Algebra: Part II Rings and Fields

Lothar Göttsche

Contents

CHAPTER 1

Rings

1. Definition of a ring, examples and first properties

From high school and even from earlier, everybody is aware of the integers. Integers can be added, subtracted and multiplied, and there is the distributive law relating addition and multiplication. This is formalized in the notion of a ring. Basically the axioms of a ring will guarantee that in a ring one can compute more or less like with integers. When dealing with rings we always want to keep the integers in mind as a basic example.

DEFINITION 1.1. A *ring* R is a nonempty set R together with two binary operations

$$\begin{aligned} + :& R \times R \to R \\ \cdot :& R \times R \to R \end{aligned}$$

and a distinguished element $0 \in R$, such that

(1) $(R, +)$ is an abelian group with neutral element 0.
(2) $\cdot$ is associative

$$(a \cdot b) \cdot c = a \cdot (b \cdot c) \qquad \text{for all } a, b, c \in R.$$

(3) We have the distributive laws

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

We call $(R, +)$ the *additive group* of R. Usually we write ab for $a \cdot b$.

EXAMPLE 1.2. The integers $\mathbb{Z}$ are a ring with the usual addition and multiplication.

A subset of a ring is a subring if, with the restriction of $+$ and $\cdot$, it is itself a ring.

DEFINITION 1.3. Let R be a ring. A subset $A \subset R$ is called a *subring* of R if $0 \in A$ and with the restriction of $+$ and $\cdot$ to $A \times A$ it is a ring.

Obviously this is equivalent to the following conditions

(1) A is a subgroup of the additive group of R.
(2) For all $a, b \in A$ we have $ab \in A$.

EXAMPLE 1.4. Let

$$2\mathbb{Z} := \{2n \mid n \in \mathbb{Z}\}$$

be the set of even integers. Then $2\mathbb{Z}$ is a subring of $\mathbb{Z}$. Note that the set of odd integers is not a subring of $\mathbb{Z}$.

The integers have some further useful properties:

(1) $ab = ba$ for all $a, b \in \mathbb{Z}$,
(2) $1a = a1 = a$ for all $a \in \mathbb{Z}$.

DEFINITION 1.5. A ring R is called *commutative* if

$$ab = ba \qquad \text{for all } a, b \in R$$

i.e. if the multiplication is commutative. (Note that the addition is already commutative).

DEFINITION 1.6. An element $1 \in R \setminus \{0\}$ is called a *unit element* of R if

$$a \cdot 1 = 1 \cdot a \qquad \text{for all } a \in R.$$

Note that we are assuming that $1 \neq 0$. (Note also that 1 is then unique). If R contains a unit element, it is called a ring with unit element or shorter a ring with 1.

In this course we will mostly be interested in commutative rings with 1. After the first few lectures we will restrict our attention to them.

REMARK 1.7. In a commutative ring R the distributive law just becomes

$$a \cdot (b + c) = a \cdot b + a \cdot c.$$

EXAMPLE 1.8. (1) The integers $\mathbb{Z}$, the rational numbers $\mathbb{Q}$, the real numbers $\mathbb{R}$ and the complex numbers $\mathbb{C}$ with the usual addition and multiplication are all commutative rings with 1. We briefly recall the complex numbers. $\mathbb{C} = \{a + ib \mid a, b \in \mathbb{R}\}$ where the ring structure is given by

$$(a + ib) + (c + id) = a + c + i(b + d); \quad (a + ib)(c + id) = (ac - bd) + i(ad + bc),$$

i.e. one has that $i^2 = (-1)$ and otherwise computes normally.

(2) Let $\mathbb{Z}[i] := \{n + im \in \mathbb{C} \mid n, m \in \mathbb{Z}\}$. $\mathbb{Z}[i]$ is a subring of $\mathbb{C}$. $\mathbb{Z}[i]$ is called the ring of *Gaussian integers*.

(3) $2\mathbb{Z}$ is a commutative ring (as it is a subring of $\mathbb{Z}$, but it is not a ring with 1).

(4) Let $M_{n\times n}(\mathbb{R})$ be the set of $n \times n$ matrices with entries in $\mathbb{R}$, with the usual addition and multiplication of matrices. Then $M_{n\times n}(\mathbb{R})$ is a ring with 1. The unit element is the diagonal matrix with 1 on the diagonal. But $M_{n\times n}(\mathbb{R})$ is not commutative for $n \geq 2$. $M_{n\times n}(\mathbb{R})$ is the most common example of a noncommutative ring. If we have to think of noncommutative rings we always first look at matrices.

(5) Let $\mathbb{C}^\infty(\mathbb{R})$ be the set of $\mathbb{C}^\infty$ functions $f : \mathbb{R} \to \mathbb{R}$. Then $\mathbb{C}^\infty(\mathbb{R})$ is a commutative ring with 1 with pointwise addition and multiplication:

$$(f + g)(x) = f(x) + g(x), \qquad (fg)(x) = f(x)g(x).$$

The zero element is the constant function 0 and the unit element is the constant function 1.

(6) Let $k \in \mathbb{Z}$ and let $\mathbb{Z}/k\mathbb{Z}$ be the set of residue classes modulo k. The elements of $\mathbb{Z}/k\mathbb{Z}$ are $0, 1, \ldots, k - 1$, and addition and multiplication are defined by

$$\underline{n} + \underline{m} = \underline{n + m}, \quad \underline{n}\,\underline{m} = \underline{nm}$$

where $\underline{n}$ denotes the rest of n after division by k. Check that $\mathbb{Z}/k\mathbb{Z}$ is a commutative ring with zero element 0 and unit element 1.

Polynomial rings. We come now to the most important examples of rings. These are the rings $R[X]$ of polynomials with coefficients in a ring R. Already from high school we are familiar with polynomials $f = \sum_{i=0}^{n} a_i X^i$ with coefficients in the real numbers $\mathbb{R}$. One learns how they can be added, subtracted and multiplied in the usual way. We need to give a precise definition. The easiest way is to view a polynomial $f = \sum_{i=0}^{n} a_i X^i$ as the sequence $(a_0, a_1, \dots)$ of its coefficients.

DEFINITION 1.9. Let R be a ring with 1. A *polynomial* f in X with coefficients in R is a sequence $(a_i)_{i\in\mathbb{Z}_{\geq 0}}$ of elements $a_i \in R$, such that $a_i = 0$ for all but a finite number of i. $a_i \in R$ is called the coefficient of X^i. We denote by $R[X]$ the set of polynomials in X with coefficients in R. We identify the polynomial $(a_0, 0, 0\dots)$ with the element $a_0 \in R$. A polynomial of this form is called a *constant polynomial.* Thus R is the subset of $R[X]$ of constant polynomials.

We define a ring structure on $R[X]$. If $f = (a_i)_{i\in\mathbb{Z}_{\geq 0}}$ and $g = (b_i)_{i\in\mathbb{Z}_{\geq 0}}$ then we define

$$\begin{aligned} f+g &:= (a_i + b_i)_{i\in\mathbb{Z}_{\geq 0}}, \\ fg &:= (c_k)_{k\in\mathbb{Z}_{\geq 0}}, \quad c_k = \sum_{i+j=k} a_i b_j. \end{aligned}$$

It is easy to check (and left as an exercise) that with these operations $R[X]$ is a ring with 1. The zero element is $0 \in R$ and the unit element is $1 \in R$ (which we have identified with $(0, 0, \dots)$ and $(1, 0, \dots)$). If R is commutative, then also $R[X]$ is commutative. We also see that R is a subring of $R[X]$. We call $R[X]$ the *polynomial ring* in X over R.

It is much better to work with the usual notation for polynomials.

REMARK 1.10. Let R be a ring, $a \in R$. We denote $a^0 := 1$. For all $n \in \mathbb{Z}_{>0}$ we denote $a^n := \underbrace{a \cdot \ldots \cdot a}_{n \text{ times}}$. It is easy to check (and left as an exercise) that

$$a^n a^m = a^{n+m}, \quad (a^n)^m = a^{nm}.$$

NOTATION 1.11. We write X for the polynomial $(0, 1, 0, 0, \dots)$. Then by the rule for the product X^i is the sequence whose i^{th} term is 1 and all others are 0. Similarly for $a \in R$, we have $aX^i = (0, \dots, 0, a, 0, \dots)$ with a as the i^{th} term. Thus we see that

$$(a_0, \dots, a_n, 0\dots) = a_0 + a_1 X + \ldots + a_n X^n.$$

REMARK 1.12. We see that with this definition the addition is defined as usual by adding the corresponding coefficients. The multiplication is as usual given by putting $(aX^i)(bX^j) = abX^{i+j}$ and formally applying the distributive law.

DEFINITION 1.13. Let R be a commutative ring with 1, and let $f \in R[X] \setminus \{0\}$. Write $f = \sum_{i=0}^{n} a_i X^i$ with $a_n \neq 0$. Then n is called the *degree* of f and denoted by $deg(f)$. a_n is called the *leading coefficient* of f.

By induction we can also define polynomial rings in several variables.

DEFINITION 1.14. Inductively we define

$$R[X_1, \dots, X_n] := R[X_1, \dots, X_{n-1}][X_n].$$

Thus the ring of polynomials in $X_1, \dots, X_n$ is the ring of polynomials in X_n with coefficients polynomials in $X_1, \dots, X_{n-1}$. Inductively we can check that every $f \in R[X_1, \dots, X_n]$ can be written as a finite sum

$$f := \sum_{i_1, \dots, i_n} a_{i_1, \dots, i_n} X_1^{i_1} \dots X_n^{i_n}$$

with $a_{i_1, \dots, i_n} \in R$.

We want to be able to do computations in a ring in much the same way as in the integers. The following lemma shows that many of the basic rules carry over.

LEMMA 1.15. *Let R be a commutative ring with* 1. *Then for all $a, b \in R$*

(1) $a0 = 0$,
(2) $(-a)b = -(ab)$,
(3) $(-a)(-b) = ab$,
(4) $(-1)a = -a$,

PROOF. These are straightforward verifications. (1) $a0 = a(0 + 0) = (a0 + a0)$. As $(R, +)$ is a group, we can subtract $a0$ from both sides to get $a0 = 0$. (2) $ab + (-a)b = (a + (-a))b = 0b = 0$. (3) follows by appling (2) twice. (4) $(-1)a = -(1a) = -a$. □

NOTATION 1.16. In future we will write $a - b$ for $a + (-b)$.

We come to some further properties of rings. In the integers $\mathbb{Z}$ we know that the only way how a product ab can be 0 is that $a = 0$ or $b = 0$. This is a very useful property of the integers. It implies that we can cancel factors: if $ab = ac$ with $a \neq 0$, then $b = c$ (This is because $0 = ab - ac = a(b - c)$ thus $(b - c) = 0$.)

DEFINITION 1.17. Let R be a commutative ring. An element $a \in R \setminus \{0\}$ is called a *zero divisor* if there is an element $b \in R \setminus \{0\}$ with $ab = 0$. A commutative ring R with 1, which contains no zero divisors is called an *integral domain.*

REMARK 1.18. (Cancellation) Let R an integral domain, then the cancellation law is valid, i.e., if $ab = ac$ and $a \neq 0$, then $b = c$.

PROOF. Let $ab = ac$ and $a \neq 0$. Then $a(b - c) = 0$. As R contains no zero divisors it follows that $(b - c) = 0$. □

Integral domains are particularly nice rings. Most of the rings we will study are integral domains.

EXAMPLE 1.19. (1) $\mathbb{Z}$ is an integral domain.
(2) In $Z/6\mathbb{Z}$ we see that $\underline{2}\,\underline{3} = \underline{0}$. Thus $\mathbb{Z}/6\mathbb{Z}$ is not an integral domain.
(3) Let $p \in \mathbb{Z}_{>0}$ be a prime number. If $0 < n, m < p$ then nm is not divisible by p. Thus $\mathbb{Z}/pZ$ is an integral domain.

DEFINITION 1.20. Let R be a ring with 1. An element $a \in R$ is called a *unit* of R if it has a multiplicative inverse, i.e. if there exists an element $b \in R$ with $ab = ba = 1$. The set of units in R is denoted by R^*.

EXAMPLE 1.21. (1) $\mathbb{Z}^* = \{1, (-1)\}$.
(2) $\mathbb{Z}[i]^* = \{1, (-1), i, -i\}$.
(3) $\mathbb{Q}^* = \mathbb{Q} \setminus \{0\}$.
(4) The units in $M_{n\times n}(\mathbb{R})$ or $M_{n\times n}(\mathbb{C})$ are the matrices with nonzero determinant.

It is easy to see that in each of the above $(R^*, \cdot)$ is a group. This is always true:

PROPOSITION 1.22. *Let R be a ring with 1. Then $(R^*, \cdot)$ is a group.*

PROOF. This is easy. We know that the multiplication is associative. We know that $1 \in R^*$ is the neutral element for the multiplication, and that for every $a \in R^*$ there is an element $b \in R^*$ with $ab = ba = 1$, i.e. an inverse element. We only need to show that for $a, b \in R^*$ also $ab \in R^*$. Let $c, d \in R$ with $ac = ca = 1$, $bd = db = 1$. Then $(ab)(dc) = ac = 1$ and $(dc)(ab) = db = 1$. Thus R^* is a group. □

DEFINITION 1.23. $(R^*, \cdot)$ is called the *multiplicative group* of R.

If R is an integral domain, then also $R[X]$ is an integral domain and its units are just the units of R.

REMARK 1.24. Let R be an integral domain, $f, g \in R[X] \setminus \{0\}$.
(1) $fg \neq 0$ and $deg(fg) = deg(f) + deg(g)$.
(2) $R[X]$ is an integral domain.
(3) $R[X]^* = R^*$, the units in $R[X]$ are just the units in R.

PROOF. (1) Let $f = \sum_{i=0}^m a_i X^i$ with $a_m \neq 0$, $g = \sum_{i=0}^n b_i X^i$ with $b_n \neq 0$, then

$$fg = \sum_{i=0}^{n+m} c_i X^i, \qquad c_i = \sum_{k+l=i} a_k b_l,$$

in particular $c_{n+m} = a_m b_n \neq 0$. (2) is part of (1). (3) It is clear that every unit of R is a unit of $R[X]$. Let $f \in R[X]^*$. Then there exists $g \in R[X]$ with $fg = 1$. By (1) this implies $deg(f) + deg(g) = 0$, i.e. $deg(f) = 0$. Thus $f \in R$ and thus $f \in R^*$. □

Note that 0 is never a unit (because $0a = a0 = 0$ for all a). A ring with 1 in which every nonzero element is a unit is called a division ring.

DEFINITION 1.25. A ring R with 1 is called a *division ring*, if every nonzero element is a unit, i.e. $R^* = R \setminus \{0\}$.

The most important case are the commutative division rings: the fields.

DEFINITION 1.26. A commutative division ring R is called a *field*. Explicitly this means the following. A set R with two binary operations $+, \cdot$ and two distinguished elements $0 \neq 1$ is called a field if
(1) $(R, +)$ is a commutative group with neutral element 0,
(2) $(R \setminus \{0\}, \cdot)$ is a commutative group with neutral element 1,
(3) the distributive law holds:

$$a(b + c) = ab + ac \quad \text{for all } a, b, c \in R.$$

The concept of field is of fundamental importance. While the notion of a ring formalizes that one is able to compute like in the integers, the notion of a field formalizes that one can compute like in the rational numbers. The second half of this course will be devoted to the theory of fields.

EXAMPLE 1.27. $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ are fields. In $\mathbb{C}$ the inverse to $a+bi$, $a, b \in \mathbb{R}$ is $(a-bi)/(a^2+b^2)$.

We want to give one example of a division ring which is not a field, the Quaternions.

EXAMPLE 1.28. (Quaternions) The most famous example of a noncommutative division ring are the Quaternions. Recall that in $\mathbb{C}$ we have *complex conjugation* i.e. for $c = a + ib$ $(a, b \in \mathbb{R})$ we have $\overline{c} = a-ib$, which fulfils the folloing properties: $\overline{\overline{c}} = c$, $c\overline{c} = a^2+b^2$, which is a positive real number unless $c = 0$, $\overline{cd} = \overline{c}\overline{d}$ $\overline{c+d} = \overline{c}+\overline{d}$. Let $Q := \left\{ \begin{pmatrix} a & b \\ -\overline{b} & \overline{a} \end{pmatrix} \middle| \, a, b \in \mathbb{C} \right\}$. Q is called the ring of quaternions. It is easy to check that Q is a subring of $M_{2\times 2}(\mathbb{C})$. We have $det \begin{pmatrix} a & b \\ -\overline{b} & \overline{a} \end{pmatrix} = a\overline{a} + b\overline{b}$, thus if the element is nonzero its inverse is given by $\frac{1}{a\overline{a}+b\overline{b}} \begin{pmatrix} \overline{a} & -b \\ \overline{b} & a \end{pmatrix}$. On the other hand $\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} = -\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$. So Q is a division algebra but not a field.

