3 = a^3 + 3kab^2 + (3a^2b + kb^3)\sqrt{k}.$$

In particular,  $1 = b(3a^2 + kb^2)$ , so that  $b = \pm 1$  and  $-k = 3a^2 - b$ . If  $k$  is not of this type, we deduce that there is no solutions. And if  $-k = 3a^2 - b$ , only one value of  $b$  and two values of  $a$  are possible. We have  $y = a^3 + 3ka$  and  $x = a^2 - k$ .  $\square$

#### 4.4.4 An example of computation of $Cl(\mathcal{O}_K)$

In this section we discuss an example of computation of  $Cl(\mathcal{O}_K)$ . We will show that

$$Cl(\mathbb{Z}[\frac{1 + \sqrt{-47}}{2}]) \simeq \mathbb{Z}/5\mathbb{Z}.$$

Note that  $A = \mathbb{Z}[\frac{1 + \sqrt{-47}}{2}]$  is the ring of integers of the field  $\mathbb{Q}(\sqrt{-47})$ , since  $-47$  has no square factors and  $\equiv 1 \pmod{4}$ . We have  $disc(\mathcal{O}_K) = -47$  and Minkowski's theorem insures that any ideal class has an ideal containing an integer  $1 \leq N \leq C(1, 2)\sqrt{47} < 5$ .

Hence we are interested to determine all ideals  $I$  containing 1, 2, 3 and 4. If  $1 \in I$ , we have  $I = A$ , so that we only need to determine ideals dividing 3 and 4. We first determine prime ideals containing 2 and 3. Let  $\alpha = \frac{1 + \sqrt{-47}}{2}$  and  $P = P_{\alpha, \mathbb{Q}} = x^2 - x + 12$ . We have

$$P \equiv x(x - 1) \pmod{2, 3},$$

so that we have two prime ideals containing 2 and 3:  $D = (2, \alpha)$  and  $D' = (2, \alpha - 1)$ , as well as  $T = (3, \alpha)$  and  $T' = (3, \alpha - 1)$ . We have  $(2) = DD'$  and  $(3) = TT'$ . Hence the ideals containing  $(4) = D^2D'^2$  are  $4A, A, D^2, (D')^2, D^2D' = 2D'$  and  $2D$ . The

ideals containing (3) are  $3A, A, T, T'$ . We then deduce that  $Cl(A)$  is generated by the classes of  $D, D', T, T'$ , but  $D'$  is inverse of  $D$  and  $T'$  is an inverse of  $T$ , so that we finally have that  $Cl(A)$  is generated by the classes of  $D$  and  $T$ .

In order to find a relation between these classes, we will first search for norms of type  $2^a 3^b$  and then decompose them into prime ideals.

Recall that  $N(x + y\alpha) = x^2 + xy + 12y^2 = \frac{1}{4}(2x + y)^2 + \frac{47}{4}y^2$ . In particular,  $N(\alpha) = 12$  and  $N(4 + \alpha) = 32$ . Note that  $\alpha \in D, \alpha \in T$ , but  $\alpha \notin D'$  (if not, we would have  $D = D'$ ). We deduce that  $(\alpha) = D^2 T$ , hence  $[T] = [D]^{-2}$ . Similarly,  $(4 + \alpha) = D^5$ , hence  $[D]$  is of order dividing 5, hence 1 or 5, but it is not principle since 2 is not a norm. We then deduce, as claimed:

$$Cl(\mathbb{Z}[\frac{1 + \sqrt{-47}}{2}]) \simeq \mathbb{Z}/5\mathbb{Z}.$$

## 4.5 The Dirichlet formula

For the last chapter we come back to the analytic techniques and we introduce the Dedekind  $\zeta$ -functions associated to number fields.

**Definition 4.5.1.** Let  $K$  be a number field. The **Dedekind  $\zeta$ -function** associated to  $K$  is the function

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

where  $I$  runs over all nonzero ideals of  $\mathcal{O}_K$ .

The Euler product argument leads to the following expression:

**Proposition 4.5.2.**  $\zeta_K$  converges absolutely for  $Re(s) > 1$  and defines in particular a holomorphic function on this half-plan. If  $Re(s) > 1$ , then

$$\zeta_K(s) = \prod_{\mathfrak{p}} \frac{1}{1 - N(\mathfrak{p})^{-s}},$$

where the product is over all prime ideals  $\mathfrak{p}$  of  $K$  and is absolutely convergent.

