RINGS AND GALOIS THEORY

F.Beukers, based on lecture notes by F.Oort, H.W.Lenstra, B.van Geemen, J.Top and G.Cornelissen

Block 3, 2016

Contents

Chapter 1

Rings

1.1 Definition, examples, elementary properties

Definition 1.1.1 *A ring is a set, denoted by R, two mappings $+: (a,b) \mapsto a+b$ and $\cdot: (a,b) \mapsto a \cdot b$ from $R \times R \to R$, and two elements $0, 1 \in R$ such that:*

(R1) $(R, +, 0)$ *is an abelian group, in other words:*

- *(G1)* $a+b = b+a$ *for all* $a, b \in R$.
- *(G2)* $a+(b+c) = (a+b)+c$ *for all* $a, b, c \in R$;
- *(G3)* $0+a = a+0 = a$ *for all* $a \in R$;
- *(G4) for every* $a \in R$ *there exist an opposite* $-a \in R$ *such that* $a+(-a) = (-a)+a = 0$;

(R2) $a(bc) = (ab)c$ *for all* $a, b, c \in R$ *(associativity of* $\cdot$ *);*

(R3) $a(b+c) = ab+ac$ *and* $(b+c)a = ba+ca$ *for all* $a, b, c \in R$ *(the distributive laws).*

(R4) $1 \cdot a = a \cdot 1 = a$ *for all* $a \in R$.

(R5) $ab = ba$ *for all* $a, b \in R$ *(commutative law for multiplication).*

The expressions $a+b$ and ab are called *sum* and *product* of a and b. Note that we have already abbreviated $a \cdot b$ by ab several times. The maps $+$ and $\cdot$ are called *addition* and *multiplication* in R. The element 0 is called the *zero element* and 1 is called the *one element*.

In the literature one sometimes considers rings without axiom (R4). Very occasionally we will encounter such rings and call them *ring without* 1. In the literature one may also consider rings that do not satisfy (R5). We speak of a *non-commutative ring* in this case, as opposed to the *commutative rings* we consider by default in these lectures. In Section 1.4 we will collect some examples of non-commutative rings.

A ring is called a *field* if in addition

(R6) $1 \neq 0$ and for all non-zero $a \in R$ there is an inverse $a^{-1} \in R$ such that $a \cdot a^{-1} = a^{-1} \cdot a = 1$.

(Dutch: lichaam; Flemish: veld; French: corps; German: Körper)

Remark 1.1.2 *Notice that for any $a, b \in R$ there exists a unique element $x \in R$ such that $a + x = b$, namely $x = b + (-a)$. This follows from $a + x = b$ by addition of $-a$ on both sides. We denote this element by $b - a$, the difference of b and a. The important feature of a ring is that it is a set with multiplication and addition according to the above rules. In particular a ring is closed with respect to taking differences.*

Remark 1.1.3 *Let $a, b \in R$. Then $(a + 0)b = ab$. Working out the left hand side by the distributive law we get $ab + 0 \cdot b = ab$. Hence we conclude that $0 \cdot b = 0$ for every $b \in R$.*

Example 1.1.4. Usually we have that $1 \neq 0$. The equality $1 = 0$ would imply that $a = 1 \cdot a = 0 \cdot a = 0$ for every element $a \in R$. Hence the corresponding ring consists only of the 0-element (also 1-element). This is called the *trivial ring*.

♢

Example 1.1.5. The sets $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ of integers, rational numbers, real numbers and complex numbers (respectively) are examples of rings when equipped with the usual addition and multiplication. Moreover, $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ are fields, whereas $\mathbb{Z}$ is not a field (property (R6) does not hold).

♢

Example 1.1.6. Let $n \in \mathbb{Z}_{>1}$. Let us consider two integers a, b equivalent modulo n if $a - b$ is divisible by n. Notation: $a \equiv b (\text{mod } n)$. Another way of phrasing this is that a, b are equivalent modulo n if they have the same remainder after division by n. There are finitely many equivalence classes modulo n (also called residue classes modulo n) corresponding to the remainders $0, 1, 2, \ldots, n - 1$. We denote this set by $\mathbb{Z}/n\mathbb{Z}$. On this set we have the usual addition and multiplication modulo n. With these operations $\mathbb{Z}/n\mathbb{Z}$ is a commutative ring with 1.

In Theorem 1.2.9 we shall see that $\mathbb{Z}/n\mathbb{Z}$ is a field if and only if n is a prime number.

Note that $n = 1$ would give us the trivial ring $\mathbb{Z}/1\mathbb{Z}$.

♢

Example 1.1.7. A *polynomial* with coefficients in $\mathbb{R}$ is an expression of the form

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0,$$

where $a_i \in \mathbb{R}$ for $i = 0, 1, \ldots, n$. We can add and multiply polynomials with the usual rules and thus obtain a ring. Notation: $\mathbb{R}[x]$. Later on we shall deal with polynomial rings extensively.

♢

There are many more examples, as well as constructions to produce rings out of existing ones, see Section 1.3.

Definition 1.1.8 *A subset R' of a ring R is called subring of R if the following conditions are satisfied:*

(D0) $1, 0 \in R'$.

(D1) $a - b \in R'$ *for all* $a, b \in R'$.

(D2) $ab \in R'$ *for all* $a, b \in R'$.

Note that D0, D1 imply that R' is a subgroup of R with respect to addition. A subring R' of a ring R is itself ring, if we endow it with the addition and multiplication from R.
Every ring R has R itself as subring.

Example 1.1.9. The set $\mathbb{Z}[i] = \{a + bi : a, b \in \mathbb{Z}\} \subset \mathbb{C}$ is a subring of $\mathbb{C}$. The set $\mathbb{Z}[i]$ is also called the *ring of Gauss integers*. It is a commutative ring with 1, but not a field. The set $\mathbb{Q}[i] = \{a + bi : a, b \in \mathbb{Q}\}$ is also a subring of $\mathbb{C}$ but this time it is a field as well. The inverse of $a + bi$ ($\neq 0$) is given by $\frac{a}{a^2+b^2} + \frac{-b}{a^2+b^2} i$. Similar remarks apply to

$$\mathbb{Z}[\sqrt{m}] = \{a + b\sqrt{m} : a, b \in \mathbb{Z}\},$$

$$\mathbb{Q}[\sqrt{m}] = \{a + b\sqrt{m} : a, b \in \mathbb{Q}\},$$

where m is an integer which is not a square. In particular $m = -1$ yields $\mathbb{Z}[i]$ and $\mathbb{Q}[i]$.

◊

1.2 Units and zero divisors

Definition 1.2.1 *Let R be a ring and $a, b \in R$. We say that b divides a if there exists $c \in R$ such that $a = bc$. Notation: $b|a$.*

Definition 1.2.2 *Let R be a ring. An element $a \in R$ is called a unit (or invertible) if it divides 1, i.e. there exists $b \in R$ with $ab = 1$. The set of units in R is denoted by R^* and is called the unit group of R (the fact that it is a group is shown in Theorem 1.2.4).*

Example 1.2.3. $\mathbb{Z}^* = \{1, -1\}$, $\mathbb{Q}^* = \mathbb{Q} - \{0\}$, $\mathbb{R}^* = \mathbb{R} - \{0\}$, $\mathbb{C}^* = \mathbb{C} - \{0\}$. In general, axiom R6 for fields implies that in a field all non-zero elements are units.

◊

Theorem 1.2.4 *The set R^* of units in a ring R is a group with respect to multiplication in R.*

Proof: Let $a, b \in R^*$. Then the element $b^{-1}a^{-1}$ is an inverse of ab. Verification: $ab \cdot b^{-1}a^{-1} = a \cdot a^{-1} = 1$. Therefore $ab \in R^*$.
Associativity follows from axiom (R2).
R^* has a neutral element, namely $1 \in R^*$. Axiom (R4) asserts that $1 \cdot a = a$ for every $a \in R^*$.
Finally, for every $a \in R^*$ there exists $b \in R$ with $ab = ba = 1$. But then automatically $b \in R^*$. So a has an inverse in R^*.
Thus the 4 group axioms have been verified for R^*.

□

Example 1.2.5. Let $R = \mathbb{Z}[\sqrt{m}] = \{a + b\sqrt{m} | a, b \in \mathbb{Z}\}$ as in Example 1.1.9, where m is an integer, not a square. We define the *norm*

$$N : R \longrightarrow \mathbb{Z}, \quad N(a + b\sqrt{m}) = (a + b\sqrt{m}) \cdot (a - b\sqrt{m}) = a^2 - mb^2.$$

One easily verifies that: $N(\alpha\beta) = N(\alpha) \cdot N(\beta)$ for all $\alpha, \beta \in R$. Furthermore, $N(1) = 1$ and

$$N(\alpha) = 0 \iff \alpha = 0.$$

We assert:

$$\alpha \in R^* \iff N(\alpha) = \pm 1.$$

$\Leftarrow$: If $\alpha = a + b\sqrt{m}$ and $N(\alpha) = \pm 1$, then $(a + b\sqrt{m}) \cdot (a - b\sqrt{m}) = \pm 1$, hence $\pm(a - b\sqrt{m})$ is an inverse of α.
$\Rightarrow$: If $\alpha\beta = 1$ then $N(\alpha) \cdot N(\beta) = N(\alpha\beta) = N(1) = 1$, and $N(\alpha)$, $N(\beta) \in \mathbb{Z}$. So $N(\alpha)$ is an integer dividing 1, hence $N(\alpha) = \pm 1$.
Thus we see that the determination of units in $\mathbb{Z}[\sqrt{m}]$ comes down to solving the equation

$$a^2 - m \cdot b^2 = \pm 1$$

in integers a, b.
There is a major difference between the cases $m > 0$ and $m < 0$. When $m < 0$ things are easy, $a^2 - m \cdot b^2 = a^2 + |m| \cdot b^2$. Because squares are positive the norm can only be 1. In addition we conclude that either $a^2 = 1, |m|b^2 = 0$ or $a^2 = 0, |m|b^2 = 1$. Thus we find that $\mathbb{Z}[i]^* = \{\pm 1, \pm i\}$ and $\mathbb{Z}[\sqrt{m}]^* = \{\pm 1\}$ when $m < -1$.
When $m > 0$ (but not a square) the equation $a^2 - mb^2 = \pm 1$ in $a, b \in \mathbb{Z}$ is far more interesting. It is referred to as *Pell's equation* and it has an infinite number of solutions. The group $\mathbb{Z}[\sqrt{m}]^*$ is an infinite group with two generators, namely -1 and the smallest unit $\epsilon = a_0 + b_0\sqrt{m}$ strictly larger than 1. In other words, $\mathbb{Z}[\sqrt{m}]$ consists of

$$..., \pm\epsilon^{-2}, \pm\epsilon^{-1}, \pm 1, \pm\epsilon, \pm\epsilon^2, ...$$

When $m = 2$ we can take $\epsilon = 1 + \sqrt{2}$. This unit corresponds to the solution $a = 1, b = 1$ of $a^2 - 2b^2 = 1$. Looking at the powers of $1 + \sqrt{2}$ we get other solutions $x_n + y_n\sqrt{2} = (1+\sqrt{2})^n$,

$$\begin{array}{llll} x_0 = 1 & y_0 = 0 & x_5 = 41 & y_5 = 29 \\ x_1 = 1 & y_1 = 1 & x_6 = 99 & y_6 = 70 \\ x_2 = 3 & y_2 = 2 & x_7 = 239 & y_7 = 169 \\ x_3 = 7 & y_3 = 5 & & \\ x_4 = 17 & y_4 = 12 & & \end{array}$$

(In general: $x_{n+1} = 2x_n + x_{n-1}$, $y_{n+1} = 2y_n + y_{n-1}$.)
Sometimes the smallest unit ϵ can be quite large. When $m = 67$ it is given by $48842 + 5967\sqrt{67}$, corresponding to the solution $a = 48842$, $b = 5967$ of $a^2 - 67b^2 = 1$.

$\diamondsuit$

In a ring it may happen that $a \cdot b = 0$ whereas $a \neq 0$, $b \neq 0$. For example in $\mathbb{Z}/6\mathbb{Z}$ we have $2 \cdot 3 \equiv 0 (\text{mod } 6)$ and $2, 3 \not\equiv 0 (\text{mod } 6)$. In $\mathbb{Z}/8\mathbb{Z}$ we have $2^3 \equiv 0 (\text{mod } 8)$. We call such elements zero-divisors.

Definition 1.2.6 *Let R be a ring. A non-zero element $a \in R$ is called a zero divisor if there exists non-zero $b \in R$ such that $ab = 0$.*
An element $a \in R$ is called nilpotent if $a \neq 0$ and there exists a positive integer n such that $a^n = 0$. In particular a nilpotent element is a zero-divisor.
An element $a \in R$ is called idempotent if $a \neq 0, 1$ and $a^2 = a$. An idempotent element is always a zero divisor since $a^2 = a$ implies $a(a-1) = 0$ and $0 \neq a \neq 1$ implies a, $a - 1 \neq 0$.

Theorem 1.2.7 *A unit cannot be a zero divisor.*

Proof: Suppose $a \in R$ is both a unit and a zero divisor. So there exist $b, c \in R$, non-zero such that $ba = 0$ and $ac = 1$. Multiply the latter relation by b. We obtain $bac = b$ and because $ba = 0$ this yields $0 = b$, which contradicts the fact that $b \neq 0$.

□

Consequence 1.2.8 *A field has no zero divisors.*

Proof: This follows immediately from Theorem 1.2.7, since all non-zero elements of a divison ring are units by definition.

□

Theorem 1.2.9 *Let $n \in \mathbb{Z}_{>1}$. Then, $\mathbb{Z}/n\mathbb{Z}$ is a field if and only if n is a prime number.*

Remark 1.2.10 *When p is a prime we often denote $\mathbb{Z}/p\mathbb{Z}$ by $\mathbb{F}_p$.*

Proof: Suppose n is not a prime. Then there exist integers $a, b > 1$ such that $n = ab$. In particular $ab \equiv 0(\text{mod } n)$, hence a, b are zero divisors in $\mathbb{Z}/n\mathbb{Z}$.

Suppose that n is prime. Let $a \in \mathbb{Z}$ not divisible by n, so $a(\text{mod } n)$ is a non-zero residue class. Consider the map $\phi : \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ given by $\phi : x(\text{mod } n) \mapsto ax(\text{mod } n)$. The map ϕ is injective. Namely it follows from $ax \equiv ay(\text{mod } n)$ that n divides $ax - ay = a(x - y)$. Since n is a prime which does not divide a we infer that it divides $x - y$. Hence $x \equiv y(\text{mod } n)$ and so ϕ is injective. An injective map from a finite set to itself is also surjective. Hence there exists x such that $ax \equiv 1(\text{mod } n)$. Hence $a(\text{mod } n)$ has an inverse.

