Algebra 3

Winter Term 2013

Université du Luxembourg

Sara Arias-de-Reyna
sara.ariasdereyna@uni.lu

Contents

Preface

These lecture notes correspond to the course *Algebra 3* from the Bachelor en Sciences et Ingénierie, Filière mathématiques, of the University of Luxembourg. This course was taught in the Winter Term 2013 and it consists of 14 lectures of 90 minutes each. This lecture belongs to the third semester of the Bachelor, and it builds on the lectures *Algebra 1* and *Algebra 2*, belonging to the first and second semester respectively (cf. [4], [5]).

The aim of this course is to introduce the students to the theory of algebraic extensions of fields, and culminates with the application of the theory to the solution (negative solution, in fact) of the three classical Greek problems concerning constructions with ruler and compass. This lecture is also a preliminary step towards Galois theory, which is taught in the fourth semester of the Bachelor.

The main reference I used to prepare this lecture are the lecture notes [6] written by Gabor Wiese for the course *Àlgebra 3*, taught in the Winter term 2012. For the constructions with ruler and compass we used the reference [2]. Other sources we have used to prepare the lecture are [1] and [3].

Luxembourg, December 2013.
Sara Arias-de-Reyna,

‘e E e‘e‘eÃ¨

1 Basic definitions

In this section we recall some definitions from the lectures Algebra 1 and 2, which shall be at the core of the material we will address during the course.

Definition 1.1. A *group* is a pair $(G, *)$, where G is a nonempty set and $* : G \times G \to G$ is a map satisfying:

1. *Associativity*: for all $g_1, g_2, g_3 \in G$, $(g_1 * g_2) * g_3 = g_1 * (g_2 * g_3)$;

2. *Existence of neutral element*: there exists $e \in G$ such that, for all $g \in G$, $e * g = g * e = g$;

3. *Existence of inverse element*: for all $g \in G$ there exists $h \in G$ such that $h * g = g * h = e$.

If $(G, *)$ satisfies in addition that

4. *Commutativity*: For all $g_1, g_2 \in G$, $g_1 * g_2 = g_2 * g_1$;

then we say that $(G, *)$ is *abelian* (also *commutative*).

Example 1.2. 1. $(\mathbb{Z}, +)$ is a commutative group. The neutral element is 0.

2. Let p be a prime number. $(\mathbb{Z}/p\mathbb{Z}, +)$ is a commutative group. The neutral element is $0 + p\mathbb{Z}$.

3. $(\mathbb{Q}, +)$ is a commutative group. The neutral element is 0.

4. $(\mathbb{Q} \setminus \{0\}, \cdot)$ is a commutative group. The neutral element is 1.

5. Let p be a prime number. $((\mathbb{Z}/p\mathbb{Z}) \setminus \{0 + p\mathbb{Z}\}, \cdot)$ is a commutative group. The neutral element is $1 + p\mathbb{Z}$.

6. The group consisting of all permutations of the set $\{1, 2, \ldots, n\}$, together with the composition of permutations, is a group which is not commutative if $n > 2$ (see Algebra 1). It is denoted by $(S_n, \circ)$.

Definition 1.3. Let $(G, *_G)$ and $(H, *_H)$ be two groups. A map $f : G \to H$ is called a *group (homo)morphism* if for all $g_1, g_2 \in G$ it holds that $f(g_1 *_G g_2) = f(g_1) *_H f(g_2)$.

Definition 1.4. Let $(G, *)$ be a group with neutral element e, and $H \subset G$ a subset. We say that H is a *subgroup* of G if

1. $e \in H$.

2. For all $h_1, h_2 \in H$, $h_1 * h_2 \in H$.

3. For all $h \in H$, the inverse of h belongs to H.

Remark 1.5. Let $(G, *)$ be a group and $H \subset G$ a subgroup. Denote by $*|_H : H \times H \to H$ the restriction of $*$ to $H \times H$. Then $(H, *|_H)$ is a group.

Definition 1.6. A *ring* is a tuple $(A, +_A, \cdot_A)$, where A is a nonempty set, $+_A : A \times A \to A$, $\cdot_A : A \times A \to A$ are maps satisfying:

1. $(A, +_A)$ is a commutative group.
2. $(A, \cdot_A)$ satisfies the three properties
   (a) *Associativity*: For all $a_1, a_2, a_3 \in A$, $(a_1 \cdot_A a_2) \cdot_A a_3 = a_1 \cdot_A (a_2 \cdot_A a_3)$;
   (b) *Existence of neutral element*: There exists $1_A \in A$ such that, for all $a \in G$, $1_A \cdot_A a = a \cdot_A 1_A = a$;
   (c) *Commutativity*: For all $a_1, a_2 \in A$, $a_1 \cdot_A a_2 = a_2 \cdot_A a_1$;
3. *Distributivity*: For all $a_1, a_2, a_3 \in A$, $a_1 \cdot_A (a_2 +_A a_3) = a_1 \cdot_A a_2 +_A a_1 \cdot_A a_3$.

Given a ring $(A, +_A, \cdot_A)$, one usually denotes by 0_A the neutral element for $+_A$ and by 1_A the neutral element for $\cdot_A$. The map $+_A$ is called *addition* or *sum*, and the map $\cdot_A$ is called *multiplication* or *product*. If $a \in A \setminus \{0_A\}$, we will denote by $-a$ the inverse of a with respect to the addition, and, if a has an inverse with respect to the multiplication, we will denote it by a^{-1}. If there is no risk of misunderstanding, we will drop the subindex $_A$ from $+_A$, $\cdot_A$, 0_A and 1_A.

Example 1.7.

1. $(\mathbb{Z}, +, \cdot)$ is a ring. The neutral element for the sum is 0 and the neutral element for the product is 1.
2. $(\mathbb{Z}/p\mathbb{Z}, +, \cdot)$ is a ring. The neutral element for the sum is $0 + p\mathbb{Z}$ and the neutral element for the product is $1 + p\mathbb{Z}$.
3. $(\mathbb{Q}, +)$ is a ring. The neutral element for the sum is 0 and the neutral element for the product is 1.

Definition 1.8. Let $(A, +_A, \cdot_A)$ and $(B, +_B, \cdot_B)$ be two rings. A map $f : A \to B$ is called a *ring (homo)morphism* if

1. For all $a_1, a_2 \in A$, $f(a_1 +_A a_2) = f(a_1) +_B f(a_2)$;
2. For all $a_1, a_2 \in A$, $f(a_1 \cdot_A a_2) = f(a_1) \cdot_B f(a_2)$;
3. $f(1_A) = 1_B$.

Definition 1.9. Let $(A, +, \cdot)$ be a ring and $B \subset A$ a subset. We say that B is a *subring* of A if

1. B is a subgroup of $(A, +)$.
2. $1_A \in B$ and for all $b_1, b_2 \in B$, $b_1 \cdot b_2 \in B$.

Remark 1.10. Let $(A, +, \cdot)$ be a ring and and $B \subset A$ a subring. Denote by $+|_B : B \times B \to B$ and $\cdot|_B : B \times B \to B$ the restrictions of $+$ and $\cdot$ to $B \times B$. Then $(B, +|_B, \cdot|_B)$ is a ring.

There are other subsets of a ring which are of particular importance in the theory of rings, namely ideals.

Definition 1.11. Let $(A, +, \cdot)$ be a ring and $I \subset A$ a subset. We say that I is an *ideal* of A if

1. I is a subgroup of $(A, +)$.

2. For all $i \in I$, for all $a \in A$, $i \cdot a \in I$.

Given a ring A and a subset $S \subset A$, one can consider the intersection of all ideals $J \subset A$ that contain S. This set is an ideal, which is the minimal ideal that contains S, and is called the *ideal generated by* S and denoted by (S). If $S = \{a_1, \dots, a_n\}$ is a finite set, we denote by $(a_1, \dots, a_n)$ the ideal generated by S.

Some elements in A have special properties; we gather some of them in the following definition.

Definition 1.12. Let $(A, +, \cdot)$ be a ring, $a \in A$ a nonzero element.

1. We say that a is a *unit* (or an *invertible* element) if there is $b \in A$ such that $a \cdot b = b \cdot a = 1$. The set of all units of A is denoted $A^\times$.

2. We say that a is *irreducible* if a is not a unit and, for all $b_1, b_2 \in A$ such that $b_1 \cdot b_2 = a$, then $b_1 \in A^\times$ or $b_2 \in A^\times$.

3. We say that a is *prime* if a is not a unit and, for all $b_1, b_2 \in A$ such that $a | b_1 \cdot b_2$, then $a | b_1$ or $a | b_2$.

Example 1.13.

1. The units of $\mathbb{Z}$ are 1 and -1. The irreducible elements of $\mathbb{Z}$ coincide with the prime elements of $\mathbb{Z}$, that is, the union of the set of prime numbers $\{3, 5, 7, 11, \dots\}$ and their opposites $\{-3, -5, -7, -11, \dots\}$.

2. The units of $\mathbb{Q}$ are all nonzero elements. $\mathbb{Q}$ does not have prime elements nor irreducible elements.

Lemma 1.14. *Let $(A, +, \cdot)$ be a ring. Then $(A^\times, \cdot)$ is a group.*

There are also some ideals that have special properties.

Definition 1.15. Let $(A, +, \cdot)$ be a ring, $I \subset A$ an ideal.

1. We say that I is a *principal ideal* if it can be generated by a unique element, that is to say, there exists $a \in A$ such that $I = (a)$.

2. We say that I is *prime* if for all $b_1, b_2 \in A$ with $b_1 \cdot b_2 \in I$, then $b_1 \in I$ or $b_2 \in I$.

Remark 1.16. Let $(A, +, \cdot)$ be a ring and $a \in A \setminus A^\times$ be nonzero. Then (a) is prime if and only if a is a prime element.

The following definitions gathers some special properties of rings.

Definition 1.17. Let $(A, +, \cdot)$ be a ring.

1. We say that A is an *integral domain* if, for all $a_1, a_2 \in A$, $a_1 \cdot a_2 = 0$ implies that $a_1 = 0$ or $a_2 = 0$.

2. We say that A is a *factorial domain* (or *unique factorization domain*, *UFD* for short) if it is an integral domain such that, for all nonzero $a \in A \setminus A^{\times}$, there exist a number $n \in \mathbb{N}$ and prime elements $p_1, \cdots, p_n \in A$ satisfying $a = \prod_{i=1}^{n} p_i$.

3. We say that A is a *principal ideal domain* (in short, *PID*) if it is an integral domain such that all ideals of A are principal.

Example 1.18. $(\mathbb{Z}, +, \cdot)$ is an integral domain which is a factorial domain and a principal ideal domain.

Remark 1.19. 1. Let A be an integral domain, $a \in A$. If a is prime element, then a is an irreducible element.

2. Let A be a factorial domain, $a \in A$. Then a is a prime element if and only if a is an irreducible element.

3. Let A be a ring. The relation $\sim$ (*association*) is defined as

$$a \sim b \Leftrightarrow \exists u \in A^{\times} : a = ub$$

is an equivalence relation. If $a \sim b$ we say that a and b are *associated*.

Assume now that A is a factorial domain, and $\mathbb{P}$ a system of representatives of all irreducible elements modulo $\sim$. Then for all nonzero $a \in A \setminus A^{\times}$ there exists a unique unit $u \in A^{\times}$, a unique $n \in \mathbb{N}$ and unique (except for a rearrangement) $p_1, \ldots, p_n \in \mathbb{P}$ such that $a = u \prod_{i=1}^{n} p_i$. Hence the name *unique* factorization domain.

Finally we arrive at the definition of field, which is the central object in this course.

Definition 1.20. We say that a ring $(A, +, \cdot)$ is a *field* if

1. $0_A \neq 1_A$.
2. For all $a \in A \setminus \{0\}$ there exists $b \in A$ such that $a \cdot b = b \cdot a = 1_A$.

Example 1.21. 1. Let p be a prime number. Then $(\mathbb{Z}/p\mathbb{Z}, +, \cdot)$ is a field.

2. $(\mathbb{Q}, +, \cdot)$ is a field. $(\mathbb{R}, +, \cdot)$ and $(\mathbb{C}, +, \cdot)$ are fields.

3. Let A be an integral domain. Then the *field of fractions* of A, denoted $\mathrm{Frac}(A)$, is a field. Let us recall how it is defined: we consider the set $A \times (A \setminus \{0\})$ and define an equivalence relation $\sim$ on this set by

$$(a, x) \sim (b, y) \Leftrightarrow a \cdot y = b \cdot x.$$

The set Frac(A) is the quotient of $A \times (A \setminus \{0\})$ by this equivalence relation; the class of $(a, x) \in A \times (A \setminus \{0\})$ is denoted by $\frac{a}{x}$. We endow $\mathrm{Frac}(A)$ with operations $+$ and $\cdot$ as follows

$$\begin{aligned} + &: \mathrm{Frac}(A) \times \mathrm{Frac}(A) \to \mathrm{Frac}(A) & \cdot &: \mathrm{Frac}(A) \times \mathrm{Frac}(A) \to \mathrm{Frac}(A) \\ \frac{a}{x} + \frac{b}{y} &:= \frac{a \cdot y + b \cdot x}{x \cdot y} & \frac{a}{x} \cdot \frac{b}{y} &:= \frac{a \cdot b}{x \cdot y}. \end{aligned}$$

We have an embedding $A \hookrightarrow \mathrm{Frac}(A)$ given by $a \mapsto \frac{a}{1}$. If $a \in A$ we will denote for convenience by a the element $\frac{a}{1}$ of $\mathrm{Frac}(A)$.

Definition 1.22. Let $(K,+_K,\cdot_K)$ and $(L,+_L,\cdot_L)$ be two fields. We say that a map $f:K\to L$ is a *field (homo)morphism* if it is a ring morphism between the rings $(K,+_K,\cdot_K)$ and $(L,+_L,\cdot_L)$.

Definition 1.23. Let $(E,+,\cdot)$ be a field and $K\subset E$. We say that K is a subfield of E if K is a subring of E such that, for all $a\in K\setminus\{0\}$, $a^{-1}\in K$. If K is a subfield of E, we say that E *is an extension of* K, and we denote it by E/K.

Remark 1.24. If $(E,+,\cdot)$ is a field and $K\subset E$ is a subfield, then the subring $(K,+|_K,\cdot|_K)$ is a field.

Example 1.25. $\mathbb{R}/\mathbb{Q}$ is a field extension.

Definition 1.26. Let $(K,+_K,\cdot_K)$ be a field. A *vector space* is a tuple $(V,+_V,\cdot_V)$, where V is a set, $+_V:V\times V\to V$, $\cdot_V:K\times V\to V$ are maps satisfying:

1. $(V,+_V)$ is a commutative group.
2. For all $v\in V$, $1_K\cdot v=v$.
3. For all $a\in K$, for all $v,w\in V$, $a\cdot_V(v+_V w)=a\cdot_V v+_V a\cdot_V w$.
4. For all $a,b\in K$, for all $v\in V$, $(a+_K b)\cdot_V v=a\cdot_V v+_V b\cdot_V v$.
5. For all $a,b\in K$, for all $v\in V$, $(a\cdot_K b)\cdot_V v=a\cdot_V(b\cdot_V v)$.

Example 1.27. 1. Let $(E,+,\cdot)$ be a field and $K\subset E$ a subfield. Then E is a K-vector space.

2 Polynomial Rings

Let A be a ring. The *ring of polynomials in one variable with coefficients in* A is defined as the tuple $(A[X],+,\cdot)$, where

$$A[X]:=\left\{\sum_{i=0}^{n}a_iX^i : n\in\mathbb{N}\cup\{0\}, a_i\in A \text{ for all } i=0,\dots,n\right\},$$

and the sum and the product are given respectively by the following rules: If $f(X)=\sum_{i=0}^{n}a_iX^i$ and $g(X)=\sum_{j=0}^{m}b_jX^j$, then

$$\begin{aligned} f(x)+g(x) &= \sum_{i=0}^{\max\{n,m\}}(a_i+b_i)X^i;\\ f(x)\cdot g(x) &= \sum_{k=0}^{n+m}\left(\sum_{i=0}^{k}a_ib_{k-i}\right)X^k. \end{aligned}$$

with the convention that $a_r=0$ for $r>n$ and $b_s=0$ for all $s>m$.

The degree of a polynomial is a map defined as

$$\begin{aligned} \deg : A[X]\setminus\{0\} &\to \mathbb{N}\cup\{0\}\\ \sum_{i=0}^{n}a_iX^i &\mapsto \max\{i : a_i\neq 0\}. \end{aligned}$$

We extend $\deg$ by setting $\deg(0) = -\infty$, where, by convention, $-\infty$ is a symbol satisfying that, for all $n \in \mathbb{N} \cup \{0\}$, $-\infty < n$, $-\infty + n = -\infty$, and $(-\infty) + (-\infty) = -\infty$. The degree map satisfies that, for all $f, g \in A[X]$, $\deg(f + g) \leq \max\{\deg(f), \deg(g)\}$ and $\deg(f \cdot g) \leq \deg(f) + \deg(g)$. When A is an integral domain, then we have the equality $\deg(f \cdot g) = \deg(f) + \deg(g)$. Using this equality one sees easily that if A is an integral domain, then $A[X]$ is also an integral domain. Indeed, if $f \cdot g = 0$, then $\deg(f) + \deg(g) = \deg(f \cdot g) = -\infty$, and this can only happen if at least one of $\deg(f)$, $\deg(g)$ is equal to $-\infty$. Another consequence of this equality is that if A is an integral domain, then $(A[X])^\times = A^\times$ (see Algebra 2, Cor. 48).

When the ring of coefficients is a field K, the ring $K[X]$ has a particularly nice structure. As you saw in Algebra 2, with the help of the $\deg$ map one can define an Euclidean division, which in turn allows one to prove that $K[X]$ is a principal ideal domain and thus a factorial domain. In this section we are going to be concerned with polynomial rings whose coefficient ring is not necessarily a field. We start with an example.

Example 2.1. Consider the polynomial ring $\mathbb{Z}[X]$.