We want to give one less standard example of a field, in particular we want to see that there are also fields with finitely many elements.

PROPOSITION 1.29. *Every finite integral domain is a field.*

PROOF. By the above we have to show that $R^* = R \setminus \{0\}$, that is we have to find an inverse for any nonzero $a \in R$. Note that the distributive law $a(b+c) = ab + ac$ says that the map

$$\cdot a : (R, +) \to (R, +), b \mapsto ab$$

of multiplication by a is a group homomorphism. As R has no zero divisors, the kernel of m_a is 0. Thus $\cdot a$ is injective. But it is one of the fundamental properties of finite sets, that an injective map of a finite set onto itself is also bijective. So let $c \in R$ be an element with $\cdot a(c) = 1$. Then $ac = 1$, thus a is a unit. $\square$

In particular $\mathbb{Z}/p\mathbb{Z}$ is a field for a p prime number. It is called the finite field with p elements and denoted F_p.

Quotient fields. We know that we can construct the rational numbers as the set of all quotients $\frac{a}{b}$ of integers. We want to see that the same can be done for any integral domain R. We can construct the quotient field $Q(R)$ consisting of all quotients $\frac{a}{b}$ with $a, b \in R$, and R is a subring of $Q(R)$. Thus every integral domain is in a natural way a subring of a field.

THEOREM AND DEFINITION 1.1. *Let R be an integral domain. On the set of pairs $(a, b) \in R \times (R \setminus \{0\})$ we define an equivalence relations by*

$$(a, b) \sim (c, d) \text{ if and only if } ad = bc.$$

We denote by $\frac{a}{b}$ the equivalence class of (a, b). Let

$$Q(R) := \left\{ \frac{a}{b} \,\middle|\, a \in R,\ b \in R \setminus \{0\} \right\}.$$

We define operations

$$\cdot, + : Q(R) \times Q(R) \to \mathbb{Q}(R); \frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}; \quad \frac{a}{b} + \frac{c}{d} = \frac{ad+bc}{bd},$$

i.e. multiplication and addition are defined in the usual way that one learns for rational numbers at elementary school. We identify $a \in R$ with $\frac{a}{1}$. With these operations $Q(R)$ is a field and R is a subring of $Q(R)$.

Proof. In the above there are a few things to check.

(1) First we have to check that $\sim$ is an equivalence relation. obviously $\sim$ is reflexive $(a,b) \sim (a,b)$ and $\sim$ is symmetric. We check that $\sim$ is transitive. Let $(a,b) \sim (c,d)$ and $(c,d) \sim (e,f)$. Then $ad = bc$ and $cf = de$. Thus $(ad)f = (bc)f$ and $b(cf) = b(de)$. Thus $(af)d = (be)d$. As R is an integral domain, and $d \neq 0$ it follows that $af = be$.

(2) We have to check that $+$ and $\cdot$ are well-defined. Thus assume $\frac{a'}{b'} = \frac{a}{b}$ and $\frac{c'}{d'} = \frac{c}{d}$. We have to show that $\frac{a}{b} + \frac{c}{d} = \frac{a'}{b'} + \frac{c'}{d'}$, $\frac{a}{b} \cdot \frac{c}{d} = \frac{a'}{b'} \cdot \frac{c'}{d'}$. We have $a'b = ab'$ and $c'd = cd'$, thus $acb'd' = a'c'bd$, i.e. $\frac{ac}{bd} = \frac{a'c'}{b'd'}$. For $+$ the computation is slightly more complicated, but similar.

(3) We want to show that $Q(R)$ is a field. It is easy to check that $(Q(R), +, \cdot)$ is a commutative ring with $1 = \frac{1}{1}$ as unit element and $0 = \frac{0}{1}$ as zero element and $\frac{-a}{b}$ as additive inverse to $\frac{a}{b}$. So we need to find a multiplicative inverse to $\frac{a}{b} \in Q(R) \setminus \{0\}$. Then $a \neq 0$, thus $\frac{b}{a} \in Q(R)$ and obviously $\frac{a}{b}\frac{b}{a} = 1$. Thus $Q(R)$ is a field. Obviously R is a subring of $Q(R)$. □

Example 1.30.
(1) $\mathbb{Q}$ is the quotient field of $\mathbb{Z}$.
(2) Let k be a field. Then $k[X]$ is an integral domain. The quotient field $Q(k[X])$ is usually denoted by $k(X)$ and called the *field of rational functions* in X with coefficients in k. Elements of $k(X)$ are quotients $\frac{f}{g}$ of polynomials $f, g \in k[X]$ with $g \neq 0$.

2. Homomorphisms and Ideals

When one studies sets with a structure, like groups and rings, the most important thing to look at are the maps between them which are compatible with the structure. For groups G, H these are the group homomorphisms, i.e. the maps $\varphi : G \to H$ which are compatible with the multiplication in the groups $\varphi(g_1g_2) = \varphi(g_1)\varphi(g_2)$. For rings we have two operations $+$ and $\cdot$, so a ring homomorphism has to be compatible with both of them.

Definition 2.1. Let A, B be rings. A map $\varphi : A \to B$ is called a *ring homomorphism* if for all $a, b \in A$

(1) $\varphi(a+b) = \varphi(a) + \varphi(b)$,
(2) $\varphi(ab) = \varphi(a)\varphi(b)$.

The *image* of φ is $\varphi(A)$ and the *kernel* is $ker(\varphi) := \varphi^{-1}(0)$. A bijective ring homomorphism $\varphi : A \to B$ is called an *isomorphism*, in this case we say that A and B are *isomorphic*. A ring isomorphism $\varphi : A \to A$ is called an *automorphism*.

Remark 2.2.
(1) The identity map $id : R \to R; a \mapsto a$ is a ring isomorphism.

(2) If $\varphi : A \to B$ is a is an isomorphism, then also the inverse map φ^{-1} is a ring isomorphism.
(3) The composition of ring homomorphisms $\varphi : A \to B$, $\psi : B \to C$ is a ring homomorphism $\psi \circ \varphi : A \to C$.

REMARK 2.3. A ring homomorphism $\varphi : A \to B$ is in particular a homomorphism from the additive group of A to that of B. Thus results about group homomorphisms apply to ring homomorphisms. In particular φ is injective if and only of $ker(\varphi) = 0$.

EXAMPLE 2.4. (1) The multiplication $\cdot 2 : \mathbb{Z} \to \mathbb{Z}, n \mapsto 2n$ by 2 is an injective ring homomorphism. The image is $2\mathbb{Z}$.
(2) The residue map $mod_m : \mathbb{Z} \to \mathbb{Z}/m\mathbb{Z}, n \mapsto \underline{n}$ sending n to the rest after division by m is a surjective ring homomorphism. The kernel is $m\mathbb{Z}$.

Like in these two examples, the image of a ring homomorphism $\varphi : A \to B$ is always a subring of B.

REMARK 2.5. Let $\varphi : A \to B$ be a ring homomorphism. Then $\varphi(A)$ is a subring of B.

PROOF. As φ is a homomorphism for the additive groups, we know that $\varphi(A)$ is a subgroup of the additive group of B. If $x = \varphi(a)$, $y = \varphi(b)$ are elements of $\varphi(A)$, then $xy = \varphi(a)\varphi(b) = \varphi(ab) \in \varphi(A)$. □

The kernel of a homomorphism of groups $\varphi : A \to B$ is not just a subgroup of A, but has a stronger structure: it is a normal subgroup. Also for homomorphisms of rings the kernel is not just a subring but something better; it is an ideal.

DEFINITION 2.6. A subset I of a ring R is called an *ideal*, if I is a subgroup of the additive group of R and for each $a \in I$ and all $x \in R$, we have $xa \in I$ and $ax \in I$. In particular ideals are subrings.

It is easy to see that 0, R are ideals in R and that the intersection of a set of ideals in R is an ideal in R.

DEFINITION 2.7. Let R be a commutative ring and let $a_1, \dots, a_n \in R$. The *ideal generated by* $a_1, \dots, a_n$ is the set

$$\langle a_1, \dots, a_n \rangle := \big\{a_1r_1 + \ldots + a_nr_n \bigm| r_1, \dots, r_n \in R\big\}.$$

It is easy to check that this is an ideal. In particular $\langle a \rangle = aR = \big\{ar \bigm| x \in r\big\}$ for $a \in R$ is called a *principal ideal*. E.g. $\langle n \rangle = n\mathbb{Z}$ is an ideal in $\mathbb{Z}$.

Ideals are much more important then subrings. The first reason is that the kernel of a homomorphism of rings is an ideal.

LEMMA 2.8. *Let $\varphi : A \to B$ be a homomorphism of rings. Then $ker(\varphi)$ is an ideal in A. More generally $\varphi^{-1}(I)$ is an ideal in A for any ideal $I \subset B$.*

PROOF. The first statement follows from the second by putting $I = \{0\}$.
As I is a subgroup of the additive group of B, also $\varphi^{-1}(I)$ is a subgroup of the additive group of A. For each $a \in \varphi^{-1}(I)$ and each $x \in R$ we have $\varphi(a) \in I$, $\varphi(xa) = \varphi(x)\varphi(a) \in I$ and $\varphi(ax) = \varphi(a)\varphi(x) \in I$. Thus $ax, xa \in \varphi^{-1}(I)$. □

Thus the kernel of a ring homomorphism is an ideal. We shall soon show that every ideal is the kernel of some ring homomorphism. Thus *the ideals are precisely the kernels of ring homomorphisms.* We know an analogous statement for normal subgroups: the normal subgroups are precisely the kernels of group homomorphisms. In fact *ideals in rings are like normal subgroups in groups.*

REMARK 2.9. (1) Let I be an ideal in a commutative ring R with 1. If I contains a unit, then $I = R$.
(2) The only ideals of a field K are $\{0\}$ and K.
(3) Let K be a field and $\varphi : K \to R$ be a ring homomorphism, then $\varphi = 0$ or φ is injective.

PROOF. This is all very easy. (1) Let $a \in I \cap R^*$. Then there is an element $b \in R$ with $ab = 1$. Thus $1 \in I$ and for all $x \in R$ we have $x = x1 \in I$. (2) follows from (1) because $K^* = K \setminus \{0\}$. (3) follows from (2) because $ker(\varphi)$ is an ideal of K. □

3. Quotient Rings

We said before that ideals in rings are analogous to normal subgroups in groups. If $N \subset G$ is a normal subgroup in a group, then one can form the quotient group G/N. Thus for an ideal I in a ring R we want to form the quotient ring R/I.

Let R be a ring and $I \subset R$ an ideal. As the additive group of R is commutative, I is a normal subgroup in $(R, +)$. Thus we can form the quotient group

$$R/I := \{x + I \mid x \in R\}.$$

Here $x + I = \{x + a \mid a \in I\}$. As R is commutative, R/I is a commutative group with addition

$$(x + I) + (y + I) = (x + y) + I$$

and neutral element $0 + I$.

We will look at this in a slightly different way. On R we have an equivalence relation

$$x \sim y \iff x - y \in I.$$

(Check this is an equivalence relation). The equivalence classes are precisely the $x + I$, $x \in R$. Thus R/I is the set of equivalence classes for this equivalence relation. We will usually write $[x]$ instead of $x + I$ for the equivalence class. We have a canonical surjective group homomorphism $\pi : R \to R/I; x \mapsto [x]$ for the additive group. Now we want to see that R/I is also in a natural way a ring and π is a surjective ring homomorphism.

THEOREM 3.1. *Let R be a ring and let $I \subset R$ be an ideal. R/I with the operations*

$$[x] + [y] = [x + y], \quad [x][y] = [xy]$$

is a ring and the natural projection $\pi : R \to R/I; x \to [x]$ is a ring homomorphism.

PROOF. We already know that $(R/I, +)$ is a commutative group, and π a group homomorphism. We need to show that the product is well defined (i.e. does not depend on the choice of x and y). If $[x] = [x']$ and $[y] = [y']$, then $x - x', y - y' \in I$. Thus

$$xy - x'y' = (x - x')y - x'(y' - y) \in I,$$

thus $[x'y'] = [xy]$ and the product is well-defined. The associativity and the distributive law for R/I follow directly from that for R. I check this for the associativity:

$$[x]([y][z]) = [x][yz] = [xyz] = [xy][z] = ([x][y])[z].$$

We already know that π is a group homomorphism for the additive groups and we see $\pi(xy) = [xy] = [x][y] = \pi(x)\pi(y)$. $\square$

So we have seen that we can quotient out a ring by an ideal. As $\pi : R \to R/I$ is a ring homomorphism, this shows that the ideals of R are precisely the kernels of ring homomorphisms starting from R. So ideals have two important functions: They are the kernels of ring homomorphisms, and we can take the quotient ring by ideals.

Now we want to show that if we have a homomorphism $\varphi : A \to B$ of rings, and I an ideal contained in the kernel of φ, then we get an induced ring homomorphism $\overline{\varphi} : A/I \to B$. We say that φ can be *factored through* A/I. This property in fact determines A/I up to isomorphism.

THEOREM 3.2. *(Universal property). Let $\varphi : A \to B$ be a ring homomorphism and let $I \subset A$ be an ideal contained in $ker(\varphi)$. Then there exists a unique ring homomorphism $\overline{\varphi} : A/I \to B$, so that the diagram*

$$\begin{tikzcd} A \arrow[r, "\varphi"] \arrow[d, "\pi"] & B \\ A/I \arrow[ur, "\overline{\varphi}"'] & \end{tikzcd}$$

commutes. Furthermore $ker(\overline{\varphi}) = ker(\varphi)/I$ and $\overline{\varphi}(R/I) = \varphi(R)$.

PROOF. Again this is very easy. $\varphi = \overline{\varphi} \circ \pi$ means that $\overline{\varphi}([x]) = \varphi(x)$ for all $x \in A$. Thus $\overline{\varphi}$ is unique. Now we need to show that this formula gives a well-defined ring homomorphism, which is again a straightforward verification. If $[x] = [y]$, then $x - y \in I$. Thus $\varphi(x - y) = 0$ and $\varphi(x) = \varphi(y)$. This shows $\overline{\varphi}$ is well-defined.

$$\overline{\varphi}([x] + [y]) = \overline{\varphi}([x + y]) = \varphi(x + y) = \varphi(x) + \varphi(y) = \overline{\varphi}([x]) + \overline{\varphi}([y]),$$

and similarly $\overline{\varphi}([x][y]) = \overline{\varphi}([x])\overline{\varphi}([y])$. Thus $\overline{\varphi}$ is a ring homomorphism. $\overline{\varphi}(A) = \varphi(A)$ is obvious from the definition and

$$[x] \in ker(\overline{\varphi}) \iff x \in ker(\varphi) \iff [x] \in ker(\varphi)/I.$$

$\square$

The most important case of this is the Homomorphism Theorem: for a surjective ring homomorphism $\varphi : A \to B$, B is isomorphic to A/I.

THEOREM 3.3. *(Homomorphism Theorem) Let $\varphi : A \to B$ be a surjective ring homomorphism with kernel I. Then the map*

$$\overline{\varphi} : A/I \to B; [x] \mapsto \varphi(x)$$

is a ring isomorphism.

PROOF. By the universal property of the quotient $\overline{\varphi}$ is a surjective ring homomorphism, with kernel $I/I = [0]$. Thus it is an isomorphism. $\square$

The Homomorphism Theorem is very useful, we will apply it very often. Now we can compare the ideals in A and A/I.

THEOREM 3.4. *Let K be an ideal in a ring R, then the map $I \mapsto I/K$ is a bijection between ideals in R containing K and the ideals in R/K. It inverse is $J \mapsto \pi^{-1}(J)$.*

PROOF. As the canonical projection $\pi : R \to R/K$ is a ring homomorphism, $\pi^{-1}(J)$ is an ideal in R containing K for any ideal $J \subset R/K$.

Now let $I \subset R$ be an ideal containing K. Then I/K is an ideal in R/K: As π is a homomorphism of the additive groups, $\pi(I) = I/K$ is a subgroup if R/I. Now let $a \in I$ and $x \in R$. Then $xa \in I$ and thus $[x][a] = [xa] \in I/K$ and similarly $bx \in I/K$. Thus I/K is an ideal in R/I.

For any ideal J in R/I we have $\pi(\pi^{-1}(J)) = J$. Let $I \subset R$ be an ideal with $K \subset I$, then

$$\pi^{-1}(I/K) = \{a \mid [a] \in I/K\} = \{a + k \mid a \in I,\ k \in K\} = I.$$

□

4. Prime Ideals and Maximal Ideals

From now on until the end of the lectures a ring is a commutative ring with 1 and a homomorphism $\varphi : A \to B$ of rings satisfies $\varphi(1) = 1$.