□

Definition 1.2.11 *A domain is a ring without zero divisors.*

Example 1.2.12. Fields are examples of domains (because of Consequence 1.2.8), such as

$$\mathbb{Q},\ \mathbb{R},\ \mathbb{C},\ \mathbb{F}_{59},$$

as well as subrings of fields, such as $\mathbb{Z}$, $\mathbb{Z}[i]$. In Section 1.3.4 we shall see that any domain can be embedded in a field. In the next section we will also see that a polynomial ring with coefficients in a domain again forms a domain.

The ring $\mathbb{Z}/n\mathbb{Z}$ is not a domain if n is not prime. When $n = ab$ with $a, b > 1$ the residueclasses a, b are zerodivisors.

♢

Theorem 1.2.13 *Let R be a domain and $a, b, c \in R$. Then:*

a. $\quad ab = 0 \iff a = 0$ *or* $b = 0$,

b. $\quad ab = ac \iff a = 0$ *or* $b = c$.

Proof: The first statement is trivial, if a, b were both non-zero, they would be zero divisors.

The second statement follows from $ab - ac = 0 \Rightarrow a(b - c)$ and the first statement which implies $a = 0$ or $b - c = 0$.

□

1.3 Ring constructions

We describe some standard constructions to manufacture new rings out of given ones.

1.3.1 Polynomial rings

Let R be a ring. A *polynomial* in X with coefficients in R is an expression of the form

$$a_nX^n + a_{n-1}X^{n-1} + \cdots + a_1X + a_0, \qquad n \in \mathbb{Z}_{\geq 0},\ a_i \in R.$$

Here X is a formal symbol. We could also have used Y, T or another symbol. The elements $a_i \in R$ are called the *coefficients* of the polynomial. The coefficient a_0 is called the *constant term.* If $a_i = 0$ for all $i > 0$ we call the polynomial a *constant polynomial.* Two polynomials $\sum_{i=0}^n a_iX^i$ and $\sum_{i=0}^n b_iX^i$ are considered the same if and only if $a_i = b_i$ for $i = 0, 1, \ldots, n$.
The set of polynomials in X with coefficients in R is denoted by $R[X]$. Let $f(X) = \sum_{i=0}^n a_iX^i$ and $g(X) = \sum_{j=0}^m b_jX^j$. Then the sum and product of $f(X)$ and $g(X)$ are defined in the obvious way. That is,

$$f(X) + g(X) = \sum_{i=0}^{n}(a_i + b_i)X^i$$

and

$$f(X)g(X) = \sum_{i=0}^{n}\sum_{j=0}^{n} a_ib_jX^{i+j} = \sum_{k=0}^{2n}(\sum_{i+j=k} a_ib_j)X^k.$$

We have taken $m = n$ here, which we can do without loss of generality. Multiplication is carried out by applying the distributive laws and the rule $X^i \cdot X^j = X^{i+j}$ (in other words, just eliminate parentheses in the usual way). With this addition and multiplication $R[X]$ forms a commutative ring. The zero element is the polynomial with all its coefficients 0. Clearly R itself embeds as the subring consisting of the constant polynomials.
Let $f(X) = \sum_{i=0}^n a_iX^i \in R[X]$. Suppose that $f(X)$ is not 0. For every $a_i \neq 0$ the term a_iX^i in $f(X)$ is called the term of degree i. The largest degree which occurs is called the *degree* of $f(X)$. Notation: $\deg(f)$. In particular, if $a_n \neq 0$ we have $\deg(f) = n$. The coefficient of the highest degree term is called the *leading coefficient* of f. If the leading coefficient of f is 1, we call the polynomial f *monic.*
Consider the example $\mathbb{Z}[X]$ and $f = 2X^4 + 3X^2 - 1, g = X^3 - X + 1$. Then we have $\deg(f) = 4, \deg(g) = 3$. The leading coefficient of f is 2 and the leading coefficient of g is 1

Theorem 1.3.2 *Let R be a domain. Then $R[X]$ is a domain. Moreover, for any $f, g \in R[X]$ with $f \neq 0, g \neq 0$ we have*

$$\deg(f + g) \leq \max(\deg(f), \deg(g)) \qquad \deg(fg) = \deg(f) + \deg(g).$$

Proof: Let $f = \sum_{i=0}^i a_iX^i$ and $g = \sum_{j=0}^m b_jX^m$ be non-zero elements in $R[X]$. We can assume that n, m are chosen in such way that $a_n \neq 0$ and $b_m \neq 0$. Then the highest degree term of $f(X)g(X)$ is given by $a_nb_mX^{m+n}$. Since R is a domain and we have $a_nb_m \neq 0$. Hence $fg \neq 0$. So $R[X]$ has no zero divisors.

In our example we also have $\deg(f) = n, \deg(g) = m$. Since $a_n b_m X^{m+n}$ is the highest degree term of $f(X)g(X)$ we see that $\deg(fg) = m + n = \deg(f) + \deg(g)$, as asserted.
Since $f + g$ consists of sums of terms from f and g, the degree of $f + g$ cannot be larger than that of f or g. In other words $\deg(f + g) \le \max(\deg(f), \deg(g))$.
□

1.3.3 Product of rings

Let R_1 and R_2 be rings. We define the product ring $R_1 \times R_2$ as the set of pairs (r_1, r_2) with $r_1 \in R_1, r_2 \in R_2$ with addition and multiplication defined by

$$(r_1, r_2) + (s_1, s_2) = (r_1 + s_1, r_2 + s_2), \qquad (r_1, r_2) \cdot (s_1, s_2) = (r_1 s_1,\ r_2 s_2)$$

for all $r_1, s_1 \in R_1$ and $r_2, s_2 \in R_2$. One easily verifies the ring axioms for $R_1 \times R_2$. It has the following properties

1. The zero element of $R_1 \times R_2$ is given by $(0, 0)$. Note the first component is the zero of R_1, the second is the zero of R_2.
2. The 1-element of $R_1 \times R_2$ is given by $(1, 1)$.
3. $(R_1 \times R_2)^* = R_1^* \times R_2^*$.
4. When R_1, R_2 are non-trivial rings, the ring $R_1 \times R_2$ has zero divisors because $(a, 0) \cdot (0, b) = (0, 0)$ for all a, b.
5. The elements $(1, 0)$ en $(0, 1)$ are idempotent in $R_1 \times R_2$.

1.3.4 Quotient fields.

Let R be a domain. We will construct the smallest field which contains R and call it the *quotient field* of R. Notation: $Q(R)$. This construction mimicks the construction of the rational numbers out of the integers.
Let $S = R \setminus \{0\}$. On the set $R \times S = \{(a, s) : a, s \in R, s \neq 0\}$ we define the equivalence relation $\sim$ by

$$(a, s) \sim (b, t) \quad \Longleftrightarrow \quad at = bs.$$

The properties of an equivalence relation are easily checked.

1. reflexivity $((a, s) \sim (a, s))$
2. $((a, s) \sim (b, t) \Rightarrow (b, t) \sim (a, s))$ so we have symmetry.
3. To show transitivity let $(a, s) \sim (b, t)$ and $(b, t) \sim (c, u)$. Then $at = bs$ and $bu = ct$. Multiply the first by u and use the second. We get $uat = ubs = cts$. Hence $t(ua - cs) = 0$. Since $t \neq 0$ we conclude that $ua = cs$ (we work in a domain here). Hence $(a, s) \sim (c, u)$.

Let now $Q(R)$ be the set of equivalence classes with respect to $\sim$. The equivalence class of the pair (a, s) is denoted by $\frac{a}{s}$, which is a very suggestive notation for the quotient of a divided by s. We now define addition and multiplication on $Q(R)$ as follows

$$\frac{a}{s} + \frac{b}{t} = \frac{at + bs}{st} \qquad \frac{a}{s} \cdot \frac{b}{t} = \frac{ab}{st}.$$

Note that in both cases $st \neq 0$ because $s, t \neq 0$ and R is a domain. Of course one needs to verify that the addition and multiplication just defined is not dependent on the choice of representatives of the classes $\frac{a}{s}$ and $\frac{b}{t}$. We leave this to the reader. The verification that $Q(R)$ satisfies the axioms (R1) up to (R6) is time consuming but very straightforward. In brief, $Q(R)$ turns out to be a field. Notice that the inverse of $\frac{a}{s}$. when $a \neq 0$, is given by $\frac{s}{a}$.

The ring R can be considered as a subring of $Q(R)$ by identifying $a \in R$ with $\frac{a}{1} \in Q(R)$:

$$R \to Q(R), \qquad r \mapsto \frac{r}{1}.$$

The most well known example is of course $\mathbb{Q} = Q(\mathbb{Z})$. Another common example is the ring of polynomials with coefficients in a field K denoted by $K[X]$. Since K is a field $K[X]$ is a domain. The quotient field of $K[X]$ is denoted by $K(X)$. This is the field of rational functions with coefficients in K. Sample elements of $K(X)$ are $\frac{1}{1+X} = \frac{X}{X+X^2}$ or $\frac{1-X^2}{1-X+X^3}$.

1.3.5 Rings of functions.

Let V be a set, R a ring, and consider the set of maps $f : V \to R$. This set is denoted by R^V. Any two maps $f, g : V \to R$ yield a new map given by $x \in V \mapsto f(x) + g(x)$. We denote this map by $f + g$. Similarly the map given by $x \mapsto f(x)g(x)$ is denoted by fg. The addition and muliplication just defined turn the R-valued functions on V into a ring. The function which is 0 everywhere serves as the zero element in this new ring. The constant function with value 1 is the 1-element.

Function rings become interesting if one requires the functions to have extra properties. For example the set of continous functions $f : [0, 1] \to \mathbb{R}$ on the closed interval is such an example. It is denoted by $C([0, 1])$. The ring of functions $f : \mathbb{R} \to \mathbb{R}$ which are n times continuously differentiable is denoted by $C^n(\mathbb{R})$. Note that all examples above have zero divisors. In the case $C(\mathbb{R})$ one simply takes an arbitrary continuous $f : \mathbb{R} \to \mathbb{R}$ with $f(x) = 0$ for all $x \leq 0$ and $g : \mathbb{R} \to \mathbb{R}$ with $g(x) = 0$ for all $x \geq 0$. Then clearly fg vanishes on all of $\mathbb{R}$.

1.4 Non-commutative rings

In the literature there are many instances where Axiom (R5) is not assumed. Such rings are called *non-commutative rings*. Although we shall deal almost exclusively with commutative rings in these lecture notes, we like to make some remarks and give some examples for the non-commutative case in this section.

A *division ring* (or *skew field*) is a (non)-commutative ring R which, besides (R1) to (R4) also satisfies:

- To every non-zero $a \in R$ there exists $b \in R$ such that $ab = ba = 1$.

Example 1.4.1. Let $n \in \mathbb{Z}_{\geq 0}$. The set $M(n, \mathbb{R})$ of $n \times n$-matrices with real coefficients, with the usual matrix-addition and matrix-multiplication, is a ring with 1 (the $n \times n$ identity matrix). When $n \geq 2$ this ring is not commutative. Analogously one can define $M(n, R)$ for any ring R.

$\diamondsuit$

Example 1.4.2. The *quaternions* are expressions of the form:

$$a + bi + cj + dk, \qquad \text{with} \quad a,\, b,\, c,\, d \in \mathbb{R}.$$

We define a component-wise addition

$$(a+bi+cj+dk)+(a'+b'i+c'j+d'k) = (a+a')+(b+b')i+(c+c')j+(d+d')k.$$

Multiplication of quaternions is based on the following rules:

$$ij = -ji = k, \quad i^2 = -1, \quad j^2 = -1, \qquad \text{and}$$

$$x(a + bi + cj + dk) = (a + bi + cj + dk)x = ax + bxi + cxj + dxk,$$

where $x = x + 0 \cdot i + 0 \cdot j + 0 \cdot k \in K$. To obtain a ring associativity must certainly hold. This implies in particular that:

$$\begin{aligned} k^2 &= (ij)(ij) = ((ij)i)j = (i(ji))j = -i^2j^2 = -1, \\ ik &= i(ij) = -j, \\ ki &= (-ji)i = j, \\ jk &= j(-ji) = i, \end{aligned}$$
$$kj = (ij)j = -i.$$

Elaboration using the distributive laws gives us

$$\begin{aligned} &(a + bi + cj + dk) \cdot (a' + b'i + c'j + d'k) = \\ &\quad (aa' - bb' - cc' - dd') \\ &\quad +(ab' + ba' + cd' - dc')i \\ &\quad +(ac' - bd' + ca' + db')j \\ &\quad +(ad' + bc' - cb' + da')k \end{aligned}$$

A straightforward (but elaborate) verification shows that the quaternions form a (non-commutative) ring. We denote the set of quaternion with the operations above by $\mathbb{H}$.
For any quaternion $q = a + bi + cj + dk$ we write

$$\overline{q} := a - bi - cj - dk.$$

We define

$$N(q) := q\overline{q} = (a+bi+cj+dk)(a-bi-cj-dk) = a^2+b^2+c^2+d^2.$$

In particular $N(q) \in \mathbb{R}_{>0}$ for every non-zero $q \in \mathbb{H}$. Note that it follows from $q\overline{q} = \overline{q}q = N(q)$ that

$$q\frac{\overline{q}}{N(q)} = \frac{\overline{q}}{N(q)}q = 1.$$

In other words, every non-zero $q \in \mathbb{H}$ has an inverse, namely $\overline{q}/N(q)$. So $\mathbb{H}$ is a divison algebra or skew field.

The quaternions were discovered in 1843 by Sir William Rowan Hamilton (1805-1865).

♢

When defining units in a non-commutative ring R we need to make a distinction. An element $a \in R$ is called *left unit* if there exists $b \in R$ such that $ab = 1$, and a *right unit* if there exists $c \in R$ such that $ca = 1$.

If $a \in R$ is both left unit and right unit then a is simply called a *unit*. In that case the inverse elements b, c given above coincide as can be seen from

$$ab = 1, \quad ca = 1 \quad \Longrightarrow \quad cab = c \Longrightarrow b = c.$$

In a non-commutative ring a left unit need not be a right unit and vice versa. Of course, in a commutative ring the concepts left unit, right unit, and unit coincide.

If R is commutative, then of course R^* is abelian. The converse need not hold, there exist non-commutative rings R for which R^* is abelian, see exercise 7.

Example 1.4.3. Suppose $A \in M(n,\mathbb{R})$ is an invertible $n \times n$-matrix with inverse B. Then $AB = BA = I_n$, where I_n is the $n \times n$ identity matrix. Moreover:

A is a left unit $\Longleftrightarrow$ A is a right unit $\Longleftrightarrow$ $\det(A) \neq 0$.