1. Since $\mathbb{Z}$ is an integral domain, we have that $\mathbb{Z}[X]$ is an integral domain and $(\mathbb{Z}[X])^\times = \mathbb{Z}^\times = \{1, -1\}$.

2. $\mathbb{Z}[X]$ is not a PID. Namely, the ideal $I = (2, X)$ is not principal.

3. We have a natural ring morphism $\iota : \mathbb{Z}[X] \to \mathbb{Q}[X]$ (the *inclusion*), defined as $\iota(\sum_{i=0}^n a_i X^i) = \sum_{i=0}^n a_i X^i$. And we know that $\mathbb{Q}[X]$ is a factorial domain. The polynomial $2X + 2 \in \mathbb{Z}[X]$ is not irreducible, since it can be decomposed as $2X + 2 = 2 \cdot (X + 1)$, and neither 2 nor $X + 1$ belong to $(\mathbb{Z}[X])^\times$. However, $\iota(2X + 2)$ is irreducible in $\mathbb{Q}[X]$ (recall that $2 \in \mathbb{Q}^\times$).

4. Is $\mathbb{Z}[X]$ a factorial domain?

The aim of this section is to show that, if A is a factorial domain, then $A[X]$ is also a factorial domain.

Definition 2.2. Let A be a factorial domain, and fix a system of representatives $\mathbb{P}$ of all irreducible elements modulo association.

By Remark 1.19-3 we know that for any $a \in A \setminus \{0\}$ there exist a unique $u \in A^\times$, $n \in \mathbb{N} \cup \{0\}$ and $p_1, \ldots, p_n \in \mathbb{P}$ such that $a = up_1 \cdot \cdots \cdot p_n$. The p_i need not be all distinct: we can collect together all those that coincide, and we obtain a factorisation

$$a = u \cdot \prod_{i=1}^m p_i^{s_i}$$

for some $m \leq n$. We will define the p_i-adic valuation of a as $v_{p_i}(a) := s_i$ for all $i = 1, \ldots, m$, and we define $v_p(a) := 0$ if $p \in \mathbb{P}$ is different from p_i for all $i = 1, \ldots, m$.

Remark 2.3. Let $A, \mathbb{P}$ be as in Definition 2.2. Then for each $a \in A \setminus \{0\}$ we can write $a = u \cdot \prod_{p \in \mathbb{P}} p^{v_p(a)}$ for a certain (unique) element $u \in A^\times$.

In the rest of the section, A will always denote a factorial domain and $\mathbb{P}$ a fixed system of representatives of irreducible elements modulo association. Note that, if $\mathbb{P}'$ is another system of representatives of irreducible elements modulo association then whenever we have elements $p \in \mathbb{P}$, $p' \in \mathbb{P}'$ such that $p \sim p'$, it holds that, for all $a \in A \setminus \{0\}$, $v_p(a) = v_{p'}(a')$ (See Exercise Sheet 3). In other words, the p-valuation depends only on the class of p modulo association, so the set of valuation maps $\{v_p : p \in \mathbb{P}\}$ and $\{v_{p'} : p' \in \mathbb{P}'\}$ coincide.

Lemma 2.4. *Let A be a factorial domain, and fix a system of representatives $\mathbb{P}$ of all irreducible elements modulo the association. In particular, A is an integral domain, so we may consider the field of fractions $K := \mathrm{Frac}(A)$.*

Then, for all $z \in K \setminus \{0\}$, there exist a unique $u \in A^\times$, a unique $n \in \mathbb{N} \cup \{0\}$, unique $p_1, \ldots, p_n \in \mathbb{P}$ and unique $r_i \in \mathbb{Z}$ so that

$$z = u \cdot \prod_{i=0}^{n} p_i^{r_i}.$$

Proof.

- **Existence:** Let $z \in K \setminus \{0\}$; then $z = \frac{a}{b}$ for some $a, b \in A \setminus \{0\}$. Since A is a unique factorisation domain, we have that there exist (unique) $u_a, u_b \in A^\times$ such that

$$a = u_a \cdot \prod_{p \in \mathbb{P}} p^{v_p(a)}, b = u_b \cdot \prod_{p \in \mathbb{P}} p^{v_p(b)}.$$

Thus

$$x = \frac{a}{b} = \frac{u_a \cdot \prod_{p \in \mathbb{P}} p^{v_p(a)}}{u_b \cdot \prod_{p \in \mathbb{P}} p^{v_p(b)}} = \frac{u_a \cdot u_b^{-1}}{1} \cdot \prod_{p \in \mathbb{P}} p^{v_p(a) - v_p(b)}.$$

where $u_a \cdot u_b^{-1} \in A^\times$ and $v_p(a) - v_p(b) \in \mathbb{Z}$ and is equal to zero for all save finitely many $p \in \mathbb{P}$.

- **Uniqueness** Assume we have two such factorisations of z, namely

$$z = u \cdot \prod_{p \in \mathbb{P}} p^{r_p} = v \cdot \prod_{p \in \mathbb{P}} p^{s_p}$$

for some $u, v \in A^\times$, $r_p, s_p \in \mathbb{Z}$ for all $p \in \mathbb{P}$.

It follows that

$$uv^{-1} \cdot \prod_{p \in \mathbb{P}} p^{r_p - s_p} = 1.$$

This is an equality in $\mathrm{Frac}(A)$; we want to express it as an equality within the ring A. For all $p \in \mathbb{P}$, let us write

$$\begin{aligned} \mathbb{P}_{z,+} &:= \{p \in \mathbb{P} \text{ such that } r_p - s_p > 0\} \\ \mathbb{P}_{z,0} &:= \{p \in \mathbb{P} \text{ such that } r_p - s_p = 0\} \\ \mathbb{P}_{z,-} &:= \{p \in \mathbb{P} \text{ such that } r_p - s_p < 0\} \end{aligned}$$

We have clearly that $\mathbb{P} = \mathbb{P}_{z,+} \cup \mathbb{P}_{z,0} \cup \mathbb{P}_{z,-}$, and consequently

$$\frac{uv^{-1} \cdot \prod_{p \in \mathbb{P}_{z,+}} p^{r_p - s_p}}{\prod_{p \in \mathbb{P}_{z,-}} p^{-r_p + s_p}} = \frac{1}{1}.$$

Note that $uv^{-1}\cdot\prod_{p\in\mathbb{P}_{z,+}}p^{r_p-s_p}$ and $\prod_{p\in\mathbb{P}_{z,-}}p^{-r_p+s_p}$ belong to A. By the definition of the fraction field (see Remark 1.213) this is equivalent to the following equality in A:

$$uv^{-1}\cdot\prod_{p\in\mathbb{P}_{z,+}}p^{r_p-s_p}=\prod_{p\in\mathbb{P}_{z,-}}p^{-r_p+s_p}. \tag{2.1}$$

Hence all p appearing in the left-hand-side must appear in the right hand side as well. But $\mathbb{P}_{z,+}\cap\mathbb{P}_{z,-}=\emptyset$. Therefore there cannot be any prime in the product of the left-hand-side. Similarly, there cannot be any prime appearing in the product of the right-hand-side. That is to say, $\mathbb{P}_{z,+}=\mathbb{P}_{z,-}=\emptyset$, so $\mathbb{P}=\mathbb{P}_{z,0}$. In other words, for all $p\in\mathbb{P}$, $r_p=s_p$.

Moreover, Equation (2.1) reduces to $uv^{-1}=1$, so $u=v$. We conclude that the two factorisations of z were actually the same.

□

The previous lemma allows us to extend Definition 2.2 to the field of fractions of factorial domains.

Definition 2.5. Let A be a factorial domain, and fix a system of representatives $\mathbb{P}$ of all irreducible elements modulo association. Let $K:=\mathrm{Frac}(A)$.

By Lemma 2.4 we know that for any $z\in K\setminus\{0\}$ there exist a unique $u\in A^\times$, and for all $p\in\mathbb{P}$, a unique $r_p\in\mathbb{Z}$, such that

$$z=u\cdot\prod_{p\in\mathbb{P}}p^{r_p}$$

For all $p\in\mathbb{P}$, we define the p-adic valuation of z as $v_p(z):=r_p$.

Remark 2.6. Given a factorial domain A and an irreducible element p, the definition above gives us a map

$$v_p:\mathrm{Frac}(A)\setminus\{0\}\to\mathbb{Z}.$$

We would like to extend this definition to the whole $\mathrm{Frac}(A)$, in a coherent way. In order to do so, we introduce a symbol, $+\infty$, and we extend operation $+$ of $\mathbb{Z}$ to $\mathbb{Z}\cup\{+\infty\}$ by setting

- For all $a\in\mathbb{Z}$, $a+(+\infty):=+\infty$.
- $(+\infty)+(+\infty)=+\infty$.

We also make the convention that, for all $a\in\mathbb{Z}$, $a<+\infty$.

Lemma 2.7. *Let A be a factorial domain, $K=\mathrm{Frac}(A)$ and let p be an irreducible element of A. For all $a,b\in K$, the following properties hold:*

1. $v_p(a)=+\infty$ *if and only if* $a=0$.
2. $v_p(a\cdot b)=v_p(a)+v_p(b)$.
3. $v_p(a+b)\geq\min\{v_p(a),v_p(b)\}$.

Proof. Exercise. □

Remark 2.8. Let K be a field. A map $v : K \to \mathbb{Z} \cup \{\infty\}$ satisfying 1, 2 and 3 of the previous lemma is called a *discrete valuation*. This notion will play a central role in the lecture *Local Fields* of the master in Mathematics.

Remark 2.9. Let A be a factorial domain and $\mathbb{P}$ a set of representatives of all irreducible elements of A modulo association. Let $a \in \operatorname{Frac}(A)$, say

$$a = u \cdot \prod_{p \in \mathbb{P}} p^{v_p(a)} \tag{2.2}$$

be the unique factorisation of a with respect to $\mathbb{P}$. Clearly if $a \in K$ is such that $v_p(a) \geq 0$ for all $p \in \mathbb{P}$, then $a \in A$ because it is a product of elements of A. Reciprocally, if $a \in A$, then all $v_p(a)$ must be greater than or equal to zero. Indeed, we can always get an equality in A from Equation (2.2), namely

$$a \cdot \prod_{p \in \mathbb{P}_{a,-}} p^{-v_p(a)} = \prod_{p \in \mathbb{P}_{a,+}} p^{v_p(a)}.$$

Here we are using the notation introduced in the proof of Lemma 2.4. Now any prime $p \in \mathbb{P}_{a,-}$ appears in the left hand side with strictly positive exponent, therefore it must divide the right hand side, and consequently must belong to $\mathbb{P}_{a,+}$. Since $\mathbb{P}_{a,-} \cap \mathbb{P}_{a,+} = \emptyset$, we conclude that $\mathbb{P}_{a,-} = \emptyset$, as we wished to show.

Analogously we can prove that $a \in A^\times$ if and only if for all $p \in \mathbb{P}$, $v_p(a) = 0$ (see Exercise Sheet 4).

Definition 2.10. Let A be a factorial domain and K its fraction field. Let $p \in A$ be irreducible and let $f(X) := a_n X^n + a_{n-1} X^{n-1} + \cdots + a_1 X + a_0 \in K[X]$. We define the *p-valuation* of $f(X)$ as

$$v_p(f) := \min_{i=0,\ldots,n} \{v_p(a_i)\}.$$

We say that $f(X)$ is *primitive* if $v_p(f) = 0$ for all irreducible elements $p \in A$.

Remark 2.11. Note that, if $f(X) \in K[X]$ is a primitive polynomial, in particular $f(x) \in A[X]$. Indeed, by Remark 2.9 it suffices to show that, for all $p \in A$ irreducible, for all $i = 0, \ldots, \deg(f)$, $v_p(a_i) \geq 0$. Let $p \in A$ be irreducible. Then $v_p(f) = 0$. That is to say, $\min\{v_p(a_i) : i = 0, \ldots, \deg(f)\} = v_p(f) = 0$. Thus for all $i = 0, \ldots, \deg(f)$, the coefficient a_i satisfies that $v_p(a_i) \geq \min\{v_p(a_i) : i = 0, \ldots, \deg(f)\} = 0$.

Example 2.12. Let us consider the ring $\mathbb{Z}[X]$.

1. The polynomial $f(X) = X^2 + 2X + 4$ is primitive. Indeed, let p be an irreducible element and call $a_2 = 1$, $a_1 = 2$, $a_0 = 4$. On the one hand, $v_p(a_i) \geq 0$ for all $i = 0, 1, 2$ (because the coefficients lie in $\mathbb{Z}$). That means that $v_p(f) = \min\{v_p(a_i) : i = 0, 1, 2\} \geq 0$. On the other hand, $v_p(a_2) = v_p(1) = 0$, hence $\min\{v_p(a_i) : i = 0, 1, 2\} \leq 0$.

2. The polynomial $f(X) = 10X^2 + 2X + 4$ is not primitive. Indeed, and call $a_2 = 10$, $a_1 = 2$, $a_0 = 4$. There exists an irreducible element of $\mathbb{Z}$, namely 2, such that $v_2(a_2) = 1 > 0$, $v_2(a_1) = 1 > 0, v_2(a_0) = 2 > 0$. Therefore $v_2(f) = \min\{v_2(a_2), v_2(a_1), v_2(a_0)\} = 1 > 0$.

3. The polynomial $f(X) = \frac{1}{2}X^2 + 2X + 4$ is not primitive. Indeed, call and call $a_2 = 1/2$, $a_1 = 2$, $a_0 = 4$ .There exists an irreducible element of $\mathbb{Z}$, namely 2, such that $v_2(a_2) = -1 < 0$, therefore $v_2(f) = \min\{v_2(a_2), v_2(a_1), v_2(a_0)\} \leq v_2(a_2) = -1 < 0$.

Recall that, if A is a factorial domain and we have elements $a_1, \dots, a_n \in A$, a *greatest common divisor of* $a_1, \dots, a_n$, denoted $\gcd(a_1, \dots, a_n)$, is defined as an element $d \in A$ such that $d|a_i$ for all $i = 1, \dots, n$ and for all $b \in A$ satisfying that $b|a_i$ for all $i = 1, \dots, n$, then $b|d$.

Lemma 2.13. *Let A be a factorial domain, let $\mathbb{P}$ be a system of representatives of all irreducible elements modulo association, and let $f(X) = \sum_{i=0}^{n} a_i X^i \in K[X]$ be a nonzero polynomial. Then the following properties hold.*

1. *$v_p(f) = 0$ for all $p \in \mathbb{P}$ except for finitely many.*

2. *$f \in A[X] \Leftrightarrow \forall p \in \mathbb{P}, v_p(f) \geq 0$.*

3. *Assume that $a_i \in A$ for all $i = 0, \dots n$, and let $p \in \mathbb{P}$. Then $v_p(f) = v_p(\gcd(a_0, a_1, \dots, a_n))$.*

4. *Assume that $a_i \in A$ for all $i = 0, \dots n$, and let $p \in \mathbb{P}$. Then p divides f in the ring $A[X]$ if and only if $v_p(f) > 0$.*

5. *There exists $a \in K$ such that $a \cdot f(X)$ is a primitive polynomial.*

6. *Let $p \in \mathbb{P}$, $a \in K$. Then $v_p(a \cdot f) = v_p(a) + v_p(f)$.*

Proof. Exercise. □

We have seen in the lemma above that for every $a \in K$ and $f \in K[X]$, the equality $v_p(a \cdot f) = v_p(a) + v_p(f)$ holds for all irreducible element p. We would like to know what happens if we have two elements $f(X), g(X) \in K[X]$ and we multiply them together. Is $v_p(f \cdot g) = v_p(f) + v_p(g)$? The next result answers this question.

Proposition 2.14 (Gauß's Lemma)**.** *Let A be a factorial domain, K its fraction field and p an irreducible element of A. Then for all $f, g \in K[X]$,*

$$v_p(f \cdot g) = v_p(f) + v_p(g).$$

Proof. We will do the proof in three steps: In Step (1) we will prove the lemma for a particular case, namely when $f, g \in A[X]$ are both primitive. In Step (2) we will prove the lemma in another particular case, namely when $f, g \in A[X]$, using Step (1). In Step (3) we will prove the lemma in full generality using Step (2).

(1) In this step we assume $f, g \in A[X]$ and that $v_p(f) = v_p(g) = 0$. First of all, note that $v_p(f \cdot g) \geq 0$ because $f \cdot g \in A[X]$. Consider the following ring morphism

$$\begin{aligned} \pi : A &\to A/(p) \\ a &\mapsto \overline{a} := a + (p). \end{aligned}$$

Note that, since p is irreducible and A is a factorial domain, p is also prime. Hence the ideal (p) is prime. Recall that this implies that $A/(p)$ is an integral domain (indeed: if $\overline{a} \cdot \overline{b} = \overline{0}$, this implies that $\overline{a \cdot b} = \overline{0}$, or equivalently, $a \cdot b \in (p)$. By definition of prime ideal, this implies that $a \in (p)$ or $b \in (p)$. In other words, $\overline{a} = \overline{0}$ or $\overline{b} = \overline{0}$).

We can extend π to a ring morphism between $A[X]$ and $(A/(p))[X]$, by defining

$$\pi \left(\sum_{i=0}^{n} a_i X^i \right) := \sum_{i=0}^{n} \overline{a}_i X^i$$

(check!). Write $f(X) = \sum_{i=0}^{n} a_i X^i$. Since $\min\{v_p(a_i) : i = 0, \dots, n\} = v_p(f) = 0$, there exists some coefficient a_i of f with $v_p(a_i) = 0$. Thus $\pi(f) = \sum_{i=0}^{n} \overline{a}_i X^i \neq 0$. Similarly we can conclude that $\pi(g) \neq 0$.