We now want to look at the two most important classes of ideals. We have learned in the last section, that for an ideal I in a ring R the natural thing to do is to form the quotient ring R/I. The nicest ideals will be those for which the quotient is the nicest. The nicest rings are the fields, where we can compute like in $\mathbb{Q}$, followed by the integral domains, where we can compute like in $\mathbb{Z}$. A prime ideal in a ring R will be an ideal P so that R/P is an integral domain, and a maximal ideal M in R will be such that R/M is a field.

DEFINITION 4.1. Let R be a ring (remember this means commutative with 1). An ideal $P \subsetneq R$ is called a *prime ideal* if for all $a, b \in R$ with $ab \in P$, we have $a \in P$ or $b \in P$.

EXAMPLE 4.2. Let $m \in \mathbb{Z}_{>0}$. Then $\langle m \rangle = m\mathbb{Z}$ is a prime ideal in $\mathbb{Z}$ if and only if m is a prime number. This is the motivation for the name prime ideal. If m is a prime number and $kl \in \langle m \rangle$, then m divides kl, and thus m divides k or m divides l. Thus $k \in \langle m \rangle$ or $l \in \langle m \rangle$. If m is not a prime number, then either $m = 1$ and $\langle m \rangle = \mathbb{Z}$ or otherwise there are integers $1 < k, l < m$ with $kl = m$. Thus $k, l \notin \langle m \rangle$, but $kl = m \in \langle m \rangle$.

A maximal ideal is a proper ideal that does not fit into any bigger proper ideal.

DEFINITION 4.3. Let R be a ring. An ideal $M \subsetneq R$ is called *maximal* if there is no ideal I with $M \subsetneq I \subsetneq R$.

THEOREM 4.4.
(1) *Let R be a ring. An ideal $P \subset R$ is a prime ideal if and only if R/P is an integral domain.*
(2) *Let R be a ring. An ideal $M \subset R$ is a maximal ideal if and only if R/M is a field.*

Proof. (1) As R is commutative with 1 also R/P is commutative with 1. For $a, b \in R$, $[a] = 0$ if and only if $a \in P$; and $0 = [a][b] = [ab]$ only if $ab \in P$. Thus $[a]$ is a zero divisor if and only if $a \notin P$ and there is an element $b \notin P$ with $ab \in P$.

(2) Suppose first M is an ideal in R such that R/M is a field. As R/M is a field, the only ideals in R/M are 0 and R/M. By the bijection of ideals in R/M and ideals in R containing M, the only ideals in R containing M are M and R. Thus M is a maximal ideal. If on the other hand M is maximal ideal, by the same bijection, the only ideals in R/M are 0 and R/M. Thus the Theorem follows from the next Lemma. □

Lemma 4.5. *A commutative ring R with 1 whose only ideals are $\{0\}$ and R is a field.*

Proof. For every $a \in R \setminus \{0\}$ we need to find an element $b \in R$ with $ab = 1$. $\langle a \rangle := \{ab \mid b \in R\}$ is an ideal in R, thus $\langle a \rangle = 0$ or $\langle a \rangle = R$. Since $0 \neq a = 1a \in \langle a \rangle$ we see that $\langle a \rangle = R$. Therefore $1 \in \langle a \rangle$, i.e. there exists $b \in R$ with $ab = 1$. □

As a field is in particular an integral domain we also get.

Corollary 4.6. *Every maximal ideal is a prime ideal.*

Example 4.7. (1) If K is a field, then $\{0\}$ is a maximal ideal in K.
(2) We have seen that $\langle m \rangle$ is a prime ideal in $\mathbb{Z}$ if and only if m is a prime number. But in this case $\langle m \rangle$ is even a maximal ideal because $\mathbb{Z}/\langle m \rangle$ is a field.
(3) With the Lemma of Zorn one shows that every ideal in a ring is contained in a maximal ideal. We will not go into this here.

5. Polynomial rings over a field

We have already introduced the polynomial rings $R[X]$ in the first section. In this section we want to study the case of polynomial rings over a field in more detail. In this section let k be a field. We want to start out by looking at questions of divisibility in $k[X]$. So we first define this in general.

Definition 5.1. Let R be an integral domain. Let $a, b \in R$. We say that *a divides b*, denoted $a|b$, if there exists an element $c \in R$ with $b = ac$. Otherwise we write $a \not| b$.

The relation "divides" has a number of obvious properties.

Remark 5.2. $a|b$ is equivalent to $b \in \langle a \rangle$, in particular $a|b$ and $b|c$ implies $a|c$; and $a|b$ and $a|c$ implies $a|(b \pm c)$.

Questions of divisibility of polynomials over k are governed by the algorithm of division with rest, that works in $k[X]$ in a very similar way as for integers.

Theorem 5.3. *(Division with rest) Let $f, g \in k[X]$ with $g \neq 0$. There is are unique $q \in k[X]$ and $r \in k[X]$ such that*

$$f = qg + r, \quad \text{with } deg(r) < deg(g) \text{ or } r = 0.$$

PROOF. Existence: If $f = 0$ or $deg(f) < deg(g)$, we can put $q = 0$ and $r = f$. Thus assume that $m := deg(f) \geq deg(g)$ and make induction on m.

Let a be the leading coefficient of f and b the leading coefficient of g. Then the coefficient of X^m in

$$\overline{f} := f - (\frac{a}{b} X^{m-deg(g)})g$$

is $a - ab/b = 0$, thus $\overline{f}$ has degree $\leq m - 1$, and by induction we can write $\overline{f} = q'g + r'$ with $r' = 0$ or $deg(r') < deg(g)$. But then we put $r := r'$ and $q = q' + (\frac{a}{b} X^{m-deg(g)})$, and the claim is satisfied.

Uniqueness: If $qg + r = q'g + r'$ with the above properties, then $(q - q')g = r - r'$. $q \neq q'$ would imply

$$deg(r - r') = deg(q - q') + deg(g) \geq deg(g),$$

a contradiction to $deg(r), deg(r') < deg(g)$. Thus $q = q'$ and thus $r = r'$. □

In $\mathbb{Z}$ we can talk about the greatest common divisors. Now we introduce greatest common divisors in any integral domain R.

DEFINITION 5.4. Let R be an integral domain and let $a_1, \dots, a_n \in R$. An element $r \in R \setminus \{0\}$ is called a *common divisor* of $a_1, \dots, a_n$, if $r|a_i$ for $i = 1, \dots, n$. An element $r \in R$ is called a *greatest common divisor* of $a_1, \dots, a_n$, if r is a common divisor, and $s|r$ for any other common divisor. $a_1, \dots, a_r$ are called *relatively prime*, if 1 is a greatest common divisor.

By definition the greatest common divisor of $a_1, \dots, a_r$ is not unique, but it is unique up to multiplication by a unit. In $k[X]$ the division algorithm gives an algorithm to find the greatest common divisor.

REMARK 5.5. Let $f, g \in k[X]$. We put $r_0 := f$, $r_1 := g$. Then by the division algorithm we find $f = q_1 g + r_2$, $g = q_2 r_2 + r_3$ and inductively

$$r_{i-1} = q_i r_i + r_{i+1}, \quad \text{with } q_i \in k[X],\ r_{i+1} \in k[X], \quad \text{and } deg(r_{i+1}) < deg(r_i).$$

The process stops when $r_{n-1} = q_n r_n$ and $r_n \neq 0$.

We claim that then r_n is a greatest common divisor of f and g: The equation $r_{i-1} = q_i r_i + r_{i+1}$ implies sucessively that $r_n|r_{n-1},\ r_n|r_{n-2}, \dots, r_n|g,\ r_n|f$. Thus r_n is a common divisor of f and g. On the other hand, if t is a common divisor of f and g, then the same equation implies sucessively that $t|r_2,\ t|r_3, \dots, t|r_n$. Thus r_n is a greatest common divisor. This algorithm is called the *Euclidean algorithm.*

It implies one further result, which we will use in the future (in the proof of the theorem of the primitive element). If K is another field, so that $k \subset K$ is a subring then we call k a subfield of K. It is straightforward to see that then $k[X]$ is a subring of $K[X]$.

REMARK 5.6. Let K be a field and $k \subset K$ a subfield. Let $f, g \in k[X]$. Let h be the greatest common divisor of f and g in $k[X]$ computed with the Euclidean algorithm. Then h is also a greatest common divisor of f and g in $K[X]$. This is in fact obvious, because in the division with rest $f = qg + r$ the quotient q and the rest r are determined uniquely by f and g, independent of the field in which we do the computation. Thus the Euclidean algorithm in $k[X]$ and $K[X]$ are the same.

A very important property of polynomials is that we can evaluate them at any element of k and more generally at any element of a field K that contains k as a subfield, i.e. we can substitute for X any element of k or of K.

DEFINITION 5.7. let $f = \sum_{i=0}^{n} a_i X^i \in k[X]$ and let K be a field, that contains k as a subfield. The *value* of f in s is

$$f(s) := \sum_{i=0}^{n} a_i s^i \in K.$$

It is straightforward to check that

$$(f+g)(s) = f(s) + g(s), \quad fg(s) = f(s)g(s).$$

Thus we get a ring homomorphism

$$ev_s : k[X] \to K, f \mapsto f(s).$$

ev_s is called the *evaluation homomorphism* at s. An element $s \in S$ is called a *zero* of f if $f(s) = 0$.

REMARK 5.8. For the moment the most important case will be that $K = k$, so that we can evaluate a polynomial in $k[X]$ at any element of k. But we shall see later that the general case is important.

Now we want to use the division with rest to study zeros of polynomials in $k[X]$. If $a \in k$ is a zero of $f \in k[X]$, then we can divide f by $X - a$. This is then used to show that if f has degree n, then it can have at most n zeros in k.

THEOREM 5.9. *Let $f \in k[X]$ and $a \in k$. Then a is a zero of f, if and only if $(X-a)|f$, i.e. if $f = (X-a)g$ for some $g \in k[X]$.*

PROOF. If $f = (X-a)g$, then $f(a) = (a-a)g(a) = 0$. Thus assume $f(a) = 0$. If $f = 0$, we can put $g = 0$. Otherwise by division with rest we can write $f = (X-a)g + r$ with $deg(r) = 0$ or $r = 0$. Thus $r \in k$. Evaluating at a we get

$$0 = f(a) = (a-a)g(a) + r = r,$$

i.e. $f = (X-a)g$. □

THEOREM 5.10. *Let $f \in k[X] \setminus \{0\}$. Then f has at most $deg(f)$ zeros in k.*

PROOF. We show this by induction on the degree of f. If f has degree 0, then f is constant, and thus has no zero. Now let f have degree $n+1$. If f has no zero in k, then the claim is trivially true. Otherwise let $a \in k$ be a zero of f. Then we can write $f = (X-a)g$ with $deg(g) = n$. Thus by induction g has at most n zeros in k. If $b \neq a$ is a zero of f, then $0 = (b-a)g(b)$, thus $g(b) = 0$. So f has at most $n+1$ zeros in k. □

6. Euclidean rings and principal ideal domains

An important property of the integers is that in the integers we can do division with rest, i.e. if a, b are integers, we can write

$$a = tb + r, \quad t, r \in \mathbb{Z}, \quad 0 \leq r < b.$$

r is the rest of the division of a by b. In the last section we have seen that also in polynomial rings over a field we have division with rest. Now we want to formalize this. The rest r should in some sense be smaller then b, in general rings we have to measure this by a function $d : R \setminus \{0\} \to \mathbb{Z}_{>0}$.

DEFINITION 6.1. An integral domain R is called a *Euclidean ring* if there is a function $d : R \setminus \{0\} \to \mathbb{Z}_{\geq 0}$ with the following property. For all a, b in $R \setminus \{0\}$ there exist $q, r \in R$ such that $a = qb + r$ with either $r = 0$ or $d(r) < d(b)$.

EXAMPLE 6.2.
(1) $\mathbb{Z}$ with $d(n) = |n|$ is a Euclidean ring.
(2) Let again $\mathbb{Z}[i] := \{n + mi \mid n, m \in \mathbb{Z}\} \subset \mathbb{C}$ be the ring of Gaussian integers. We claim that $\mathbb{Z}[i]$ is a Euclidean ring with $d(n + im) = n^2 + m^2$. We can extend d to $\mathbb{C}$, by $d(a + ib) = a^2 + b^2$, d is just the square of the complex absolute value. In particular $d(a + ib) \neq 0$ for $a + ib \neq 0$. We then have $d(zw) = d(z)d(w)$ for all $z, w \in \mathbb{C}$. If $z, w \in \mathbb{Z}[i] \setminus \{0\}$, let $z/w = a + ib$ the the quotient in $\mathbb{C}$. Choose $n, m \in \mathbb{Z}$ such that $|a - m| \leq 1/2$, $|b - n| \leq 1/2$. Then
$$d(z/w - (m + in)) = (a - m)^2 + (b - n)^2 \leq 1/2.$$
Thus we can put $r := z - (m + in)w$ as the rest of the division, and get $d(r) = d(w)d(z/w - (m + in)) < d(w)$.
(3) The most important example of a Euclidean ring is that of a polynomial ring over a field. Let k be a field, and for $f \in k[X] \setminus \{0\}$ put $d(f) = deg(f)$. Then $(k[X], d)$ is a Euclidean ring. This is just a reformulation of Theorem 5.3.

Euclidean rings have another nice property. The set of their ideals is very simple: every ideal $I \subset R$ is a principal ideal.

DEFINITION 6.3. Let R be an integral domain. Recall that an ideal of the form $\langle a \rangle = \{ar \mid r \in R\}$ for an $a \in R$ is called a *principal ideal. An integral domain R is called a principal ideal domain (PID), if every ideal $I \subset R$ is a principal ideal.*

THEOREM 6.4. *A Euclidean ring is a principal ideal domain.*

PROOF. Let (R, d) be a Euclidean ring. We have to show that every ideal $I \subset R$ is a principal ideal. The zero ideal $\{0\} = \langle 0 \rangle$ is principal. Thus let $I \neq \{0\}$ be an ideal. Let $a \in I \setminus \{0\}$ be such that $d(a)$ is the smallest $d(b)$ for $b \in I \setminus \{0\}$. We want to show $I = \langle a \rangle$: Otherwise there is an element $b \in I \setminus \langle a \rangle$. Then division with rest gives

$$b = ta + r, \quad t, r \in R \setminus \{0\}, \quad d(r) < d(a).$$

This is a contradiction to the choice of a, because $r = b - ta \in I$. □

COROLLARY 6.5. *$\mathbb{Z}$, $\mathbb{Z}[i]$, and $k[X]$ for a field k are principal ideal domains.*

We see in particular that the ideals in $\mathbb{Z}$ are precisely the $m\mathbb{Z}$ that we already know and the quotient rings are precisely the $\mathbb{Z}/m\mathbb{Z}$.

In general integral domains greatest common divisors do not always need to exist. However in a PID they do.

THEOREM 6.6. *Let R be a principal ideal domain. Let $a_1, \dots, a_r \in R$. Then there is greatest common divisor d of $a_1, \dots, a_r$ of the form $d = a_1x_1 + \dots + a_rx_r$ with $x_i \in R$.*

PROOF. Choose $d \in R$ such that $\langle d \rangle = \langle a_1, \dots, a_r \rangle$. Then by definition $d = a_1x_1 + \dots + a_rx_r$ with $x_i \in R$. We claim that d is a greatest common divisor. By definition $a_i \in \langle d \rangle$ for all i, thus $d|a_i$. On the other hand if $e \in R$ divides all a_i, then $e|a_1x_1 + \dots + a_rx_r = d$. □

EXAMPLE 6.7. Let $a, b \in \mathbb{Z}$, then $\langle a, b \rangle = \langle gcd(a, b) \rangle$, e.g. $\langle 4, 6 \rangle = \langle 2 \rangle$ and $\langle 4, 7 \rangle = \mathbb{Z}$.

7. Prime elements and irreducible elements

In this section we study the analogue of prime numbers in $\mathbb{Z}$ in integral domains. Let R be an integral domain.

DEFINITION 7.1. We say that $a, b \in R$ are associates if $a = ub$ for u a unit in R^*.

Being associates is obviously an equivalence relation.

EXAMPLE 7.2. In $\mathbb{Z}$ the associates of n are $\{n, -n\}$; in $Z[i]$ they are $\{n, -n, in, -in\}$.

Now we will try to define the analogues of prime numbers. For a number $p \in \mathbb{Z} \setminus \{1, -1\}$ being an associate to a prime number $p > 0$ is equivalent to any of the following.

(1) if $p = ab$, then $a = \pm 1$ or $b = \pm 1$,
(2) if p divides a product ab, then $p|a$ or $p|b$.

Unfortunately in a general integral domain these two properties are not equivalent.