So $M(n,\mathbb{R})^* = GL(n,\mathbb{R})$ (this is actually the definition of $GL(n,\mathbb{R})$, the general linear group in dimension n).

♢

Definition 1.4.4 *A non-zero element a of a ring R is called a left zero divisor if there exists a non-zero $b \in R$ such that $ab = 0$.*

It is called a right zero divisor if there exists a non-zero $b \in R$ such that $ba = 0$.

It is called a zero divisor if it is a left or right zero divisor.

An element $a \in R$ is called nilpotent if $a \neq 0$ and there exists a positive integer n such that $a^n = 0$. In particular a nilpotent element is a zero-divisor.

An element $a \in R$ is called idempotent if $a \neq 0, 1$ and $a^2 = a$. In a ring with 1 an idempotent element is always a zero divisor since $a^2 = a$ implies $a(a-1) = (a-1)a = 0$ and $0 \neq a \neq 1$ implies a, $a - 1 \neq 0$.

Example 1.4.5. In $M(2,\mathbb{R})$ we consider the following elements:

$$a := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad b := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad c := \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Notice that $ab = 0$, so a is a left zero divisor and b is a right zero divisor. Notice that $ba \neq 0$, but $ca = 0$, so a is right zero divisor. Moreover $a^2 = 0$, so a is a nilpotent element (this illustrates that a is both a left and right zero divisor). Note that $b^2 = b$ and $c^2 = c$, so b and c are idempotent elements.

♢

Here is a more advanced example of a non-commutative ring which is often used in representation theory of groups.

1.4.6 Group rings (optional)

Let R be a ring and G a finite group whose composition we write as multiplication. The group ring $R[G]$ of G over R consists of all expressions

$$\sum_{g\in G} a_g \cdot g$$

with $a_g \in R$ for all $g \in G$. Addition is defined componentwise:

$$\left(\sum_{g\in G} a_g \cdot g\right) + \left(\sum_{g\in G} b_g \cdot g\right) = \sum_{g\in G} (a_g + b_g) \cdot g.$$

Multiplication in R and in G can be combined to define multiplication on $R[G]$ as follows:

$$(a_g \cdot g) \cdot (b_h \cdot h) = (a_g b_h) \cdot gh \qquad (a_g, b_h \in R,\ g, h \in G).$$

In the last double summation we collect all terms in front of a given element $k \in G$,

$$\left(\sum_{g\in G} a_g g\right) \cdot \left(\sum_{h\in G} b_h h\right) = \sum_{k\in G} \left(\sum_{g,h,gh=k} a_g b_h\right) k.$$

We leave the proof that $R[G]$ is indeed a ring to the reader.
Since R has a 1 element we can consider G as a subgroup of $R[G]^*$. There are also zero divisors. For example let $g \in G$ have order n and consider the element $1+g+g^2+\cdots+g^{n-1}$. It is clearly non-zero and we have $(1-g)(1+g+\cdots+g^{n-1}) = 1 - g^n = 0$ (note that $1 - g \neq 0$).
When G is an infinite group, we can also define a group ring by taking finite sums of terms of the form $a_g g$, $g \in G$.

1.5 Exercises

1. Let R be a ring and let $1' \in R$ have the property that $1'a = a$ for all $a \in R$. Prove that $1' = 1$.

2. Let R be a ring and $a \in R^*$. Show that there is exactly one element $b \in R$ such that $ab = 1$.

3. Let m be an integer, not the square of another integer. Let $\alpha := \frac{1+\sqrt{m}}{2} \in \mathbb{C}$.

    a. For which m is $R_m := \{a + b\alpha : a, b \in \mathbb{Z}\}$ a subring of $\mathbb{C}$?

    b. Sketch the points of R_{-3} in the complex plane.

4. Let R be a ring with 1 and H an additive subgroup of R. Let $R_0 = \{x \in R | \forall h \in H : xh \in H\}$. Prove that R_0 is a subring of R.

5. Let R be a ring with the property that $x^3 = x$ for all $x \in R$. Prove: $x + x + x + x + x + x = 0$ for all $x \in R$.

6. Let R be a ring consisting of 10 elements. Prove that R is commutative.

7. (Newton's *binomial expansion*). Let R be a ring. We use the notation nr for any $n \in \mathbb{Z}, r \in R$ as in 1. Prove that

$$(*) \qquad (a+b)^n = \sum_{k=0}^{n} \left(\begin{array}{c} n \\ k \end{array} \right) \cdot a^k b^{n-k}$$

    for all $a, b \in R$ and $n \in \mathbb{Z}_{>0}$.

8. Let $\alpha = 1,3247...$ be the real number such that $\alpha^3 = \alpha + 1$. Prove that $\mathbb{Z}[\alpha] := \{a + b\alpha + c\alpha^2 | a, b, c \in \mathbb{Z}\}$ is a subring of $\mathbb{R}$ and that $\alpha, \alpha - 1, \alpha^2 - 1, \alpha^3 - 1 \in \mathbb{Z}[\alpha]^*$. (You may freely use the fact that $a + b\alpha + c\alpha^2 = 0 \iff a = b = c = 0$.)

9. Let m be a positive integer which is not the square of an integer.

    a. Let $\epsilon = a + b\sqrt{m} \in \mathbb{Z}[\sqrt{m}]^*$. Prove: $\{\epsilon, \epsilon^{-1}, -\epsilon, -\epsilon^{-1}\} = \{\pm a \pm b\sqrt{m}\}$. Conclude from this: $\epsilon > 1 \Leftrightarrow a, b > 0$.

    b. We are given that $\mathbb{Z}[\sqrt{m}]^* \neq \{\pm 1\}$. Prove that $\mathbb{Z}[\sqrt{m}]$ contains a smallest unit ϵ_1 with $\epsilon_1 > 1$. Now show that $\mathbb{Z}[\sqrt{m}]^* =< -1, \epsilon_1 >\cong (\mathbb{Z}/2\mathbb{Z}) \times \mathbb{Z}$.

10. Give an example of a ring R which contains an element a with the following properties: $a \neq 0$, a is not a unit in R, and a is not a zero divisor in R.

11. Give an example of an infinite ring with zero divisors.

12. Let R be a commutative ring and R' a subring of R. Provide a proof or counter example to each of the following statements:

    a. If R is a field, then R' is a field.

    b. If R is a domain, then R' is a domain.

    c. If R' is a domain, then R is a domain.

13. Let R_1 and R_2 be rings. Prove that $R_1 \times R_2$ cannot be a domain.

14. An *arithmetic function* is a map $f : \mathbb{Z}_{>0} \to \mathbb{C}$. The *sum* $f_1 + f_2$ of two arithmetic functions f_1 and f_2 is defined by

$$(f_1 + f_2)(n) = f_1(n) + f_2(n).$$

The *convolution product* $f_1 * f_2$ of two arithmetic functions f_1, f_2 is defined by

$$(f_1 * f_2)(n) = \sum_{d|n} f_1(d) f_2\left(\frac{n}{d}\right)$$

where the summation is taken over all positive divisors d of n.

   a. Show that the set R of arithmetic functions with these two operations is a domain.

   b. Let $f \in R$. Prove: $f \in R^* \Leftrightarrow f(1) \neq 0$.

15. a. Let R be a domain and R' a subring of R. Show that $Q(R')$ can be embedded as subring in $Q(R)$.

   b. Prove that for any domain R:

$$R = Q(R) \iff R \text{ is a field.}$$

   c. Let m be an integer and not the square of another integer. Prove that $\mathbb{Q}[\sqrt{m}]$ is the same ring as $Q(\mathbb{Z}[m])$.

   d. Let R be a domain and K a field such that $R \subset K \subset Q(R)$. Show that $K = Q(R)$.

16. Let R be a ring and $S \subset R$ a non-empty multiplicative subset, that is $s, t \in S \Longrightarrow st \in S$.

   a. Prove that the relation $\sim$ defined by

$$(a, s) \sim (b, t) \iff \exists u \in S : \ atu = bsu$$

   is an equivalence relation on $R \times S$.

   b. Let $S^{-1}R = (R \times S)/\sim$, denote by $\frac{a}{s} \in S^{-1}R$ the equivalence class represented by (a, s). Prove that $S^{-1}R$ becomes a ring with 1 with the following addition and multiplication:

$$\frac{a}{s} + \frac{b}{t} = \frac{at + bs}{st}, \qquad \frac{a}{s} \cdot \frac{b}{t} = \frac{ab}{st}.$$

   c. Prove that $S^{-1}R$ is the trivial ring if and only if $0 \in S$.

17. Prove that $\{f \in C[0, 1])$: f is three times continuously differentiable$\}$ is a subring of $C([0, 1])$.

18. A *Boolean ring* (named after the English mathematician George Boole, (1815-1864)) is a (not necessarily commutative) ring R where $x^2 = x$ for all $x \in R$.

a. Prove: $x + x = 0$ for all x in a Boolean ring R.

b. Prove that every Boolean ring is commutative.

c. Let R be a Boolean ring which is at the same time a field. Prove that $R \cong \mathbb{Z}/2\mathbb{Z}$.

19. Let X be a set, and $R = P(X)$ the set of subsets of X. For $A, B \in R$ (so $A, B \subset X$) we define

$$A + B = (A \cup B) - (A \cap B), AB = A \cap B.$$

Prove that R becomes a ring with these operations. Prove that R is a field if and only if X consists of one element. Prove that R is a Boolean ring (Exercise 18).

The following exercises concern rings for which Axiom (R4) need not hold, i.e. R need not contain a 1-element.

1. Let R be a ring, not necessarily with 1, and define on $\mathbb{Z} \times R$ the following addition and multiplication:

$$(n, r) + (m, s) = (n + m, r + s), \qquad (n, r) \cdot (m, s) = (nm, ns + mr + rs)$$

for all $n, m \in \mathbb{Z}$ and $r, s \in R$. We have denoted $2s = s+s,\ 3s = s+s+s, \ldots$

a. Prove that $\mathbb{Z} \times R$ is a ring with 1.

b. Prove that every ring can be embedded in a ring with 1.

2. Let A be an abelian group with the composition written *additively* (with a plus). Define multiplication by $a \cdot b = 0$ for all $a, b \in A$. Prove that A becomes a ring with this multiplication, not necessarily with 1. Does this ring have a unit?

3. Let R be a ring, not necessarily containing 1, with $R^+ \cong \mathbb{Q}/\mathbb{Z}$. Prove that $ab = 0$ for all $a, b \in R$.

1.6 Exercises in non-commutative rings

1. Let R be a non-commutative ring, and $a, b \in R$ such that $ab = 0$. Prove that $(ba)^2 = 0$ and $1 + ba \in R^*$.

2. Let $M(2, 2\mathbb{Z})$ be the set of 2×2-matrices with coefficients in $2\mathbb{Z}$. Prove: with the usual addition and multiplication of matrices $M(2, 2\mathbb{Z})$ is a non-commutative ring *without* 1.

3. Let R be a ring. Define on R a new multiplication $*$ by $a * b = ba$ for all $a, b \in R$. Prove that R provided with the original addition and multiplication $*$ is again a ring. This ring is called the *opposite* ring for R. Notation: R^0.

4. Let R be a ring. The *center* of R is defined by
$$Z(R) = \{a \in R : \forall x \in R : ax = xa\}.$$
Prove that $Z(R)$ is a commutative subring of R.

5. (Newton's *binomial expansion*). Let R be a ring. We use the notation nr for any $n \in \mathbb{Z}, r \in R$ as in 1.

   a. Suppose R is commutative. Prove that
   $$(*) \qquad (a+b)^n = \sum_{k=0}^{n} \left(\begin{array}{c} n \\ k \end{array} \right) \cdot a^k b^{n-k}$$
   for all $a, b \in R$ and $n \in \mathbb{Z}_{>0}$.

   b. Suppose conversely that $(*)$ holds for all $a, b \in R$ and $n \in \mathbb{Z}_{>0}$. Show that R is commutative.

6. Let R be a domain and $n \in \mathbb{Z}_{>1}$. For $A \in M(n, R)$ one defines the determinant $\det(A)$ through the well-known formula
$$\det(A) = \sum_{\sigma \in S_n} \epsilon(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)} \qquad \text{als } A = [a_{ij}]_{1 \leq, i,j \leq n}.$$
Prove: $A \in M(n, R)^* \Longleftrightarrow \det(A) \in R^*$.

7. Let R be a domain. Let $T = \left\{ \left(\begin{array}{cc} a & b \\ c & d \end{array} \right) \in M(2, R) : c = 0 \right\}$.

   a. Prove that T is a subring of $M(2, R)$ and that T is not commutative.

   b. Prove: $[\begin{array}{cc} a & b \\ 0 & d \end{array}] \in T^* \Leftrightarrow a \in R^*$ and $d \in R^*$.

   c. Prove: T^* is commutative $\Longleftrightarrow R^* = \{1\}$.

   d. Suppose that $R = \mathbb{Z}/2\mathbb{Z}$. Prove that T is a non-commutative ring with commutative unit group.

8. Let R be a ring and $a \in R$. Define
$$S = \{x \in R : \ ax = xa\}.$$

   a. Prove that S is a subring of R.

   b. Prove: $S^* = R^* \cap S$.

9. Let $A \in M(n, \mathbb{R})$. Prove: A is a left zero divisor $\Leftrightarrow$ A is a right zero divisor $\Leftrightarrow$ $A \neq 0$ and $\det(A) = 0$.

10. Let K be a field and define on $R = K \times K$ the following addition and multiplication:
$$(x, y) + (u, v) = (x + u, y + v),$$
$$(x, y) \cdot (u, v) = (xu, xv).$$

a. Prove that R is a non-commutative ring without 1.

b. Determine the left zero divisors and the right zero divisors of R.

11. (G. Higman, Proc. London Math. Soc. 46 (1940), 231-248).

a. Let $R = \mathbb{Z}[S_3]$, $a = (13) \cdot \{1 - (12)\}, b = 1 + (12) \in R$. Prove that $ab = 0$, and find a unit in $\mathbb{Z}[S_3]$ not of the form $\pm\sigma$, with $\sigma \in S_3$.

b. Let G be a group and $g \in G$ an element of G of finite order such that $< g >$ is not a normal subgroup of G. Prove that $\mathbb{Z}[G]$ contains a unit not of the form $\pm h$ with $h \in G$.

c. Let G be a group, and suppose $g \in G$ has order 5. Prove that $1 - g - g^{-1} \in \mathbb{Z}[G]^*$.