Now we use the fact that $A/(p)$ is an integral domain, and conclude that $(A/(p))[X]$ is also an integral domain; thus $\pi(f) \cdot \pi(g) \neq 0$. But $\pi(f) \cdot \pi(g) = \pi(fg)$. So if we write $fg = \sum_{i=0}^{m} b_i X^i$ for some $b_i \in A$, it holds that $\sum_{i=0}^{m} \overline{b}_i X^i \neq 0$. So there must be some $j \in \{0, \dots, m\}$ such that $b_j \notin (p)$. Therefore $v_p(b_j) = 0$, and we conclude that $v_p(f \cdot g) = \min\{v_p(b_i) : i = 1, \dots m\} = 0 = v_p(f) + v_p(g)$.

(2) In this step we assume that $f(X), g(X) \in A[X]$ (but we do not assume that $v_p(f) = 0$ and $v_p(g) = 0$). Let d_f (resp. d_g) denote the greatest common divisor of all coefficients of f (resp. of g). Then $\tilde{f} := \frac{1}{d_f} f$ and $\tilde{g} := \frac{1}{d_g} g$ are polynomials in $A[X]$ (by definition of greatest common divisor). Moreover, by Lemma 2.13-(3 and 6), we have $v_p(\tilde{f}) = v_p(\frac{1}{d_f} f) = v_p(\frac{1}{d_f}) + v_p(f) = -v_p(d_f) + v_p(d_f) = 0$. Similarly $v_p(\tilde{g}) = 0$.

Thus we can apply the result of Step 1 and conclude that $v_p(\tilde{f}\tilde{g}) = v_p(\tilde{f}) + v_p(\tilde{g})$. Hence

$$\begin{aligned} v_p(f) + v_p(g) = v_p(d_f \tilde{f}) + v_p(d_g \tilde{g}) &= v_p(d_f) + v_p(\tilde{f}) + v_p(d_g) + v_p(\tilde{g}) \\ &= v_p(d_f) + v_p(d_g) + v_p(\tilde{f}\tilde{g}) = v_p(d_f \cdot d_g \cdot \tilde{f} \cdot \tilde{g}) = v_p(f \cdot g). \end{aligned}$$

(3) In this step we do not make any extra assumptions on f and g; that is to say, $f, g \in K[X]$ are arbitrary. Choose $a, b \in A$ such that $af, bg \in A[X]$. Then we can apply Step 2 and conclude that $v_p(af \cdot bg) = v_p(af) + v_p(bg)$. Thus, using again 2.13-(6), we obtain

$$\begin{aligned} v_p(f) + v_p(g) = v_p\left(\frac{1}{a}(af)\right) + v_p\left(\frac{1}{b}(bg)\right) &= v_p\left(\frac{1}{a}\right) + v_p(af) + v_p\left(\frac{1}{b}\right) + v_p(bg) \\ &= v_p(af \cdot bg) + v_p\left(\frac{1}{a}\right) + v_p\left(\frac{1}{b}\right) = v_p\left(\frac{1}{a} af \frac{1}{b} bg\right) = v_p(fg). \end{aligned}$$

□

Definition 2.15. Let B be a ring. We say that $f(X) = \sum_{i=0}^{\deg(f)} b_i X^i \in B[X]$ is *monic* (*unitaire* in French) if $b_{\deg f} = 1$.

Corollary 2.16. *Let A be a factorial domain and $K = \mathrm{Frac}(A)$. Let $f, g \in K[X]$ be monic polynomials. Assume that $f \cdot g \in A[X]$. Then we have that $f, g \in A[X]$.*

Proof. By Lemma 2.13-(2), it suffices to prove that, for all irreducible element p in A, $v_p(f) \geq 0$ and $v_p(g) \geq 0$.

Let us fix an irreducible element p. Write $f(X) = \sum_{i=0}^{\deg(f)} a_i X^i$. Since $f \in K[X]$ is monic, $a_{\deg(f)} = 1$, thus $0 = v_p(a_{\deg(f)}) \geq \min\{v_p(a_i) : i = 0, \dots, \deg(f)\} = v_p(f)$. Similarly $0 \geq v_p(g)$.

Since $f \cdot g \in A[X]$, Lemma 2.13-(2) implies that $v_p(f \cdot g) \geq 0$. Collecting together this information and making use of Gauß Lemma, we obtain that

$$0 \leq v_p(f \cdot g) = v_p(f) + v_p(g) \leq 0$$

Thus $0 = v_p(f) + v_p(g)$. But both integers $v_p(f)$ and $v_p(g)$ are less than or equal to zero, so the only possibility is that both of them are zero. In particular, they are greater than or equal to zero, as was to be proved. □

Let A be a factorial domain, and $K = \operatorname{Frac}(A)$. We have that both A and $K[X]$ are factorial domains, and we have that $A[X]$ sits between them as $A \subset A[X] \subset K[X]$. This inclusions point out two different kinds of special elements of $A[X]$; on the one hand, we have the prime elements of A (which are contained in $A[X]$), and on the other hand we can consider those elements of $A[X]$ that are prime as elements of $K[X]$. In fact, it is not difficult to see every element in $A[X]$ can be written as a product of some of these special elements. The following lemma gives a precise formulation.

Lemma 2.17. *Let A be a factorial domain and $K = \operatorname{Frac}(A)$. Then every nonzero element $f \in A[X]$ can be written as a product*

$$f = u \cdot \left(\prod_{i=1}^{r} p_i \right) \cdot \left(\prod_{j=1}^{s} g_j \right)$$

for some $r, s \in \mathbb{N} \cup \{0\}$, where $u \in A^\times$, $p_1, \dots, p_r$ are prime elements of A and $g_1, \dots, g_s \in A[X]$ are primitive polynomials that are prime elements of $K[X]$.

Proof. Let us fix some nonzero $f = \sum_{i=0}^{\deg f} a_i X^i \in A[X]$. Let $a = \gcd\{a_0, \dots, a_{\deg f}\}$. To use it later on, we will decompose a as a product of prime elements of A, say $a = p_1 \cdot \dots \cdot p_r$. We have that $\widetilde{f} = \frac{1}{a} f$ is a primitive polynomial of $A[X]$. We can view it inside $K[X]$, and write it as a product of prime elements of $K[X]$, namely $\widetilde{f} = h_1 \cdot \dots \cdot h_s$, where $h_i \in K[X]$ for all $i = 1, \dots, s$. For each i, we can multiply h_i by the product of all the denominators appearing in the coefficients of h_i, say M_i, so we get that $M_i h_i \in A[X]$. Now we can consider the greatest common divisor of the coefficients of $M_i h_i$, say d_i, and look at the polynomial $\widetilde{h}_i = (M_i/d_i) h_i$. This polynomial belongs to $A[X]$ and is primitive. We can thus write

$$\widetilde{f} = \prod_{i=1}^{s} h_i = \prod_{i=1}^{s} (d_i/M_i) \widetilde{h}_i = u \cdot \prod_{i=1}^{s} \tilde{h}_i,$$

where $u = \prod_{i=1}^{s} (d_i/M_i)$ is an element in K. Actually, we can see that it belongs to $A^\times$: for any irreducible element $p \in A$, we have that

$$v_p(\widetilde{f}) = v_p(u \cdot \prod_{i=1}^{s} \tilde{h}_i) = v_p(u) + v_p \left(\prod_{i=1}^{s} \tilde{h}_i \right) = v_p(u) + \sum_{i=1}^{s} v_p(\tilde{h}_i) = v_p(u)$$

because all the $\tilde{h}_i$ are primitive. But $\widetilde{f}$ is also primitive, therefore $v_p(u) = v_p(\widetilde{f}) = 0$. Since this holds for all irreducible p in A, we conclude by Remark 2.9 that $u \in A^\times$. Now we just need to put all together:

$$f = a \cdot \widetilde{f} = \left(\prod_{i=1}^{r} p_i\right) \cdot \left(u \prod_{i=1}^{s} \tilde{h}_i\right).$$

□

So we have that the two types of special elements of $A[X]$, namely prime elements of A and primitive polynomials in $A[X]$ that are prime elements of $K[X]$ behave very much as though they were the prime elements of $A[X]$. The next proposition shows that, in fact, this is true.

Proposition 2.18. *Let A be a factorial domain, $K = \mathrm{Frac}(A)$ and let $f(X) \in A[X]$ nonzero. Then following assertions are equivalent:*

(i) $f(X)$ is a prime element in $A[X]$.

(ii) One of the following two possibilities holds:

(I) f is a constant polynomial (i.e., $f \in A$) and f is a prime element of A.

(II) f is primitive and f is a prime element in $K[X]$.

Proof. ($\Leftarrow$) Recall that, to prove that an element $f \in A[X]$ is prime, we have to see that:

(1) f is nonzero

(2) f is not a unit

(3) if $f | g_1 \cdot g_2$ for some $g_1, g_2 \in A[X]$, then either $f | g_1$ or $f | g_2$.

Since $(A[X])^\times = A^\times$, it is clear that the elements of type (I) and (II) are not units of $A[X]$ (and clearly they are also nonzero).

(I) Assume first that f is a prime element of A, and assume that $f | g_1 \cdot g_2$ for some $g_1, g_2 \in A[X]$. Since f is a prime element of A, it is an irreducible element of A. Thus we can consider the f-adic valuation v_f. By Lemma 2.13-(4), we have that $f | g_1 \cdot g_2$ is equivalent to saying that $v_f(g_1 \cdot g_2) > 0$. By Gauß's Lemma (Proposition 2.14) it holds that $v_f(g_1 \cdot g_2) = v_f(g_1) + v_f(g_2)$. Therefore $v_f(g_1)$ and $v_f(g_2)$ are two numbers which are greater than or equal to zero, and $v_f(g_1) + v_f(g_2) > 0$, so either $v_f(g_1) > 0$ or $v_f(g_2) > 0$. By Lemma 2.13-(4), this is equivalent to the assertion that either $f | g_1$ or $f | g_2$.

(II) Assume now that $f \in A[X]$ is primitive and is a prime element of $K[X]$. Assume also that $f | g_1 \cdot g_2$ for some $g_1, g_2 \in A[X]$. That means there exists $f_0 \in A[X]$ such that $f f_0 = g_1 g_2$. We can look at this equality in $K[X]$; we have then that $f | g_1 \cdot g_2$ in the ring $K[X]$. Since f is a prime element of this ring, this implies that either $f | g_1$ or $f | g_2$ in $K[X]$. Let us assume that $f | g_1$ (the other case is analogous). Then there exists $h \in K[X]$ such that $f \cdot h = g_1$. We will prove that, in fact, $h \in A[X]$. To see this, we will show that, for any irreducible element $p \in A$, $v_p(h) \geq 0$ and apply Lemma 2.13-(2).

Fix an irreducible element p of A; we have that $v_p(g_1) = v_p(f \cdot h) = v_p(f) + v_p(h) = v_p(h)$ (recall that $v_p(f) = 0$ because f is primitive). Thus $v_p(h) = v_p(g_1) \geq 0$ because $g_1 \in A[X]$.

$(\Rightarrow)$ Let $f \in A[X]$ be a prime element. We have to prove that it is of type (I) or (II). By Lemma 2.17, we can write f as

$$f = u \cdot \left(\prod_{i=1}^{r} p_i \right) \cdot \left(\prod_{j=1}^{s} g_j \right)$$

for some $r, s \in \mathbb{N} \cup \{0\}$, where $u \in A^\times$, $p_1, \ldots, p_r$ elements of type (I) and $g_1, \ldots, g_s \in A[X]$ are elements of type (II). But prime elements are always irreducible elements (this is valid in any integral domain, not necessarily a factorial domain). Thus f is irreducible, and it is written as a product of a unit and elements that are not units. Therefore there can only be one element (up to product by unities) in the product, that is to say, either $f = vp_i$ for some i and $v \in A^\times$ or $f = vg_j$ for some j and $v \in A^\times$. In the first case, f is of type (I) and in the second case f is of type (II). □

Collecting together Lemma 2.17 and Proposition 2.18, we get that if A is a factorial domain, then every element in $A[X] \setminus \{0\}$ can be written as a product of a unit times several prime elements. In other words, we obtain that $A[X]$ is a factorial domain. We write this statement in the following corollary.

Corollary 2.19 (Gauß). *Let A be a factorial domain. Then $A[X]$ is a factorial domain.*

Proof. If A is a factorial domain, A is in particular an integral domain, thus $A[X]$ is also an integral domain. By Lemma 2.17 and Proposition 2.18, every nonzero element of $A[X]$ can be written as a product of prime elements times a unit. Hence $A[X]$ is a factorial domain. □

Example 2.20.

1. The ring $\mathbb{Z}[X]$ is a factorial domain.
2. Let K be a field. The ring $K[X, Y] := (K[X])[Y]$ is a factorial domain.
3. The ring $K[X_1, \ldots, X_n] := (\cdots((K[X_1])[X_2])\cdots)[X_n]$ is a factorial domain.

We state below a corollary that will be very useful for us in the next lectures.

Corollary 2.21. *Let A be a factorial domain, $K = \operatorname{Frac}(A)$ and $f \in A[X]$ a primitive polynomial which is not a constant (i.e., $f \notin A$). Then the following statements are equivalent:*

(i) f is an irreducible element of $A[X]$.

(ii) f is a prime element of $A[X]$.

(iii) f is an irreducible element of $K[X]$.

(iv) f is a prime element of $K[X]$.

Proof. $(i) \Leftrightarrow (ii)$ and $(iii) \Leftrightarrow (iv)$ follow from the fact that in a factorial domain the prime elements coincide with the irreducible element. $(ii) \Leftrightarrow (iv)$ follow from Proposition 2.18, taking into account that by hypothesis f cannot be of type (I). □

This corollary tells us, for instance, that in order to check if a primitive polynomial $f \in \mathbb{Q}[X]$ is irreducible, it suffices to check if it is irreducible in $\mathbb{Z}[X]$. And this should be easier, because in $\mathbb{Z}[X]$ we have less polynomials than in $\mathbb{Q}[X]$ (hence less possibilities for f to break into a product of two polynomials). In what follows, we are going to be interested in determining when a polynomial $f \in A[X]$ is irreducible (for some factorial domain A). We will see two criteria for this; the *reduction criterion* and the *Eisenstein criterion.*

Proposition 2.22 (Reduction Criterion). *Let A be a factorial domain and $f = \sum_{i=1}^{\deg f} a_i X^i \in A[X]$ be a non constant polynomial which is primitive. Let $p \in A$ be a prime element of A, and consider the map*

$$\pi : A[X] \to (A/(p))\,[X]$$
$$\sum_{i=0}^{r} b_i X^i \mapsto \sum_{i=0}^{r} \bar{b}_i X^i$$

where $\bar{b} := b + (p) \in A/(p)$. If $p \nmid a_{\deg f}$ and $\pi(f)$ is irreducible in $(A/(p))\,[X]$, then f is irreducible in $A[X]$.

Proof. Assume that f is not irreducible, say $f = gh$ for some $g, h \in A[X]$, $g, h \notin (A[X])^{\times} = A^{\times}$. Since f is primitive, it follows from Gauß's Lemma (Proposition 2.14) that g, h are also primitive. This immediately implies that $g, h \notin A$ (if $g \in A$ is primitive, it must lie in $A^{\times}$, and this is not the case. The same applies to h). That is to say, $\deg g, \deg h > 0$. Note that $\deg f = \deg g + \deg h$.

Let us apply π to f, g, h. We have clearly that $\deg f \geq \deg \pi(f)$, $\deg g \geq \deg \pi(g)$ and $\deg h \geq \deg \pi(h)$. In fact, we have that $\deg f = \deg \pi(f)$ because $p \nmid a_{\deg f}$. Recall that $A/(p)$ is an integral domain (because p is a prime element of A), thus $\deg \pi(f) = \deg \pi(gh) = \deg(\pi(g)\pi(h)) = \deg \pi(g) + \deg \pi(h)$. Collecting all this information together we have that

$$\deg g + \deg h = \deg f = \deg \pi(f) = \deg \pi(g) + \deg \pi(h),$$

with $\deg g \geq \deg \pi(g)$ and $\deg h \geq \deg \pi(h)$. We conclude that $\deg g = \deg \pi(g)$ and $\deg h = \deg \pi(h)$

Now we will use the hypothesis that $\pi(f)$ is irreducible. Namely, this imples that either $\pi(g)$ or $\pi(h)$ is a unit in $(A/(p))[X]$. But the units of $(A/(p))[X]$ are the units of $A/(p)$ (because $A/(p)$ is an integral domain). In particular, such units have degree zero. So either $\deg \pi(g) = 0$ or $\deg \pi(h) = 0$. That is to say, either $\deg g = 0$ or $\deg h = 0$. This is a contradiction. □

Example 2.23. 1. Let $f_1(X) = X^2 + X + 1$ and $f_2(X) = X^2 + 27X + 43$ be polynomials in $\mathbb{Z}[X]$. These polynomials are monic, hence they are primitive, and no prime p of $\mathbb{Z}$ divides their *leading coefficient*. So if for some prime p we obtain that their images by $\pi_p : \mathbb{Z}[X] \to \mathbb{F}_p[X]$ are irreducible, then we can apply the reduction criterion and conclude that they are irreducible in $\mathbb{Z}[X]$.

Let us take $p = 2$. Then $\pi_2(f_1) = \pi_2(f_2) = X^2 + X + 1$. This polynomial is irreducible in $\mathbb{F}_2[X]$ (because it is of degree smaller than 4 and it has no roots in $\mathbb{F}_2$), hence we can apply the reduction criterion with $p = 2$ and conclude that f_1 and f_2 are irreducible.

2. Let $f_3 = X^2 + 2X - 1$. Again f_3 is monic, hence primitive, and no prime p of $\mathbb{Z}$ divides its leading coefficient. If we take $p = 2$, we obtain $\pi_2(f_3) = X^2 + 1 = (X + 1)^2$, which is not irreducible, so we cannot apply the reduction criterion with $p = 2$. But let us try with $p = 3$. We obtain that $\pi_3(f_3) = x^2 + 2X + 2$, which is irreducible (since it has degree smaller than 4, it suffices to check it has no root in $\mathbb{F}_3$). So we can apply the reduction criterion for $p = 3$ and conclude that f_3 is irreducible.