DEFINITION 7.3. Let R be an integral domain. An element $p \in R \setminus 0$ is called *prime element* if p is not a unit and for a, b in R if $p|ab$, then $p|a$ or $p|b$.

An element $q \in R \setminus \{0\}$ is called *irreducible* if q is not a unit and for any $a, b \in R$ with $q = ab$ either a or b is a unit. An element of R is called reducible, if it is not irreducible.

EXAMPLE 7.4. (1) A field has no prime elements and no irreducible elements.
(2) $m \in \mathbb{Z}$ is a prime element if and only if it is an irreducible element if and only if $|m|$ is a prime number.
(3) If k is a field and $a \in k \setminus \{0\}$, $b \in k$, then $aX + b$ is an irreducible element of $k[X]$: If $aX + b = fg$, then $deg(f) + deg(g) = 1$, thus $deg(f) = 0$ or $deg(g) = 0$, thus $f \in k^*$ or $g \in k^*$.
(4) The irreducible elements in $R[X]$ are called irreducible polynomials.

An irreducible element is one that cannot in a nontrivial way be written as a product of other elements, and a prime element has to divide one factor if it divides a product. Prime elements are more important then irreducible elements, because they generate prime ideals. Later we will study unique factorization domains, in which prime elements and irreducible elements are the same.

REMARK 7.5. (1) Every prime element is irreducible.

(2) An element $p \in R \setminus \{0\}$ is a prime element if and only if $\langle p \rangle$ is a prime ideal.

PROOF. (1) Let p be a prime element. Assume $p = ab$. Then, as p is a prime element $p|a$ or $p|b$. Assume $p|a$. Then $a = pc$ for $c \in R$, thus $p = pcb$. Thus $cb = 1$ and b is a unit.

(2) p is not a unit if and only if $\langle p \rangle \neq R$. We know that $c \in \langle p \rangle$ if and only if $p|c$. Thus $ab \in \langle p \rangle$ if and only if $p|ab$ and $a, b \neq \langle p \rangle$ if and only if $p \not| a$ and $p \not| b$. □

Now we want to see that in a principal ideal domain prime elements and irreducible elements are the same, and generate even maximal ideals.

PROPOSITION 7.6. *Let R be a principal ideal domain and $p \in R$ an irreducible element. Then p is a prime element and $\langle p \rangle$ is a maximal ideal.*

PROOF. Let $p \in R$ be an irreducible element. It is enough to show that $\langle p \rangle$ is a maximal ideal, then it is in particular a prime ideal and thus p is a prime element.

We know that every ideal in R is of the form $\langle a \rangle$. Thus let $\langle p \rangle \subset \langle a \rangle$. Then $p \in \langle a \rangle$, thus $p = ab$ for some $b \in R$. Then, as p is irreducible, either a is a unit and $\langle a \rangle = R$ or b is a unit and $\langle a \rangle = \langle p \rangle$. Thus $\langle p \rangle$ is maximal. □

This gives us a way to construct new fields. We will use this in the chapter on fields.

COROLLARY 7.7. *Let k be a field and let f in $k[X]$ be an irreducible polynomial.*

Then $\langle f \rangle$ is a maximal ideal in $k[X]$ and $k[X]/\langle f \rangle$ is a field containing k as a subfield (as the image of the constant polynomials).

Now we want to see an example where not every irreducible element is a prime element.

EXAMPLE 7.8. Let $R := \mathbb{Z}[\sqrt{-5}] := \{m + in\sqrt{5} \in \mathbb{C} \mid m, n \in \mathbb{Z}\}$. This is a subring of $\mathbb{C}$, and thus it is an integral domain. We can restrict the square of the complex absolute value to $\mathbb{Z}[i\sqrt{5}]$ to get $|\ |^2 : \mathbb{Z}[i\sqrt{5}] \to \mathbb{Z}; m + in\sqrt{5} \mapsto m^2 + 5n^2$. Then $|xy|^2 = |x|^2|y|^2$ for $x, y \in R$.

(1) First we show that $R^* = \{1, -1\}$. Obviously $1, -1$ are units in R. Let $u = m + in\sqrt{5}$ by a unit. Then $|u|^2 = m^2 + 5n^2$ divides $|1|^2 = 1$. Thus $n = 0$ and $m^2 = 1$, i.e. $u = \pm 1$.

(2) We want to show that 3, $2 + i\sqrt{5}$, $2 - i\sqrt{5}$ are pairwise nonassociated irreducible elements. Let z be one of them. Then $|z|^2 = 9$. Thus if $x|z$, then we can write $z = xy$ and $|x|^2|y|^2 = 9$. As $|x|^2 = 3$ is impossible, it follows that $|x|^2 = 1$ or $|y|^2 = 1$, and thus $x = \pm 1$ or $x = \pm z$. Thus z is irreducible.

On the other hand $9 = 3 \cdot 3 = (2 + i\sqrt{5})(2 - i\sqrt{5})$. In particular $3|(2 + i\sqrt{5})(2 - i\sqrt{5})$ but $3 \not| (2 + i\sqrt{5})$ and $3|(2 - i\sqrt{5})$. Thus 3 is not a prime element in R. Similarly $(2 + i\sqrt{5})$ and $(2 - i\sqrt{5})$ are not prime elements. This example also shows something else. We can write 9 in two completely different ways as product of irreducible elements.

8. Irreducibility of polynomials over $\mathbb{Q}$.

Now we study irreducibility of polynomials over $\mathbb{Q}$ and over $\mathbb{Z}$. For the last part of the course about fields, we mostly need the case of $\mathbb{Q}$.

DEFINITION 8.1. Let R be an integral domain. A polynomial $f = \sum_{i=0}^{n} a_i X^i \in R[X]$ is called *primitive* if the coefficients $a_0, \dots, a_n$ are relatively prime.

If $f \in \mathbb{Q}[X]$, we can first multiply f with the denominators of all the coefficients, and then divide by the greatest common divisor of the new coefficients. In this way we get a primitive $af \in \mathbb{Z}[X]$.

LEMMA 8.2. *(Lemma of Gauss) Let $f, g \in \mathbb{Z}[X]$ be primitive. Then also fg is primitive.*

PROOF. Let $f = \sum_{i=0}^{n} a_i X^i$, $g = \sum_{i=0}^{m} b_i X^i$. Suppose that fg is not primitive. Then there is a prime number p which is a common divisor of the coefficients of fg. Let a_i be the lowest coefficient of f that p does not divide and b_j the lowest coefficient of g that p does not divide. Then in fg the coefficient of X^{i+j} is

$$c_{i+j} = a_i b_j + \sum_{k>i} a_k b_{i+j-k} + \sum_{k<i} a_k b_{i+j-k}.$$

Then p divides b_{i+j-k} for $k > i$ and p divides a_k for $k < i$. Thus p divides both sums. As p does not divide $a_i b_j$, we get p does not divide c_{i+j}. □

THEOREM 8.3. *(Gauss Lemma) Let $f \in \mathbb{Z}[X]$ be a non-constant primitive polynomial. Then f is irreducible in $\mathbb{Z}[X]$ if and only if it is irreducible in $\mathbb{Q}[X]$.*

PROOF. If f is reducible in $\mathbb{Z}[X]$, then, because f is primitive, we have $f = gh$ with $deg(g) > 0, deg(h) > 0$. Thus it is also reducible in $\mathbb{Q}[X]$.

Suppose $f = gh$ where $g, h \in \mathbb{Q}[X]$ are polynomials of positive degree. By clearing denominators and dividing by the greatest common divisor of the coefficients, we can write $f = \frac{a}{b} g' h'$, where g', h' are primitive polynomials in $\mathbb{Z}[X]$ and $a, b \in \mathbb{Z}$ are relatively prime. Thus $bf = ag'h'$. But both f and $g'h'$ are primitive. So on the left hand side the greatest common divisor of the coefficients is a and on the right hand side it is b. Thus $a = \pm b$, and thus $f = \pm g'h'$ is reducible in $\mathbb{Z}[X]$. □

Now we give a useful criterion for the irreducibility of polynomials.

THEOREM 8.4. *(Eisenstein's criterion) Let $f = \sum_{i=0}^{n} a_i X^i$ be a primitive element in $\mathbb{Z}[X]$ of positive degree. Assume there is a prime number p with $p|a_0, p|a_1, \dots, p|a_{n-1}$ but $p \not| a_n$ and $p^2 \not| a_0$. Then f is irreducible over $\mathbb{Z}$ and thus over $\mathbb{Q}$.*

PROOF. Assume $f = gh$ with $g, h \in \mathbb{Z}[X]$. We have to show $g \in \mathbb{Z}^*$ or $h \in \mathbb{Z}^*$. Let $g = \sum_{i=0}^{k} b_i X^i$ with $b_k \neq 0$ and $h = \sum_{i=0}^{l} c_i X^i$ with $c_l \neq 0$. As $a_0 = b_0 c_0$ and $p|a_0$ and $p^2 \not| a_0$, we see that p divides exactly one of b_0, c_0. Assume $p|b_0$ and $p \not| c_0$. As $a_n = b_k c_l$ and $p \not| a_n$ we see that $p \not| b_k$. Thus there exists a maximal $j \in \{1, \dots, k\}$ such that $p|b_i$ for $i < j$ and $p \not| b_j$. Putting $c_i := 0$ for $i > l$, we get

$$a_j = b_j c_0 + b_{j-1} c_1 + \ldots + b_0 c_j.$$

By definition p does not divide $b_j c_0$, but it divides each other summand. Thus $p \not| a_j$. Thus $j = n$ and thus $k = n$. Thus $deg(g) = n$ and therefore h is a constant polynomial. As f is primitive, we get $h \in \mathbb{Z}^*$. □

EXAMPLE 8.5.
(1) $f := X^5 - 9X + 3$ is irreducible in $\mathbb{Z}[X]$ by Eisenstein's criterion with $p = 3$. Thus by Gauss Lemma it is irreducible in $\mathbb{Q}[X]$.
(2) For p a prime number and $n \in \mathbb{Z}_{>0}$ the polynomial $X^n - p$ is irreducible over $\mathbb{Q}$.

9. Unique factorization domains

One of the standard facts of elementary mathematics is that every integer $n \geq 2$ is in a unique way a product of prime numbers $p_j > 1$. We want to formalize this property.

DEFINITION 9.1. An integral domain R is called a *unique factorization domain* (UFD) if every nonunit $a \in R \setminus \{0\}$ is a product $a = q_1 \cdot \ldots \cdot q_r$ of irreducible elements, and furthermore this decomposition is unique up to order and replacing some of the q_i by associates.

EXAMPLE 9.2. (1) We have seen that $\mathbb{Z}$ is a UFD.
(2) Because in $\mathbb{Z}[\sqrt{-5}]$ we have $9 = 3 \cdot 3 = (2 + i\sqrt{5})(2 - i\sqrt{5})$, we see that $\mathbb{Z}[\sqrt{-5}]$ is not a UFD.

We have formulated this definition in terms of irreducible elements. But in fact in a UFD prime elements and irreducible elements are the same.

PROPOSITION 9.3. *In a UFD every irreducible element is a prime element.*

PROOF. Let q be an irreducible element in a UFD R, and assume $q|ab$ with $a, b \in R$. We have to show $q|a$ or $q|b$. If $a = 0$, $b = 0$, $a \in R^*$ or $b \in R^*$ this is obvious. Assume there is a nonunit $c \in R \setminus \{0\}$ with such that $ab = qc$. Then we can write

$$a = p_1 \cdot \ldots \cdot p_r, \; b = p'_1 \cdot \ldots \cdot p'_s, \; c = p''_1 \cdot \ldots \cdot p''_t$$

for p_i, p'_i, p''_i irreducible elements. By the uniqueness of the decomposition q must be an associate of one of p_i or one of the p'_i. Thus $q|a$ or $q|b$. □

Thus in a UFD every nonunit has a decomposition into prime elements.

EXERCISE 9.4. Check that for R to be a UFD one does not need the uniqueness of the decomposition into prime elements. If every nonunit $a \in R \setminus \{0\}$ is a product $a = p_1 \cdot \ldots \cdot p_r$ of prime elements, then R is a UFD.

THEOREM 9.5. *Every principal ideal domain is a UFD.*

PROOF. Let R be a PID. We know that in a PID every irreducible element is a prime element. Thus by the exercise above it is enough to show that every nonunit $a \in R \setminus \{0\}$ can be written as a product of irreducible elements. First we show that every nonunit $a \in R \setminus \{0\}$ has an irreducible divisor. We argue by contradiction. Assume b does not have an irreducible divisor. Then b is reducible and we can write $b = b_1 b'_1$ with b_1 again reducible, and we can write $b_1 = b_2 b'_2$ with b_2 reducible and so forth. Thus we get a infinite chain of ideals $\langle b \rangle \subsetneq \langle b_1 \rangle \subsetneq \ldots \subsetneq \langle b_n \rangle \subsetneq \ldots$. Let $I = \bigcup_{i \in \mathbb{Z}_{>0}} \langle b_i \rangle$. Then I is an ideal. Thus as R is a PID, we can write $I = \langle c \rangle$, for some element $c \in I$. But then c must lie in one of the $\langle b_i \rangle$, thus $I = \langle b_i \rangle$ and $\langle b_j \rangle = \langle b_i \rangle$ for $j > i$, a contradiction.

Now assume there is a nonunit $a \in R \setminus \{0\}$, which is not a product of irreducible elements. Write $a = q_1 b_1$ for q_1 an irreducible divisor of a. Then b_1 is not a unit and not irreducible. So we write $b_1 = q_2 b_2$ with q_2 irreducible and b_2 not a unit and not irreducible. Again we have an infinite chain of ideals $\langle b \rangle \subsetneq \langle b_1 \rangle \subsetneq \ldots \subsetneq \langle b_n \rangle \subsetneq \ldots$. Let $I = \bigcup_{i \in \mathbb{Z}_{>0}} \langle b_i \rangle$. Then I is an ideal. Thus as R is a PID, we can write $I = \langle c \rangle$, for some element $c \in I$. But then c must lie in one of the $\langle b_i \rangle$, thus $I = \langle b_i \rangle$ and $\langle b_j \rangle = \langle b_i \rangle$ for $j > i$, a contradiction. □

COROLLARY 9.6. (1) *Every Euclidean ring is a UFD.*

(2) *In particular if k is a field, then $k[X]$ is a UFD.*

We now want to see that also in a UFD greatest common divisors exist.

THEOREM 9.7. *Let R be a UFD and $a_1, \dots, a_n \in R$. Then $a_1, \dots, a_n$ have a greatest common divisor.*

PROOF. We can assume that $a_1, \dots, a_r$ are nonunits in $R \setminus \{0\}$. As R is a UFD we can write each a_i as $a_i = u_i p_1^{k_{1i}} \dots p_r^{k_{ri}}$ for u_i a unit $p_1, \dots, p_r$ irreducible elements and $k_{ji} \in \mathbb{Z}_{\geq 0}$. For each j put $m_j = min\{k_{ji} \mid i = 1, \dots n\}$. Then $d := p_1^{m_1} \dots p_r^{m_r}$ is a greatest common divisor: Obviously $d|a_i$ for all i. On the other hand if $c|a_i$ for all i, then by unique factorization $c = up_1^{l_1} \dots p_r^{l_r}$, with u a unit and $l_i \leq m_i$ for all i. Thus $c|d$. $\square$

Using that in UFDs there are greatest common divisors, we can now generalize the results about irreducibility of polynomials over $\mathbb{Z}$ and $\mathbb{Q}$ to any UFD R and its quotient field. The proofs are precisely the same as in the case of $\mathbb{Z}$ and $\mathbb{Q}$. We just have to use prime elements instead of prime numbers and replace $\mathbb{Q}$ by the quotient field of R.

LEMMA 9.8. *(Lemma of Gauss) Let R be a UFD and let $f, g \in R[X]$ be primitive. Then also fg is primitive.*

THEOREM 9.9. *(Gauss Lemma) Let R be a UFD and let $K := Q(R)$ be its quotient field. Let $f \in R[X]$ be a non-constant primitive polynomial. Then f is irreducible in $R[X]$ if and only if it is irreducible in $K[X]$.*

THEOREM 9.10. *(Eisenstein's criterion) Let R be a UFD and let $f = \sum_{i=0}^n a_i X^i$ be a primitive element in $R[X]$ of positive degree. Assume there is a prime element $p \in R$ with $p|a_0, p|a_1, \dots, p|a_{n-1}$ but $p \nmid a_n$ and $p^2 \nmid a_0$.*
Then f is irreducible over R.

Using the Theorem of Gauss it is not difficult to prove that for a UFD R also $R[X]$ is a UFD. We will only give a sketch of the proof.