12. Let R be a ring. Let $v \in R$ be a right inverse of $u \in R$: $uv = 1$. Prove the equivalence of the following three statements:

a. u has more than one right inverse

b. u is not a unit;

c. u is a left zero divisor, that is $\exists x \neq 0 : \; ux = 0$.

13. (Kaplansky) Let R be a ring. Prove that if u has more than one right inverse, then it has infinitely many right inverses. (Hint: if $uv = 1$ and $vu \neq 1$, consider the elements $v + (1 - vu)u^n$.)

14. Let R be a finite ring with 1 and let $u \in R$ with $u \neq 0$. Prove that the following statements are equivalent:

a. u has a right inverse;

b. u has a left inverse;

c. u is not a left zero divisor;

d. u is not a right zero divisor;

e. u is a unit.

15. Let R be a ring with 1. Prove that for any a, $b \in R$:

$$1 - ab \in R^* \iff 1 - ba \in R^* \iff \begin{pmatrix} 1 & a \\ b & 1 \end{pmatrix} \in M(2, R)^*.$$

Chapter 2

Ideals and ring homomorphisms

2.1 Ring homomorphisms

Definition 2.1.1 *Let R_1, R_2 be two rings. A map $f : R_1 \to R_2$ is called ring homomorphism if*

a. $f(1) = 1$.

b. $f(a+b) = f(a) + f(b)$ *for all* $a, b \in R_1$.

c. $f(ab) = f(a)f(b)$ *for all* $a, b \in R_1$.

A *bijective* ring homomorfism is called a *ring isomorphism*, its inverse map is then also a ring isomorphism. Two rings R_1 and R_2 are called *isomorphic* if there is an isomorphism $R_1 \to R_2$. Notation: $R_1 \cong R_2$. An isomorphism of a ring R to itself is called a *(ring) automorphism* of R.

2.1.2 Examples.

a. For any subring R' of a ring R the inclusion $R' \to R$ is an injective ring homomorphism.

b. Let n be a positive integer. The natural map $\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ given by $a \mapsto a(\text{mod } n)$ is a surjective ring homomorphism.

c. Let $\mathbb{R}[X]$ be the ring of polynomials with coefficients in $\mathbb{R}$ and $a \in \mathbb{R}$. Then the map $\mathbb{R}[X] \to \mathbb{R}$ given by $f(X) \mapsto f(a)$ is a homomorphism, called the *evaluation homomorphsim* at a.

d. Let $f : R_1 \to R_2$ be a ring homomorphism. Then the induced map $F : R_1[X] \to R_2[X]$ given by

$$F : a_nX^n + \cdots + a_1X + a_0 \mapsto f(a_n)X^n + \cdots + f(a_1)X + f(a_0)$$

is a homomorphism.

An often used example will be the map $\mathbb{Z}[X] \to (\mathbb{Z}/n\mathbb{Z})[X]$ induced by reduction modulo n on the coefficients.

e. In case R is a non-commutative ring, let $s \in R^*$. Then the map (conjugation by s):

$$\gamma_s : R \longrightarrow R, \qquad r \mapsto srs^{-1}$$

is a bijective ringhomomorphism. When R is commutative the map γ_s is the identity map for all $s \in R^*$. In case $R = M(n, \mathbb{R})$ we know from linear algebra that a change of basis in $\mathbb{R}^n$ induces conjugation on $M(n, \mathbb{R})$.

f. Let R_1, R_2 be rings, then the natural projection $f : R_1 \times R_2 \to R_1$ given by $f((a, b)) = a$, is a ring homomorphism.

Definition 2.1.3 *Let $f : R_1 \to R_2$ be a ring homorphism. Then the image of f is given by*

$$f(R_1) := \{f(x) : \ x \in R_1\}.$$

The kernel of f is defined by

$$\ker(f) := \{x \in R_1 : \ f(x) = 0\}.$$

There are several simple properties of ring homomorphisms which are analogous to homomorphisms between groups or linear maps between vector spaces. We state two of them here.

Proposition 2.1.4 *Let $f : R_1 \to R_2$ be a ring homomorphism. Then*

(i) $f(0) = 0$.

(ii) the image $f(R_1)$ of f is a subring of R_2.

(iii) $\ker(f) = \{0\} \iff f$ is injective.

Proof: (i) Clearly $f(a) = f(a + 0) = f(a) + f(0)$ for any $a \in R_1$. Hence $f(0)$ is the 0-element in R_2.
(ii) Suppose $A, B \in f(R_1)$. Then there exist $a, b \in R_1$ such that $f(a) = A, f(b) = B$. Hence $A - B = f(a) - f(b) = f(a - b) \in f(R_1)$ and $AB = f(a)f(b) = f(ab) \in f(R_1)$. So $f(R_1)$ is a subring of R_2.
(iii) Suppose f is injective. Then clearly, $\ker f = \{0\}$. Suppose conversely that $\ker(f)$ is trivial. Then it follows that $f(x) = f(y) \Rightarrow f(x - y) = 0 \Rightarrow x - y = 0 \Rightarrow x = y$ for all $x, y \in R_1$. Hence f is injective.

□

Definition 2.1.5 *Let R be a commutative ring. An ideal of R is a subset $I \subset R$ satisfying the following properties:*

(I0) $0 \in I$,

(I1) $a - b \in I$ *for all* $a, b \in I$,

(I2) *for all $r \in R$ and $a \in I$ we have $ra \in I$.*

Example 2.1.6. Trivial examples of ideals are $\{0\}$ and R itself.
In $R = \mathbb{Z}$ the multiples of a given integer n (denoted by $n\mathbb{Z}$), form an ideal.
More generally, the set $\{ra | \ r \in R\}$ consisting of the multiples of a given element $a \in R$, form an ideal.

♢

Theorem 2.1.7 *Let $f : R_1 \to R_2$ be a ring homomorphism. Then* $\ker(f)$ *is an ideal in R_1.*

Proof: Let $\ker(f)$ be the kernel of a ring homomorphism $f : R_1 \to R_2$. First of all we know that $f(0) = 0$. Hence $0 \in \ker(f)$.
Secondly, suppose $a, b \in \ker(f)$. Then $f(a - b) = f(a) - f(b) = 0 - 0 = 0$. Hence $a - b \in \ker(f)$.
Thirdly, let $a \in \ker(f)$ and $r \in R_1$. Then $f(ra) = f(r)f(a) = f(r) \cdot 0 = 0$. Hence $ra \in \ker(f)$.

□

Later on we shall see that any ideal is the kernel of a suitably chosen ring homomorphism.

Example 2.1.8. Let n be a positive integer and consider the mod n mapping $\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$. The kernel consists of the set of multiples of n, which is of course an ideal.

♢

Example 2.1.9. Let $a \in \mathbb{R}$ and consider the evaluation homorphsim $\mathbb{R}[X] \to \mathbb{R}$ given by $f(X) \mapsto f(a)$. The kernel consists of all polynomials with a zero in a. It is not hard to verify that this is also an ideal.

♢

Example 2.1.10. Let $\mathbb{R}[X, Y]$ be the ring of polynomials in two variables. Consider the map $\mathbb{R}[X, Y] \to \mathbb{R}$ given by $f(X, Y) \mapsto f(0, 0)$. The kernel consists of all polynomials whose constant term is zero. One easily verifies that this is an ideal.

♢

Definition 2.1.11 *Let R be a ring and let a_1, $a_2 \ldots, a_n \in R$. Then the ideal generated by $a_1, \ldots, a_n$ is defined as the ideal given by*

$$\{r_1a_1 + r_2a_2 + \ldots + r_na_n : \quad r_1, \, r_2, \ldots, r_n \in R\}$$

Notation: $(a_1, a_2, \ldots, a_n)$. In particular, since the number of generators is finite we call the ideal finitely generated. When $n = 1$ we speak of a principal ideal.

Example 2.1.12. The multiples of an integer $n > 1$ form a principal ideal in $\mathbb{Z}$, namely (n).

♢

Example 2.1.13. Consider $R = \mathbb{Z}[i] = \{a + bi |\ a, b \in \mathbb{Z}\}$ and I the subset given by all $a + bi$ with $a \equiv b(\text{mod } 2)$. To see that this is an ideal we need to verify two things

(I1) $x, y \in I \Rightarrow x - y \in I$. This is clear, let $x = a_1 + b_1 i, y = a_2 + b_2 i$ then $x - y = (a_1 - a_2) + (b_1 - b_2)i$ and $(a_1 - a_2) - (b_1 - b_2) \equiv (a_1 - b_1) - (a_2 - b_2) \equiv 0 - 0(\text{mod } 2)$.

(I2) $x \in I, r \in \mathbb{Z}[i] \Rightarrow rx \in I$. Let $r = u + vi, x = a + bi$. Then $rx = ua - bv + (ub + va)i$ and $ua - bv - (ub + va) \equiv u(a - b) - v(a - b) \equiv 0 - 0 = 0(\text{mod } 2)$.

It turns out that $I = (1 + i)$. Notice that $1 + i \in I$, hence $(1 + i) \subset I$. Conversely suppose that $a + bi \in I$, in other words $a \equiv b(\text{mod } 2)$. Now note that $a + bi = (\frac{a+b}{2} + \frac{b-a}{2}i)(1 + i)$ and $\frac{a+b}{2} + \frac{a-b}{2}i \in \mathbb{Z}[i]$ because a, b have the seame parity. We conclude that $a + bi \in (1 + i)$.

♢

Example 2.1.14. Let R be a ring and $a \in R$. Let I be the ideal of polynomials in $R[X]$ with a zero in a. First of all notice that $X - a \in I$. Also every multiple of $X - a$ is in I. Hence $(X - a) \subset I$. But in fact we have equality, every polynomial which vanishes in a is a multiple of $X - a$. We can see this as follows. Let $p(X) = \sum_{i=0}^{n} p_i X^i$. Then

$$p(X) - p(a) = \sum_{i=0}^{n} p_i (X^i - a^i).$$

Every term $X^i - a^i$ is divisible by $X - a$, namely

$$X^i - a^i = (X - a)(X^{i-1} + aX^{i-2} + \cdots + a^{i-2}X + a^{i-1}).$$

Hence $p(X) - p(a)$ is divisible by $X - a$. In particular, if $p(a) = 0$ then $p(X)$ is divisible by $X - a$, hence $p(X) \in (X - a)$.

♢

Example 2.1.15. Let I be the ideal of polynomials in $\mathbb{R}[X, Y]$ with a vanishing constant term. They do not form a principal ideal. However, this ideal can be described by (X, Y). We can see this as follows. First of all any polynomial in (X, Y) has the form $Xp(X, Y) + Yq(X, Y)$. So its constant term is zero. Conversely, suppose $f(X, Y)$ is a polynomial with zero constant term. Then it is a sum of terms of the form cX^iY^j with $\max(i, j) \geq 1$. If $i \geq 1$ the term is divisible by X, hence in (X, Y). If $j \geq 1$ the term is again in (X, Y). By the additive property of ideals we have $f(X, Y) \in (X, Y)$.

♢

Theorem 2.1.16 *Let R be a ring and $r \in R^*$. Then $(r) = R$.*
The only ideals in a field are R and (0).

Proof: The ideal generated by 1 of course equals R. For any unit $r \in R^*$ the ideal also contains $r^{-1}r = 1$. Hence $(r) = R$.

Let R be a field and $I \subset R$ and ideal. Suppose I contains a non-zero element a. Since R is a field a is a unit. And hence $(a) = R$. Since $(a) \subset I$ this also implies $I = R$.

□

Consequence 2.1.17 *A ring homomorphism $f : K \to R$ from a field K to a nontrivial ring R is injective.*

Proof: The kernel $\ker(f)$ of f is an ideal in K, hence by Theorem 2.1.16 $\ker(f)$ is either (0) or K. In the latter case $f(1) = 0 \neq 1$, hence f is not a homomorphism. In the first case f is injective.

□

2.2 The factor ring R/I.

Let R be a ring and $I \subset R$ an ideal. We call two elements $a, b \in R$ *equivalent modulo* I if $a - b \in I$. One easily verifies that this is an equivalence relation. The equivalence classes with respect to this relation are called (residue) classes modulo I. A class which contains $a \in R$ is denoted by $a + I$, $a(\text{mod } I)$ or $\overline{a}$. Of course two classes $a + I, b + I$ are equal if and only if $a - b \in I$. Addition of two classes is defined by:

$$(a + I) + (b + I) := (a + b) + I, \quad \text{i.e.} \quad \bar{a} + \bar{b} = \overline{a + b}.$$

and multiplication by:

$$(a + I) \cdot (b + I) := ab + I, \qquad \text{d.w.z.} \quad \bar{a} \cdot \bar{b} = \overline{ab}.$$

It remains to check that these operations are well-defined, that is: the outcome is independent of the choice of a, b in the classes $a + I, b + I$. So we need to verify that for any $a' \in a + I$ and $b' \in b + I$ we get $a' + b' \in a + b + I$ and $a'b' \in ab + I$.
To see the first statement notice that $a' + b' - (a + b) = a' - a + b' - b \in I$ because $a' - a, b' - b \in I$. To see the second statement notice that $a'b' - ab = (a' - a)b' + a(b' - b) \in I$.
The set of classes modulo I is denoted by R/I.

Theorem 2.2.1 *With the addition and multiplication just defined the set R/I forms a ring. The zero element is given by the class $\overline{0} = 0 + I$.*

The proof is a straightforward verification of the ring axioms.

Example 2.2.2. The ring $\mathbb{Z}/n\mathbb{Z}$ is the first example of the factor ring construction. Taking $n = 6$, we see that R/I may have zero divisors even if R is a domain.

♢

Theorem 2.2.3 *Let R be a ring and $I \subset R$ an ideal. The natural map which maps R to R/I by assigning to $a \in R$ the class $a+I$ is a ring homomorphism. It is surjective and its kernel is I.*

Proof: The homorphism property follows directly from the definition of addition and multiplication of classes modulo I. Surjectivity follows from the fact that R/I is by definition the set of all classes modulo I. The fact that the kernel is I comes from the fact that $0+I$ is the zero element in R/I.

□

It turns out that a ring homomorphism $f : R_1 \to R_2$ can in fact be thought of as a canonical map $R_1 \to R_1/\ker(f)$ followed by an embedding (=injective map) into R_2. This is described in more detail by the homomorphism and isomorphism theorems.
The following theorem will be the most heavily used tool in proving isomorphisms between rings.