3. Let $f_4 = X^2 - 3X + 2$. It is monic, hence primitive, and no prime p of $\mathbb{Z}$ divides its leading coefficient. If we take $p = 2$, we obtain $\pi_2(f_3) = X^2 + X$, which is not irreducible, so we cannot apply the reduction criterion with $p = 2$. If we take $p = 3$, we obtain $\pi_2(f_3) = X^2 + 2$, which is not irreducible, so we cannot apply the reduction criterion with $p = 3$. In fact, we are not going to be able to apply the criterion with any p! The reason is that $f_4 = (x - 2)(x - 1)$ is reducible, so we cannot prove that it is irreducible!

Proposition 2.24 (Eisenstein's Criterion)**.** *Let A be a factorial domain, $K = \mathrm{Frac}(A)$ and let $f = \sum_{i=0}^{d} a_i X^i \in A[X]$ a polynomial of degree $d > 0$. Assume that f is primitive. Let $p \in A$ be a prime element such that:*

$$p \nmid a_d,\ p|a_i \text{ for all } i \in \{0, \ldots, d-1\} \text{ and } p^2 \nmid a_0.$$

Then f is an irreducible element of $A[X]$ (and also of $K[X]$).

Proof. It suffices to prove that f is irreducible in $A[X]$ (indeed: since f is nonconstant and primitive, by Corollary 2.21 if f is irreducible in $A[X]$, it is also irreducible in $K[X]$). Assume that f is not irreducible in $A[X]$, say $f = gh$ with $g, h \in A[X] \setminus (A[X])^\times$. Write $g = \sum_{i=0}^{r} b_i X^i$ and $h = \sum_{i=0}^{s} c_i X^i$, with $b_r, c_s \neq 0$. Note that, since f is primitive, g and h cannot be elements of A. In other words, $r, s > 0$.

On the one hand we have that $d = \deg f = \deg g + \deg h = r + s$ (because A is an integral domain), thus $a_d = b_r \cdot c_s$. Since $p \nmid a_d$, it follows that $p \nmid b_r$ and $p \nmid c_s$. On the other hand, we have that $p|a_0 = b_0 c_0$, so $p|b_0$ or $p|c_0$. In fact, it cannot happen that p divides both b_0 and c_0, because in this case a_0 would be divisible by p^2, and by hypothesis we know this is not the case. So p divides one of b_0 or c_0 and does not divide the other. Let us assume that $p|b_0$ and $p \nmid c_0$ (the other case is analogous).

We have that $p|b_0$ but $p \nmid b_r$. So there must exist some $t \in \{1, \ldots, r\}$ which is the smallest satisfying that $p \nmid b_t$. Note that since $s > 0$ and $r + s = d$, we have $r < d$, and therefore also $t < d$. Now we compute a_t: with the usual convention that $c_i = 0$ if $i > s$, we obtain

$$a_t = \underbrace{b_0 c_t + b_1 c_{t-1} + \cdots + b_{t-1} c_1}_{\text{divisible by } p} \quad + \quad \underbrace{b_t c_0}_{\text{not divisible by } p}.$$

This contradicts the fact that $p|a_i$ for all $i < d$. □

Example 2.25. 1. Let $f = 9X^7 + 8X^5 + 20X + 6 \in \mathbb{Z}[X]$. Since $\gcd\{9, 8, 20, 6\} = 1$ we have that f is primitive. We can apply Eisenstein's criterion for $p = 2$ and conclude that it is irreducible both in $\mathbb{Q}[X]$ and in $\mathbb{Z}[X]$.

2. Let $f = 9X^7 + 24X^5 + 60X + 6 \in \mathbb{Z}[X]$. It is not primitive because $\gcd\{9, 24, 60, 6\} = 3$. But $(1/3)f = 3X^7 + 8X^5 + 20X + 2$ is primitive, and we can apply Eisenstein's criterion for $p = 2$. We conclude that it is irreducible in $\mathbb{Q}[X]$. Since $3 \in \mathbb{Q}^\times = (\mathbb{Q}[X])^\times$, we get that f is irreducible in $\mathbb{Q}[X]$. But it is not irreducible in $\mathbb{Z}[X]$.

3. Let K be a field and $A = K[T]$. Let $n \in \mathbb{N}$. The polynomial $F = X^n - T \in A[X]$ is monic, hence primitive. Moreover we can apply Eisenstin's criterion to the prime element $T \in A$. It follows that F is irreducible in $A[X]$.

4. Let p be a prime, and consider the polynomial $X^p - 1 \in \mathbb{Z}[X]$. This is clearly not irreducible, because
$$X^p - 1 = (X - 1)(X^{p-1} + X^{p-2} + \cdots + X + 1).$$
The polynomial $\Phi_p(X) := X^{p-1} + X^{p-2} + \cdots + X + 1$ is called the *p-th cyclotomic polynomial* (*p-ième polynôme cyclotomique* in French). To prove that it is irreducible, we cannot apply directly Eisenstein's criterion. But note that
$$\Phi_p(X+1) = \frac{(X+1)^p - 1}{(X+1) - 1} = \frac{\left(\sum_{i=0}^p \binom{p}{i} X^i\right) - 1}{X} = \frac{\sum_{i=1}^p \binom{p}{i} X^i}{X}$$
$$= \sum_{i=1}^{p} \binom{p}{i} X^{i-1} = X^p + \sum_{i=1}^{p-1} \binom{p}{i} X^{i-1}.$$
Note that $\Phi_p(X+1)$ is primitive and p does not divide the leading term. However, p divides all the other terms because $p | \binom{p}{i}$ for all $i = 1, \dots, p-1$. Furthermore, $p^2 \nmid \binom{p}{1} = p$. Thus we can apply Eisenstein's criterion and conclude that $\Phi_p(X+1)$ is irreducible. But this implies that $\Phi_p(X)$ is irreducible (if $\Phi_p(X) = f(X)g(X)$, then $\Phi_p(X+1) = f(X+1)g(X+1)$ would also be reducible).

3 Characteristic of an integral domain

In this short chapter we will attach to each integral domain a number, the *characteristic*. This notion will be especially useful when we consider field extensions in the next chapters.

Lemma 3.1. *Let A be a ring. There exists a unique ring morphism $\varphi : \mathbb{Z} \to A$.*

Proof. We will first prove the uniqueness. Let $\varphi : \mathbb{Z} \to A$ be a ring morphism. We have that

1. $\varphi(1) = 1_A$.
2. $\varphi(a + b) = \varphi(a) + \varphi(b)$.
3. $\varphi(a \cdot b) = \varphi(a) \cdot \varphi(b)$.

From (2) we obtain that $\varphi(0) = \varphi(0 + 0) = \varphi(0) + \varphi(0)$, hence $\varphi(0) = 0_A$. For all $n \in \mathbb{N}$, we have that
$$\varphi(n) = \varphi(\underbrace{1 + \cdots + 1}_{n \text{ times}}) = \underbrace{\varphi(1) + \cdots + \varphi(1)}_{n \text{ times}} = \underbrace{1_A + \cdots + 1_A}_{n \text{ times}} = n \cdot 1_A.$$

From (2) we obtain that $\varphi(0) = \varphi(1 + (-1)) = \varphi(1) + \varphi(-1)$, hence from (1) we obtain that $1 + \varphi(-1) = \varphi(0) = 0_A$, so that $\varphi(-1) = -1_A$.

Finally, for all $n \in \mathbb{N}$, $\varphi(-n) = \varphi((-1)n) = \varphi(-1)\varphi(n) = -(n \cdot 1_A)$.

Thus for all $a \in \mathbb{Z}$ we have determined $\varphi(a)$. This implies that there cannot be two different ring morphisms $\varphi_1, \varphi_2 : \mathbb{Z} \to A$.

Moreover it is easy to check that the map $\varphi : \mathbb{Z} \to A$ defined as

$$\varphi(a) = \begin{cases} a \cdot 1_A & \text{if } a \in \mathbb{N} \cup \{0\}, \\ -(-a) \cdot 1_A & \text{otherwise} \end{cases}$$

is a ring morphism. This proves the existence.

Lemma 3.2. *Let A be an integral domain and let $\varphi_A : \mathbb{Z} \to A$ be the only ring homomorphism. The kernel of φ_A is a prime ideal of $\mathbb{Z}$.*

Let $I = \ker \varphi_A \subset \mathbb{Z}$. This set is an ideal of $\mathbb{Z}$ (see Example 1.26 of [Algebra 2]). It remains to show that it is prime. In order to do this, we will show that the ring $\mathbb{Z}/I$ is an integral domain.

Let $B = \varphi_A(\mathbb{Z}) \subset A$. The set B is a subring of A (see again Example 1.2 of [Algebra 2]). Since A is an integral domain, it follows that B must also be an integral domain (Assume $b_1, b_2 \in B$ satisfy $b_1 \cdot b_2 = 0$. Then b_1, b_2 are elements of A satisfying $b_1 \cdot b_2 = 0$, and A is an integral domain, thus either $b_1 = 0$ or $b_2 = 0$).

Finally, recall that by the First Isomorphism Theorem (Theorem 2.9 of [Algebra 2]), we have that φ_A induces an isomorphism of rings between $\mathbb{Z}/I$ and B. Hence $\mathbb{Z}/I$ is an integral domain, as we wanted to show. □

Remark 3.3. Let $I \subset \mathbb{Z}$ be an ideal. Since $\mathbb{Z}$ is a principal ideal domain, there exists $a \in \mathbb{Z}$ such that $I = (a)$.

Assume now that I is a nonzero prime. Then a is a prime element of $\mathbb{Z}$. So $a = p$ or $a = -p$ for some prime number p. Note moreover that the ideal (0) is a prime ideal of $\mathbb{Z}$. Indeed, $\mathbb{Z}/(0) \simeq \mathbb{Z}$ is an integral domain. Therefore the prime ideals of $\mathbb{Z}$ are the ideal (0) and the ideals (p) with p a prime number.

Definition 3.4. Let A be an integral domain and let $\varphi_A : \mathbb{Z} \to A$ be the unique ring morphism. By Lemma 3.2 and Remark 3.3, we have that $\ker \varphi - A$ is equal to (a), where $a = 0$ or $a = p$ is a prime number. We will say that A has *characteristic* a, and we will write $\operatorname{char}(A) = a$.

Example 3.5.

- $\operatorname{char}\mathbb{Z} = \operatorname{char}\mathbb{Q} = \operatorname{char}\mathbb{R} = \operatorname{char}\mathbb{C} = 0$.
- Let p be a prime number. $\operatorname{char}(\mathbb{Z}/p\mathbb{Z}) = p$.

Remark 3.6. Let A be an integral domain of characteristic p. Then

$$p \cdot 1_A = \underbrace{1 + \cdots + 1}_{p \text{ times}} = \underbrace{\varphi_A(1) + \cdots + \varphi_A(1)}_{p \text{ times}} = \varphi_A(p) = 0_A.$$

Lemma 3.7. *Let A, B be two integral domains and $f : A \to B$ an injective ring morphism. Then* $\operatorname{char}(A) = \operatorname{char}(B)$.

Proof. Let $\varphi_A : \mathbb{Z} \to A$ (resp. $\varphi_B : \mathbb{Z} \to B$) be the unique ring morphism from $\mathbb{Z}$ to A (resp. to B). To prove that $\mathrm{char} A = \mathrm{char} B$, it suffices to show that $\ker \varphi_A = \ker \varphi_B$.

First, note that the map $f \circ \varphi_A : \mathbb{Z} \to B$ is a ring morphism. Therefore, by the uniqueness of φ_B, we must have $\varphi_B = f \circ \varphi_A$.

Therefore, for all $x \in \mathbb{Z}$,

$$\begin{aligned} x \in \ker \varphi_A &\Leftrightarrow \varphi_A(x) = 0 \\ &\Leftrightarrow f(\varphi_A(x)) = 0 \\ &\Leftrightarrow \varphi_B(x) = 0 \\ &\Leftrightarrow x \in \ker \varphi_B, \end{aligned}$$

where the equivalence in the second line is due to the fact that f is injective. □

Corollary 3.8. *1. Let K, L be two fields such that $\mathrm{char} K \neq \mathrm{char} L$. Then there does not exist any field morphism $f : K \to L$.*

2. Let A be an integral domain and $K = \mathrm{Frac}(A)$. Then $\mathrm{char} A = \mathrm{char} K$.

Proof. 1. Recall that a morphism $f : K \to L$ is a field morphism if it is a ring morphism. Assume that there exists a ring morphism $f : K \to L$. Then $\ker f \subset K$ is an ideal. But since K is a field, it has only two ideals, namely (0) and K. We know that $\ker f \neq K$ because $f(1_K) = 1_L \neq 0$, so $1_K \notin \ker f$. Therefore we conclude that $\ker f = (0)$. In other words, f is an injective ring morphism. Lemma 3.7 implies that $\mathrm{char} K = \mathrm{char} L$, which is a contradiction.

2. It is easy to check that the map

$$\begin{aligned} f : A &\to K \\ a &\mapsto \frac{a}{1} \end{aligned}$$

is an injective ring morphism. Therefore by Lemma 3.7 we obtain that $\mathrm{char} A = \mathrm{char} K$.

□

Lemma 3.9. *Let A be an integral domain.*

1. *If $\mathrm{char} A = 0$, there is an injective ring morphism $f : \mathbb{Z} \to A$.*

2. *If $\mathrm{char} A = p \neq 0$, there is an injective ring morphism $f : \mathbb{Z}/p\mathbb{Z} \to A$.*

Proof. Let $\varphi_A : \mathbb{Z} \to A$ the unique ring morphism.

1. Since $\mathrm{char} A = 0$, we have that $\ker \varphi_A = (0)$, so $f = \varphi_A$ is an injective ring morphism.

2. Since $\mathrm{char} A = p$, we have that $\ker \varphi_A = (p)$. Thus we can define the map

$$\begin{aligned} f : \mathbb{Z}/p\mathbb{Z} &\to A \\ x + p\mathbb{Z} &\mapsto \varphi_A(x) \end{aligned}$$

This map is well defined, injective, and it is a ring morphism.

□

Lemma 3.10. *Let K be a field.*

1. *If $\mathrm{char} K = 0$, then there is a field morphism $f : \mathbb{Q} \to K$.*
2. *If $\mathrm{char} K = p \neq 0$, then there is a field morphism $f : \mathbb{F}_p \to K$.*

Proof. 1. Let $\varphi_K : \mathbb{Z} \to K$ be the unique ring morphism from $\mathbb{Z}$ to K. By Lemma 3.9–(1), φ_K is injective. We define a map $f : \mathbb{Q} \to K$ by the following rule: let $x \in \mathbb{Q}$. Write $x = a/b$ for some $a, b \in \mathbb{Z}$, $b \neq 0$. Then we define $f(x) := \varphi_K(a)/\varphi_K(b)$. Note that, since φ_K is injective and $b \neq 0$, $\varphi_K(b) \neq 0$, hence it makes sense to divide by $\varphi_K(b)$. Moreover, $f(x)$ does not depend on the representation of $x = a/b$ chosen, and it is easy to show that f is a ring morphism (hence a field morphism).

2. Take $A = K$ in (2) of Lemma 3.9.

□

Remark 3.11. Let K be a field.

1. The field morphisms in Lemma 3.10 are *unique*. In other words, if K is a field of characteristic zero, there is a unique field morphism $f : \mathbb{Q} \to K$, and if K is a field of characteristic p, there is a unique field morphism $f : \mathbb{F}_p \to K$.
2. The fields $\mathbb{Q}$ and $\mathbb{F}_p$ satisfy that they contain no subfields. Moreover, they are the only fields with this property.

Since the intersection of a collection of subfields of a field K is always a field, it makes sense to speak of the “smallest subfield of K” (by which we mean the intersection of all subfields of K).

Definition 3.12. Let K be a field. We call the *prime field* of K (in French: *corps premier* de K) the smallest subfield of K.

Remark 3.13. Let K be a field of characteristic zero (resp. of characteristic $p \neq 0$), and let $f : \mathbb{Q} \to K$ (resp. $f : \mathbb{F}_p \to K$) the unique field morphism. Then the prime field of K coincides with the subfield $f(\mathbb{Q}) \subset K$ (resp. $f(\mathbb{F}_p) \subset K$).

Lemma 3.14. 1. *Let A be an integral domain of characteristic $p \neq 0$. Then the map*

$$\begin{aligned} \mathrm{Frob} : A &\to A \\ x &\mapsto x^p \end{aligned}$$

is a ring morphism.

2. *Let K be a finite field of characteristic p. Then the map* Frob *defined above is an automorphism of K (in other words, it is an isomorphism from K to K).*

Proof. 1. To show that Frob is a ring homomorphism, we have to check three properties:

(a) $\mathrm{Frob}(x + y) = \mathrm{Frob}(x) + \mathrm{Frob}(y)$ for all $x, y \in K$.

(b) $\mathrm{Frob}(xy) = \mathrm{Frob}(x)\mathrm{Frob}(y)$ for all $x, y \in K$.

(c) $\mathrm{Frob}(1) = 1$.

We will prove them one by one.

(a) Let $x, y \in K$. Then

$$\mathrm{Frob}(x+y) = (x+y)^p = \sum_{i=0}^{p} \binom{p}{i} x^i y^{p-i}.$$

Recall the definition of the symbol $\binom{p}{y}$:

$$\binom{p}{i} = \frac{p!}{i!(p-i)!}$$

where by convention $0! = 1$.
We have that $\binom{p}{0} = \binom{p}{p} = 1$, and for each $i = 1, \dots, p-1$, $\binom{p}{i} = \frac{a}{b}$ where $a, b \in \mathbb{Z}$ are such that $p|a$ and $p \nmid b$. Therefore p divides the integer $a/b = \binom{p}{i}$, say $\binom{p}{i} = c_i \cdot p$. But since K has characteristic p, by Remark 3.6 we conclude that $\binom{p}{i} \cdot 1_K = 0$. Therefore

$$\mathrm{Frob}(x+y) = \sum_{i=0}^{p} \binom{p}{i} x^i y^{p-i} = x^p + \underbrace{\sum_{i=1}^{p-1} \binom{p}{i} x^i y^{p-i}}_{=0} + y^p = x^p + y^p.$$

(b) Let $x, y \in K$. Then

$$\mathrm{Frob}(xy) = (xy)^p = x^p y^p = \mathrm{Frob}(x)\mathrm{Frob}(y).$$

(c) $\mathrm{Frob}(1) = 1^p = 1$.