THEOREM 9.11. *Let R be a UFD. Then also $R[X]$ is a UFD. By induction $R[X_1, \dots, X_n]$ is a UFD. In particular for k a field $k[X_1, \dots, X_n]$ is a UFD.*

PROOF. (Sketch) Let Q be the quotient field of R. Let $f \in R[X]$. We write $f = ag$ with $a \in R$ and g primitive. As $Q[X]$ is a UFD, g can in $Q[X]$ be written in a unique way $g = g_1 \dots g_n$ as a product of irreducible elements. By the the Gauss lemma we can assume that all g_i are primitive elements in $R[X]$. As also R is a UFD, a can uniquely be written $a = a_1 \dots a_m$ for irreducible elements in R. Then it is easy to show that $f = a_1 \dots a_m g_1 \dots g_n$ is the unique way of writing f as a product of irreducible elements. $\square$

10. More criteria for irreducibility

To study examples of field extensions it is important to be able to check the irreducibility of polynomials in $k[X]$. The most important case is again that of $k = \mathbb{Q}$. We give a few more methods to check irreducibility.

REMARK 10.1. If $f \in k[X]$ is a polynomial of degree at more 3, then f is irreducible if and only if it has no zero in k. Let in particular $f = X^n + \sum_{i=0}^{n-1} b_i X^i \in \mathbb{Z}[X]$ be a monic polynomial. Then you have shown in an exercise that any zero $a \in \mathbb{Q}$ of f must lie in $\mathbb{Z}$. It is easy to check (exercise) that then $f/(X-a) \in \mathbb{Z}[X]$ and that $a|b_n$. So if $n \leq 3$ we can check the irreducibility very fast by showing that none of the divisors of b_n is a zero of f.

Another possiblity is to substitute something else for the variable of f.

REMARK 10.2. Let $f = \sum_i a_i X^i \in \mathbb{Z}[X]$ be a primitive polynomial, let $a \in \mathbb{Z}$. Then f is irreducible in $\mathbb{Z}[X]$ (and in $\mathbb{Q}[X]$) if and only if $f(X+a) := \sum_i a_i (X+a)^i$ is irreducible in $\mathbb{Z}[X]$ (and equivalently in $\mathbb{Q}[X]$):

It is easy to check that the map $\sigma_a : \mathbb{Q}[X] \to \mathbb{Q}[X], g \mapsto g(X+a)$ is an isomorphism of rings. Thus f is irreducible if and only if $f(X+a)$ is irreducible.

EXAMPLE 10.3. Let p be a prime number. Then $f = X^{p-1} + X^{p-2} + \ldots + X + 1$ is irreducible in $\mathbb{Q}[X]$.

Note that $(X-1)f = X^p - 1$. Thus $X\sigma_1(f) = (X+1)^p - 1$ and thus $\sigma_1(f) = \sum_{i=1}^{p} \binom{p}{i} X^{i-1}$. It is easy to check that $p | \binom{p}{i}$ for $1 \leq i \leq p-1$ and $p^2 \not| \binom{p}{1} = p$. Thus $\sigma_1(f)$ is irreducible by Eisensteins criterion, and thus f is irreducible over $\mathbb{Q}$.

Finally we can use reduction modulo p.

REMARK 10.4. Let p be a prime number. Let $f = \sum_{i=0}^{n} a_i X^i \in \mathbb{Z}[X]$ be primitive polynomial, and assume $p \not| a_n$. Denote $[a_i]$ the class of a_i in F_p. Let $\bar{f} := \sum_{i=0}^{n} [a_i] X^i \in F_p[X]$. If $\bar{f}$ is irreducible in $F_p[X]$, then f is irreducible in $\mathbb{Q}[X]$.

PROOF. If f was reducible, then as it is primitive, we could write $f = gh$ with $deg(g), deg(h) < n$. Thus $\bar{f} = \bar{g}\bar{h}$, with $deg(\bar{g}), deg(\bar{h}) < n$. As $deg(\bar{f}) = n$, it follows that $deg(\bar{g}), deg((\bar{h}) \geq 1$, thus $\bar{f}$ is reducible. $\square$

The advantage of reducing modulo p is that irreducibility over F_p is easy to check: If $f \in F_p[X]$ of degree n is reducible, then there exists a monic polynomial $g \in F_p[X]$ with $1 \leq deg(g) \leq deg(f)/2$. There is only a finite number of such g (if p and n are small it is a rather small number), and one can check them all with the division algorithm.

CHAPTER 2

Fields

In this last part of the lecture we want to study fields. We have already said before that the most important class of rings are the fields. Recall that a field is a commutative ring with 1 in which every nonzero element is a unit. This means that we can compute in a field basically like in the rational numbers $\mathbb{Q}$. We will look at algebraic extensions of a field k, i.e. we will study fields $L \supset k$ such that every element of L is the zero of a polynomial $f \in k[X]$. This is closely related to the following question: Given a polynomial f, find a field L with $L \supset k$, so that f has a zero in L, or even so that f can be written as a product $f = b(X - a_1)\dots(X - a_n)$ with $b, a_1, \dots, a_n \in L$ (then in particular $a_1, \dots, a_n$ are the zeros of f in L). Finally we will show the principal theorem of Galois theory, which relates algebraic extensions of a field to groups.

1. The characteristic of a field

Let R be a ring, let $a \in R$ and let $n \in \mathbb{Z}_{>0}$. Recall that we write $n \cdot a = \underbrace{a + \ldots + a}_{n\ times}$. In the field of rational numbers $\mathbb{Q}$ we have $n \cdot a = 0$ for all $a \in \mathbb{Q}$, $n \in \mathbb{Z}_{\geq 0}$. On the other hand, if p is a prime number, then in the finite field F_p we have $p \cdot 1 = \underline{p} = 0$.

Definition 1.1. Let k be a field. The *characteristic* of k, denoted $char(k)$, is the smallest $n \in \mathbb{Z}_{>0}$ with $n \cdot 1 = 0$ in k, if such an n exists, and otherwise $char(k) = 0$.

Example 1.2. (1) $char(\mathbb{Q}) = char(\mathbb{R}) = char(\mathbb{C}) = 0$.
(2) $char(F_p) = p$.

Definition 1.3. A subring $k \subset K$ of a field K is called a *subfield* if is is a field. In this case we call K a *field extension* of k. We write K/k is a field extension.

Remark 1.4. Let k be a subfield of a field K. Then $char(k) = char(K)$ (because the 1 of K is also the 1 of k).

Remark 1.5. The characteristic of a field is 0 or a prime number.

Proof. If $char(k) = nm$ with $n, m \in \mathbb{Z}_{>0}$, then $0 = (nm) \cdot 1 = (n \cdot 1)(m \cdot 1)$. Thus $(m \cdot 1) = 0$ or $(n \cdot 1) = 0$. Assume $(m \cdot 1) = 0$, then $m \geq char(k) = mn$, thus $n = 1$. □

2. Degree of a field extension

If K/k is a field extension, then K is in particular a k-vector space. Its dimension will be called the degree of the field extension.

Remark 2.1. Let K/k be a field extension. Then $(K, +)$ is an abelian group and the restriction of the multiplication in K defines a scalar multiplication

$$k \times K \to K;\ (a, x) \mapsto ax$$

so that the distributive laws $(a+b)x = ax + bx$, $a(x+y) = ax + ay$ and the associative law $(ab)x = a(bx)$ hold and furthermore $1x = x$. Thus K is a k-vector space.

Definition 2.2. The *degree* $[K : k]$ of the field extension K/k is the dimension of K as a k-vector space (with $[K : k] = \infty$ if this is not a finite dimensional vector space). The field extension K/k is called *finite* if $[K : k] < \infty$.

Remark 2.3. $[K : k] = 1$ if and only if $K = k$.

Proof. $[K : k] = 1$ if and only if 1 is a k-basis of K if and only if $K = k \cdot 1 = k$. □

The degree of field extensions behaves multiplicatively.

Definition 2.4. Let K/k be a field extension. Then a subfield $L \subset K$ with $k \subset L$ is called an intermediate field of K/k.

Theorem 2.5. *(Degree theorem) If L is an intermediate field of a field extension K/k, then*

$$[K : k] = [K : L][L : k].$$

(This is with the convention $n\infty = \infty n = \infty\infty$ for all $n \in \mathbb{Z}_{>0}$.) In particular K/k is a finite extension if and only if K/L and L/k are finite extensions.

Proof. If K/L or L/k is not finite, then K/k is trivially not finite.

So assume L/k and K/L are finite field extensions. Let $(x_1, \ldots, x_n)$ be a basis of L over k and $(y_1, \ldots, y_m)$ be a basis of K over L. We will show that $\{x_i y_j \mid i = 1, \ldots, n,\ j = 1, \ldots, m\}$ is a basis of K over k. They generate: If $y \in K$, then we can write

$$y = \sum_{j=1}^{m} b_j y_j, \quad b_j \in L,$$

and for all j we can write

$$b_j = \sum_{i=1}^{n} a_{ij} x_i, \quad a_{ij} \in k.$$

Thus we get $y = \sum_{i,j} a_{ij} x_i y_j$.

They are linearly independent: If

$$\sum_{i,j} a_{ij} x_i y_j = 0, \quad a_{ij} \in k,$$

then for all j we have $\sum_i a_{ij} x_i = 0$, because the y_j are linearly independent over L, and thus for all i, j we get $a_{ij} = 0$, because the x_i are linearly independent. □

Corollary 2.6. *If L is an intermediate field of a finite field extension K/k, then $[L : k]$ divides $[K : k]$. In particular if $[K : k]$ is a prime number, the only intermediate fields are k and K.*

3. Algebraic extensions and simple algebraic extensions

In this section we fix a field extension K/k. We are interested in elements of K which are zeros of polynomials in $k[X]$.

DEFINITION 3.1. An element $a \in K$ is called *algebraic* over k if there is a nonzero polynomial $f = \sum_{i=1}^{n} b_i X^i \in k[X]$ with $f(a) = \sum_i b_i a^i = 0$. $a \in K$ is called *transzendental* over k if it is not algebraic over k. K is called an *algebraic extension* of k if all elements of K are algebraic over k.

Thus an element $a \in K$ is transzendental if it is not the root of a nonzero polynomial in $k[X]$. In this course we will not deal with transzendental elements.

DEFINITION 3.2. Let $a_1, \dots, a_n \in K$. The *extension of k generated by $a_1, \dots, a_n$* is the intersection of all subfields L of K which contain k and $a_1, \dots, a_n$. It is denoted $k(a_1, \dots, a_n)$. By definition $k(a_1, \dots, a_n)$ is obviously a subfield of K containing k and $a_1, \dots, a_n$.

Of special importance are the extensions generated by one algebraic element.

DEFINITION 3.3. If $a \in K$ is algebraic over k, then $k(a)$ is called a *simple algebraic extension* of k.

There are two reasons for the importance of simple algebraic extensions. First they can be understood very well. Secondly "almost all" finite algebraic extensions are simple algebraic extensions (see the Theorem of the Primitive Element below). We start by giving an explicit description of a simple algebraic extension $k(a)$ in terms of the minimal polynomial of a, which is the unique irreducible monic polynomial $f \in k[X]$ with $f(a) = 0$.

DEFINITION 3.4. A polynomial $f \in k[X]$ is called monic, if its leading coefficient is 1. Note that for $f \in k[X] \setminus \{0\}$ there is a unique monic polynomial g with $\langle f \rangle = \langle g \rangle$. (If $b \in k^*$ is the leading coefficient of f, just put $g = f/b$).

Let $a \in K$ be algebraic over k. Let $ev_a : k[X] \to K, g \mapsto g(a)$ be the evaluation homomorphism. Then, by definition, a is algebraic if and only if $ker(ev_a) \neq 0$. As $k[X]$ is a PID, there is a unique monic polynomial $f_a \in k[X]$ with $ker(ev_a) = \langle f_a \rangle$. f_a is called the *minimal polynomial* of a over k.

PROPOSITION 3.5. *Let $a \in K$ be algebraic over k. The minimal polynomial f_a of a over k is the unique irreducible monic polynomial $f \in k[X]$ with $f(a) = 0$.*

PROOF. We want to show f_a is irreducible. Assume $f_a = gh$ with $g, h \in k[X]$. We have to show $g \in k^*$ or $h \in k^*$. We have $0 = f_a(a) = g(a)h(a)$, thus $g(a) = 0$ or $h(a) = 0$. If, say, $g(a) = 0$, then $g \in \langle f_a \rangle$, thus $g = lf_a$ for some $l \in k[X]$. Thus $f_a = gh = f_a lh$. Therefore $lh = 1$ and $l \in k^*$.

On the other hand let g be a monic irreducible polynomial with $g(a) = 0$. Then $g \in \langle f_a \rangle$, thus $g = lf_a$ for some $l \in k[X]$. As g is irreducible, we get $l \in k^*$. As both g and f_a are monic, this implies $l = 1$. □

EXAMPLE 3.6. For each prime number p and all $n \in \mathbb{Z}_{>1}$, we know that $X^n - p$ is irreducible in $\mathbb{Q}[X]$ by the criterion of Eisenstein. Thus $X^n - p$ is the minimal polynomial for $\sqrt[n]{q}$ over $\mathbb{Q}$.

Now let $k(a)/k$ be a simple algebraic extension. We can give a very explicit description of $k(a)$ in terms of the minimal polynomial of a over k.

THEOREM 3.7. *Let $a \in K$ be algebraic over k and let f be the minimal polynomial of a over k, and $m := deg(f)$. Then*

(1) $k(a) \simeq k[X]/\langle f\rangle$,
(2) $[k(a) : k] = m$ *and* $(1, a, \dots a^{m-1})$ *is a basis of* $k(a)$ *over* k.

PROOF. (1) $ev_a : k[X] \to k(a);\ g \mapsto g(a)$ is a ring homomorphism with kernel $\langle f\rangle$. Let L be the image of ev_a. Then by the homomorphism theorem $L \simeq k[X]/\langle f\rangle$. As f is irreducible and $k[X]$ is a PID, we know that $\langle f\rangle$ is a maximal ideal in $k[X]$ and $k[X]/\langle f\rangle$ is a field. Thus L is a subfield of $k(a)$. Furthermore $k \subset L$ as the image of the constant polynomials and $a = ev_a(X) \in L$. Thus $L = k(a)$.

(2) We have shown in (1) that $ev_a : k[X] \to k(a)$ is surjective, thus

$$k(a) = \{g(a) \mid g \in k[X]\}.$$

We first show that $1, a, \dots, a^{m-1}$ generate $k(a)$ over k. Let $b \in k(a)$, then $b = g(a)$ for some $g \in k[X]$. If $deg(g) > deg(f)$, then division with rest gives

$$g = qf + r, \quad q, r \in k[X], \quad deg(r) < deg(f).$$

Then by $f(a) = 0$ we get $g(a) = r(a)$. Thus $b = r(a)$ for a polynomial $r = \sum_{i=0}^{m-1} b_i X^i$, i.e. $1, a, \dots, a^{m-1}$ generate $k(a)$. Assume $1, a, \dots, a^{m-1}$ are not linearly independent, then we can write $0 = \sum_{i=0}^{m-1} b_i a^i$, with $b_i \in k$ not all 0. Thus $h(a) = 0$ for $h = \sum_{i=0}^{m-1} b_i X^i$. On the other hand $h \in ker(ev_a) = \langle f\rangle$ implies $deg(h) \geq deg(f) = m$, a contradiction. □

REMARK 3.8. Let $k(b)/k$ be a simple algebraic extension of degree n and let $f = \sum_{i=0}^{n} a_i X^i$ be the minimal polynomial of b over k. Then by the Theorem $k(b)$ can be explicitly described as follows:

$$k(b) = \Big\{\sum_{i=0}^{n-1} c_i b^i \;\Big|\; c_i \in k\Big\}.$$

Addition and multiplication are the usual ones as if these were polynomials in the indeterminate b together with the rule that we eliminate any power of b bigger than $n-1$ by $b^n = -\sum_{i=0}^{n-1} a_i b^i$.

EXAMPLE 3.9. The minimal polynomial of i over $\mathbb{R}$ is X^2+1. Thus $\mathbb{C} = \mathbb{R}(i) = \{a+bi \mid a, b \in \mathbb{R}\}$, with addition and multiplication

$$(a+bi)+(c+di) = a+c+(b+d)i,\ (a+bi)(c+di) = ac+(ad+bc)i+bdi^2 = ac-bd+(ad+bc)i.$$

Now we want to see that $k(a)$ is in a strong sense uniquely determined up to isomorphism by the minimal polynomial of a over k. We state a more general result, because that is what we will need.