Theorem 2.2.4 (First isomorphism Theorem) *Let $f : R_1 \to R_2$ be a ring homomorphism. Then the rings $R_1/\ker(f)$ and $f(R_1)$ are isomorphic. In particular, if f is surjective, then $R_1/\ker(f) \cong R_2$.*

Proof: In this proof we let $I = \ker(f)$. We define the map $g : R_1/I \to R_2$ as follows. Let A be a congruence class modulo I. Choose $a \in A$, and then define $g(A) = f(a)$. This map is well-defined, that is: independent of the choice of the representing element $a \in A$. Namely, let $a' \in A$. Then we would get $g(A) = f(a')$ which is the same as $f(a)$ because $a - a' \in \ker(f)$. The map g satisfies the homomorphism axioms:

$$\begin{aligned} g(a+b+I) &= f(a+b) = f(a)+f(b) = g(a+I)+g(b+I) \\ g(ab+I) &= f(ab) = f(a)f(b) = g(a+I)g(b+I) \end{aligned}$$

Furthermore $g(a+I) = 0 \Rightarrow f(a) = 0 \Rightarrow a \in I$. So g has trivial kernel, it is injective. Since g is surjective on its image we see that $g : R_1/\ker(f) \to f(R_1)$ is a ring isomorphism.

□

For the examples of ideals given earlier we can now give a description of factor rings through the isomorphism theorem.

Example 2.2.5. Let $R = \mathbb{Z}[i]$ and $I = \{a+bi| \ a \equiv b (\text{mod } 2)\}$. Consider the map $\phi : \mathbb{Z}[i] \to \mathbb{Z}/2\mathbb{Z}$ given by $\phi : a+bi \mapsto a-b (\text{mod } 2)$. It is a homomorphism, namely

$$\begin{aligned} \phi(a_1+b_1i+a_2+b_2i) &= a_1+a_2-b_1-b_2 (\text{mod } 2) \\ &= (a_1-b_1)+(a_2-b_2)(\text{mod } 2) = \phi(a_1+b_1i)+\phi(a_2+b_2i) \\ \phi((a_1+b_1i)(a_2+b_2i)) &= (a_1a_2-b_1b_2)-(a_1b_2+a_2b_1)(\text{mod } 2) \\ &= a_1a_2+b_1b_2-a_1b_2-a_2b_1(\text{mod } 2) \\ &= (a_1-b_1)(a_2-b_2)(\text{mod } 2) = \phi(a_1+b_1i)(a_2+b_2i) \end{aligned}$$

It is of course surjective and its kernel is I. We have seen before that $I = (1+i)$, hence we conclude $\mathbb{Z}[i]/(1+i) \cong \mathbb{Z}/2\mathbb{Z}$.

♢

Example 2.2.6. We use the first isomorphism theorem to prove that for any positive integer N:

$$\mathbb{Z}[X]/N\mathbb{Z}[X] \cong (\mathbb{Z}/N\mathbb{Z})[X].$$

Here $(\mathbb{Z}/N\mathbb{Z})[X]$ is the ring of polynomials in X with coefficients in $\mathbb{Z}/N\mathbb{Z}$. Define the map $\psi : \mathbb{Z}[X] \to (\mathbb{Z}/N\mathbb{Z})[X]$:

$$\psi : \sum_{i=0}^{n} a_i X^i \mapsto \sum_{i=0}^{n} \overline{a_i} X^i,$$

where $\overline{a_i} \in \mathbb{Z}/N\mathbb{Z}$. In other words ψ is the operation which considers the coefficients of a polynomial modulo N. Verify that ψ is surjective ring homomorphism. A polynomial $f(X)$ in the kernel of ψ are characterised by the fact that $\overline{a_i} = 0$, in other words all coefficients a_i are divisible by N. Hence $f(X)$ is N times $f(X)/N$, which is in $\mathbb{Z}[X]$. Conversely, any polynomial of the form $Ng(X)$ with $g(X) \in \mathbb{Z}[X]$ is in $\ker(\psi)$. Hence $\ker(f)$ is the ideal $N\mathbb{Z}[X]$. The isomorphism theorem yields our assertion.

♢

There are also a second and third isomorphism theorem. But their statement and proof have been relegated to Problems 17 and 18.

2.3 Ideal arithmetic

In Definition 2.1.11 we defined

$$(a_1, \ldots, a_n) = \{r_1 a_1 + \ldots + r_n a_n : \quad r_i \in R\},$$

the ideal generated by $a_1, \ldots, a_n$ in a ring R. A *principal ideal* is an ideal which can be generated by one element.

Definition 2.3.1 *A domain R is called a principal ideal domain if every ideal in R is a principal ideal.*

Theorem 2.3.2 *The ring $\mathbb{Z}$ is a principal ideal domain. In particular, any ideal $\neq (0)$ is generated by its smallest positive element.*

Proof: Let $I \subset \mathbb{Z}$ be an ideal. If I contains only zero, then clearly $I = (0)$ in other words it is principal. Suppose that I contains a non-zero element a. Then it also contains $-a$. So I contains positive elements. Let d be the smallest positive element. We claim that $I = (d)$. To see this note that obviously $(d) \subset I$. Suppose now that $a \in I$. Write $a = qd + r$ with integers q, r with $0 \leq r < d$ (division by d with remainder r). Notice that, since $a, d \in I$ we have that $r \in I$. If r is non-zero, then by $r < d$ this would contradict the minimality of d. Therefore we conclude that $r = 0$ and $a = qd$ is a multiple of d. Hence $I \subset (d)$.

This proves our theorem.

□

Example 2.3.3. As we indicated before, the ideal (X,Y) in $\mathbb{R}[X,Y]$ is not principal. In the proof we will use the fact that a polynomial $f \in \mathbb{Q}[X,Y]$ has a degree in X and a degree in Y. Furthermore a polynomial p such that $\deg_X(p) = \deg_Y(p) = 0$ is constant. Suppose that (X,Y) is principal and generated by the two variable polynomial $f(X,Y)$. Since $X,Y \in (f)$ there exist polynomials h,k such that $X = fh$ and $Y = fk$. From $X = fh$ we infer $0 = \deg_Y(X) = \deg_Y(f) + \deg_Y(h)$. Hence $\deg_Y(f) = 0$. Similarly we find that $0 = \deg_X(Y) = \deg_X(f) + \deg_X(k)$. Hence $\deg_X(f) = 0$. Hence we conclude that f is a (non-zero) constant. But then f is a unit and $(f) = \mathbb{R}[X,Y]$. Since $R[X,Y] \neq (X,Y)$ we get a contradiction, so (X,Y) cannot be principal.

♢

Example 2.3.4. Let $R = \{a + b\sqrt{-6}|\ a,b \in \mathbb{Z}\}$ and $I = (2, \sqrt{-6})$. First we make the observation that the norm (=absolute value squared) of any element in I is an even integer. Namely,

$$|(a+b\sqrt{-6})2+(c+d\sqrt{-6})\sqrt{-6}|^2 = |2a-6d+(2b+c)\sqrt{-6}|^2 = (2a-6d)^2+6(2b+c)^2.$$

We now show that I is not a principal ideal. Suppose there exist $u,v \in \mathbb{Z}$ such that $I = (u + v\sqrt{-6})$. Then there exist $x,y \in \mathbb{Z}$ such that $2 = (x + y\sqrt{-6})(u + v\sqrt{-6})$. Take the norm on both sides. We get $4 = (x^2 + 6y^2)(u^2 + 6v^2)$. When y or v is non-zero at least one of the factors on the right is at least $6 > 4$. So we get a contradiction. Hence $y = v = 0$. So $I = (u)$ with $u \in \mathbb{Z}$. Furthermore there exist $x,y \in \mathbb{Z}$ such that $\sqrt{-6} = (x + y\sqrt{-6})u$. From this we see that $ux = 0$ and $uy = 1$. Therefore $u = \pm 1$. However, ± 1 cannot be in I because its norm is 1, which is odd.

♢

Later on we shall see that $K[X]$ is a principal ideal ring whenever K is a field.

Definition 2.3.5 *Let R be a ring and I,J ideals in R. The sum of I and J is defined by*

$$I + J = \{x + y : x \in I, y \in J\}.$$

The product of I and J is defined as the set of finite sums with terms of the form ij with $i \in I, j \in J$. More explicitly,

$$IJ = \{i_1 j_1 + \cdots + i_r j_r|\ i_k \in I,\ j_l \in J\}.$$

Clearly the sum and product of two ideals are again ideals. In an analogous way we can also define sum and product of more than two ideals. These are again ideals.

Remark 2.3.6 *Let R be a ring and I,J two ideals. Then $I \cap J$ is again an ideal and $IJ \subset I \cap J$.*

Proof: The fact that $I \cap J$ is an ideal follows directly from the properties of I and J. Furthermore any element ij with $i \in I$ and $j \in J$ is contained in $I \cap J$. Hence any finite sum of such terms is in $I \cap J$.

□

We say that I and J are *relatively prime* or *co-prime* when $I + J = R$.
In particular, there exist elements $i \in$ and $j \in J$ such that $i + j = 1$. Conversely, the existence of such i, j implies that $ri + rj = r$ for any $r \in R$, hence $I + J = R$.

Example 2.3.7. To get a better intuition we consider sums and product of ideals in $\mathbb{Z}$. Since $\mathbb{Z}$ is a principal ideal domain, every ideal in $\mathbb{Z}$ has the form (n) for some integer n. Suppose $I = (m)$, $J = (n)$ where m, n are non-zero. Again $I + J$ is principal that is, there exists an integer d such that $(d) = (m) + (n)$. We may assume d to be positive and assert that $d = \gcd(m, n)$, the greatest common divisor of m, n. To see this notice that $m, n \in (d)$. Hence m, n are multiples of d. In particular d is a common divisor of m, n. Furthermore $(d) = (m) + (n)$ implies the existence of $x, y \in \mathbb{Z}$ such that $d = mx + ny$. The greatest common divisor of m, n also divides $mx + ny$ and hence d. Therefore $d = \gcd(m, n)$.
In particular the statement $(m) + (n) = \mathbb{Z}$ is equivalent to $\gcd(m, n) = 1$.
The intersection of the ideals $(n), (m)$ is the principal ideal generated by $\operatorname{lcm}(m, n)$, the *least common multiple* of m, n. This can be seen as follows. There exists a positive integer N such that $(m) \cap (n) = (N)$. Note that an integer is contained in $(m) \cap (n)$ if and only if it is a multiple of both m and n. Since N is the smallest such positive element in (N) it must be the smallest common multiple.
Finally, the product of the ideals (m) and (n) consists of multiples of mn. Furthermore mn itself is contained in $(m)(n)$. Hence $(m)(n) = (mn)$.

♢

Example 2.3.8. From the definitions it follows that

$$(a_1, \ldots, a_n) \cdot (b_1, \ldots b_m) = (a_1 b_1, \ldots, a_i b_j, \ldots, a_n b_m).$$

Furthermore: $(a,\, b) = (a + rb,\, b)$ for all $a,\, b,\, r \in R$ (please check), in other words you can 'sweep' in ideals, the same way we do with systems of linear equations. In $R = \mathbb{Z}[\sqrt{-5}]$ we have for example:

$$\begin{aligned} & (2,\, 1+\sqrt{-5}) \cdot (3,\, 1+\sqrt{-5}) \\ = \;& (6,\, 2+2\sqrt{-5},\, 3+3\sqrt{-5},\, -4+2\sqrt{-5}) \\ = \;& (6,\, 2+2\sqrt{-5},\, 1+\sqrt{-5}, -6) \\ = \;& (6,\, 1+\sqrt{-5}) \\ = \;& (1+\sqrt{-5}) \end{aligned}$$

The ideals $(2,\, 1+\sqrt{-5})$ and $(3,\, 1+\sqrt{-5})$ are not principal (see exercise 7), Nevertheless their product is principal. Multiplication of ideals plays a crucial role in *algebraic number theory.*

♢

Theorem 2.3.9 (Chinese remainder theorem) *Let R be a ring. Let I, J be two relatively prime ideals, that is: $I + J = R$. Then to any $a, b \in R$ there exists $r \in R$ such that $r \equiv a (\mathrm{mod}\ I)$ and $r \equiv b(\mathrm{mod}\ J)$.*
More precisely, the ring homomorphism

$$\phi : R \to (R/I) \times (R/J)$$

given by

$$r \mapsto (r + I, r + J)$$

is surjective with kernel $I \cap J$. Furthermore, $I \cap J = I \cdot J$ and consequently (by the first isomorphism theorem),

$$R/(I \cdot J) \cong R/I \times R/J.$$

Proof: Since $I + J = R$ there exist $x \in I,\ y \in J$ such that $x + y = 1$. Let $a, b \in R$ be as in the statement of the theorem. Choose $r = bx + ay$. Then $r = bx+ay = bx+a(1-x) = a+(b-a)x$. Since $x \in I$ this implies $r \equiv a(\mathrm{mod}\ I)$. Similarly, $r = bx + ay = b(1 - y) + ay = b + (a - b)y$ and hence $r \equiv b(\mathrm{mod}\ J)$. This result implies that ϕ is a surjective homomorphism. Notice that $\phi(a) = 0$ if and only if $a \in I$ and $a \in J$. In other words, $\ker(\phi) = I \cap J$.
It remains to show that $I \cap J = I \cdot J$. Since $I \cdot J$ is contained in both I and J we have that $I \cdot J \subset I \cap J$. Let $z \in I \cap J$. Let $x \in I, y \in J$ be as above. Then $z = z(x + y) = zx + zy$. Since $zx \in J \cdot I$ and $zy \in I \cdot J$ we see that $z = zx + zy \in I \cdot J$. Hence $z \in I \cdot J$ and we conclude that $I \cap J \subset I \cdot J$. The equality $I \cap J = I \cdot J$ is now proven.

□

Consequence 2.3.10 *Let $n, m \in \mathbb{Z}$ be relatively prime. Then the ring homomorphism $\mathbb{Z}/mn\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$ given by $a(\mathrm{mod}\ mn) \mapsto (a(\mathrm{mod}\ n), a(\mathrm{mod}\ m)$ is an isomorphism.*

Proof: This follows immediately from Theorem 2.3.9 because $\gcd(m, n) = 1$ is equivalent to $(m) + (n) = \mathbb{Z}$.

□

Notice that relative primality of m, n cannot be omitted. For example, when $m = n = 2$ we see that $a(\mathrm{mod}\ 4) \mapsto (a(\mathrm{mod}\ 2), a(\mathrm{mod}\ 2))$ hence $(0(\mathrm{mod}\ 2), 1(\mathrm{mod}\ 2))$ cannot be in the image.