2. By (1) above, $\mathrm{Frob} : K \to K$ is a field morphism. In particular, it is injective. Thus it suffices to show that Frob is surjective.

   But an injective map between two finite sets of the same cardinality is always surjective (see Proposition 3.32–(c) of [Algebra 1]. Hence we have that Frob is an automorphism of K.

□

Definition 3.15. Let A be an integral domain of characteristic $p \neq 0$. The ring morphism

$$\begin{aligned} \mathrm{Frob} : &A \to A \\ &x \mapsto x^p \end{aligned}$$

is called the *Frobenius morphism*.

Example 3.16. Let p be a prime number and $K = \mathbb{F}_p(X) := \mathrm{Frac}(\mathbb{F}_p[X]) = \{f/g : f, g \in \mathbb{F}_p[X], g \neq 0\}$.

Then $\mathrm{Frob} : K \to K$ is a field morphism but it is not surjective. For example, the element X does not belong to the image of Frob.

Proposition 3.17. *Let K be a field of characteristic $p \neq 0$. Then the prime field of K coincides with the set $\{y \in K : y^p = y\}$.*

Proof. Let $f : \mathbb{F}_p \to K$ be the unique field morphism. Recall that the prime field is the subfield $f(\mathbb{F}_p) \subset K$. We will first proof that, for all $y \in f(\mathbb{F}_p)$, $y^p = y$, and then that no other element $z \in K$ satisfies that $z^p = z$.

Let us first take $x \in \mathbb{F}_p$ a nonzero element. Then x belongs to the multiplicative group $\mathbb{F}_p^\times$, which has $p-1$ elements. Therefore we have that the order of $x \in \mathbb{F}_p$ divides $p-1$ (Lagrange's Theorem, see Theorem 3.2 of [Algebra 2]). This implies that $x^{p-1} = 1$, and thus $x^p = x$. Moreover, we also have that $0^p = 0$. Thus, for all $x \in \mathbb{F}_p$, $x^p = x$.

Let us take any $y \in f(\mathbb{F}_p)$. There exists $x \in \mathbb{F}_p$ such that $y = f(x)$. Since f is a field morphism, we obtain that

$$y^p = f(x)^p = f(x^p) = f(x) = y.$$

We have thus $f(\mathbb{F}_p) \subset \{y \in K : y^p = y\}$. To see the other inclusion, it suffices to prove that the cardinality of the set $\{y \in K : y^p = y\}$ is less than or equal to the cardinality of $f(\mathbb{F}_p)$; that is to say, less than or equal to p. But this follows immediately from the fact that the polynomial $P = X^p - X \in K[X]$ cannot have more that $\deg P = p$ roots in K (see Exercise Sheet 7). □

4 Algebraic Field Extensions

Definition 4.1. Let L/K be a field extension. Recall that L has a K-vector space structure, and denote by $\dim_K(L) \in \mathbb{N} \cup \{\infty\}$ its dimension. We define the *degree* of L/K as $[L : K] := \dim_K(L)$.

If $[L : K] < \infty$, we say that L/K is a *finite field extension*.

Remark 4.2.

1. The expression "finite field extension" can be ambiguous, since it is not clear whether the adjective "finite" is attached to "field" or to "field extension". We have defined it in the second way; that is to say, a finite field extension is a field extension which is finite, and not an extension of finite fields.

2. When we say that $K_1 \subseteq K_2 \subseteq K_3$ are field extensions, we mean that K_2/K_1 is a field extension and K_3/K_2 is a field extension. In this situation, K_3/K_1 is also a field extension ($K_1 \subseteq K_3$ is a field, and the addition and multiplication in K_1 are the restriction to K_1 of the addition and multiplication of K_3).

Example 4.3.

1. $\mathbb{C}/\mathbb{R}$ is a field extension of degree 2. Namely, $\mathbb{C}$ is isomorphic to $\mathbb{R} \times \mathbb{R}$ as $\mathbb{R}$-vector space; one isomorphism is given by $a + b\sqrt{-1} \in \mathbb{C} \mapsto (a, b) \in \mathbb{R} \times \mathbb{R}$.

2. Let K be a field. K/K is a finite field extension of degree $[K : K] = 1$.

Proposition 4.4. *Let $K \subseteq L \subseteq M$ be field extensions. Then*

$$[M : K] = [M : L] \cdot [L : K].$$

(where $\infty \cdot \infty = \infty$ and, for all $n \in \mathbb{N}$, $n \cdot \infty = \infty \cdot n = \infty$).

Proof. We have that L/K and M/L are field extensions. Then M/K is also a field extension.

- Assume that $[M : L] < \infty$ and $[L : K] < \infty$. Then

$$\begin{aligned} M &\simeq L^{[M:L]} \text{ as } L\text{-vector spaces,} \\ L &\simeq K^{[L:K]} \text{ as } K\text{-vector spaces,} \\ M &\simeq K^{[M:K]} \text{ as } K\text{-vector spaces.} \end{aligned}$$

  Therefore

$$M \simeq (K^{[L:K]})^{[M:L]} \simeq K^{[L:K]\cdot[M:L]}$$

  as K-vector spaces. But two K-vector spaces are isomorphic if and only if their dimensions coincide, so $[M : K] = [L : K] \cdot [M : L]$.

- Assume $[M : L] = \infty$. Then there exists an infinite sequence $(x_n)_{n\in\mathbb{N}}$ of different elements of M which are linearly independent over L. In particular, $(x_n)_{n\in\mathbb{N}}$ are linearly independent over K, thus $[M : K] = \infty$.

- Assume $[L : K] = \infty$. Then there exists an infinite sequence $(y_n)_{n\in\mathbb{N}}$ of different elements of L which are linearly independent over K. In particular, each $y_n \in M$, so we have that $(y_n)_{n\in\mathbb{N}}$ is an infinite sequence of different elements of M which are linearly independent over K. Thus $[M : K] = \infty$.

□

Corollary 4.5. *Let $K \subseteq L \subseteq M$ be field extensions. If $[M : K] = p$ is a prime number, then either $M = L$ or $L = K$.*

Proof. We have that $p = [M : K] = [M : L] \cdot [L : K]$, with $[M : L], [L : K] \in \mathbb{N}$. Therefore $[M : L] = 1$ or $[L : K] = 1$. That is to say, either $M = L$ or $L = K$. □

Definition 4.6. Let L/K be a field extension and $a \in L$. We define the *subfield of L generated by a over K* as

$$K(a) := \bigcap_{\substack{K\subseteq E\subseteq L \\ \text{field extensions} \\ a\in E}} E.$$

$K(a)$ is the minimal subfield of L that contains both K and a.

Definition 4.7. Let L/K be a field extension and $a \in L$. We define the map *evaluation at a* as

$$\begin{aligned} \mathrm{ev}_a : K[X] &\to L \\ \sum_{i=0}^{n} a_i X^i &\mapsto \sum_{i=0}^{n} a_i a^i. \end{aligned}$$

Remark 4.8. 1. $\mathrm{ev}_a : K[X] \to L$ is a ring morphism. Thus the image $\mathrm{ev}_a(K[X])$ is a subring of L containing K and a.

2. $\mathrm{ev}_a(K[X]) = \{f(a) : f(X) \in K[X]\}$.

3. $\mathrm{ev}_a(K[X]))$ is the smallest subring of L containing K and a. In other words, if $B \subseteq L$ is a subring containing K and a, then $\mathrm{ev}_a(K[X]) \subseteq B$.

Definition 4.9. Let L/K be a field extension and $a \in L$. We denote $K[a] := \mathrm{ev}_a(K[X])$.

Remark 4.10. Let L/K be a field extension and $a \in L$.

1. Since $K(a) \subseteq L$ is a subfield containing K and a, we have that
$$K[a] \subseteq K(a). \tag{4.3}$$

2. From Equation (4.3), we obtain that $\mathrm{Frac}(K[a]) \subseteq K(a)$.

3. $\mathrm{Frac}(K[a]) \subseteq L$ is a subfield containing a and K. Thus $K(a) \subseteq \mathrm{Frac}(K[a])$.

4. From (3) and (4) above we can conclude that
$$K(a) = \mathrm{Frac}(K[a]) = \left\{ \frac{f(a)}{g(a)} : f(X), g(X) \in K[X], g(a) \neq 0 \right\}.$$

Definition-Lemma 4.11. Let L/K be a field extension, $a_1, \dots, a_n \in L$.

1. The *evaluation map* defined as
$$\begin{aligned} \mathrm{ev}_{a_1,\dots,a_n} : K[X_1, \dots, X_n] &\to L \\ f(X_1, \dots, X_n) &\mapsto f(a_1, \dots, a_n) \end{aligned}$$
is a ring morphism. The image of $\mathrm{ev}_{a_1,\dots,a_n}$, denoted as $K[a_1, \dots, a_n]$, is the smallest subring of L containing K and the set $\{a_1, \dots, a_n\}$.

2. We define the *subfield of L generated by* $a_1, \dots, a_n$ *over* K as
$$K(a_1, \dots, a_n) := \bigcap_{\substack{K \subseteq E \subseteq L \\ \text{field extensions} \\ a_1,\dots,a_n \in E}} E.$$
$K(a_1, \dots, a_n)$ is the minimal subfield of L that contains both K and the set $\{a_1, \dots, a_n\}$.

Example 4.12. 1. $K = \mathbb{Q}$, $L = \mathbb{R}$, $\sqrt{2} \in L$. $\mathbb{Q}[\sqrt{2}] = \{f(\sqrt{2}) : f(X) \in \mathbb{Q}[X]\}$.

- We claim that
$$\mathbb{Q}[\sqrt{2}] = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}.$$
The inclusion $\supseteq$ is easy: it suffices to take $f(X) = a + bX \in \mathbb{Q}[X]$. Reciprocally, let $f(X) = \sum_{i=0}^n a_i X^i \in \mathbb{Q}[X]$; we have to prove that $f(\sqrt{2}) = a + b\sqrt{2}$ for some $a, b \in \mathbb{Q}$. Write
$$\begin{aligned} f(\sqrt{2}) &= \sum_{i=0}^n a_i(\sqrt{2})^i = \sum_{\substack{0 \le i \le n \\ i \text{even}}} a_i(\sqrt{2})^i + \sum_{\substack{0 \le i \le n \\ i \text{odd}}} a_i(\sqrt{2})^i \\ &= \sum_{\substack{0 \le i \le n \\ i \text{even}}} a_i(\sqrt{2})^i + \sqrt{2} \sum_{\substack{0 \le i \le n \\ i \text{odd}}} a_i(\sqrt{2})^{i-1} = \underbrace{\sum_{\substack{0 \le i \le n \\ i \text{even}}} a_i 2^{i/2}}_{\in \mathbb{Q}} + \sqrt{2} \underbrace{\sum_{\substack{0 \le i \le n \\ i \text{odd}}} a_i 2^{(i-1)/2}}_{\in \mathbb{Q}}. \end{aligned}$$

- By definition, we have that $\mathbb{Q}[\sqrt{2}] \subseteq \mathbb{Q}(\sqrt{2})$. We claim that $\mathbb{Q}[\sqrt{2}] = \mathbb{Q}(\sqrt{2})$. Indeed, it suffices to show that $\mathbb{Q}[\sqrt{2}]$ is a field. Let us take $a + b\sqrt{2} \in \mathbb{Q}[\sqrt{2}]$ nonzero; we have to prove that $a + b\sqrt{2}$ is invertible in $\mathbb{Q}[\sqrt{2}]$. That is to say, we have to find $c, d \in \mathbb{Q}$ such that
$$(a + b\sqrt{2}) \cdot (c + d\sqrt{2}) = 1. \tag{4.4}$$
Note that $c := a/(a^2 - 2b^2)$ and $d := -b/(a^2 - 2b^2)$ (which are well defined because $a^2 - 2b^2$ cannot be zero, since 2 is not a square in $\mathbb{Q}$) satisfy Equation (4.4).

2. Let $K = \mathbb{Q}$, $L = \mathbb{R}$, $\pi \in \mathbb{R}$. Then $\mathbb{Q}[\pi] \neq \mathbb{Q}(\pi)$. Indeed, if $\mathbb{Q}[\pi]$ were a field, then π would be invertible in $\mathbb{Q}[\pi]$. That is to say, there would exist $g(X) = \sum_{i=0}^{n} b_i X^i \in \mathbb{Q}[X]$ such that $\pi \cdot g(\pi) = 1$. Then
$$b_n\pi^{n+1} + b_{n-1}\pi^n + b_{n-2}\pi^{n-1} + \cdots + b_0\pi - 1 = 0.$$
But Lindemann has proven that π does not satisfy any such equation!

Remark 4.13. Let L/K be a field extension and $a \in L^\times$. Then if $K[a] = K(a)$, there exists some polynomial $f(X) \in K[X]$ with $f(a) = 0$.

Indeed, if $K[a] = K(a)$, in particular $K[a]$ is a field. Thus a must be invertible; that is to say, there exists $g(X) \in K[X]$ with $a \cdot g(a) = 1$. Take $f(X) := Xg(X) - 1 \in K[X]$; we have that $f(a) = 0$.

The converse is also true; this is shown in Remark 4.18.

Definition 4.14. Let L/K be a field extension and $a \in L$. We say that a is *algebraic over* K if there exists a nonzero polynomial $f(X) \in K[X]$ such that $f(a) = 0$. We say that a is *trascendental over* K if a is not algebraic over K.

Example 4.15. 1. $\sqrt{2} \in \mathbb{R}$ is algebraic over $\mathbb{Q}$, since $f(X) = X^2 - 2 \in \mathbb{Q}[X]$ satisfies $f(\sqrt{2}) = 0$.

2. $\pi \in \mathbb{R}$ is trascendental over $\mathbb{Q}$.

3. $\pi \in \mathbb{R}$ is algebraic over $\mathbb{R}$, since $f(X) = X - \pi \in \mathbb{R}[X]$ satisfies $f(\pi) = 0$.

Remark 4.16. Let L/K be a field extension, $a \in L$. Then a is algebraic over K if and only if the set $\{1, a, a^2, a^3, \dots\}$ is linearly dependent over K. Equivalently, a is trascendental over K if and only if the set $\{1, a, a^2, a^3, \dots\}$ is linearly independent over K (in particular, $K[a]$ has infinite dimension as a K-vector space).

Proposition 4.17. *Let L/K be a field extension and $a \in L$.*

1. $\mathrm{ev}_a : K[X] \to L$ *is injective if and only if a is trascendental over K.*

For the rest of the proposition, assume that a is algebraic over K.

2. *There exists a unique monic polynomial* $\mathrm{minpoly}_a(X) \in K[X]$ *such that*
$$\ker \mathrm{ev}_a = (\mathrm{minpoly}_a(X)).$$
We call $\mathrm{minpoly}_a(X)$ *the* minimal polynomial of a over K.

3. $\mathrm{minpoly}_a(X)$ *is an irreducible element of* $K[X]$.

4. *The map*
$$\begin{aligned} \overline{\mathrm{ev}}_a : K[X]/\ker \mathrm{ev}_a &\to L \\ f(X) + \ker \mathrm{ev}_a &\mapsto f(a) \end{aligned} \tag{4.5}$$
induced by ev_a *is a field morphism. The image of* ev_a *is* $K[a]$ *(thus* $K[a]$ *is a field).*

Proof. 1. We will show that $\mathrm{ev}_a : K[X] \to L$ is not injective if and only if a is algebraic over K. Indeed:
$$\begin{aligned} \mathrm{ev}_a\text{is not injective } &\Leftrightarrow \ker \mathrm{ev}_a \neq (0) \\ &\Leftrightarrow \exists f(X) \in K[X] \setminus \{0\} \text{ such that } f(X) \in \ker \mathrm{ev}_a \\ &\Leftrightarrow \exists f(X) \in K[X] \setminus \{0\} \text{ such that } f(a) = 0. \end{aligned}$$

2. $K[X]$ is a principal ideal domain, thus the ideal $\ker \mathrm{ev}_a$ is a principal ideal, say $\ker \mathrm{ev}_a = (f(X))$ for $f(X) \in K[X]$. Since a is algebraic over K, $\ker \mathrm{ev}_a \neq (0)$, thus $f \neq 0$. Therefore we can write $f(X) = \sum_{i=0}^{n} a_i X^i$ with $a_n \neq 0$. The polynomial $\widetilde{f} = \frac{1}{a_n} f \in K[X]$ is monic, and $(f(X)) = (\widetilde{f}(X))$.

   To see the unicity, assume there exists $g(X)$ monic with $(g(X)) = (\widetilde{f}(X))$. Then g and $\widetilde{f}$ are associated; there exists a unit $u \in (K[X])^\times = K^\times$ such that $g = u \cdot \widetilde{f}$. Comparing the leading coefficients of g and $u \cdot \widetilde{f}$, we conclude that $u = 1$.

3. We have that $\ker \mathrm{ev}_a = (\mathrm{minpoly}_a(X))$. Therefore for all $f(X) \in K[X]$, it holds that
$$f(a) = 0 \Leftrightarrow \mathrm{minpoly}_a(X) | f(X). \tag{4.6}$$
   Assume that $\mathrm{minpoly}_a(X)$ is not irreducible, say $\mathrm{minpoly}_a(X) = g_1(X)g_2(X)$ with $g_1(X) \notin (K[X])^\times$ and $g_2(X) \notin (K[X])^\times$. Then
$$g_1(a) \cdot g_2(a) = \mathrm{minpoly}_a(a) = 0,$$
   thus $g_1(a) = 0$ or $g_2(a) = 0$. By Equation (4.6), we have that either $\mathrm{minpoly}_a(X)|g_1(X)$ or $\mathrm{minpoly}_a(X)|g_2(X)$. But clearly $g_1(X)|\mathrm{minpoly}_a(X)$ and $g_2(X)|\mathrm{minpoly}_a(X)$, so either $g_1(X) \sim \mathrm{minpoly}_a(X)$ or $g_2(X) \sim \mathrm{minpoly}_a(X)$. In other words, either $g_2(X)$ is a unit or $g_1(X)$ is a unit; we have a contradiction.