REMARK 3.10. Let $\varphi : k \to k'$ be an isomorphism of fields. Then φ defines an isomorphism

$$\widehat{\varphi} : k[X] \to k'[X]; f = \sum_i a_i X^i \mapsto \sum_i \varphi(a_i) X^i.$$

THEOREM 3.11. *Let $\varphi : k \to k'$ be an isomorphism of fields. Let K/k and K'/k' be field extensions. Let $a \in K$ be algebraic over k with minimal polynomial f. Let $\widehat{f} = \widehat{\varphi}(f)$ and let $a' \in K'$ be a zero of $\widehat{f}$.*

Then there is a unique isomorphism $\Phi : k(a) \to k'(a')$ with $\Phi|_k = \varphi$ and $\Phi(a) = a'$.

PROOF. (Uniqueness). Let $\Phi : k(a) \to k'(a')$ with $\Phi|k = \varphi$ and $\Phi(a) = a'$. Then $k(a) = \{g(a)|g \in k[X]\}$, and for $g = \sum_i b_i X^i \in k[X]$, Φ is uniquely determined by:

$$\Phi(g(a)) = \sum_i \varphi(b_i)\Phi(a)^i = \sum_i \varphi(b_i)(a')^i = \widehat{\varphi}(g)(a').$$

(Existence) We look at the composition

$$k[X] \xrightarrow{\widehat{\varphi}} k'[X] \xrightarrow{ev_{a'}} k'(a').$$

By definition $\widehat{f}$ is the minimal polynomial of a' over k'. Thus $ev_{a'}$ is surjective with kernel $\langle\widehat{f}\rangle$. As $\widehat{\varphi}$ is an isomorphism, we get that the composition is surjective with kernel $\langle f\rangle$. By the homomorphism theorem the induced map $k[X]/\langle f\rangle \to k'(a'); [g] \mapsto \widehat{\varphi}(g)(a')$ is an isomorphism. Composing with the isomorphism

$$k(a) \to k[X]/\langle f\rangle; g(a) \mapsto [g]$$

we get an isomorphism

$$\Phi : k(a) \to k(a'); g(a) \mapsto \widehat{\varphi}(g)(a').$$

For $u \in k$ a we get $\Phi(u) = \widehat{\varphi}(u) = \varphi(u)$. And we get $\Phi(a) = \widehat{\varphi}(X)(a') = a'$. □

For the future the following corollary will be important:

COROLLARY 3.12. *Let K/k be a field extension and let $a, a' \in K$ be algebraic with the same minimal polynomial.*

Then there exists a unique isomorphism $\varphi : k(a) \to k(a')$ with $\varphi|_k = id$ and $\varphi(a) = a'$.

We have seen that a simple algebraic extension is a finite extension. Now we want to see that conversely finite field extensions are algebraic extensions.

THEOREM 3.13. *Let K/k be a field extension. K/k is a finite extension if and only if K/k is algebraic and there are finitely many elements $a_1, \dots, a_n \in K$ such that $K = k(a_1, \dots, a_n)$.*

PROOF. "$\Longrightarrow$" Let $m := [K : k]$. Then for any $a \in K$ the elements $1, a, \dots, a^m$ are linearly dependent. Thus there exists a nonzero polynomial $f \in k[X]$ with $f(a) = 0$. If $a_1, \dots, a_n$ is a basis of K over k, then $K = k(a_1, \dots, a_n)$.

"$\Longleftarrow$" We prove this by induction on n, the case $n = 0$ being trivial. Assume that $K = k(a_1, \dots, a_{n+1})$ and that for $L := k(a_1, \dots, a_n)$ the degree $m := [L : k]$ is finite. Let g be the minimal polynomial of a_{n+1} over L. Then $[K : L] = deg(f)$ and $[K : k] = [K : L][L : k] = deg(f)m$ is finite. □

This result implies that the elements of K which are algebraic over k form an intermediate field of K/k.

COROLLARY 3.14. *Let $L := \{a \in K \mid a \text{ is algebraic over } k\}$. Then L is a subfield of K/k and L/k is an algebraic extension.*

PROOF. We know $k \subset L$. If $a, b \in L$, then $k(a,b)/k$ is a finite algebraic extension, and by the previous theorem also $a - b$, a/b lie in L. Thus L is a field, and by definition L/k is an algebraic extension. □

EXAMPLE 3.15. The set $\overline{\mathbb{Q}} := \{a \in \mathbb{C} \mid a \text{ is algebraic over } \mathbb{Q}\}$ is the field of algebraic numbers. It is an (infinite) algebraic extension of $\mathbb{Q}$.

We also get that the composition of algebraic extensions is an algebraic extension.

COROLLARY 3.16. *Let K/L and L/k be algebraic field extensions. Then $K \subset k$ is an algebraic field extension.*

PROOF. Let $u \in K$. We need to show that u is algebraic over k. As u is algebraic over L, $g(u) = 0$ for a nonzero $g = \sum_{i=0}^{m} a_i X^i \in L[X]$. We put $M := k(a_0, \ldots, a_m)$. Then $M \subset k$ is a finite extension, and $M(u)/M$ is a finite extension. Thus $[M(u) : k] = [M(u) : M][M : k]$ is finite. Therefore u is algebraic over k. □

4. Algebraic closure

We will briefly without proofs introduce the algebraic closure of a field. The results will not be used in the rest of the course. A field K is called algebraically closed, if every nonconstant $f \in K[X]$ has a zero in K. An algebraic closure of a field k is an algebraically closed field which is an algebraic extension of k.

DEFINITION 4.1. Let K be a field. K is called *algebraically closed* if the following equivalent statements hold.

(1) Every nonconstant polynomial $f \in K[X]$ has a zero in K.
(2) Every nonconstant polynomial $f \in K[X]$ splits into linear factors i.e. there exist $a_1, \ldots, a_n, b \in K$ such that
$$f = b(X - a_1) \cdot \ldots \cdot (X - a_n).$$

It is easy to see that the two statements are equivalent: (2)$\Longrightarrow$(1) is obvious. Now assume (1). Let $a \in K$ be a zero of a nonconstant polynomial $f \in K[X]$. Then we can write $f = (X - a)g$ with $g \in K[X]$ and $deg(g) = deg(f) - 1$. The result follows by induction on $deg(f)$.

EXAMPLE 4.2. (1) $\mathbb{Q}$ is not algebraically closed, e.g. $X^2 - 2$ has no zero in $\mathbb{Q}$.
(2) $\mathbb{R}$ is not algebraically closed, e.g. $X^2 + 1$ has no zero in $\mathbb{R}$.
(3) $\mathbb{C}$ is algebraically closed. This is the *Fundamental Theorem of Algebra.* It is usually proved in Complex Analysis. At the end of this course we will give another proof based on the fundamental theorem of Galois theory.
(4) $\overline{\mathbb{Q}}$ is algebraically closed.

DEFINITION 4.3. A field extension K/k is called an *algebraic closure* of k if K/k is an algebraic extension and K is algebraically closed.

With the Lemma of Zorn one shows that every field has up to isomorphism a unique algebraic closure. We will not go into the proof.

THEOREM 4.4. (1) *Every field k has an algebraic closure.*

(2) *If K, L are algebraic closures of k, then there exists an isomorphism $\Phi : K \to L$ with $\Phi|_k = id_k$.*

EXAMPLE 4.5. (1) The algebraic closure of $\mathbb{R}$ is $\mathbb{C}$.
(2) The algebraic closure of $\mathbb{Q}$ is $\overline{\mathbb{Q}}$. Note that $\mathbb{C}$ is not an algebraic closure of $\mathbb{Q}$, because it contains transzendental elements, e.g. e, π.

5. Splitting fields

In this section we fix a field k. Until now we discussed algebraic elements in field extensions K/k, i.e. elements $a \in K$ which are zeros of nonzero polynomials in $k[X]$. Now we want to look at the problem the other way round: given a nonconstant polynomial $f \in k[X]$ we want to find a finite field extension K/k such that k has a zero in K or even such that f splits into linear factors over K. First we show that for a nonconstant polynomial $f \in k[X]$ we can always find a root of f in a finite extension K of k.

THEOREM 5.1. *Let $f \in k[X]$ be irreducible of degree $n > 0$. There exists a simple algebraic extension K/k with $[K : k] = n$, such that f has a zero in K.*

PROOF. If L is a field and $a \in L$ is algebraic over k, then we know that $k(a) = k[X]/\langle g\rangle$ where g is the minimal polynomial of a, and that $[k(a) : k] = deg(g)$. We just turn this around. By dividing f by its leading coefficient, we can assume that f is monic. We want to show that there is a simple extension $k(a)/k$ such that f is the minimal polynomial of a over k. Then we can put $K = k(a)$, and the result follows.

As f is irreducible and $k[X]$ is a PID, we know that $K := k[X]/\langle f\rangle$ is field. Let $\pi : k[X] \to K$ be the canonical projection. $\pi|_k$ is injective and the image is a subfield of K (which we identify with k). Thus K/k is a field extension. We claim that f has a zero in K, namely the class $[X]$: Write $f := \sum_{i=0}^n a_i X^i$ with $a_i \in k$. Then

$$0 = [f] = \sum_{i=0}^n [a_i][X^i] = \sum_{i=0}^n a_i [X]^i = f([X]).$$

Thus $[X] \in K$ is algebraic over k and f is an irreducible monic polynomial with $f([X]) = 0$. Thus f is the minimal polynomial of $[X]$ over k and $K = k([X])$. □

DEFINITION 5.2. In the situation of the Theorem we say that K is obtained from k *by adjoining a root of f*.

COROLLARY 5.3. *Let $f \in k[X]$ be a polynomial of degree $n > 0$. Then there exists a field extension K/k with $[K : k] \leq n$ such that f has a zero in K.*

PROOF. Just apply the previous theorem to any irreducible factor of f. □

Given a polynomial $f \in k[X]$ we can adjoin successively more roots until f splits into linear factors. The smallest such extension of k will be called a splitting field of f over k. It is uniquely determined up to isomorphism.

DEFINITION 5.4. Let $f \in k[X]$ by a polynomial of degree $n > 0$. A finite extension K/k is called a *splitting field* of f over k. if

(1) f splits over K into linear factors, i.e. there exist $a_1, \dots, a_n, b \in K$ such that $f = b(X - a_1) \cdot \ldots \cdot (X - a_n)$.
(2) f does not split over any intermediate field $K \supsetneq L \supset k$.

Corollary 5.5. *Let $f \in k[X]$ be a polynomial of degree $n > 0$. Then there exists a splitting field K of f over k with $[K : k] \leq n!$.*

If L/k is an extension of k over which f splits into linear factors $X - a_1, \dots, X - a_n$ then $k(a_1, \dots, a_n)$ is a splitting field of f over k.

Proof. First by the previous corollary we find K_1/k with $[K_1 : k] \leq n$ such that f has a root $a_1 \in K_1$. Then $f = (X - a_1)g$ with $deg(g) \leq n - 1$. By induction there is a finite field extension K/k of degree $\leq n!$ such that f splits over K.

Now let L/k be a field extension, such that over L we have $f = b(X - a_1) \dots (X - a_n)$ with $b \in k$, $a_i \in L$. We claim that $L := k(a_1, \dots, a_n)$ is a splitting field of f. Obviously f splits over L into linear factors. Assume $L_1 \subset L$ is a subfield over which f factors into linear factors $f = b(X - c_1) \dots (X - c_n)$. Then, as $L[X]$ is UFD, we have that up to reordering $a_i = c_i$ for all i. Thus $a_1, \dots, a_n \in L_1$ and thus $L = L_1$. □

Now we want to show that the splitting field is uniquely determined up to isomorphism. We again prove a stronger result that we will use in the future.

Theorem 5.6. *Let $\varphi : k \to k'$ be an isomorphism of fields. Let $f \in k[X]$ be a nonzero polynomial and $f' := \widehat{\varphi}(f)$. Let K be the splitting field of f over k and K' the splitting field of $\widehat{f}$ over K'. Then there is an isomorphism $\Phi : K \to K'$ with $\Phi|_k = \varphi$.*

In particular if K, K' are splitting fields of f over k then there is an isomorphism $\Phi : K \to K'$ with $\Phi|_k = id$.

Proof. We use induction over the degree $[K : k]$. If $[K : k] = 1$, then $K = k$ and $K' = k'$, so there is nothing to show.

If $[K : k] > 1$, then f contains an irreducible factor g of degree $deg(g) > 1$ and similarly $\widehat{g} = \widehat{\varphi}(g)$ is an irreducible factor of $\widehat{f}$ of degree > 1. Let a be a zero of g in K and a' a zero of $\widehat{g}$ in K'. Then by the previous theorem there is an isomorphism $\varphi' : k(a) \to k'(a')$ with $\varphi'|_k = \varphi$ and $[K : k(a)] < [K : k]$.

On the other hand K is the splitting field of f over $k(a)$ (because f splits over K into linear factors, and there is no intermediate field between $k(a)$ and K where it splits). Similarly K' is the splitting field of $\widehat{f}$ over $k'(a')$. Thus by induction there is an isomorphism $\Phi : K \to K'$ with $\Phi|_{k(a)} = \varphi'$ and thus $\Phi|_k = \varphi$. □

Example 5.7. (1) $\mathbb{C}$ is a splitting field of $X^2 + 1$ over $\mathbb{R}[X]$ and $\mathbb{Q}[\sqrt{2}]$ is a splitting field of $X^2 - 2$ over $\mathbb{Q}$.
(2) More generally let $f \in k[X]$ be an irreducible polynomial of degree 2 and K/k is an extension of degree 2 such that f has a zero in K. Then K is a splitting field of f over k.
(3) (Splitting field of $X^4 + 1$ over $\mathbb{Q}$). Let α be a root of $X^4 + 1$ in an extension of $\mathbb{Q}$. Then also $-\alpha$, $\frac{1}{\alpha}$ and $-\frac{1}{\alpha}$ are roots of $X^4 + 1$. These roots are distinct: $\alpha \neq -\alpha$

because $\alpha \neq 0$ and if $\alpha = \pm\frac{1}{\alpha}$, then $\alpha^2 = \pm 1$; thus $\alpha^4 + 1 = 2 \neq 1$. Thus over $\mathbb{Q}(\alpha)$ we get a splitting

$$X^4 - 1 = (X - \alpha)(X + \alpha)(X - \frac{1}{\alpha})(X + \frac{1}{\alpha}).$$

Therefore $\mathbb{Q}(\alpha)$ is the splitting field of $X^4 + 1$.

(4) (Splitting field of $X^3 - 2$). Over the complex numbers we have

$$X^3 - 2 = (X - \sqrt[3]{2})(X - e^{2\pi i/3}\sqrt[3]{2})(X - e^{4\pi i/3}\sqrt[3]{2}).$$

Clearly $X^3 - 2$ splits over $\mathbb{Q}(\sqrt[3]{2}, e^{2\pi i/3})$, but it does not split over $\mathbb{Q}(\sqrt[3]{2})$ or $\mathbb{Q}(e^{2\pi i/3})$, because $\sqrt[3]{2}$ is real and $\mathbb{Q}(e^{2\pi i/3})$ does not contain $\sqrt[3]{2}$. $[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = 3$ and $[\mathbb{Q}(e^{2\pi i/3}) : \mathbb{Q}] = 2$. Thus $[\mathbb{Q}(\sqrt[3]{2}, \sqrt{-3}) : \mathbb{Q}] = 6$.

6. Normal extensions

Related to the notion of splitting field is that of a normal field extension. An algebraic field extension K/k is normal if every irreducible polynomial in $k[X]$ that has a zero in K splits over K into linear factors. We will see that a finite extension K/k is normal if and only if K is the splitting field of a polynomial in $k[X]$.

DEFINITION 6.1. A field extension K/k is called *normal* if it is an algebraic extension and every irreducible polynomial $f \in k[X]$ that has a zero in K splits over K into linear factors.

EXAMPLE 6.2. In the above example $\mathbb{Q}(\sqrt[3]{2})/\mathbb{Q}$ is not a normal extension, because $X^3 - 2$ has a root in $\mathbb{Q}(\sqrt[3]{2})$, but does not split into linear factors.

PROPOSITION 6.3. *A finite extension K/k is normal if and only if K is the splitting field of a polynomial in $k[X]$.*

PROOF. "$\Longleftarrow$" Let K be a splitting field of a polynomial $f \in k[X]$. Let $g \in k[X]$ be an irreducible polynomial with a zero $\alpha \in K$. We have to show that g splits over K. Let β be another zero of g in an extension of K. We will show $\beta \in K$. Then the result follows by induction.

Since g is irreducible, we know that there is an isomorphism $\varphi : k(\alpha) \to k(\beta)$ with $\varphi|_k = id$. K is also a splitting field of f over $k(\alpha)$, and $K(\beta)$ is a splitting field of f over $k(\beta)$. Thus we get an isomorphism $\Phi : K \to K(\beta)$ with $\Phi|_{k(\alpha)} = \varphi$ and thus $\Phi|_k = id$. Thus $[K : k] = [K(\beta) : k]$ and therefore $[K(\beta) : K] = 1$, which implies $K(\beta) = K$. Thus $\beta \in K$.