Example 2.3.11. Let $R = \mathbb{Q}[X]$ and $I = (X - 1)$, $J = (X + 1)$. Notice that $\frac{1}{2}(X + 1) - \frac{1}{2}(X - 1) = 1$, so I and J are relatively prime ideals. The product ideal equals $IJ = (X^2 - 1)$. So the Chinese remainder theorem implies that

$$\mathbb{Q}[X]/(X^2 - 1) \cong (\mathbb{Q}[X]/(X - 1)) \times (\mathbb{Q}[X]/(X + 1)).$$

Furthermore, the evaluation homomorphism $Q[X] \to \mathbb{Q}$ given by $f(X) \mapsto f(1)$ is surjective with kernel $(X - 1)$. The first isomorphism theorem then implies $\mathbb{Q}[X]/(X - 1) \cong \mathbb{Q}$. Similarly, $\mathbb{Q}[X]/(X + 1) \cong \mathbb{Q}$. Hence

$$\mathbb{Q}[X]/(X^2 - 1) \cong \mathbb{Q} \times \mathbb{Q}.$$

Remark: on the contrary, $\mathbb{Z}[X]/(X^2-1) \not\cong \mathbb{Z}\times\mathbb{Z}$ (see Exercise 13).

$\diamondsuit$

Example 2.3.12. Let $R = R_1 \times R_2$, where R_1, R_2 are rings. Then $(1,0)$ and $(0,1)$ are idempotents of R, that is: elements $e \in R$ with the property that $e^2 = e$. We shall now prove that any idempotent in a ring arises in this way.
Let R be a ring and $e \in R$ an idempotent. We apply the Chinese remainder theorem to the ideals (e) and $(1-e)$. Clearly they are relatively prime, because $e + (1-e) = 1$. Their product ideal equals $(e(1-e)) = (e - e^2) = (0)$. Hence we get

$$R \cong R/(0) \cong R/(e) \times R/(1-e).$$

The corresponding isomorphism is given by $r \mapsto (r + (e), r + (1-e))$. The element e is then mapped to $(e + (e), e + (1-e)) = ((e), 1 + (1-e)) = (\overline{0}, \overline{1})$. Similarly $1-e$ is mapped to $(\overline{1}, \overline{0})$.
We conclude that there is one to one correspondence between the idempotents of a ring with 1 and the ways in which R can be written as a direct product of rings.

$\diamondsuit$

2.4 Exercises

1. Check whether the following maps between rings are homomorphisms.

   (a) $f : \mathbb{C} \to \mathbb{C}$ given by $f : z \mapsto 2z$.

   (b) $f : \mathbb{C} \to \mathbb{C}$ given by $f : z \mapsto \overline{z}$.

   (c) $f : \mathbb{Z}[X] \to \mathbb{Z}[X]$ given by $f : P(X) \mapsto P(X^2)$.

   (d) $f : \mathbb{Z}[X] \to \mathbb{Z}[X]$ given by $f : P(X) \mapsto P(X)^2$.

   (e) $f : \mathbb{Q}[X] \to \mathbb{Q} \times \mathbb{Q}$ given by $f : P(X) \mapsto (P(1), P(-1))$.

   (f) $f : \mathbb{Q}[X] \to \mathbb{Q}$ given by $f : P(X) \mapsto P(1)P(-1)$.

   (g) Let R be the boolean ring consisting of subsets of a set X (see Exercise 19). Let $V \subset X$ be non-empty and take $f : R \to R$ given by $f : A \mapsto A \cap V$.

   (h) Same ring as before, but f given by $f : A \mapsto A \cup V$.

2. Let R be a ring with 1. Prove that there exists precisely one ringhomomorphism $f : \mathbb{Z} \to R$.
   N.B. The non-negative generator of $ker(f)$ is called the *characteristic* of R. Notation: char(R).

3. Prove that the characteristic of a domain is either 0 or a prime number.

4. Prove that the following rings have only the identity map as automorphism:

$$\mathbb{Z}, \quad \mathbb{Z}/n\mathbb{Z}, \quad \mathbb{Q}.$$

5. Let σ be a ring automorphism of $\mathbb{R}$.

a. Prove: $x > 0 \Rightarrow \sigma(x) > 0$.

b. Prove: $\sigma = id_{\mathbb{R}}$.

6. Define $\phi : \mathbb{Z}[X] \to \mathbb{Z}/2\mathbb{Z}$ by $\phi : f(X) \mapsto f(0) + 2\mathbb{Z}$.

   a. Prove that ϕ is a surjective homomorphism and show that $\ker(\phi) = (2, X)$.

   b. Prove that $(2, X)$ is not a principal ideal.

7. Define $\phi : \mathbb{Z}[\sqrt{-5}] \to \mathbb{Z}/3\mathbb{Z}$ by $\phi(a + b\sqrt{-5}) = a + b (\text{mod } 3)$.

   a. Prove that ϕ is a surjective ring homomorphism.

   b. Prove that $ker(\phi) = (3, 1 - \sqrt{-5})$

   c. Prove that $ker(\phi)$ is not a principal ideal. (Hint: suppose $ker(\phi) = (x)$, with $3 = xy$ and $1 - \sqrt{-5} = xz$, then consider $N(xy)$ and $N(xz)$ with $N(a + b\sqrt{-5}) = a^2 + 5b^2$ as in 1.2.5.)

   d. Show that $(2, 1 + \sqrt{-5})$ is not a principal ideal.

   e. Is the ideal $(3, 1 - \sqrt{-5}) \cdot (3, 1 - \sqrt{-5})$ principal?

8. Define $\varphi : \mathbb{Z}[i] \to \mathbb{F}_{13}$ by $\varphi(a + bi) = a + 5b (\text{mod } 13)$.
   Prove that φ is a homomorphsim, en show that $ker(\varphi)$ is the ideal generated by 13 and $i - 5$. Find a single generator for $ker(\varphi)$.

9. Let R_1 and R_2 be rings, and $I = \{0\} \times R_2 \subset R_1 \times R_2$.

   a. Prove that I is an ideal in $R_1 \times R_2$.

   b. Prove that I is a principal ideal.

10. Let R_1 and R_2 be rings. Prove that all ideals in $R_1 \times R_2$ have the shape $I_1 \times I_2$ where I_i is an ideal in R_i $(i = 1, 2)$.

11. Let R be a non-trivial ring with 1, and suppose that $f : R \to R$, $f(x) = x^2$ is a ring homomorphism. Prove that R has characteristic 2 (see exercise 2).
    Prove also that $1 + x \in R^*$ for all $x \in \ker f$.

12. Let $I \subset R$ be an ideal and $\phi : R \to R/I$ the natural homomorphism.

    a. Let $J' \subset R/I$ be an ideal. Prove that $\phi^{-1}(J')$ is an ideal in R. Notice in particular that $I \subset \phi^{-1}(J')$.

    b. Prove that $J' \mapsto \phi^{-1}(J')$ provides a bijection between the ideals J' in R/I and the ideals J in R with $I \subset J$.

    c. Prove that for any ideal $J \subset R$ with $I \subset J$ we have $(R/I)/\phi(J) \cong R/J$.

13. Show that $\mathbb{Z}[X]/(X^2 - 1) \ncong \mathbb{Z} \times \mathbb{Z}$. (hint: determine the set of solutions to $a^2 = 1$ for both rings).

14. Let K be a field. The ring of *dual numbers* over K, notation: $K[\epsilon]$, consists of expressions $a + b\epsilon$, with $a, b \in K$, with the following addition and multiplication,

$$(a + b\epsilon) + (c + d\epsilon) = (a + c) + (b + d)\epsilon$$

$$(a + b\epsilon) \cdot (c + d\epsilon) = ac + (ad + bc)\epsilon)$$

(so $\epsilon^2 = 0$), for all $a, b, c, d \in K$.

   a. Show that $K[\epsilon]$ is a ring isomorphic to $K[X]/(X^2)$.
   b. Prove that $K[\epsilon]$ has precisely *three* ideals.
   c. Prove: $K[\epsilon]^* \cong K^* \times K^+$ (as groups).

15. Let R be a ring and $I = R - R^*$. Suppose that for every $x \in I$ there exists $n \in \mathbb{Z}_{>0}$ such that $x^n = 0$. Prove that I is an ideal in R and show that R/I is a field.

16. Let R be a ring and $I \subset R$ an ideal. Show that the set

$$\{x \in \mid \exists n \in \mathbb{Z}_{>0} \text{ such that } x^n \in I\}$$

is an ideal in R. This ideal is known as the *radical ideal* of I. Notation $\sqrt{I}$.

17. Let R be a ring, $I \subset R$ an ideal, and $R' \subset R$ a subring. Prove:

   a. $R' \cap I$ is an ideal in R';
   b. $R' + I = \{r + s : \ r \in R', \ s \in I\}$ is a subring of R;
   c. $R'/(R' \cap I) \cong (R' + I)/I$.

   This statement is known as the second isomorphism theorem for rings.

18. Let R be a ring, $I \subsetneq R$ an ideal and $\phi : R \to R/I$ the canonical map.

   (a) Let $J \subset R$ be an ideal. Show that $\phi(J)$ is an ideal in R/I.
   (b) Show that $J \mapsto \phi(J)$ gives a 1-1 correspondence between the ideals $J \subset R$ with $I \subset J$ and the ideals in R/I.
   (c) Show that for every ideal $J \subset R$ containing I we have $R/J \cong (R/I)/\phi(J)$.

   This statement is known as the third isomorphism theorem for rings.

19. Let $R = \mathbb{Z}[X]$ and consider the ideal $I = (2, X) \subset R$. Prove that $X^2+4 \in I \cdot I$, but that $X^2 + 4$ cannot be written as xy, where $x, y \in I$. Conclude that $\{xy : \ x, y \in I\}$ is not an ideal in R.

20. Let R be a ring, and I, J ideals in R. Prove that

$$(I + J) \cdot (I \cap J) \subset (I \cdot J).$$

Show that we have equality when $R = \mathbb{Z}$.

21. Let R be a ring and I_1, I_2, I_3 ideals in R. Prove:

$$I_1 + I_3 = I_2 + I_3 = R \iff (I_1 \cdot I_2) + I_3 = R.$$

22. (Chinese remainder theorem for several ideals). Let R be a ring with 1, and $I_1, I_2, ..., I_t$ in R. Suppose that these ideals are pairwise relatively prime, in other words: $I_i + I_j = R$ voor $1 \leq i < j \leq t$. Prove:

$$R/(\prod_{i=1}^{t} I_i) \cong \prod_{i=1}^{t} (R/I_i).$$

(Hint: prove $(I_1 \cdot I_2 \cdot ... \cdot I_{t-1}) + I_t = R$ as in exercise 21, and use induction on t.)

23. Let R be a ring such that $1 + 1 \in R^*$. Prove:

$$R[X]/(X^2 - 1) \cong R \times R.$$

24. Let $R = \{(a, b) \in \mathbb{Z} \times \mathbb{Z} : a \equiv b (\text{mod } 2)$.

   a. Prove that R is a subring of $\mathbb{Z} \times \mathbb{Z}$.

   b. Prove: $\mathbb{Z}[X]/(X^2 - 1) \cong R$.

   c. Prove: $\mathbb{Z}[X]/(X^2 - 1)$ is *not* isomorphic with $\mathbb{Z} \times \mathbb{Z}$ (hint: determine the idempotents in R and $\mathbb{Z} \times \mathbb{Z}$).

25. Let R be a ring with 1. Let $w_1, w_2, ..., w_m \in R$ be such that $w_i - w_j \in R^*$ for all i, j, $1 \leq i < j \leq m$. Let $f = \prod_{i=1}^{m} (X - w_i) \in R[X]$. Prove: $R[X]/(f) \cong R \times R \times ... \times R$ (m factors).

26. Prove:

$$\mathbb{Q}[X]/(X^3 + X) \cong \mathbb{Q} \times \mathbb{Q}[X]/(X^2 + 1),$$

and

$$\mathbb{R}[X]/(X^4 - 1) \cong \mathbb{R} \times \mathbb{R} \times \mathbb{C}.$$

27. Let R be a ring with 1, and $Id(R)$ the set of idempotents in R (including 0, 1). Suppose that e_1, e_2 are idempotents. Show that $e_1 + e_2 - 2e_1e_2$ and e_1e_2 are idempotents.
Show that $Id(R)$ is a ring when we define addition $\oplus$ and multiplication $\circ$ by

$$e_1 \oplus e_2 = e_1 + e_2 - 2e_1e_2 \qquad e_1 \circ e_2 = e_1e_2.$$

Under what circumstances is $Id(R)$ a subring in R?

Chapter 3

Polynomials, unique factorisation

3.1 Polynomials

Previously we have discussed polynomial rings $R[X]$ where R is any ring. In Example 2.1.14 we have proven the following theorem.

Theorem 3.1.1 *Let R be a ring. Let $P(X) \in R[X]$ and $a \in R$. Then $P(a) = 0$ (in other words, a is a zero of P) if and only if $P(X)$ is divisible by $X - a$.*

Furthermore we recall: when R is a domain then $R[X]$ is also a domain.
In this subsection we discuss a few further facts. For example,

Theorem 3.1.2 *Let R be a domain. Let $P(X) \in R[X]$ and suppose P has r distinct zeros $a_1, \ldots, a_r \in R$. Then there exists $Q(X) \in R[X]$ such that*

$$P(X) = Q(X)(X - a_1) \cdots (X - a_r).$$

Proof: We use induction on r. We have seen that $P(a) = 0$ implies the existence of Q such that $P(X) = Q(X)(X - a)$. So the case $r = 1$ is true.
Now let $r > 1$ and suppose our theorem is proven for the case of $r - 1$ zeros. Then, by the induction hypothesis there exists Q such that $P(X) = Q(X)(X - a_1) \cdots (X - a_{r-1})$. From $P(a_r) = 0$ it follows that $Q(a_r)(a_r - a_1) \cdots (a_r - a_{r-1}) = 0$. Since we work in a domain the distinctness of the a_i implies that $(a_r - a_{r-1}) \cdots (a_r - a_1) \neq 0$. Hence we conclude $Q(a_r) = 0$. So there exists Q_1 such that $Q(X) = Q_1(X)(X - a_r)$. Hence $P(X) = Q_1(X)(X - a_1) \cdots (X - a_r)$ and our induction step is completed.

□

Corollary 3.1.3 *Let R be a domain and $P \in R[X]$ a non-trivial polynomial of degree r. Then P has at most r distinct zeros in R.*

Proof: Suppose P has t zeros $a_1, \ldots, a_t$. Then there exists $Q \in R[X]$ such that $P(X) = Q(X)(X - a_1) \cdots (X - a_t)$. Since P and hence Q is non-trivial, we have the degree inequality $r = \deg(P) \geq t$. Hence our assertion follows.

□

Note that the condition that R has no zero divisors is essential. For example, the polynomial $X^2 - 1 \in (\mathbb{Z}/24\mathbb{Z})[X]$ has 8 zeros: $1, 5, 7, 11, 13, 17, 19, 23 (\mathrm{mod}\ 24)$. In general, if in a ring R there exist non-zero and distinct $a, b \in R$ such that $ab = 0$, then the quadratic polynomial $(X - a)(X - b) = X^2 - (a + b)X$ has at least three distinct zeros, namely $0, a, b$.
The following result is a classical one, first proven by Gauss. It deals with the existence of a complex zero of a polynomial with complex (or real) coefficients.