4. By the First Isomorphism Theorem, we know that the map
$$\begin{aligned} \overline{\mathrm{ev}}_a : K[X]/\ker \mathrm{ev}_a &\to L \\ f(X) + \ker \mathrm{ev}_a &\mapsto f(a) \end{aligned} \tag{4.7}$$
   is an injective ring morphism.

   Since $K[X]$ is a principal ideal domain and $\mathrm{minpoly}_a(X)$ is irreducible, we know (cf. Proposition 7.4 of [Algebra 2]) that $(\mathrm{minpoly}_a(X)) = \ker \mathrm{ev}_a$ is a maximal ideal of $K[X]$. Therefore (cf. Proposition 7.6 of [Algebra 2]) $K[X]/\ker \mathrm{ev}_a$ is a field.

It remains to compute the image of $\overline{\mathrm{ev}}_a$. But this image coincides with the image of ev_a, which is by definition $K[a]$.

□

Remark 4.18. Let L/K be a field extension, $a \in L$. Then a is algebraic over K if and only if $K[a]$ is a field (equivalently, if and only if $K[a] = K(a)$).

Proposition 4.19. *Let L/K be a field extension and $a \in L$ be algebraic over K. Then $K(a)/K$ is a finite field extension and $[K(a) : K] = \deg(\mathrm{minpoly}_a(X))$.*

Call $d := \deg(\mathrm{minpoly}_a(X))$. Then the set $\{1, a, a^2, \dots, a^{d-1}\}$ is a basis of $K(a)$ as K-vector space.

Proof. It suffices to prove the last assertion, namely that $\{1, a, a^2, \dots, a^{d-1}\}$ is a basis of $K(a)$ as K-vector space.

- First we note that $1, a, a^2, \dots, a^{d-1}$ are linearly independent over K. Indeed, if they were not linearly independent, there would exist $b_0, \dots, b_{d-1}$, not all zero, such that $\sum_{i=0}^{d-1} b_i a^i = 0$. Call $g(X) = \sum_{i=0}^{d-1} b_i X^i = 0$. Then $\mathrm{minpoly}_a(X)|g(X)$, but the degree of $g(X)$ is strictly less than $d = \mathrm{degminpoly}_a(X)$. This is a contradiction.

- Now we want to prove that $1, a, a^2, \dots, a^{d-1}$ generate $K(a)$ as K-vector space. Since a is algebraic, $K(a) = K[a]$, so it suffices to see that $1, a, a^2, \dots, a^{d-1}$ generate $K[a]$ as K-vector space.

  Take any $f(X) \in K[X]$. We will show that there exist $b_0, \dots, b_{d-1} \in K$ such that $f(a) = b_0 \cdot 1 + b_1 \cdot a + b_2 \cdot a^2 + \dots + b_{d-1} \cdot a^{d-1}$. Indeed, divide $f(X)$ by $\mathrm{minpoly}_a(X)$ in the Euclidean ring $K[X]$; there exist $g(X), h(X) \in K[X]$, with $\deg h(X) < \deg \mathrm{minpoly}_a(X) = d$ such that
  $$f(X) = g(X) \cdot \mathrm{minpoly}_a(X) + h(X).$$
  Write $h(X) = \sum_{i=0}^{d-1} b_i X^i$ for some $b_i \in K$. Then
  $$f(a) = g(a) \cdot \underbrace{\mathrm{minpoly}_a(a)}_{=0} + h(a) = h(a) = \sum_{i=0}^{d-1} b_i a^i.$$

□

Example 4.20. 1. Let K be a field, $a \in K$. Then a is algebraic over K. $\mathrm{minpoly}_a(X) = X - a \in K[X]$.

2. Let $K = \mathbb{Q}$, $L = \mathbb{R}$, $a = \sqrt{2}$. Then $\sqrt{2}$ is algebraic over $\mathbb{Q}$. $\mathrm{minpoly}_{\sqrt{2}}(X) = X^2 - 2 \in \mathbb{Q}[X]$. (note that $X - \sqrt{2} \notin \mathbb{Q}[X]$).

3. Let $K = \mathbb{Q}$, $\pi \in \mathbb{R}$. Then π is trascendental over $\mathbb{Q}$.

Definition 4.21. Let L/K be a field extension, $f(X) \in K[X]$ a nonzero, irreducible polynomial. We say that L is a *rupture field* of $f(X)$ over K if there exists $a \in L$ such that

1. $f(a) = 0$.

2. $L = K(a)$.

Example 4.22. Let $f(X) = X^3 - 2 \in \mathbb{Q}[X]$. We can view $f(X) \in \mathbb{Q}[X] \subset \mathbb{R}[X] \subset \mathbb{C}[X]$. Let α be the only real root of $f(X)$, and $\zeta_3 = e^{2\pi i/3} \in \mathbb{C}$. Then the roots of $f(X)$ in $\mathbb{C}$ are $\alpha_1 = \alpha$, $\alpha_2 = \zeta_3\alpha$, $\alpha_3 = \zeta_3^2\alpha$. For each $i = 1, 2, 3$, let $L_i = \mathbb{Q}(\alpha_i) \subset \mathbb{C}$. Then each of the L_i is a rupture field of $f(X)$ over $\mathbb{Q}$. However, they are not all the same. Namely, $L_1 \subset \mathbb{R}$ but $L_2 \not\subset \mathbb{R}$, thus $L_1 \neq L_2$.

Proposition 4.23. *Let K be a field and $f(X) \in K[X]$ be a non-zero, irreducible polynomial. There exists a field extension L/K such that L is a rupture field over K.*

Proof. Let $L := K[X]/(f(X))$. Since $K[X]$ is a principal ideal domain and $f(X) \in K[X]$ is an irreducible, non-zero polynomial, the ideal $(f(X)) \subset K[X]$ is maximal. Therefore $L = K[X]/(f(X))$ is a field.

The map

$$\begin{aligned} j : K &\to K[X]/(f(X)) \\ a &\mapsto a + (f(X)) \end{aligned}$$

is an injective ring morphism, so that $j(K) \subset L$ is a subfield isomorphic to K. We identify K with the image of j in L.

Let $\alpha = X + (f(X)) \in L$. To prove that L is a rupture field of $f(X)$ over K, we will prove that $f(\alpha) = 0$ and $L = K(\alpha)$.

1. $f(\alpha) = f(X + (f(X))) = f(X) + (f(X)) = 0 + (f(X)) \in L$.

2. The inclusion $K(\alpha) \subset L$ is clear because $K \subset L$ and $\alpha \in L$. To see the other inclusion, let us take some $y \in L = K[X]/(f(X))$. Then there exists $g(X) = \sum_{i=0}^n b_i X^i \in K[X]$ such that $y = g(X) + (f(X))$. Then

$$y = g(X) + (f(X)) = \sum_{i=0}^n b_i X^i + (f(X)) =$$
$$\sum_{i=0}^n b_i (X + (f(X)))^i = \sum_{i=0}^n b_i \alpha^i \in K[\alpha] \subseteq K(\alpha).$$

□

Definition 4.24. Let L/K be a field extension.

1. We say that L/K is *algebraic* (equivalently, we say that L is an *algebraic extension* of K) if, for all $a \in L$, a is algebraic over K.

2. We say that L/K is *trascendental* if it is not algebraic.

Proposition 4.25. *Let L/K be a finite extension of fields. Then L/K is algebraic. Moreover, L can be generated over K (as a field) by a finite set of elements (which are algebraic over K).*

Proof. Let us pick any $a \in L$. We have to prove that a is algebraic over K.

Consider the ring $K[a] \subset L$. In fact, $K[a]$ is a sub-K-vector space of L. Therefore, we have that $\dim_K(K[a]) \leq \dim_K(L)$, and $\dim_K(L) < \infty$ by hypothesis. Thus the dimension of $K[a]$ as K-vector space is finite. Therefore, the set $\{1, a, a^2, \ldots, a^n, \ldots\}$ must be linearly dependent over K. By Remark 4.16, this implies that a is algebraic over K.

Now we want to prove that there exist a finite set of elements $\{a_1, \ldots, a_n\} \subset L$ such that $L = K(a_1, \ldots, a_n)$. Let $d = \dim_K L < \infty$. If $d = 1$, then $L = K$ and there is nothing to prove. Assume $d > 1$, so that $L \neq K$.

- Choose some $a_1 \in L \setminus K$. Then we have the inclusions $K \subset K(a_1) \subset L$, and $d = [L : K] = [L : K(a_1)] \cdot [K(a_1) : K]$. Since $K(a_1) \neq K$, we know that $[K(a_1) : K] > 1$, thus $[L : K(a_1)] < d$. If $[L : K(a_1)] = 1$, then $L = K(a_1)$ and we are done. If not, then $1 < [L : K(a_1)] < d$.
- Choose some $a_2 \in L \setminus K(a_1)$. Then we have again $K(a_1) \subset K(a_1, a_2) \subset L$, and $[L : K(a_1)] = [L : K(a_1, a_2)] \cdot [K(a_1, a_2) : K(a_1)]$. Since $K(a_1, a_2) \neq K(a_1)$, we know that $[K(a_1, a_2) : K(a_1)] > 1$, thus $[L : K(a_1, a_2)] < [L : K(a_1)]$. If $[L : K(a_1, a_2)] = 1$, then $L = K(a_1, a_2)$ and we are done. If not, then $1 < [L : K(a_1, a_2)] < [L : K(a_1)] < d$.
- ...

It is clear that this process will finish in at most d steps, producing a finite set $\{a_1, \ldots, a_n\} \subset L$ (with $n \leq d$) such that $L = K(a_1, a_2, \ldots, a_n)$. □

Proposition 4.26. *Let L/K be a field extension and let $a_1, \ldots a_n \in L$. Then the following are equivalent:*

(i) For all $i = 1, \ldots, n$, a_i is algebraic over K.

(ii) $K(a_1, \ldots, a_n)/K$ is a finite extension of fields.

Proof. See Exercise Sheet 11. □

Proposition 4.27. *Let $K \subseteq L \subseteq M$ be field extensions.*

1. *Let $a \in M$ be algebraic over L, and assume that L/K is algebraic. Then a is algebraic over K.*
2. *M/K is algebraic if and only if both L/K and M/L are algebraic.*

Proof. 1. Since $a \in M$ is algebraic over L, we can consider the minimal polynomial of a over L, say

$$\text{minpoly}_a(X) = X^n + a_{n-1}X^{n-1} + \cdots + a_1 X + a_0 \in L[X].$$

To prove that a is algebraic over K, it suffices to prove that the field extension

$$K(a_0, \ldots, a_{n-1}, a)/K$$

is finite. Indeed, if this is the case, then by Proposition 4.25, we can conclude that the extension $K(a_0, \dots, a_{n-1}, a)/K$ is algebraic. That is to say, all the elements in $K(a_0, \dots, a_{n-1}, a)$ are algebraic over K; in particular, a is algebraic over K.

Call $N = K(a_0, \dots, a_{n-1})$. Since L/K is algebraic, then for all $i = 0, \dots n-1$, a_i is algebraic over K. Thus by Proposition 4.26, the field extension N/K is a finite extension of fields. Note, moreover, that a satisfies an algebraic equation with coefficients in N, namely $a^n + a_{n-1}a^{n-1} + \cdots + a_1X + a_0 = 0$, thus a is algebraic over N. Therefore, by Proposition 4.26 the field extension $N(a)/N$ is finite. Thus

$$[K(a_0, \dots, a_{n-1}, a)/K] = [N(a)/K] = [N(a) : N] \cdot [N : K] < \infty.$$

2. $\Rightarrow$ Assume first that M/K is algebraic. To prove that M/L is algebraic, it suffices to show that, for all $a \in M$, there exists a polynomial $f(X) \in L[X]$ with $f(a) = 0$. But, since M/K is algebraic, we know that there exists a polynomial $g(X) \in K[X]$ such that $g(a) = 0$; then it suffices to take $f(X) = g(X)$. To prove that L/K is algebraic, it suffices to prove that every $a \in L$ is algebraic. Take $a \in L$; then $a \in M$, and since M/K is algebraic, then a is algebraic over K.

   $\Leftarrow$ Assume that M/L and L/K are both algebraic field extensions. To prove that M is algebraic over K, it suffices to see that, for all $a \in M$, a is algebraic over K. Take $a \in M$. Then a satisfies the hypothesis of Part 1 of this proposition, hence a is algebraic over M, as we wanted to prove.

□

Definition 4.28. Let L/K be a field extension and $a_1, \dots, a_n \in L$.

1. We say that $a_1, \dots, a_n$ are *algebraically independent* over K if

$$\begin{aligned} \mathrm{ev}_{a_1,\dots,a_n} : K[X_1, \dots, X_n] &\to L \\ f(x_1, \dots, X_n) &\mapsto f(a_1, \dots, a_n) \end{aligned}$$

   is injective.

2. We say that $a_1, \dots, a_n$ are *algebraically dependent* over K if they are not algebraically independent over K.

Remark 4.29. Let L/K be a field extension, and let $a_1, \dots, a_n$ be algebraically independent over K. Then $a_1, \dots, a_n$ are linearly independent over K.

Example 4.30. Consider the field extension $\mathbb{R}/\mathbb{Q}$.

1. π, π^2 are not algebraically independent over $\mathbb{Q}$. Ideed, the polynomial $f(X) = X^2 - X \in \mathbb{Q}[X]$ satisfies that

$$\mathrm{ev}_{\pi,\pi^2}(f(X)) = \pi^2 - \pi^2 = 0.$$

2. It is not known if π and $e = \lim_{n\to\infty}(1 + 1/n)^n$ are algebraically independent over $\mathbb{Q}$.

5 Constructions with ruler and compass

In this section we turn to a different topic, namely plane geometry. This topic has a very long history, since everyday-life situations have posed problems in plane geometry throughout the history of mankind. Around 300 B.C. the study of geometry and basic arithmetic had acquired an essence of their own, independently of their applications, and had been developed as a field of study. The monumental treatise of Euclid, the *Elements*, develops the whole theory from a series of axiom, and constitutes the first mathematics textbook in history.

The first book starts with a list of definitions, where the objects that are going to be studied are introduced, like *points*, *lines*, *circles*, and so on.

Then the *construction axioms* are introduced. From a given set of objects, we will be able to construct other objects by using two tools:

1. **Ruler:** Given two points P_1 and P_2, construct the line that passes through P_1 and P_2.
2. **Compass:** Given two points P_1 and P_2, construct the circle with center P_1 and radius equal to the segment with ends P_1 and P_2.

Their interpretation in terms of geometric tools is clear: those are the constructions that we can do using a ruler and a compass. To be more precise; a ruler which has no mark in it and a compass with a spike in one side and a pencil in the other. Given two points P_1 and P_2, we can fix the spike in P_1 and open the compass to situate the pencil in P_2, and then draw the circle. But, the moment we lift the compass from the plane, it closes (so that it does not remember the distance between P_1 and P_2).

These two axioms allow us to construct lines and circles. There is also a way to construct points, namely:

3. **Intersection:**
   (a) Given two lines L_1 and L_2, construct the point P where they meet (if they meet).
   (b) Given a line L_1 and a circle C_1, construct the set of points where they meet (if they meet).
   (c) Given two circles C_1 and C_2, construct the set of points where they meet (if they meet).

We will say that a geometric object *can be constructed with ruler and compass* if it can be constructed using the three rules ruler, compass and intersection above. The question arises as to which constructions can be carried out with ruler and compass. The three classical problems, which defied the mathematicians during many centuries, are the following:

- **Trisecting the angle:** Given two lines L_1 and L_2 meeting at a point P, construct a third line, L_3, passing through P and such that the angle between L_1 and L_3 is one third of the angle between L_1 and L_2.
- **Squaring the circle:** Given a circle, construct a square with the same area as the circle.
- **Duplicating the cube:** Given a segment $\overline{P_1P_2}$, construct another segment $\overline{P_1P_3}$ such that the volume of the cube with side $\overline{P_1P_3}$ is the double of the volume of the cube with side $\overline{P_1P_2}$.

We will see that the theory of field extensions that we have developed so far allows us to prove that these three constructions cannot be carried out by ruler and compass alone. The first step is to prove that some constructions can be done by using ruler and compass. Let us introduce some notation.

Definition 5.1. Let P, Q be points, L_1, L_2 be lines, C_1, C_2 be circles.

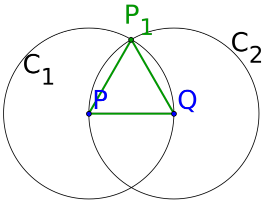
1. We denote by $L(P, Q)$ the line passing through P and Q constructed with the ruler.
2. We denote $\overline{PQ}$ the segment with end points P and Q.
3. We denote by $C(P, \overline{PQ})$ the circle with centre P and radius $\overline{PQ}$ constructed with the compass.
4. We denote by $L_1 \cap L_2$ the intersection point of L_1 and L_2 (if L_1 and L_2 meet).
5. We denote by $L_1 \cap C_1$ the set of intersection points of L_1 and C_1 (if L_1 and C_1 meet).
6. We denote by $C_1 \cap C_2$ the set of intersection points of C_1 and C_2 (if C_1 and C_2 meet).