*Proof.* Let  $a_n$  be the number of ideals of  $\mathcal{O}_K$  with norm  $n$ . Then  $a_1 = 1$  and the multiplicativity of the norm implies that  $a_{nm} = a_n a_m$  if  $m$  and  $n$  are relatively prime. Hence, for any prime number  $p$  and for  $Re(s) > 0$ , we have

$$\prod_{\mathfrak{p}|p} \frac{1}{1 - N(\mathfrak{p})^{-s}} = \sum \frac{a_{p^i}}{p^{is}}. \quad (4.1)$$

We now show the absolute convergence: recall that for a given prime  $p$  we have at most  $[K : \mathbb{Q}]$  prime ideals  $\mathfrak{p}$  of  $\mathcal{O}_K$  containing  $p$ , and for such ideal  $\mathfrak{p}$  we have

$N(\mathfrak{p}) = p^f$  with  $f \geq 1$ . In particular,  $|(1 - N(P)^{-s})^{-1} - 1| \leq 2p^{-\sigma}$ , if  $\sigma = \operatorname{Re}(s) > 0$ . Hence for  $\sigma = \operatorname{Re}(s) > 1$  we have

$$\sum_p \sum_{\mathfrak{p}|p} |(1 - N(\mathfrak{p})^{-s})^{-1} - 1| \leq 2[K : \mathbb{Q}] \sum_p p^{-\sigma} < 2[K : \mathbb{Q}] \zeta(\sigma) < \infty$$

and the theorem follows (see also proposition (1.2.9).)

□

As a corollary, we obtain that for  $\operatorname{Re}(s) > 1$  the function  $\zeta_K(s)$  coincides with the Dirichlet series

$$\zeta_K(s) = \sum \frac{a_n}{n^s},$$

where  $a_n$  is the number of nonzero ideals of  $\mathcal{O}_K$  of norm  $n$ . For example, for  $K = \mathbb{Q}(i)$ , we have  $\mathcal{O}_K = \mathbb{Z}[i]$  and any ideal of  $\mathcal{O}_K$  is principal, containing exactly  $4 = |\mathbb{Z}[i]^*|$  distinct generators. In addition  $N((a + bi)) = a^2 + b^2$  and hence

$$\zeta_{\mathbb{Q}(i)}(s) = \sum_{n \geq 1} \frac{a_n}{n^s}$$

where  $4a_n$  is the number of possibilities to write  $n$  as a sum of two squares.

Let now  $K$  be an imaginary quadratic field and  $D = \operatorname{disc}(\mathcal{O}_K)$ . Let, as before,  $h_K$  be the number of ideal classes of  $\mathcal{O}_K$ . Also denote

$$w_K = |\mathcal{O}_K^*|.$$

**Theorem 4.5.3.** (*Dirichlet*) *If  $K$  is an imaginary quadratic field, then  $\zeta_K(s)$  extends (uniquely) to a meromorphic function for  $\operatorname{Re}(s) > \frac{1}{2}$  and has a unique pole at  $s = 1$ , this pole is simple and*

$$\operatorname{Res}_{s=1} \zeta_K(s) = \frac{2\pi h_K}{w_K |\operatorname{disc}(\mathcal{O}_K)|^{\frac{1}{2}}}.$$

*Proof.* Let  $a_n$  be the number of ideals of  $\mathcal{O}_K$  of norm  $n$  and, for  $C \in \operatorname{Cl}(\mathcal{O}_K)$  we denote

$$a_{n,C} = |\{I \in C, N(I) = n\}|.$$

Then  $a_n = \sum_{C \in \operatorname{Cl}(\mathcal{O}_K)} a_{n,C}$ . In order to count the ideals we use lemmas 4.5.4, 4.5.5 below. We claim that then

$$\sum_{i=1}^n a_{i,C} = \frac{2\pi n}{w_K |D|^{\frac{1}{2}}} + \mathcal{O}(n^{\frac{1}{2}}) \tag{4.2}$$

when  $n \rightarrow \infty$ . Indeed, let  $J$  be an ideal in the class  $C^{-1}$ . By lemma 4.5.4 below, we want to count principal ideals in  $J$  of norm  $\leq nN(J)$ : any such ideal has  $w_K$  generators and, since  $N(z) = N((z))$ , we just need to count elements  $z$  of the lattice  $J \subset \mathbb{C}$  such that  $|z| \leq \sqrt{nN(J)}$ . Then the claim follows from lemma 4.5.5 below

and the fact that  $J$  is of covolume  $\frac{1}{2}N(J)|D|^{\frac{1}{2}}$ . Now the theorem follows easily from the asymptotics (4.2). Indeed, we obtain