Theorem 3.1.4 (Main theorem of algebra) *Every non-constant polynomial with coefficients in $\mathbb{C}$ has a zero in $\mathbb{C}$.*

As a consequence we have

Corollary 3.1.5 *Let $P(z) \in \mathbb{C}[z]$ be a monic polynomial of degree $r \geq 1$. Then there exist r complex numbers $z_1, z_2, \ldots, z_r$ such that*

$$P(z) = (z - z_1)(z - z_2) \cdots (z - z_r).$$

Proof: By the main Theorem P has a complex zero, say z_1. Then there exists P_1 such that $P(z) = P_1(z)(z - z_1)$. Again, by the main Theorem, P_1 has a complex zero, say z_2. Then there exists P_2 such that $O(z) = P_1(z)(z - z_1) = P_2(z)(z - z_1)(z - z_2)$. We repeat the argument until we end with a factorisation $P(z) = P_r(z)(z - z_1) \cdots (z - z_r)$. By degree count we see that P_r is constant and since we work with monic polynomials we conclude that $P_r = 1$.

□

Theorem 3.1.6 *Let K be a field and $f, g \in K[X]$ with $g \neq 0$. Then there exist unique q, $r \in K[X]$ such that*

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

One calls q and r the quotient and remainder for division by g.

Sometimes one adopts the convention that $\deg(0) = -\infty$. In that case the inequality $\deg(r) < \deg(g)$ includes the possibility that $r = 0$.
Proof: We first prove existence of q, r. Unicity follows later.
For fixed g we carry out induction on $\deg(f)$.
When $\deg(f) < \deg(g)$ we can take $q = 0,\ r = f$. This is the initialisation of our induction.
Now let $n \geq \deg(g)$ and assume the existence of q, r is proven for all polynomials f of degree $< n$. Suppose now that f has degree n and g has degree m. Let a be the leading coefficient of f, b the leading coefficient of g. Notice now that

$f(X)-(a/b)X^{n-m}g(X)$ has degree $< n$. The induction hypothesis then implies the existence of polynomials q, r such that

$$f(X)-(a/b)X^{n-m}g(X)=q(X)g(X)+r(X)$$

with $\deg(r)<\deg(g)$ or $r=0$. Hence $f=(q+(a/b)X^{n-m})g+r$ and our assertion is shown.
It remains to show unicity of q and r. Suppose that beside q, r we have other polynomials q', r' satisfying $f=q'g+r'$ and $\deg(r')<\deg(g)$. Then, after taking the difference: $0=f-f=(q-q')g+r-r'$. So g divides the difference $r-r'$. But since $\deg(r-r')<\deg(g)$ this is only possible if $r-r'=0$. Hence $r=r'$ and automatically, $q=q'$.

□

Example 3.1.7. In the so-called long divsion for polynomials we basically carry out the steps of our induction procedure in a schematic way. Let $f, g \in \mathbb{Q}[X]$ be the following polynomials

$$f=X^4-X^3-2X^2+3X-4, \qquad g=X^2-1.$$

The quotient is determined as follows

$$\begin{array}{rccccccl}
X^2-1\,/ & X^4 & -X^3 & -2X^2 & +3X & -4 & \backslash X^2-X-1 \\
 & X^4 & & -X^2 & & & \\ \cline{2-4}
 & & -X^3 & -X^2 & & & \\
 & & -X^3 & & +X & & \\ \cline{3-5}
 & & & -X^2 & +2X & & \\
 & & & -X^2 & & +1 & \\ \cline{4-6}
 & & & & 2X & -5 &
\end{array}$$

So $q=X^2-X-1,\ r=2X-5$.

◊

The residue classes in $\mathbb{Z}$ modulo n are often represented by the integers $0,1,2,\ldots,n-1$. In the same way we represent the residue classes of polynomials in $K[X]$ modulo g by the polynomials of degree $<\deg(g)$. According to Theorem 3.1.6 any polynomial $f \in K[X]$ is modulo g equivalent to a polynomial r of degree $<\deg(g)$.

Theorem 3.1.8 *Let K be a field. Then every ideal in $K[X]$ is a principal ideal.*
More precisely, a non-trivial ideal $I \subset K[X]$ is generated by any non-zero element $g \in I$ of minimal degree.

Proof: Let $I \subset K[X]$ be an ideal. Suppose it is non-trivial, i.e I contains non-zero elements. From these non-zero elements we choose an element g of minimal degree. We assert that $I=(g)$.
Clearly $(g) \subset I$ because $g \in I$. Now suppose $f \in I$. Determine q, r such that $f=qg+r$ with $\deg(r)<\deg(g)$. Since $f, g \in I$ we also have that $r \in I$. If r

were non-zero, the inequality $\deg(r) < \deg(g)$ would contradict the minimality of $\deg(g)$. Therefore we must have $r = 0$. Hence $f = qg$ and $f \in (g)$. Thus we conclude that $I \subset (g)$, which together with $I \subset (g)$ finishes our proof.

□

Remark 3.1.9 *The condition in Theorem 3.1.8 that K is a field is essential. For example, the ideal $(2, X) \subset \mathbb{Z}[X]$ is not a principal ideal in $\mathbb{Z}[X]$, see exercise 6, page 34.*
Another such example is the polynomial ring $\mathbb{R}[X, Y]$ which contains the ideal (X, Y), which is non-principal.

Example 3.1.10. Let Φ_i be the evaluation homomorphism $\mathbb{R}[X] \to \mathbb{C}$ given by $\Phi_i : f \mapsto f(i)$ (we use the natural inclusion $\mathbb{R} \subset \mathbb{C}$) Note that Φ_i is surjective. Since $i \notin \mathbb{R}$, there are no linear polynomials in $\ker(\Phi_i)$. Note also that $X^2 + 1 \in \ker(\Phi_i)$. Then $X^2 + 1$ is a polynomial of minimal degree in the kernel, hence $\ker(\Phi_i) = (X^2 + 1)$. From the first isomorphism theorem it now follows that $\mathbb{R}[X]/(X^2 + 1) \cong \mathbb{C}$.
For generalisations see exercise 8 of this chapter.

♢

3.2 Irreducible elements

As is well-known, every integer > 1 can be written as a product of prime numbers and, up to ordering of the factors, this factorisation is uniquely determined. It turns out that in polynomial rings $K[X]$, where K is a field, we also have unique factorisation. To explain this we discuss the concept of irreducible and elements in a general domain R.

Definition 3.2.1 *Let R be a domain and let $r \in R$ be an element which is not 0 and not a unit.*
The element r is called irreducible if any factorisation $r = ab$ with $a, b \in R$ implies that either a or b is a unit in R.
If r is not irreducible, i.e it can be written as a product of two non-units in R, it is called reducible.

Example 3.2.2. In $\mathbb{Z}$ the irreducible elements are precisely the numbers $\pm p$ where p is a prime number. Note that 1 is not considered as prime number.

♢

Example 3.2.3. Consider the polynomial ring $K[X]$ where K is a field. Notice that the unit group in $K[X]$ is precisely the group K^*, in other words: the non-trivial constant polynomials in $K[X]$. Suppose $p(X) \in K[X]$. Then p is reducible if and only if there is a factorisation $p = ab$ where $a, b \in K[X]$ with $\deg(a), \deg(b) < \deg(p)$. In particular this means that a linear polynomial in $K[X]$ is always irreducible.

A quadratic polynomial $p(X)$ is reducible if and only if it can be written as a product of two linear polynomials. Since any linear polynomial in $K[X]$ has a zero in K, reducibility of a quadratic polynomial is equivalent to existence of a zero $a \in K$ of $p(X)$.
As application consider the polynomial $X^2 + 1$. It has no zero in $\mathbb{R}$, therefore $X^2 + 1$ is irreducible in $\mathbb{R}[X]$. Since $X^2 + 1$ has the zero $i \in \mathbb{C}$ it is reducible in $\mathbb{C}[X]$ with factorisation $(X + i)(X - i)$.
If $X^2 + 1$ is considered as element of $(\mathbb{Z}/5\mathbb{Z})[X]$ it has $2 (\mathrm{mod}\ 5)$ as zero. This gives the factorisation $X^2+1 \equiv (X-2)(X+2) (\mathrm{mod}\ 5)$. Since X^2+1 as element of $(\mathbb{Z}/3\mathbb{Z})[X]$ has no zero in $\mathbb{Z}/3\mathbb{Z}$, it is irreducible in $(\mathbb{Z}/3\mathbb{Z})[X]$.

♢

Example 3.2.4. By the main theorem of algebra the irreducible elements in $\mathbb{C}[X]$ are precisely the linear polynomials.
In $\mathbb{R}[X]$ the situation is slightly more involved. Of course the linear polynomials are irreducible. A quadratic polynomial X^2+aX+b with negative discriminant $a^2 - 4b$ has no real zero and is therefore irreducible.
Conversely, any polynomial $P(X) \in \mathbb{R}[X]$ is either divisible by a linear polynomial or a quadratic polynomial with negative discriminant. To see this, note that the main theorem of algebra asserts the existence of a zero $a \in \mathbb{C}$. If $a \in \mathbb{R}$ then $P(X)$ has a factor $X - a$. If $a \notin \mathbb{R}$ then its conjugate $\overline{a}$ must also be a zero of $P(X)$. Hence $P(X)$ is divisible by $(X - a)(X - \overline{a})$ which, after expansion, is a polynomial with real coefficients and negative discriminant.

♢

Example 3.2.5. In polynomial rings $R[X]$ where R is a domain, but not a field, the concept of irreducibility is slighty more subtle. Consider for example $\mathbb{Z}[X]$. Since a prime number p is not a unit in $\mathbb{Z}$, we see that a prime number is also irreducible in $\mathbb{Z}[X]$. A linear polynomial can also be reducible in this case. For example, $2X + 2 = 2(X + 1)$ is a factorisation into two non-units in $\mathbb{Z}[X]$.

♢

In $\mathbb{Z}$ we know that a prime number p has the property that if p divides a product ab of two integers, then it divides a or b (or both). However, in more general domains this need not be the case. In Example 3.3.7 we will see that 2 is an irreducible element of $\mathbb{Z}[\sqrt{-5}]$. It also divides the product $(1+\sqrt{-5})(1-\sqrt{-5}) = 6$, but 2 does not divide any of the factors $1 \pm \sqrt{-5}$. Luckily, in principal ideal domains the property still holds.

Theorem 3.2.6 *Let R be a principal ideal domain and $\pi \in R$ irreducible. Suppose that $a, b \in R$ and that $\pi | ab$. Then $\pi | a$ or $\pi | b$.*

Proof: Consider the ideal (a, π). Since R is a principal ideal domain this ideal is generated by a single element, which we call g. Note that $g|\pi$. Hence either g is a unit, or $g = \epsilon\pi$ with $\epsilon \in R^*$. In the second case, since $g|a$ we also get $\pi|a$ and our assertion is proven.
Now suppose that g is a unit and hence $(a, \pi) = R$. In this case there exist $x, y \in R$ such that $ax + \pi y = 1$. Multiply by b to get $abx + \pi by = b$. Both

terms on the left are divisible by π. Hence the sum b is also divisible by π which proves our theorem.

□

By induction on the number of factors we can also show the following.

Corollary 3.2.7 *Let R be a principal ideal domain and $\pi \in R$ irreducible. Let $a_1, a_2, \ldots, a_n \in R$ and suppose that $\pi | a_1 a_2 \cdots a_n$. Then there exists an index i such that $\pi | a_i$.*

3.3 Unique factorisation

Let R be a domain. Two non-zero elements $r, s \in R$ will be called *associates* of one another if there exists a unit $u \in R^*$ such that $r = us$. For example an integer $n \in \mathbb{Z}$ has $-n$ as its associate. Or $a + bi \in \mathbb{Z}[i]$ has $-a - bi$ and $\pm i(a + bi) = \pm(-b + ai)$ as associates. In a field every non-zero element is an associate of 1.

In this section we shall be interested in factorisation of elements as products of irreducible elements and, if they exist, whether they are unique. To make the latter more precsise, suppose we have $r \in R$ and suppose we have two factorisations $r = p_1 p_2 \cdots p_n = q_1 q_2 \cdots q_m$. We say that these two factorisation are the *same up to units and ordering of factors* if $m = n$ and there exists a permutaion σ of $\{1, 2, \ldots, n\}$ such that p_i and $q_{\sigma(i)}$ are associates for $i = 1, 2, \ldots, n$.

Definition 3.3.1 *A domain R is called a unique factorisation domain if every non-zero element which is not a unit can be written uniquely (up to units and ordering of factors) as product of irreducible elements.*

The best known example of a unique factorisation domain is $\mathbb{Z}$. Any integer > 1 can be written uniquely as a product of prime numbers. For example $120 = 2^3 \cdot 3 \cdot 5$. But of course a factorisation like $120 = 5 \cdot (-2)^3 \cdot (-3)$ is also valid. Clearly this doesn't differ essentially from the first factorisation. That is precisely the reason we consider factorisations up to units and ordering of factors.

Although unique factorsiation in $\mathbb{Z}$ is a commonly known fact, very few people are aware of a proof for it. If we look at such a proof we see that it is based on the fact that $\mathbb{Z}$ is a principal ideal domain. The beautiful thing is that this proof can be translated almost directly to the case of general principal ideal domains.

Theorem 3.3.2 *A principal ideal domain is a unique factorisation domain.*

To prove this Theorem we need an important property of principal ideal domains.

Proposition 3.3.3 *Let R be a principal ideal domain. Then any infinite chain of ideals $I_0 \subset I_1 \subset I_2 \subset \cdots$ has the property that there exists an index i such that $I_i = I_{i+1} = I_{i+2} = \cdots$.*

Proof: Consider the union $I = \cup_{j \geq 0} I_j$. Then I is again an ideal (please verify). Since R is a principal ideal domain there exists $d \in R$ such that $I = (d)$. Since $d \in I$ there exists i such that $d \in I_i$. So, for any $j \geq i$ we have $d \in I_j$ and $(d) \subset I_j$. On the other hand, $I_j \subset (d)$ for all j. So we conclude that $I_j = (d)$ for all $j \geq i$.