Proposition 5.2. *The following constructions can be carried out by ruler and compass.*

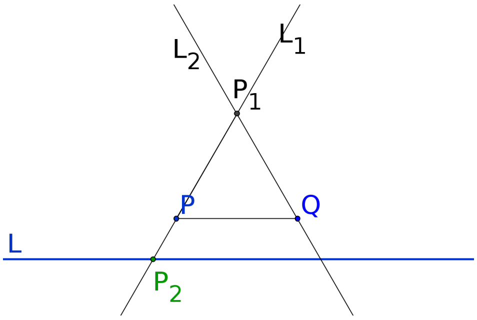
1. *Given two points $P \neq Q$, construct an equilateral triangle with side the segment $\overline{PQ}$.*
2. *Given a line L and two points $P \neq Q$, construct a point in L.*
3. *Given two points $P \neq Q$, draw the perpendicular bisector of the segment $\overline{PQ}$.*
4. *Given a line L and two points $P \neq Q$ such that $P \in L$, construct the line which passes through P and is perpendicular to L.*
5. *Given a line L and two points $P \neq Q$, such that $P \notin L$, construct the line which passes through P and is perpendicular to L.*
6. *Given a line L and two points $P \neq Q$, such that $P \notin L$, construct the line which passes through P and is parallel to L.*
7. *Given three points P, Q, R which are not alineated, and a line L passing through R, construct a point on L whose distance to R equals the length of the segment $\overline{PQ}$.*
8. *Given three points P, Q, R which lie on a line L, construct a point on L whose distance to R equals the length of the segment $\overline{PQ}$.*

Proof.

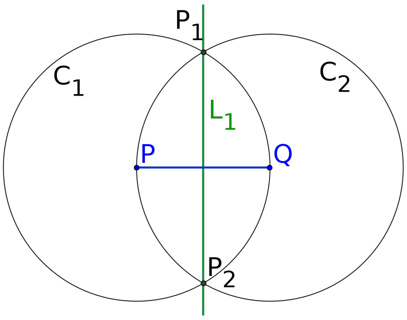
1. First we draw the circles $C_1 = C(P, \overline{PQ})$ and $C_2 = C(Q, \overline{QP})$. Let P_1 be one of the points in the intersection of C_1 and C_2. Then the points P, Q, P_1 are the vertices of an equilateral triangle.



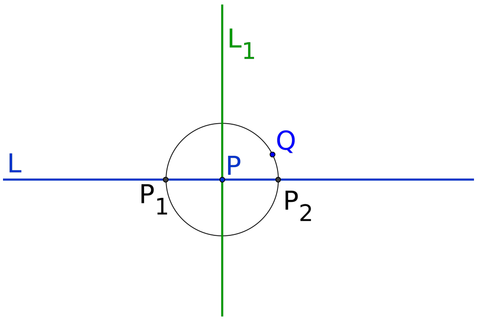
2. Using (1), we can construct a point P_1 such that the triangle P, Q, P_1 is equilateral. Then, since the two lines $L_1 = L(P_1, P)$ and $L_2 = L(P_1, Q)$ are not parallel, one of them at least must intersect with L (Here we are using that there is at most one line passing through P_1 and parallel to L). So either L_1 or L_2 intersects L. Choose the point P_2 as either the point of intersection of L and L_1 (if they meet) or L and L_2 (if they meet). P_2 is the point we wanted to construct.



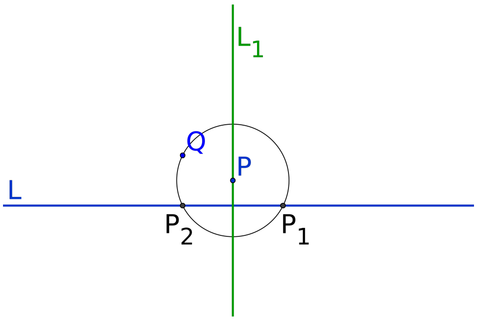
3. Draw the circles $C_1 = C(P, \overline{PQ})$ and $C_2 = C(Q, \overline{QP})$, and let P_1 and P_2 the points where they intersect. The line L_1 passing through P_1 and P_2 is the line we wanted to construct.



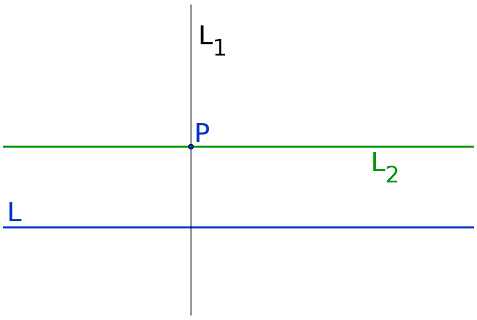
4. The circle $C(P, \overline{PQ})$ intersects L in two points, P_1 and P_2. The line L_1 constructed using (3) as the perpendicular bisector of the segment $\overline{P_1P_2}$ is the line we wanted to construct.



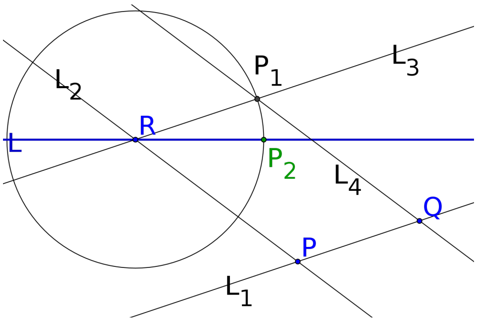
5. Using (2), we can construct a point P_1 on the line L. The circle $C(P, \overline{PP_1})$ intersect L in a point different from P_1; call it P_2. The line L_1 constructed using (3) as the perpendicular bisector of the segment P_1P_2 is the line we wanted to construct.



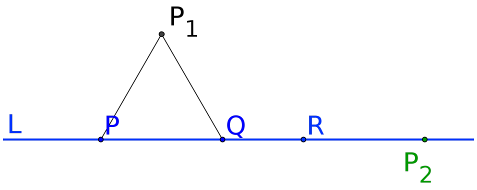
6. Using (4), we can draw the line L_1 which passes through P and is perpendicular to L. The line L_2 constructed by using (5) as the line which passes through P and is perpendicular to L_1 is the line that we wanted to construct.



7. Let $L_1 = L(P, Q)$ and $L_2(P, R)$. Using (6), draw the line L_3 which passes through R and is parallel to L_1, and the line L_4 which passes through Q and is parallel to L_2. The lines L_3 and L_4 meet in a point P_1. The circle $C_1 = C(R, \overline{RP_1})$ intersects L in two points; choose one of them, say P_2. P_2 is the point that we wanted to construct.



8. Using (1), draw an equilateral triangle with side $\overline{PQ}$. Let P_1 be the other vertex of this triangle. Then we can apply (7) to P_1, Q, R (which are not alineated) and L, to find a point P_2 in L such that the segment $\overline{RP_2}$ has the same length as the segment $\overline{PQ}$. P_2 is the point we wanted to construct.



□

We are going to be interested in the points in $\mathbb{R} \times \mathbb{R}$ that can be constructed using ruler and compass from some given subset of points. We formalise this notion in the following definition.

Definition 5.3. Let $S \subset \mathbb{R} \times \mathbb{R}$ be a subset. We say that a point $P \in \mathbb{R} \times \mathbb{R}$ is *constructible in one step from* S, if one of the following holds:

- $P \in S$.
- P is the intersection point of two different lines L_1 and L_2 constructed with the ruler by joining two points of S.
- P belongs to the intersection of a line L_1 constructed with the ruler by joining two points of S and a circle C_1 constructed with the compass using two points of S.
- P belongs to the intersection point of two different circles C_1 and C_2 constructed with the compass using two points of S.

Denote by $\mathcal{X}_1(S)$ the set of points $P \in \mathbb{R} \times \mathbb{R}$ which are constructible in one step from S, and set $\mathcal{X}_0(S) := S$.

Definition 5.4. Let $S \subset \mathbb{R} \times \mathbb{R}$ be a subset. For any $n \in \mathbb{N}$, we define $\mathcal{X}_{n+1}(S)$ as the subset of elements of $\mathbb{R} \times \mathbb{R}$ that are constructible in one step from $\mathcal{X}_n(S)$. Denote by

$$\mathcal{X}(S) := \bigcup_{n \geq 0} \mathcal{X}_n(S) \subset \mathbb{R} \times \mathbb{R}.$$

$\mathcal{X}(S)$ is the subset of elements of $\mathbb{R} \times \mathbb{R}$ that are constructible from S in a finite number of steps.

Remark 5.5. Note that, if $S \subset \mathbb{R} \times \mathbb{R}$ is finite, then $\mathcal{X}_n(S)$ is also finite for each $n \in \mathbb{N}$. Indeed, it suffices to prove the case $n = 1$. And, if we have a finite number of points, there is only a finite number of lines we can construct with the ruler which pass through two such points, and a finite number of circles with centre in one such points and passing through another such point. These finite number of lines and circles intersect only in a finite number of points. This proves that $\mathcal{X}_1(S)$ is a finite set.

We are interested in certain subfields of $\mathbb{R}$ obtained from constructible points. More precisely, we make the following definitions:

Definition 5.6. Let $S \subset \mathbb{R} \times \mathbb{R}$. For each $n \in \mathbb{N}$, consider the set

$$T_n := \bigcup_{P=(x,y) \in \mathcal{X}_n(S)} \{x, y\} \subset \mathbb{R}.$$

We define the field $\mathcal{K}_n(S)$ as the subfield of $\mathbb{R}$ generated over $\mathbb{Q}$ by the set T_n, and

$$\mathcal{K}(S) = \mathbb{Q}\left(\bigcup_{n \geq 0} T_n\right) \subset \mathbb{R}.$$

Definition 5.7. We say that a real number $a \geq 0$ is *constructible* if there exist two points P_1 and P_2, constructible from $S_0 = \{(0,0),(0,1)\} \subset \mathbb{R} \times \mathbb{R}$ in a finite number of steps, such that the length of the segment $\overline{P_1P_2}$ equals a.

Lemma 5.8. *Let $S_0 = \{(0,0),(0,1)\}$, and let $a \in \mathbb{R}_{\geq 0}$. The following are equivalent:*

(i) a is constructible.

(ii) The point $P = (0,a)$ belongs to $\mathcal{X}(S_0)$.

(iii) The point $P = (a,0)$ belongs to $\mathcal{X}(S_0)$.

Remark 5.9. Let $S_0 = \{(0,0),(0,1)\}$ and let a be constructible. As as consequence of Lemma 5.8, we obtain that $a \in \mathcal{K}(S_0)$.

Lemma 5.10. *Let $S_0 = \{(0,0),(0,1)\}$, $P = (x,y) \in \mathbb{R} \times \mathbb{R}$. The following are equivalent:*

(i) $|x|, |y|$ are constructible numbers

(ii) $P \in \mathcal{X}(S_0)$.

Proof. See Exercise 3-(b) from Exercise Sheet 12. □

The next proposition relates the notion of constructible real numbers with field theory.

Proposition 5.11. *Let $K = \{a \in \mathbb{R} : |a|$ is constructible$\}$. Then K is a subfield of $\mathbb{R}$.*

Proof. To prove that K is a subfield of $\mathbb{R}$, we have to show the following properties:

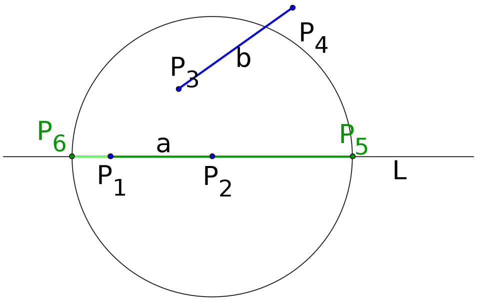
1. $0, 1 \in K$.
2. For all $a, b \in K$, $a + b$.
3. For all $a \in K$, $-a \in K$.
4. For all $a, b \in K$, $a \cdot b \in K$.
5. For all $a \in K^\times$, $a^{-1} \in K$.

Let $S_0 = \{(0,0),(0,1)\}$.

1. Since $P = (0,0)$ and $Q = (0,1)$ are constructible from S_0 in a finite (zero) number of steps, then the lengths of the segment $\overline{PP}$ and $\overline{PQ}$ (0 and 1 resp.) are constructible real numbers.

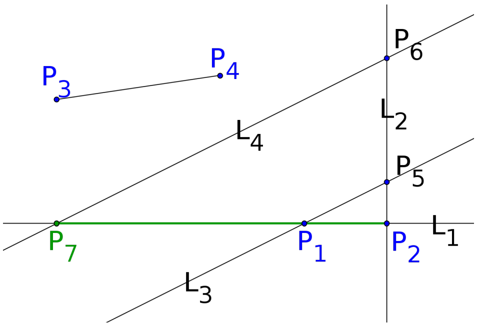
2. Let a, b be constructible. We can assume without loss of generality that both are nonzero, and $|b| \geq |a|$. Note that

$$|a+b| = \begin{cases} |a| + |b| \text{ if } a \text{ and } b \text{ have the same sign,} \\ ||a| - |b|| \text{ if } a \text{ and } b \text{ have opposite signs.} \end{cases}$$

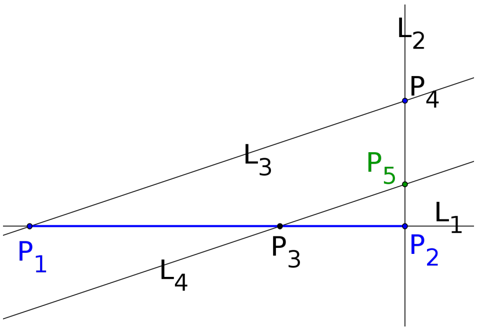


Let $P_1, P_2, P_3, P_4 \in \mathcal{X}(S_0)$ be such that $|a| = \text{length}(\overline{P_1P_2})$, $|b| = \text{length}(\overline{P_3P_4})$. Using (7) and (8) of Proposition 5.2, we can construct a point P_5 on the line $L = L(P_1, P_2)$, in such a way that the length of the segment $\overline{P_2P_5}$ equals the length of $\overline{P_3P_4}$, that is, equals $|b|$. Thus, the segment $\overline{P_1P_5}$ has lenght $|a| + |b|$. If we want to obtain a segment of length $|b| - |a|$, we can draw the circle $C = C(P_2, \overline{P_2P_5})$. Let P_6 be the other point of intersection of L and C; the the segment $\overline{P_6P_1}$ has length $|b| - |a|$.

3. Clear from the definition of K.



4. Let a, b be constructible. We can assume without loss of generality that both are nonzero. It holds that $|ab| = |a| \cdot |b|$, hence we can assume without loss of generality that $a, b > 0$. Let P_1, P_2, P_3, P_4 be points such that $a = \text{length}(\overline{P_1P_2})$, $b = \text{length}(\overline{P_3P_4})$. Let $L_1 = L(P_1, P_2)$, and let L_2 be the line passing through P_2 which is perpendicular to L_1 (cf. (4) of Prop. 5.2). Using (7) and (8) of Proposition 5.2, we can construct a point P_5 on the line $L = L(P_1, P_2)$, in such a way that the length of the segment $\overline{P_2P_5}$ is 1 (which is a constructible real number). Further, using (7) and (8) again, we can construct a point P_6 in the line L_2 such that the length of the segment $\overline{P_2P_6}$ equals b. Let $L_3 = L(P_1, P_5)$, and L_4 be a line passing through P_6 and parallel to L_3. The line L_4 intersects L_1 in a point P_7. The length of the segment P_2P_7 is ab, thus ab is constructible.



5. We may assume, without loss of generality, that $a > 0$. Assume that a is constructible, and let $P_1, P_2 \in \mathcal{X}(S_0)$ be such that $\text{length}(\overline{P_1P_2}) = a$. Let $L_1 = L(P_1, P_2)$ and L_2 be a line passing through P_2 and perpendicular to L_1. Using (7) and (8) of Proposition 5.2, we can draw a point P_3 (resp. a point P_4) on the line L_1 (resp. L_2) such that the length of P_2P_3 (resp. P_2P_4) equals 1. Let $L_3 = L(P_1, P_4)$ and L_4 be the line passing through P_3 parallel to L_3. The line L_4 meets the line L_1 in a point P_5. The length of the segment $\overline{P_2P_5}$ is a^{-1}, thus a^{-1} is constructible. □

Definition 5.12. We call the field $K = \{a \in \mathbb{R} : |a| \text{ is constructible}\}$ the *field of constructible numbers*.

Remark 5.13. Note that the field of constructible numbers K satisfies that $\mathbb{Q} \subseteq K \subseteq \mathbb{R}$.

Lemma 5.14. *Let $S_0 = \{(0,0), (0,1)\}$. The field of constructible numbers K equals $\mathcal{K}(S_0)$.*

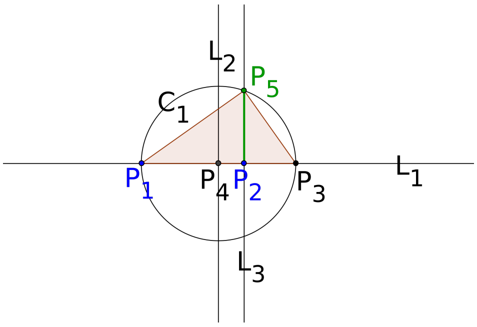
Proof. $\subseteq$ It follows from Remark 5.9.

$\supseteq$ Since K is a field containing $\mathbb{Q}$ it suffices to prove that, for all $n \in \mathbb{N}$, $T_n \subset K$. Let $n \in \mathbb{N}$ and $a \in T_n$. Then either there exists $y \in \mathbb{R}$ such that $P = (a, y) \in \mathcal{X}_n(S_0)$ or there exists $x \in \mathbb{R}$ such that $P = (x, a) \in \mathcal{X}_n(S_0)$. In both cases Lemma 5.10 implies that $a \in K$.

□

The first thing we are going to prove about the field of constructible numbers is that it is different from $\mathbb{Q}$.

Proposition 5.15. *Let $a > 0$ be constructible. Then the positive real number $\sqrt{a}$ is also constructible.*



Proof.