$$\sum_{n \leq N} a_n = \mathcal{O}(N),$$

so that  $\zeta_K$  is absolutely convergent for  $\operatorname{Re}(s) > 1$ , and holomorphic. More precisely, if we denote, for  $\operatorname{Re}(s) > 1$ ,

$$\zeta_K(s) - \frac{2\pi h_K}{w_K |D|^{\frac{1}{2}}} \zeta(s) = \sum_{n \geq 1} \frac{b_n}{n^s}, \quad (4.3)$$

then the claim implies that  $\sum_{k \leq n} b_k = \mathcal{O}(n^{\frac{1}{2}})$ . Hence the series  $\sum \frac{b_n}{n^s}$  is convergent and holomorphic for  $\operatorname{Re}(s) > \frac{1}{2}$ . We then see that the identity (4.3) is the identity between two meromorphic functions for  $\operatorname{Re}(s) > \frac{1}{2}$ . We then deduce the formula using that the  $\zeta$ -function has a simple pole with residue 1 at  $s = 1$ .  $\square$

**Lemma 4.5.4.** *Let  $C$  be a class of ideals in  $\mathcal{O}_K$ ,  $J$  be an ideal of  $\mathcal{O}_K$  in the class  $C^{-1}$  and let  $n \geq 1$  be an integer. The map*

$$I \mapsto IJ$$

*is a bijection between the set of ideals of norm  $n$  in the class  $C$  and the set of principal ideals included in  $J$  and having norm  $nN(J)$ .*

*Proof.* If  $I \in C$ , then  $[I][J] = 1$ , so that  $IJ$  is principal, of norm  $N(I)N(J)$ , from the multiplicativity of the norm. The map  $I \mapsto IJ$  is injective since  $J$  is invertible. Finally, if  $I' \subset J$  is principal of norm  $nN(J)$ , there is a unique  $I$  with  $I' = IJ$  and we have  $N(I) = n$  (from the multiplicativity of the norm), and  $[J] = [I]^{-1}$ : the map is also surjective.  $\square$

**Lemma 4.5.5.** *Let  $L \subset \mathbb{C}$  be a lattice. If  $f(r) = |\{z \in L, |z| \leq r\}|$ , then*

$$f(r) = \frac{\pi r^2}{\operatorname{covol}(L)} + \mathcal{O}(r), r \mapsto \infty.$$

*Proof.* We identify  $\mathbb{C}$  with  $\mathbb{R}^2$  via the basis  $1, i$ . Let

$$B(r) = \{z \in \mathbb{C}, |z| \leq r\},$$

where  $r$  is a real number  $r > 0$ . If  $L = L(e)$  we fix  $\Pi = \Pi(e)$ , its surface is  $\operatorname{covol}(L)$ . Let  $\delta$  be such that  $\Pi \subset B(\delta)$ . Then we have

$$\mathbb{C} = \bigsqcup_{v \in L} (v + \Pi).$$

Let  $n(r)$  be the number of translates  $v + \Pi$  with  $v \in L$ , strictly included in  $B(r)$ . If  $(v + \Pi) \cap B(r) \neq \emptyset$ , note that  $v + \Pi$  is strictly included in  $B(r + \delta)$ . In particular,

$$n(r) \leq f(r) \leq n(r + \delta),$$

and

$$\text{covol}(L)n(r) \leq \pi r^2 \leq \text{covol}(L)n(r + \delta).$$

We then obtain the result for  $r > \delta$  since

$$\pi(r - \delta)^2 \leq \text{covol}(L)f(r) \leq \pi(r + \delta)^2.$$

□

The Dirichlet theorem generalizes for any number field. The analogous formula, that we give without proof here, is due to Dedekind. We denote  $w_K$  the number of roots of unity in  $\mathcal{O}_K$ . We assume that  $K$  has  $r_1$  real embeddings and  $r_2$  pairs of complex embeddings.

**Theorem 4.5.6.** (Dedekind) *The function  $\zeta_K(s)$  is holomorphic for  $\text{Re}(s) > 1$  and extends (uniquely) to a meromorphic function for  $\text{Re}(s) > 1 - \frac{1}{[K:\mathbb{Q}]}$  and has a simple pole at  $s = 1$ , such that*

$$\text{Res}_{s=1}\zeta_K(s) = \frac{2^{r_1+r_2}\pi^{r_2}h_K R_K}{w_K|\text{disc}(\mathcal{O}_K)|^{\frac{1}{2}}},$$

where  $R_K$  is a certain constant (the regulator of  $K$ ).