"$\Longrightarrow$" Let K/k be a finite normal extension. We can write $K = k(a_1, \dots, a_n)$ with e.g. the a_i a basis of K over k. Let f_i be the minimal polynomial of a_i over k. Then because K/k is normal, each f_i splits over $k(a_1, \dots, a_n)$ into linear factors, and thus $f = f_1 \dots f_n$ does. As $K = k(a_1, \dots, a_n)$ and $a_1, \dots, a_n$ are the zeros of f, it is the splitting field of f. □

7. Separable extensions

Let k be a field. Let $f \in k[X]$ be a polynomial and let $f = b(X - a_1) \dots (X - a_n)$ be a splitting of f into linear factors over some extension of L/k. It can happen that some of the a_i coincide (e.g. $a_1 = a_2 = a$). We say in this case that a is a multiple root of f. To avoid problems coming from multiple roots, the concept of separable extension is introduced. In characteristic zero all algebraic extensions are separable.

Definition 7.1. Let $f \in k[X]$ and let

$$f = b(X - a_1)^{m_1} \dots (X - a_l)^{m_l}, \quad m_i > 0, \quad a_1, \dots, a_l \in K \text{ distinct}$$

be the splitting of f into linear factors over an extension K/k. If $m_i = 1$, then a_i is called a *simple root* in K. If $m_i > 1$, then a_i is called a *multiple root* of order m_i.

Definition 7.2. An algebraic field extension K/k is called *separable* if all irreducible polynomials $f \in k[X]$ with a zero in K have only simple roots in their splitting fields. It follows that they have only simple roots in all algebraic extensions of k. Equivalently the minimal polynomial of any $a \in K$ has only simple roots in its splitting field. If every algebraic extension K/k is separable over k, then k is called *perfect*.

We need a criterion for $f \in k[X]$ to have multiple roots in its splitting field. Like in calculus an element $a \in K$ is a multiple root of f if and only if it is both a root of f and of the derivative f'.

Definition 7.3. Let $f = \sum_{i=0}^{n} a_i X^i \in k[X]$. The *derivative* of f is $f' := \sum_{i=1}^{n} i a_i X^{i-1}$.

Remark 7.4. It is straightforward to check and left as an exercise that the usual rules for differentiation hold:

(1) $(af + bg)' = af' + bg'$ for $a, b \in k$
(2) $(fg)' = f'g + fg'$.

Lemma 7.5. *Let $f \in k[X]$ be a nonconstant polynomial and let K/k be a splitting field of f, let $a \in K$ be a zero of f. Then a is a multiple root of f if and only if $f'(a) = 0$.*

Proof. Let $r \geq 1$ be the multiplicity of the root a. Then we can write $f = (X - a)^r g$ with $g \in K[X]$ and $g(a) \neq 0$ and we get

$$f' = r(X - a)^{r-1} g + (X - a)^r g'$$

and the claim follows. □

Theorem 7.6. *Let $f \in k[X]$ be an irreducible polynomial. f has no multiple roots in its splitting field over k if and only if $f' \neq 0$.*

Proof. Let K/k be the splitting field of f. If $f' = 0$, then by the above every root of f is a multiple root. Let $a \in K$ be a multiple root of f. Then $f(a) = f'(a) \neq 0$. As f is irreducible, it is up to a constant multiple the minimal polynomial of a. Thus f' is divisible by f. If $f' \neq 0$, then $deg(f') < deg(f)$, so we get a contradiction. □

An irreducible polynomial $f \in k[X]$ can have 0 derivative only if the characteristic is finite.

COROLLARY 7.7. *Every field of characteristic* 0 *is perfect.*

PROOF. It follows from the definition that for a nonconstant polynomial $f \in k[X]$ over a field k of characteristic 0 the derivative f' is not 0. □

We finish this section with a proof of the important theorem of the primitive element. It says that every finite separable extension is a simple extension. In particular in characteristic 0 all finite extensions are simple extensions. This is very useful because we understand simple extensions much better.

THEOREM 7.8. *(Theorem of the primitive element). Let K be a finite separable field extension of k. Then there exists an element $a \in K$ with $K = k(a)$.*

PROOF. In the proof we assume for simplicity that k is infinite (e.g. has characteristic 0). Since $[K : k]$ is finite, we have $K = k(a_1, \dots, a_n)$ for some $a_i \in K$. We make induction on n. If $n > 2$, then by induction we have $k(a_1, \dots, a_{n-1}) = K(a)$ for some $a \in K$ and $K = k(a, a_n)$. Thus we can assume $n = 2$ and write $K = k(a, b)$. We will show that $K = k(c)$ for a general linear combination $c = a + zb$ with $z \in k$.

Let f and g be the minimal polynomials of a and b over k respectively. Let L/K be a field where f and g split into linear factors. Let

$$a = x_1, \dots, x_n \text{ roots of } f; \quad b = y_1, \dots, y_m \text{ roots of } g.$$

Then $b \neq y_j$ for all $j \neq 1$ (by separability). It follows that for $j \neq 1$, the element $z_{ij} := \frac{x_i - a}{b - y_j}$ is the only element of L with $a + z_{ij}b = x_i + z_{ij}y_j$. Since k is infinite, we can choose an element $z \in k$, different from all z_{ij}, thus $a + zb \neq x_i + zy_j$ unless $i = j = 1$.

Put $c = a + zb$. Then obviously $k(c) \subset k(a, b)$. We want to show $k(a, b) \subset k(c)$. We define $h \in k(c)[X]$ by $h(X) := f(c - zX)$. Then

$$h(b) = f(c - zb) = f(a) = 0.$$

As b is a zero of g in L we get that $(X - b)$ is a common factor of h and g in $L[X]$. We want to show that $X - b$ is the greatest common divisor of h and g in $L[X]$. As g splits over L into linear factors, the greatest common divisor must be a product of some linear factors of g. For $y_j \neq b$ another root of g, we get $h(y_j) = f(c - zy_j) \neq 0$ because by our choice of z, $c - zy_j \neq a_i$ for all roots a_i of f. Thus $(X - y_j)$ is not a factor of h. Thus $(X - b)$ is the greatest common divisor of h and g in $L[X]$.

However by the Euclidean algorithm the greatest common divisor of $h, g \in k(c)[X]$ must lie in $k(c)[X]$. Thus $X - b \in k(c)[X]$, which means $b \in k(c)$ and $a = c + zb \in k(c)$, thus $k(a, b) \subset k(c)$. □

8. Finite fields

Before going on with the general theory we briefly want to study finite fields. For a prime number p we already know the finite field $F_p = \mathbb{Z}/p\mathbb{Z}$ with p elements. One might think that these are the only finite fields, but this is not true: For every prime power $q = p^n$ there is up to isomorphism a unique field with q elements, the splitting field of $X^q - X$ over F_p. First we show that for any prime power $q = p^n$ there is a field with q elements. For the rest of this section we fix a prime number p, a positive integer n and put $q = p^n$.

LEMMA 8.1. *Let F be a field of characteristic p. Then $X^q - X \in F[X]$ has precisely q simple roots in its splitting field.*

PROOF. Let $a \in F$ be a multiple root. Then a is a zero of $X^q - X$ and of $(X^q - X)' = qX^{q-1} - 1 = -1$, which is impossible. □

PROPOSITION 8.2. *There exists a field with q elements.*

PROOF. Let K be the splitting field of $X^q - X$ over F_p. Let

$$F := \{a \in K \bigm| (a^q - a) = 0\},$$

be the set of roots of $(X^q - X)$. By the previous lemma F has q elements. We claim that F is a field. If $a, b \in F$, then $(ab)^q = a^q b^q = ab$, thus $ab \in F$, $(1/a)^q = 1/a^q = 1/a$, so $1/a \in F$ and $(a \pm b)^q = a^q \pm b^q = (a \pm b)$. Thus $a \pm b \in F$. Thus F is a field with q elements. □

The argument below will show that indeed $F = K$. Now we want to show that these are up to isomorphism all finite fields.

REMARK 8.3. Let F be a finite field of characteristic p. Then F contains $\{n \cdot 1 | n \in \mathbb{Z}\} \simeq F_p$. Thus F is a finite extension of F_p. Putting $n = [F : F_p]$, then F has p^n elements.

PROPOSITION 8.4. *If F is a field with $q := p^n$ elements, then F is a splitting field of the polynomial $X^q - X$ over F_p.*

PROOF. First we want to see that every element $a \in F$ satisfies $a^q - a = 0$. If $a = 0$, this is obvious. $(F \setminus \{0\}, \cdot)$ is a group of order $q - 1$. By a well known result in group theory (small theorem of Fermat) for every element g in a finite group G of order k we have $g^k = e$ where e is the neutral element of G. Thus we get $a^{q-1} = 1$ and thus $a^q - a = 0$.

As $X^q - X$ has degree q it can have at most q roots in F; on the other hand the elements of F are q distinct roots. Thus $X^q - X$ splits over F into linear factors. F is the splitting field of $X^q - X$, because it consists only of roots of $X^q - X$. □

Putting these two proposition together and recalling that the splitting field of a polynomial over a given field is unique up to isomorphism we get a complete classification of the finite fields.

THEOREM 8.5. *A finite field with n elements exists only if n is a prime power. For each prime power $q = p^m$ there is up to isomorphism a unique field with q elements: the splitting field of $X^q - X$ over F_p.*

9. Galois groups

The aim of Galois theory is to study fields via their automorphism groups. More precisely one studies finite field extensions K/k via the Galois group $Gal(K/k)$ of automorphisms of K which are the identity on k.

DEFINITION 9.1. Let K/k be a field extension. The Galois group $Gal(K/k)$ of K over k is the set of automorphisms $\varphi : K \to K$ with $\varphi|_k = id$.

EXAMPLE 9.2. Complex conjugation $\gamma : \mathbb{C} \to \mathbb{C}, a + bi \mapsto a - bi$, is an element of $Gal(\mathbb{C}/\mathbb{R})$. We will see that $Gal(\mathbb{C}/\mathbb{R}) = \{id, \gamma\}$.

At first look the Galois group looks like something very abstract that is very difficult to understand. However for a simple algebraic extension $k(a)/k$, we can identify $Gal(k(a)/k)$ with a subgroup of the groups of permutations of the roots of the minimal polynomial of a.

NOTATION 9.3. For a finite set M we denote by $S(M) := \{\sigma : M \to M \text{ bijection}\}$ the set of permutations of M. Then $S(M)$ is isomorphic to the symmetric group S_n where $n = |M|$ is the number of elements of M.

Recall that an action of a group G on a set M is *simply transitive* if for all $m_1, m_2 \in M$ there is a unique $g \in G$ with $g(m_1) = m_2$. Then obviously $|G| = |M|$.

THEOREM 9.4. *Let $k(a)/k$ be a simple algebraic extension of degree n. Let f be the minimal polynomial of a over k and let $R := \{b \in k(a) \mid f(b) = 0\}$ be the set of its roots.*

Then $Gal(k(a)/k)$ acts simply transitively on R. Thus $Gal(k(a)/k)$ is isomorphic to a subgroup of $S(R)$ of order $|R| \leq n$.

PROOF. Let $\varphi \in Gal(k(a)/k)$. If $g = \sum_i a_i X^i \in k[X]$ and $b \in k(a)$, then

$$\varphi(g(b)) = \sum_i a_i \varphi(b)^i = g(\varphi(b)).$$

Thus if b is a root of f, then $0 = f(b) = \varphi(f(b)) = f(\varphi(b))$, thus $\varphi(b)$ is also a root of f. (This elementary argument will be used very often in the future). So we see that $\varphi|_R$ is a bijection of R to itself and the map

$$res_R : Gal(k(a)/k) \to S(R), \varphi \mapsto \varphi|_R$$

is a group homomorphism. We have seen that given any two roots b_1, b_2 of f, there exists a unique element $\varphi \in Gal(k(a)/k)$ with $\varphi(b_1) = b_2$. Thus $Gal(k(a)/k)$ acts simply transitively on R, in particular res_R is injective and so $Gal(k(a)/k)$ is isomorphic to a subgroup of $S(R)$. As the action is simply transitive we get $|Gal(k(a)/k)| = |R|$. Finally $|R| \leq deg(f) = n$. $\square$

The subject of Galois theory are Galois extensions, which are finite separable normal extensions.

DEFINITION 9.5. A finite field extension K/k is called a *Galois extension* if it is separated and normal. (Note that in characteristic 0 separatedness is automatic).

COROLLARY 9.6. *Let K/k be a Galois extension of degree n. Then $Gal(K/k)$ is isomorphic to a subgroup of S_n which acts simply transitively and $|Gal(K/k)| = n$.*

PROOF. By the Theorem of the Primitive Element $K = k(a)$ for some $a \in K$, and as K/k is separated and normal the minimal polynomial of f has n distinct roots in K. $\square$

PROPOSITION 9.7. *Let K/k be a Galois extension and $a, b \in K$. $\varphi(a) = b$ for some element $\varphi \in Gal(K/k)$ if and only if a and b have the same minimal polynomial over k.*

PROOF. If a, b have the same minimal polynomial, then there is a unique isomorphism $\psi : k(a) \to k(b)$ with $\psi(a) = b$ and $\psi|_k = id_k$. K is the splitting field of a polynomial $f \in k[X]$ over k and thus also over $k(a)$. Thus by uniqueness of splitting fields ψ can be extended to an element $\varphi \in Gal(K/k)$.

For the converse let $b = \varphi(a)$, with $a, b \in K$, $\varphi \in Gal(K/k)$. Let f and g be the minimal polynomials of a, b over k. Then $0 = \varphi(f(a)) = f(\varphi(a)) = f(b)$, thus g divides

f, as g is the minimal polynomial of b. Similarly $f|g$. Therefore we get $f = g$, as f, g are monic. $\square$

A very important result for Galois theory is that k is precisely the set of elements of a Galois extension K/k fixed by all elements of $Gal(K/k)$.

THEOREM 9.8. *Let K/k be a Galois extension and $a \in K$. Then $k = \{a \in K \mid \varphi(a) = a$ for all $\varphi \in Gal(K/k)\}$.*

PROOF. "$\subset$" is trivial. "$\supset$" Assume $\varphi(a) = a$ for all $\varphi \in Gal(K/k)$. Let f be the minimal polynomial of a over k. Since K/k is normal, f splits over K. Let b be a root of f. Then there exists $\varphi \in Gal(K/k)$ with $\varphi(a) = b$. Thus $a = b$ by our assumption, i.e all the roots of f are equal. But f does not have multiple roots because K/k is separable. Thus $f = X - b \in k[X]$ and thus $b \in k$. $\square$

We will usually study field extensions given as the splitting field of some polynomial. Thus we define the Galois group of a polynomial $f \in k[X]$ as the Galois group of its splitting field K over k. In this case $Gal(f)$ can be identified with a subgroup of the permutations of the roots of f in K, and this is usually also the best way to look at it.

DEFINITION 9.9. Let $f \in k[X]$ be a nonconstant polynomial. Let K be the splitting field of f over K (unique up to isomorphism which is the identity on k). The *Galois group* of f is $Gal(f) := Gal(K/k)$.

PROPOSITION 9.10. *Let $f \in k[X]$ be a nonconstant polynomial of degree n. Let R be the set of roots of f in the splitting field K of f.*
Then $Gal(f)$ is isomorphic to a subgroup of $S(R)$ and $|Gal(f)|$ divides $n!$.

PROOF. We have $Gal(f) = Gal(K/k)$. If a is a root of f in K, then $f(\varphi(a)) = \varphi(f(a))$ for all $\varphi \in Gal(K/k)$. Thus restriction to R defines a group homomorphism $Gal(K/k) \to S(R)$. If $\varphi|_R = id$, because $K = k(R)$. Hence $Gal(f)$ is isomorphic to a subgroup of $S(R)$. As $|R| \leq n$, we get that $|Gal(f)|$ divides $|R|!$ which divides $n!$. $\square$

EXAMPLE 9.11.
(1) Let $f = X^3 - 2 \in \mathbb{Q}[X]$. Then $Gal(f) = S_3$: f is irreducible. Let $K/\mathbb{Q}$ be the splitting field, which is a Galois extension. Then $K = \mathbb{Q}[2^{\frac{1}{3}}, e^{2\pi i/3}]$, and $[K : \mathbb{Q}] = 6$. Thus $|Gal(f)| = 6$. On the other hand $Gal(f)$ is a subgroup of S_3, thus $Gal(f) = S_3$.
(2) Let $f = X^4 + 1 \in \mathbb{Q}[X]$. Then $Gal(f) = \mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z}$. If $\alpha = e^{2\pi i/8}$ is a root of f, then we have seen $f = (X - \alpha)(X + \alpha)(X - 1/\alpha)(X + 1/\alpha)$ over $\mathbb{Q}(\alpha)$, thus $\mathbb{Q}(\alpha)$ is the splitting field of f. As $\mathbb{Q}(\alpha)/\mathbb{Q}$ is a simple extension, we also know that $|Gal(f)| \leq 4$. On the other hand $\alpha \mapsto -\alpha$, $\alpha \mapsto 1/\alpha$ are obviously two commuting elements of $Gal(f)$ which generate a group isomorphic to $\mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z}$.