Let $P_1, P_2 \in \mathcal{X}(S_0)$ be such that $a = \text{length}(\overline{P_1P_2})$. Let $L_1 = L(P_1, P_2)$, and let P_3 be a point on L such that the length of P_2P_3 equals 1 (which can be constructed by means of (7) and (8) of Proposition 5.2). Let L_2 be the perpendicular bisector of the segment $\overline{P_1P_3}$ (cf. (3) of Proposition 5.2), and P_4 the intersection of L_1 and L_2. We further draw the line L_3 passing through P_2 perpendicular to L_1 (cf. (4) of Proposition 5.2). Next, draw the circle $C_1 = C(P_4, \overline{P_4P_3})$. C_1 and L_3 meet at two points; choose one of them and call it P_5. Note that the triangle with vertices in P_1, P_3, P_5 is a right triangle (Theorem of Thales). Thus, from Pitagoras's Theorem it follows that the segment $\overline{P_2P_5}$ has length $\sqrt{a}$. Thus $\sqrt{a}$ is constructible.

□

Lemma 5.16. *Let $S \subset \mathbb{R} \times \mathbb{R}$, and let $P_1, P_2 \in S$.*

1. *The line $L = L(P_1, P_2)$ can be described by an equation of the form $aX + bY + c = 0$ for some $a, b, c \in \mathcal{K}_0(S)$.*
2. *The circle $C = C(P_1, \overline{P_1P_2})$ can be described by an equation of the form $X^2 + Y^2 + aX + bY + c = 0$, for some $a, b, c \in \mathcal{K}_0(S)$.*

Proof. Let $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$; note that by definition of $\mathcal{K}_0(S)$, $x_1, x_2, y_1, y_2 \in \mathcal{K}_0(S)$.

1. If $x_1 = x_2$, then L is described by the equation $X = x_1$, which has the desired form. Assume then that $x_1 \neq x_2$. Then $(x, y) \in L$ if and only if

$$\frac{y - y_1}{x - x_1} = \frac{y_2 - y_1}{x_2 - x_1},$$

that is to say, if and only if (x, y) satisfies the equation

$$(y_2 - y_1)X + (x_1 - x_2)Y + (y_1x_2 - x_1y_2) = 0,$$

which has the desired shape.

2. A point $(x, y) \in C$ if and only if the distance from (x, y) to P_1 equals the distance of P_1 to P_2, that is to say,
$$(x - x_1)^2 + (y - y_1)^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2$$
that is to say, if and only if the point (x, y) satisfies the equation
$$X^2 + Y^2 - 2x_1X - 2y_1Y + (x_1^2 + y_1^2 - ((x_2 - x_1)^2 + (y_2 - y_1)^2) = 0,$$
which has the desired shape.

□

Lemma 5.17. *Let $S \subset \mathbb{R} \times \mathbb{R}$, and let $P_1, P_2, P_3, P_4 \in S$.*

1. *If the lines $L_1 = L(P_1, P_2)$ and $L_2 = L(P_3, P_4)$ intersect, then the point of intersection has coordinates in $\mathcal{K}_0(S)$.*

2. *If the line $L = L(P_1, P_2)$ and the circle $C = C(P_3, \overline{P_3P_4})$ intersect, then there exists $d \in \mathcal{K}_0(S)$, $d > 0$, such that the points of intersection have coordinates in $(\mathcal{K}_0(S))(\sqrt{d})$.*

3. *If the circle $C_1 = C(P_1, \overline{P_1P_2})$ and the circle $C = C(P_3, \overline{P_3P_4})$ intersect, then there exists $d \in \mathcal{K}_0(S)$, $d > 0$, such that the points of intersection have coordinates in $(\mathcal{K}_0(S))(\sqrt{d})$.*

Proof. 1. We can write L_1 and L_2 by means of equations $a_1X + b_1Y + c_1 = 0$ and $a_2X + b_2Y + c_2 = 0$ respectively, with $a_1, b_1, c_1, a_2, b_2, c_2 \in \mathcal{K}_0(S)$. The intersection point $P = (x, y)$, if it exists, satisfies the system of equations
$$\begin{cases} a_1X + b_1Y + c_1 = 0 \\ a_2X + b_2Y + c_2 = 0. \end{cases}$$
Thus x, y belong to the field $\mathbb{Q}(a_1, b_1, c_1, a_2, b_2, c_2) \subset \mathcal{K}_0(S)$.

2. We can write L by means of an equation $a_1X + b_1Y + c_1 = 0$ with $a_1, b_1, c_1 \in \mathcal{K}_0(S)$, a_1 and b_1 not both zero. Let us assume, without loss of generality, that $b_1 \neq 0$. We can write C by means of an equation $X^2 + Y^2 + a_2X + b_2Y + c_2 = 0$ with $a_2, b_2, c_2 \in \mathcal{K}_0(S)$. The intersection point $P = (x, y)$, if it exists, satisfies the system of equations
$$\begin{cases} a_1X + b_1Y + c_1 & = 0 \\ X^2 + Y^2 + a_2X + b_2Y + c_2 & = 0. \end{cases}$$
From the first equation, we can write $Y = (-a_1X - c_1)/b_1$, and replace it in the second equation. Performing the computations, we will obtain an equation of the form $AX^2 + BX + C = 0$ for some $A, B, C \in \mathcal{K}_0(S)$. Call $d = B^2 - 4AC$. Then it is clear that, if there are any solutions, they are contained in the field $\mathbb{Q}(a_1, b_1, c_1, a_2, b_2, c_2, \sqrt{d})$.

We can write C_1 and C_2 by means of equations $X^2+Y^2+a_1X+b_1Y+c_1=0$ and $X^2+Y^2+a_2X+b_2Y+c_2=0$ respectively, with $a_1,b_1,c_1,a_2,b_2,c_2\in\mathcal{K}_0(S)$. The intersection points (x,y), if they exist, satisfy the system of equations

$$\begin{cases} X^2+Y^2+a_1X+b_1Y+c_1 &=0\\ X^2+Y^2+a_2X+b_2Y+c_2 &=0.\end{cases}$$

But this system of equation is equivalent to

$$\begin{cases} (a_1-a_2)X+(b_1-b_2)Y+(c_1-c_2) &=0\\ X^2+Y^2+a_2X+b_2Y+c_2 &=0.\end{cases}$$

Thus, we can conclude as in the previous case. □

Lemma 5.18. *Let $K\subset L\subset\mathbb{R}$ be field extensions such that $[L:K]=2$. Then there exists $d\in K$ such that $L=K(\sqrt{d})$.*

Proof. Let $a\in L\backslash K$. Then $[K(a):K]>1$, and thus the equality $2=[L:K]=[L:K(a)]\cdot[K(a):K]$ implies that $L=K(a)$.

Let $f(X)\in K[X]$ be the minimal polynomial of a; since $[K(a):K]=[L:K]=2$, $f(X)$ has degree 2. Write it as $f(X)=X^2+c_1X+c_2$ with $c_1,c_2\in K$, and let $b=-c_1/2\in K$, $d=a-b$. Then

$$d^2=(a-b)^2=a^2-2ab+b^2=(-c_1a-c_2)-2a(-c_1/2)+(-c_1/2)^2=c_1^2/4-c_2\in K,$$

and the fact that $b\in K$ implies that $L=K(a)=K(\sqrt{d})$. □

Theorem 5.19. *Let $a\in\mathbb{R}$. The following are equivalent:*

(i) a is constructible.

(ii) There exist a finite chain of fields $\mathbb{Q}=F_0\subseteq F_1\subseteq\cdots\subseteq F_n$ such that

(a) $a\in F_n$.

(b) For all $i=0,\dots,n-1$, $[F_{i+1}:F_i]=1$ or 2.

Proof. Let $S_0=\{(0,0),(0,1)\}$.

- $(i)\Rightarrow(ii)$: Assume that a is constructible. By Lemma 5.14, this implies that $a\in\mathcal{K}(S_0)$. Therefore, there exists $m\in\mathbb{N}$ such that $a\in\mathcal{K}_m(S_0)$. Thus, it suffices to prove that there exists a finite chain of fields, $\mathbb{Q}=F_0\subseteq F_1\subseteq\cdots\subseteq F_n=\mathcal{K}_m(S_0)$, with $[F_{i+1}:F_i]=2$ for all $i=0,\dots,n-1$.

  We will proceed by induction on m. For $m=0$, $\mathcal{K}_0(S_0)=\mathbb{Q}(0,1)=\mathbb{Q}$, so we can take $n=0$ and the chain with just one field, $F_0=\mathbb{Q}$.

  Assume that the result is known for all $i=0,\dots,m-1$. Recall that we have defined T_i, $i\in\mathbb{N}\cup\{0\}$, as

$$T_i=\bigcup_{(x,y)\in\mathcal{X}_i(S_0)}\{x,y\}.$$

Note that each T_i is a finite set (cf. Remark 5.5). By definition, $\mathcal{K}_m(S_0)$ is generated over $\mathbb{Q}$ by the union $T_0 \cup \cdots \cup T_m$. Thus

$$\mathcal{K}_m(S_0) = \mathbb{Q}(T_0, \ldots, T_{m-1}, T_m) = (\mathbb{Q}(T_0, \ldots, T_{m-1}))\,(T_m) = (\mathcal{K}_{m-1}(S_0))\,(T_m).$$

By induction hypothesis, we have a chain $\mathbb{Q} = F_0 \subset \cdots \subset F_k = \mathcal{K}_{m-1}(S_0)$ such that each $[F_{i+1} : F_i] = 1$ or 2, $i = 0, \ldots, k-1$. Thus, to complete the proof it suffices to continue this chain up to $\mathcal{K}_m(S_0)$. We number the points in T_m as $P_1, \ldots, P_s$, and write $P_j = (x_j, y_j)$, $j = 1, \ldots, s$. For $j = 1, \ldots, s$, set

$$F'_j := F_k(x_1, \ldots, x_{j-1}, x_j, y_1, \ldots, y_{j-1}, y_j)$$

Then we have a chain

$$\mathbb{Q} = F_0 \subseteq F_1 \subseteq \cdots F_k = \mathcal{K}_{m-1}(S_0) \subseteq F'_1 \subseteq F'_s = \mathcal{K}_m(S_0),$$

where $[F_{i+1} : F_i] = 2$ for $i = 0, \ldots, k-1$, $F'_j = F'_{j-1}(x_j, y_j)$ for $j = 1, \ldots, s$. To conclude the proof, it suffices to note that, for $j = 1, \ldots, s$, there exist $Q_1, Q_2, Q_3, Q_4 \in \mathcal{X}_{m-1}(S)$ such that the point $P_j = (x_j, y_j)$ satisfies one of the following:

1. $P \in L(Q_1, Q_2) \cap L(Q_3, Q_4)$
2. $P \in L(Q_1Q_2) \cap C(Q_3, \overline{Q_3Q_4})$
3. $P \in C(Q_1, \overline{Q_1Q_2}) \cap C(Q_3, \overline{Q_3Q_4})$.

In the first case, $F'_j = F'_{j-1}$ and in the second and third cases, by Lemma 5.18, there exists $d \in F_k = \mathcal{K}_{m-1}(S)$ such that $F'_j = F'_{j-1}(\sqrt{d})$. Thus in both cases

$$[F_j : F_{j-1}] = 1 \text{ or } 2,$$

and setting $F_{k+j} := F'_j$ for all $j = 1, \ldots, s$, we obtain a chain of fields with the required properties.

- $(ii) \Rightarrow (i)$ Let

$$\mathbb{Q} = F_0 \subset F_1 \subset \cdots \subset F_n$$

be a chain of fields with $a \in F_n$ and $[F_{i+1} : F_i] = 1$ or 2 for all $i = 0, \ldots, n-1$. Let K be the field of constructible numbers. It suffices to show that $F_n \subset K$.

We proceed by induction on i: If $i = 0$, then $F_0 = \mathbb{Q} \subset K$.

$i \Rightarrow i+1$: Assume $F_i \subset K$. If $[F_{i+1} : F_i] = 1$, then $F_{i+1} = F_i \subset K$, and we are done. Thus we can assume $[F_{i+1} : F_i] = 2$. By Lemma 5.18, there exists $d_i \in F_i \subset K$ such that $F_{i+1} = F_i(\sqrt{d_i})$. By Proposition 5.15, $\sqrt{d_i} \in K$, and hence $F_{i+1} = F_i(\sqrt{d_i}) \subset K$.

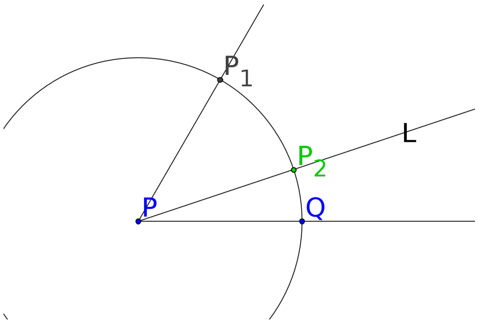
□

Corollary 5.20. *Let $S_0 = \{(0,0), (0,1)\} \subset \mathbb{R} \times \mathbb{R}$. The field extension $\mathcal{K}(S_0)/\mathbb{Q}$ is algebraic. Moreover, for each $a \in \mathcal{K}(S_0)$ the degree of the (finite) field extension $\mathbb{Q}(a)/\mathbb{Q}$ is a power of* 2.

Theorem 5.21 (Wantzel (1837)). *The trisecting of the angle cannot be performed with ruler and compass.*

Proof. Let $P = (0,0)$ and $Q = (0,1)$.

Construct an equilateral triangle with side $\overline{PQ}$ (cf. Proposition 5.2-(1)), and let P_1 be the other vertex of the triangle. Assume the angle $\theta = \pi/3$ between $\overline{PQ}$ and $\overline{PP_1}$ can be trisected. This means there exists a line L such that the angle between $L(P,Q)$ and L is $\pi/9$. Let P_2 be the intersection of L and $C(P, \overline{PQ})$. We have that $P_2 = (\cos(\pi/9), \sin(\pi/9))$ belongs to $\mathcal{X}(S_0)$. However,



$$\begin{aligned}\cos(3\theta) &= \cos(2\theta+\theta)\\ &= \cos(2\theta)\cos(\theta) - \sin(2\theta)\sin(\theta)\\ &= (\cos^2(\theta) - \sin^2(\theta))\cos(\theta) - 2(\sin(\theta)\cos(\theta))\sin(\theta)\\ &= \cos^3(\theta) - 3\sin^2(\theta)\cos(\theta)\\ &= \cos^3(\theta) - 3(1-\cos^2(\theta))\cos(\theta)\\ &= 4\cos^3(\theta) - 3\cos(\theta),\end{aligned}$$

where we have used the formulae

$$\begin{cases}\cos^2(\alpha)+\sin^2(\alpha)=1,\\ \cos(\alpha+\beta)=\cos(\alpha)\cos(\beta)-\sin(\alpha)\sin(\beta),\\ \sin(\alpha+\beta)=\sin(\alpha)\cos(\beta)+\sin(\beta)\cos(\alpha).\end{cases}$$

Since $\cos(\pi/3) = 1/2$, we have that $\alpha = \cos(\pi/9)$ satisfies

$$1/2 = 4\alpha^3 - 3\alpha.$$

In other words, α is a root of the polynomial

$$f(X) := 8X^3 - 6X - 1.$$

But this polynomial is irreducible over $\mathbb{Q}$ (e.g. apply the reduction criterion with $p = 5$). Thus $(1/8)f(X)$ is the minimal polynomial of α over $\mathbb{Q}$, and hence

$$[\mathbb{Q}(\alpha) : \mathbb{Q}] = 3,$$

which is not a power of 2. This contradicts Corollary 5.20.

Theorem 5.22. *The squaring of the circle cannot be performed with ruler and compass.*

Proof. Let $P = (0,0)$ and $Q = (0,1)$. Consider the circle $C = C(P, \overline{PQ})$. It has area π. Assume that the squaring of the circle C can be done with ruler and compass: then there are two constructible points P_1 and P_2 such that the square with side $\overline{P_1P_2}$ has the same area as the circle C, that is, π. Then $\sqrt{\pi}$ is a constructible number, and in particular π is constructible as well. But by Corollary 5.20, the field of constructible numbers is algebraic over $\mathbb{Q}$. This contradicts the fact that π is trascendental over $\mathbb{Q}$. □

Remark 5.23. The proof that π is trascendental over $\mathbb{Q}$, which implies that the squaring of the circle cannot be performed by ruler and compass, was given by Lindemann in 1882.

Theorem 5.24 (Wantzel (1837))**.** *The duplicating of the cube cannot be performed with ruler and compass.*

Proof. Let $P = (0,0)$ and $Q = (0,1)$. Consider the cube D_1 with side $\overline{PQ}$. It has volume 1. Assume that the duplicating of the cube D_1 can be done with ruler and compass: then there are two constructible points P_1 and P_2 such that the cube D_2 with side $\overline{P_1P_2}$ has volume 2. Let $\alpha = \mathrm{length}(\overline{P_1P_2})$ (which is a constructible number); α must satisfy that $\alpha^3 = 2$. That is to say, α is a root of the polynomial $f(X) := X^3 - 2$, which is irreducible over $\mathbb{Q}$ (e.g. apply Eisenstein's criterion with $p = 2$). Therefore the degree $[\mathbb{Q}(\alpha) : \mathbb{Q}] = 3$, which is not a power of 2. This contradicts Corollary 5.20. □

References

[1] Artin, Emil. *Galois theory*. Edited and with a supplemental chapter by Arthur N. Milgram. Reprint of the 1944 second edition. Dover Publications, Inc., Mineola, NY, 1998.

[2] Hadlock, Charles Robert. *Field theory and its classical problems*. Carus Mathematical Monographs, 19. Mathematical Association of America, Washington, D.C., 1978.

[3] Lang, Serge. *Algebra*. Revised third edition. Graduate Texts in Mathematics, **211**. Springer-Verlag, New York, 2002.

[4] *Algèbre 1*, Winter term 2012, Université du Luxembourg, lecture notes written by Gabor Wiese and Agnès David.

[5] *Algèbre 2*, Summer term 2013, Université du Luxembourg, lecture notes written by Gabor Wiese and Agnès David.

[6] *Algèbre 3*, Winter term 2012, Université du Luxembourg, lecture notes written by Gabor Wiese.