10. The fundamental theorem of Galois theory

Now we will prove the famous fundamental theorem of Galois theory. It relates the intermediate fields L of a Galois extension K/k to the subgroups of the Galois group $Gal(K/k)$. The importance is that it relates questions in field theory to questions in group theory and in this way often makes it possible to solve them.

DEFINITION 10.1. Let K/k be a field extension. For any subgroup $H \subset Gal(K/k)$ the *fixed field* of H is

$$Fix(H) := \{a \in K \mid \varphi(a) = a \text{ for all } \varphi \in H\}.$$

REMARK 10.2. (1) $Fix(H)$ is a subfield of K because, if $\varphi(a) = a$, $\varphi(b) = b$, then $\varphi(a \pm b) = \varphi(a) \pm \varphi(b)$, $\varphi(ab) = \varphi(a)\varphi(b)$, $\varphi(1/a) = 1/\varphi(a)$.

(2) If K/k is a Galois extension, then $Fix(Gal(K/k)) = k$.

(3) If L is an intermediate field of K/k, then

$$Gal(K/L) = \{\varphi \in Gal(K/k) \mid \varphi|_L = id\}$$

is a subgroup of $Gal(K/k)$.

THEOREM 10.3. *(Fundamental theorem of Galois theory) Let K/k be a Galois extension with Galois group G.*

(1) *The mappings $H \mapsto Fix(H)$, $L \mapsto Gal(K/L)$ are mutually inverse inclusion reversing bijections*

$$\{\textit{subgroups of } G\} \longleftrightarrow \{\textit{intermediate fields of } K/k\},$$

i.e. $Fix(Gal(L/k)) = L$, $\quad Gal(K/Fix(H)) = H$.

(2) $[K : Fix(H)] = |H|$ *and* $[Fix(H) : k] = [G : H]$ *for subgroups* $H \subset G$.

(3) $[K : L] = |Gal(K/L)|$ *and* $[L : k] = [G : Gal(K/L)]$ *for intermediate fields* $k \subset L \subset K$.

PROOF. (1) Let L be an intermediate field of K/k. Then K/L is a Galois extension (it is a splitting field of a polynomial over k and thus over L, and the minimal polynomial of any element a of K over L is a factor of the minimal polynomial of a over k). Therefore $Fix(Gal(K/L)) = L$ by the last remark. Let H be a subgroup of G. We put $F := Fix(H)$. We need to show $Gal(K/F) = H$. By the theorem of the primitive element $K = k(a)$ for some $a \in K$. Define

$$f := \prod_{h \in H} (X - h(a)) \in k(a)[X], \quad deg(f) = |H|.$$

All roots of f are distinct: $h_1(a) = h_2(a)$ implies $h_1 = h_2$, because $k(a)/k$ is a simple extension. We see that $\widehat{h}(f) = f$ for all $h \in H$, therefore $f \in F[X]$. We have $f(a) = 0$, therefore the minimal polynomial g of a over F divides f. Since $k(a) = F(a)$, we see

$$|Gal(K/F)| = [K : F] = deg(g) \leq deg(f) = |H|.$$

But obviously H is a subgroup of $Gal(K/F)$, so $H = Gal(K/F)$.

(2) Let again $F = Fix(H)$. Since K is Galois over F we have $[K : F] = |Gal(K/F)| = |H|$. The second statement follows because $[K : F][F : k] = [K : k] = |G|$.

(3) The first statement is obvious because K/L is Galois. The second statement follows because $[K : L][L : k] = [K : k] = |G|$. □

Let K/k be a Galois extensions with Galois group G. The intermediate fields L/k which are normal over k correspond to particularly nice subgroups of G. In fact L/k is normal if and only if $Gal(K/L)$ is a normal subgroup of G. This is the second part of the fundamental theorem of Galois theory.

NOTATION 10.4. Let L be an intermediate field of K/k, and let $\alpha \in G$. We write $\alpha(L) := \{\alpha(a) \mid a \in L\}$. This is clearly an intermediate field of K/k.

LEMMA 10.5. *Let L be an intermediate field of K/k and $\alpha \in G$. Then $Gal(K/\alpha(L)) = \alpha Gal(K/\alpha(L))\alpha^{-1}$.*

PROOF. Let $\varphi \in G$. Then $\varphi \in Gal(K/\alpha(L))$ if and only if for all $a \in L$ we have $\varphi(\alpha(a)) = \alpha(a)$, i.e. $\alpha^{-1}\varphi\alpha \in Gal(K/L)$, i.e. $\varphi \in Gal(K/\alpha(L))$ if and only $\varphi \in \alpha Gal(K/L)\alpha^{-1}$. □

No we can show the relation between normal subgroups and normal extensions.

THEOREM 10.6. *Let K/k be a Galois extension and let L be an intermediate field. The following are equivalent:*

(1) *L/k is normal,*
(2) *$\alpha(L) = L$ for all $\alpha \in Gal(K/k)$*
(3) *$Gal(K/L)$ is a normal subgroup of $Gal(K/k)$.*

PROOF. "(1)$\Longrightarrow$(2)" Let $a \in L$ and let f be the minimal polynomial of a over k. Since L is normal over k, all roots of f lie in L. Let $\alpha \in Gal(K/k)$, then $\alpha(a)$ is also a root of f because $f(\alpha(a)) = \alpha(f(a)) = 0$. Therefore $\alpha(a) \in L$ and $\alpha(L) \subset L$. The same argument shows $\alpha^{-1}(L) \subset L$, i.e. $L \subset \alpha(L)$.

"(2)$\Longrightarrow$(3)" Let $\alpha \in Gal(K/k)$. Then

$$\alpha Gal(K/L)\alpha^{-1} = Gal(K/\alpha(L)) = Gal(K/L).$$

Thus $Gal(K/L)$ is a normal subgroup of $Gal(K/k)$.

"(3)$\Longrightarrow$(1)" We have $Gal(K/L)$ is a normal subgroup of $Gal(K/k)$. Let $\alpha \in Gal(K/k)$. Then

$$Gal(K/L) = \alpha Gal(K/L)\alpha^{-1} = Gal(K/\alpha(L)).$$

By the fundamental theorem of Galois theory we get $L = \alpha(L)$. Let $f \in k[X]$ be irreducible with a root in L, and let b be another root of f in K. Because both a and b have f as a minimal polynomial over k, we get that there is an $\alpha \in Gal(K/k)$ with $\alpha(a) = b$. Therefore $b \in \alpha(L) = L$. Thus L is normal over k. □

COROLLARY 10.7. *Let K/k be a Galois extension. Let L/k be an intermediate field which is Galois over k. Then*

$$Gal(L/k) \simeq Gal(K/k)/Gal(K/L).$$

PROOF. Let $\alpha \in Gal(K/k)$. Then $\alpha(L) = L$, thus $\alpha|_L$ is an element of $Gal(L/k)$. The restriction map $Gal(K/k) \to Gal(L/k); \alpha \mapsto \alpha|_S$ is a homomorphism of groups with kernel $Gal(K/L)$. Thus by the first isomorphism theorem $Gal(K/k)/Gal(K/L)$ is isomorphic to a subgroup of $Gal(L/k)$. On the other hand

$$|Gal(K/k)/Gal(K/L)| = [K : k]/[K : L] = [L : k] = |Gal(L/k)|$$

because L/k is a Galois extension. Therefore $Gal(L/k) = Gal(K/k)/Gal(K/L)$. □

EXAMPLE 10.8.
(1) Let $\mathbb{Q}(a, w)$ with $a = 2^{1/3}$, $w = e^{2\pi i/3}$ be the splitting field of $X^3 - 2 \in \mathbb{Q}[X]$. We have seen that $G := Gal(\mathbb{Q}(a, w)/\mathbb{Q}) = S_3$. G has precisely 6 subgroups. Below we have on the left the diagram of subgroups H of G (a line indicates the upper group is a subgroup of the lower one) and on the right the corresponding diagram of intermediate fields $Fix(H)$ (a line indicates that the upper field is an extension of the lower one).

$$\begin{array}{cccc} & 1 & & \\ \langle\beta\rangle & \langle\alpha\rangle & \langle\gamma\rangle & \langle\delta\rangle \\ & G & & \end{array} \qquad \begin{array}{cccc} & \mathbb{Q}(a,w) & & \\ \mathbb{Q}(a) & \mathbb{Q}(w) & \mathbb{Q}(w^2a) & \mathbb{Q}(wa) \\ & \mathbb{Q} & & \end{array}$$

Here $\alpha(a) = wa$, $\alpha(w) = w$, $\beta(a) = a$, $\beta(w) = w^2$, $\gamma(a) = wa$, $\gamma(w) = w^2$, $\delta(a) = w^2a$, $\gamma(w) = w^2$. The normal subgroups of G are 1, $\langle\alpha\rangle$ and G corresponding to the normal extensions K, $\mathbb{Q}(w)$ and $\mathbb{Q}$.

(2) Let $f = X^4 + 1 \in \mathbb{Q}[X]$. Let α be a root of f, we know that over $\mathbb{Q}(\alpha)$ $f = (X - \alpha)(X + \alpha)(X - 1/\alpha)(X - 1/\alpha)$, and that $G := Gal(f) = \mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z}$, with the four elements 1, $\mu : \alpha \mapsto -\alpha$, $\delta : \alpha \mapsto 1/\alpha$, $\delta\mu : \alpha \mapsto -1/\alpha$. Thus we get the diagrams

$$\begin{array}{ccc} & 1 & \\ \langle\mu\rangle & \langle\delta\rangle & \langle\mu\delta\rangle \\ & G & \end{array} \qquad \begin{array}{ccc} & \mathbb{Q}(\alpha) & \\ \mathbb{Q}(\alpha^2) & \mathbb{Q}(\alpha + 1/\alpha) & \mathbb{Q}(\alpha - 1/\alpha) \\ & \mathbb{Q} & \end{array}$$

To show the power of the fundamental theorem of Galois theory and see how nontrivial results of group theory can be applied to fields we give a proof of the fundamental theorem of algebra. Any proof of this statement has to use some analysis. We will use only the intermediate value theorem: For polynomials it says in particular the following: Let $f \in \mathbb{R}[X]$ be monic a polynomial of odd degree, then f has a zero in $\mathbb{R}$ (this is because for a very large $f(a)$ will be positive and $f(-a)$ will be negative).

THEOREM 10.9. *(Fundamental Theorem of Algebra) Let $f \in \mathbb{C}[X]$ be a nonconstant polynomial. Then f splits over $\mathbb{C}$ into linear factors.*

PROOF. We show that if $K/\mathbb{C}$ is a finite extension, then $K = \mathbb{C}$. The result then follows: if $f \in \mathbb{C}[X]$ is a irreducible polynomial of degree ≥ 2, then $\mathbb{C}[X]/(f)$ is a finite extension of $\mathbb{C}$.

Let $K/\mathbb{C}$ be a finite extension with $K \neq \mathbb{C}$. Let $a \in K \setminus \mathbb{C}$ with minimal polynomial f over $\mathbb{R}$. By replacing K with the splitting field of $(X^2 + 1)f$ over $\mathbb{R}$ we can assume that $K/\mathbb{R}$ is a Galois extension. Let $G := Gal(K/\mathbb{R})$. Then

$$|G| = [K : \mathbb{R}] = [K : \mathbb{C}][\mathbb{C} : \mathbb{R}] = 2[K : \mathbb{C}].$$

Thus $|G|$ is even.

Let S be a Sylow 2-subgroup of G and put $F := Fix(S)$. Then $\mathbb{R} \subset F \subset K$ and $[F : \mathbb{R}] = [G : S]$ is odd. Let $a \in F$ and let g be the minimal polynomial of a over $\mathbb{R}$. Then $deg(g) = [\mathbb{R}(a) : \mathbb{R}]$ which divides $[F : \mathbb{R}]$. Thus $deg(g)$ is odd. Thus the intermediate value theorem says that g has a zero in $\mathbb{R}$. But g is irreducible so $deg(g) = 1$. Thus $a \in \mathbb{R}$ and $F = \mathbb{R}$. Thus $S = G$ and G is a 2 group.

Let $H = Gal(K/\mathbb{C})$. Then H is a subgroup of G and H is a 2-group. If $H = 1$, then $K = \mathbb{C}$, which we want to prove. Thus we assume that $H \neq 1$ and derive a contradiction. There exists a maximal proper subgroup M of H, with $[H : M] = 2$. Put $L := Fix(M)$. Then we have

$$[L : \mathbb{C}] = [Gal(K/\mathbb{C})]/[Gal(K/L)] = [H : M] = 2.$$

Thus we have found an extension L of $\mathbb{C}$ of degree 2. This is however impossible: Any $u \in L$ has a minimal polynomial $X^2 + aX + b$ $(a, b \in \mathbb{C})$ of degree at most 2. By the quadratic formula

$$u = -\frac{1}{2}(a \pm \sqrt{a^2 - 4b}) \in \mathbb{C}.$$

Thus we get $L = \mathbb{C}$. □

11. Solvability by radicals

The main problem that motivated the development of Galois theory was the question of solvability of polynomials by radicals. Given a polynomial $f = X^n + \sum_{i=0}^{n-1} a_i X^i \in \mathbb{Q}[X]$, one wants to find a formula for the roots of f in $\mathbb{C}$, in terms of the a_i. Here the operations in the formula are just $+, -, \cdot, /$ and $\sqrt[i]{\ }$. This problem has been studied over the centuries by many people. For $n = 2$ we have the classical formula for the roots of $X^2 + 2aX + b$ by $-a \pm \sqrt{a^2 - b}$. For $n = 3$ and $n = 4$ the formulas were found in the 16^{th} century. For $n \geq 5$ the problem remained open, until Galois theory was invented to give a negative solution. The problem now mostly historical interest. We restrict ourselves to sketching the results, but not give the proof. We start by giving a precise mathematical formulation of the problem.

Definition 11.1. A field extension K/k is called *radical extension* if there is a chain

$$K = L_m \supset L_{m-1} \supset \ldots \supset L_0 = k$$

of intermediate fields, so that $L_{i+1} = L_i(b_i)$ for all i, where b_i is a root of a polynomial $X^{m_i} - a_i$ with $a_i \in L_i$ (thus b_i is the m_i^{th} root of an element of L_i).

We want to express the roots of $f \in k[X]$ in terms of the usual operations of a field and by taking i^{th} roots. Thus we say that f is solvable by radicals if it splits into linear factors over a radical extension of k.

Definition 11.2. A polynomial $f \in k[X]$ is *solvable by radicals* if there is a radical extension K/k such that f splits over K into linear factors.

This problem is closely related to the structure of the Galois group of f.

Definition 11.3. A group G is called *solvable* if there is a chain

$$G = G_0 \supset G_1 \supset \ldots \supset G_n = \{1\}$$

of subgroups of G such that G_{i+1} is a normal subgroup in G_i for all i and the factor group G_i/G_{i+1} is abelian. (in particular abelian groups are solvable).

The name *solvable* obviously comes from its relation to solvability by radicals.

EXAMPLE 11.4. One can show that S_n is solvable for $n \leq 4$, but not solvable for $n \geq 5$.

The main result is now the following

THEOREM 11.5. *(Galois) Let k be a field of characteristic 0 and $f \in k[X]$. Then f is solvable by radicals if and only if $Gal(f)$ is solvable.*

We will not go into the proof because of lack of time. The proof is however quite accessible at this point.

COROLLARY 11.6. *A polynomial of degree ≤ 4 is solvable by radicals.*

A polynomial of degree $n \geq 5$ will sometimes be solvable by radicals and sometimes not. The original problem we posed is however a stronger one: We do not just want to be able to find the roots of some given polynomial in terms of radicals. We wanted a general formula in terms of the coefficients of the polynomial. This means that we should treat the coefficients of f as determinates. Let $L = k(x_0, \ldots, x_{n-1})$ be the field of rational functions in $x_0, \ldots, x_{n-1}$. Let $f_n := X^n + \sum_{i=0}^n x_i X^i \in L[X]$. That there is a universal formula for the roots of any polynomial g of degree n in terms of radicals in terms of the coefficients of g means then that f_n is solvable by radicals.

THEOREM 11.7. *(Abel) Let k be a field of characteristic 0. Then the Galois group of f_n over L is S_n.*

This immediately proves that there is no universal formula of the kind we desired for $n \geq 5$.

THEOREM 11.8. *(Abel) f_n is solvable by radicals for $n \leq 4$ and not solvable by radicals for $n \geq 5$.*