□

Remark 3.3.4 *Proposition 3.3.3 has an interesting consequence. Let $r \in R$. A divisor $d \in R$ of r is called a true divisor if neither d nor r/d is a unit in R. In particular if d is a true divisor of r then $(r) \subsetneq (d)$. A sequence $d_0, d_1, d_2, \ldots$ is called a divisor chain if d_{i+1} is a true divisor of d_i for all $i \geq 0$. For a divisor chain we have $(d_0) \subsetneq (d_1) \subsetneq (d_2) \subsetneq \cdots$. According to Proposition 3.3.3 a divisor chain has finite length.*

Remark 3.3.5 *In general, domains such that every sequence of ideals $I_0 \subset I_1 \subset \cdots$ eventually stabilises are called domains which satisfy the ascending ideal chain condition. Such rings are also called Noetherian rings (after Emmy Noether (1882-1935), one of the founders of modern algebra).*

The proof of Theorem 3.3.2 now takes several steps. Let R be a principal ideal domain.

- Any element r, not zero and not a unit, is divisible by an irreducible element in R. If r is irreducible this is clear. Suppose r is reducible. Then r has a true divisor r_1. If r_1 is irreducible we are done. Suppose r_1 is reducible, then r_1 has a true divisor which we call r_2. Continue this proces, i.e suppose r_k is reducible, then r_k has a true divisor which we call r_{k+1}. Hence $r, r_1, r_2, \cdots$ is a divisor chain which must end with r_k for some index k. This means that r_k does not have a true divisor, hence it is irreducible. This is our desired irreducible divisor of r.

- Any element r, not zero and not a unit, can be written as a product of irreducible elements. Let π_1 be an irreducible divisor. If r/π_1 is irreducible we are done. Suppose r/π_1 is reducible. Let π_2 be an irreducible divisor. If $r/(\pi_1\pi_2)$ is irreducible we are done. We continue this process and find a divisor chain $r, r/\pi_1, r/(\pi_1\pi_2), \ldots$. It must be finite, hence there is an index k such that $r/(\pi_1 \cdots \pi_k)$ is irreducible. This proves our assertion.

- Suppose we have an element $r \in R$ with two factorisations $p_1 p_2 \cdots p_n = q_1 q_2 \cdots q_m$. We can assume that there are no i, j such that p_i and q_j are associates. If there are such i, j we simply divide by p_i on both sides. Since p_1 divides $q_1 q_2 \cdots q_m$ and p_1 is irreducible, Corollary 3.2.7 implies that there exists j such that $p_1 | q_j$. Because q_j is irreducible and p_1 not a unit we conclude that p_1 and q_j are associates. This contradicts our assumption that p_i and q_j cannot be associate for any i, j.

Finally we like to point out the following theorem which generalizes Theorem 3.2.7 and can be quite useful.

Theorem 3.3.6 *Let R be a unique factorization domain and $\pi \in R$ irreducible. Let $a, b \in R$. Then $\pi | ab$ implies that $\pi | a$ or $\pi | b$.*

Proof: Let $a_1 \cdot a_2 \cdots a_n$ and $b_1 \cdot b_2 \cdots b_m$ be the factorisation of a and b into irreducible factors a_i, b_j. Let $c_1 \cdot c_2 \cdots c_l$ be the factorisation of ab/π into irreducibles. Then $a_1 \cdots a_n \cdot b_1 \cdots b_m = \pi \cdot c_1 \cdots c_l$. By the unique factorisation property the irreducible element π must be equal to a factor of the product on the left. Hence there exists i or j such that $a_i = \epsilon_i \pi$ or $b_j = \epsilon_j \pi$ with $\epsilon_i, \epsilon_j \in R^*$. In the first case π divides a, in the second π divides b.

□

Below you find two examples of rings where unique factorization does not hold.

Example 3.3.7. Let $R = \mathbb{Z}[\sqrt{-5}] = \{a + b\sqrt{-5} |\ a, b \in \mathbb{Z}\}$. Recall the norm function $N(a+b\sqrt{-5}) = a^2+5b^2$ which has the multiplicative property $N(\alpha\beta) = N(\alpha)N(\beta)$. We assert that 2 is irreducible in $\mathbb{Z}[\sqrt{-5}]$. To see this suppose that $2 = \alpha\beta$ with $\alpha, \beta \in \mathbb{Z}[\sqrt{-5}]$. Then $N(2) = N(\alpha)N(\beta)$ implies $4 = N(\alpha)N(\beta)$. Note that $N(\alpha)$ is one of the numbers $1, 2, 4$. The case $N(\alpha) = 1$ implies that α is a unit. The case $N(\alpha) = 2$ has no solution (there are no integers a, b with $a^2+5b^2 = 2$). When $N(\alpha) = 4$ we get $N(\beta) = 1$ and β is a unit. So we conclude that 2 is irreducible.
In the same way one can show that te elements $3, 1 \pm \sqrt{-5}$ are irreducible. Now note that $6 = 2 \cdot 3$ and $6 = (1 + \sqrt{-5})(1 - \sqrt{-5})$, hence two distinct irreducible factorizations of 6 in $\mathbb{Z}[\sqrt{-5}]$.

♢

Example 3.3.8. Consider the subring R of $\mathbb{Q}[X]$ of all polynomials of the form $a_0 + a_2X^2 + \cdots + a_nX^n$. In other words, R consists of all polynomials whose coefficient of X is zero. We assert that X^2 is irreducible in R.
Suppose $X^2 = (a_0+a_2X^2+\cdots+a_nX^n)(b_0+b_2X^2+\cdots+b_mX^m)$ with $a_n, b_m \neq 0$. By degree considerations we get $m, n \leq 2$. Clearly $m = n = 1$ is not possible, so $m = 2, n = 0$ or $m = 0, n = 2$. But a degree zero polynomial is constant and a unit in $\mathbb{Q}[X]$. Hence X^2 is irreducible.
In the same way one can show that X^3 is irreducible. But then note that $X^6 = X^2 \cdot X^2 \cdot X^2$ and $X^6 = X^3 \cdot X^3$ are two distinct factorizations of X^6 in R.

♢

3.4 Euclidean rings

In the previous sections we have seen that principal ideal domains are also unique factorisation domains. How do we know if a domain is a principal ideal domain? For $\mathbb{Z}$ and $K[X]$, with K a field, we have given a proof. Notice that both proofs used a divison with remainder result. We can formalise this by defining a Euclidean ring as a ring which has divison with remainder. More precisely,

Definition 3.4.1 *A domain R is called a Euclidean domain if there exists a function $N : R \to \mathbb{Z}_{\geq 0}$ such that*

1. *$N(a) = 0 \iff a = 0$,*

2. *to any $a, b \in R$ with $b \neq 0$ there exist q, r such that $a = bq + r$ with $N(r) < N(b)$.*

Notice that $\mathbb{Z}$ is a euclidean domain with $N(a) = |a|$. The ring $K[X]$, where K is a field, is a Euclidean domain with $N(P) = e^{\deg(P)}$.

Theorem 3.4.2 *A Euclidean domain is a principal ideal domain, and hence a unique factorisation domain.*

Proof: The proof is a precise copy of the proofs given for $\mathbb{Z}$ and $K[X]$. Let I be an ideal. Assume $I \neq (0)$. Let $b \in I$ be a non-zero element of I with minimal $N(b)$. Note that $N(b) > 0$. Clearly, $(b) \subset I$. On the other hand, let $a \in I$. Then there exist $q, r \in R$ such that $a = bq + r$ with $N(r) < N(b)$. If r is non-zero this would contradict the minimality of $N(b)$. Hence $r = 0$ and $a = bq \in (b)$. Therefore $I \subset (b)$ and we deduce that $I = (b)$.

□

Example 3.4.3. Let $R = \mathbb{Z}[i]$. Let $a, b \in \mathbb{Z}[i]$ with $b \neq 0$. Then, in the complex plane there exist $q \in \mathbb{Z}[i]$ such that $|q - a/b| \leq 1/\sqrt{2}$. Hence there exists $q \in \mathbb{Z}[i]$ such that $|a - bq| \leq |b|/\sqrt{2}$. Take squares on both sides, $|a - bq|^2 \leq |b|^2/2 < |b|^2$. Define $N(x) = |x|^2$ for all $x \in R$. Then $N(a - bq) < N(b)$. This proves that $\mathbb{Z}[i]$ is a euclidean domain. Hence, up to units and ordering of factors we have unique factorisation into irreducible elements in $\mathbb{Z}[i]$.

♢

3.5 Exercises

1. Determine the remainder of $X^5 + 3X^3 - 5X + 2$ after division by $X^2 - X + 1$ in $\mathbb{Q}[X]$.

2. Determine the remainder of $4X^5 + 2X^3 - X$ after division by $2X^2 + X + 2$ in $\mathbb{Q}[X]$. Same question, but now in the ring $(\mathbb{Z}/5\mathbb{Z})[X]$.

3. Let K be a field, $f \in K[X]$ and $\alpha_0, \alpha_1, \ldots \alpha_n$ an $n + 1$-tuple of distinct elements of K, where $n = \deg(f)$. Prove:

$$f = \sum_{i=0}^{n} f(\alpha_i) \frac{\prod_{j=0, j \neq i}^{n} (X - \alpha_j)}{\prod_{j=0, j \neq i}^{n} (\alpha_i - \alpha_j)}.$$

   This is known as *Lagrange's interpolation formula.*

4. Let R be a finite ring. Prove: $\exists n, m \in \mathbb{Z} : n > m > 0$, such that $x^n = x^m$ for all $x \in R$.

5. Let R be a domain, and $f, g \in R[X]$ polynomials with $\max(\deg(f), \deg(g)) < \#R$ (holds in particular if R is infinite). Prove: $(\forall x \in R : f(x) = g(x)) \Leftrightarrow f = g$.

6. Let p be a prime and $f, g \in (\mathbb{Z}/p\mathbb{Z})[X]$. Prove:
$$(\forall x \in \mathbb{Z}/p\mathbb{Z} : f(x) = g(x)) \Leftrightarrow f - g \in (X^p - X).$$
(Hint: use Fermat's little theorem $a^p \equiv a (\text{mod } p)$ for all $a \in \mathbb{Z}$.

7. We define the evaluation homomorphism:
$$\Phi : \mathbb{R}[X,Y] \longrightarrow \mathbb{R}[T], \qquad f(X,Y) \mapsto f(T^2, T^3).$$
Prove that $ker(\Phi) = (X^3 - Y^2)$ and that $\Phi(\mathbb{R}[X,Y]) = \{\sum a_i T^i : \ a_1 = 0\}$.

8. a. Let $z = a + bi \in \mathbb{C}$ and $z \notin \mathbb{R}$. Prove that the evaluation homomorphism
$$\Phi_z : \mathbb{R}[X] \longrightarrow \mathbb{C}, \qquad f \mapsto f(z),$$
(where we use the inclusion $\mathbb{R} \subset \mathbb{C}$), is surjective.

   b. Let $g = X^2 - 2aX + a^2 + b^2$. Prove that:
$$ker(\Phi_z) = (g), \qquad \text{and that} \qquad \mathbb{R}[X]/(g) \cong \mathbb{C}.$$

   c. Let $f = aX^2 + bX + c \in \mathbb{R}[X]$. Prove that:
$$\begin{array}{rcllc} \mathbb{R}[X]/(f) & \cong & \mathbb{C} & \text{if} & b^2 - 4ac < 0, \\ & \cong & \mathbb{R}[\epsilon] & \text{if} & b^2 - 4ac = 0, \\ & \cong & \mathbb{R} \times \mathbb{R} & \text{if} & b^2 - 4ac > 0, \end{array}$$

9. Let $z, w \in \mathbb{C} \setminus \mathbb{R}$ and let
$$\Phi_{z,w} : \mathbb{R}[X,Y] \longrightarrow \mathbb{C}, \qquad f \mapsto f(z,w),$$
be the evaluation homomorphism. Show that $Ker(\Phi_{z,w})$ is generated by one linear polynomial and a quadratic polynomial. Determine these polynomials explicitly when $z = 1 + i, \ w = 3 - 2i$.

10. Let K be a field and $R = K[X]/(X^n)$ where $n \in \mathbb{Z}_{\geq 1}$. We denote $x := X + (X^n) \in R$, any equivalence class r in R has a representing element of the form:
$$r = a_0 + a_1 X + \ldots a_{n-1} X^{n-1} \qquad a_i \in K.$$

    a. Show that $r \in R$ is a unit if and only if $a_0 \neq 0$. Determine the inverse of such an element.

    b. Show that every zero divisor in R is nilpotent. What is the smallest k such that $r^k = 0$ for every zero divisor r in R?

    c. For every $a \in K$ find a ring isomorphism:
$$K[X]/\left((X-a)^n\right) \cong K[X]/(X^n).$$

    d. Let $n > 1$. Find $f \in K[X]$ such that $f + (X^n)$ is a unit in R, and such thath $f + (X-1)^n \in K[X]/((X-1)^n)$ is nilpotent.

11. Show that $\mathbb{Z}[\sqrt{-2}]$ is a Euclidean ring.

Chapter 4

Factorisation of polynomials

4.1 Polynomial factorisation in practice

We shall illustrate how to factor a low degree polynomial $P(X) \in \mathbb{Q}[X]$ by hand. Let $d = \deg(P)$. Any non-trivial factor of P must have degree > 0 and $< d$. So when $d = 1$ we have a linear polynomial which is automatically irreducible. When $d = 2$ or 3 any non-trivial factorisation must have a linear factor. Let us show how to detect linear factors of P.

Write $P(X) = p_n X^n + \cdots + p_1 X + p_0$. By multiplication with a suitable integer if necessary we can assume that the coefficients p_i are in $\mathbb{Z}$. In order to simplify things we will assume that $p_n = 1$. Suppose that $P(X)$ has a linear factor $X - p/q$ where p, q are relative prime integers and $q > 0$. Then $P(X)$ has p/q as zero. Multiply the equality $P(p/q) = 0$ by q^n and expand to get

$$p^n + p_{n-1}p^{n-1}q + \cdots + p_1 p q^{n-1} + p_0 q^n = 0.$$

Every term on the left, except possibly p^n, is divisible by q. The sum is zero and thus also divisible by q. So we conclude that q divides p^n. Since q and p are relatively prime this can only happen if $q = 1$. Hence $P(X)$ has an integer zero. From $P(p) = 0$ we see that p divides the constant term p_0. This restricts p to the integer divisors of p_0 of which there are finitely many.

Take for example $P(X) = X^4 - 3X + 2$. By the above consideration any zero $a \in \mathbb{Q}$ must actually lie in $\mathbb{Z}$ and it must divide 2. Therefore $a \in \{\pm 1, \pm 2\}$. After a finite number of tries we see that 1 is a zero of $P(X)$ and we get $P(X) = (X - 1)(X^3 + X^2 + X - 2)$.

Similarly, if $X^3 + X^2 + X - 2$ is reducible it must have a linear factor and thus have a rational zero a. Then a must be an integer which divides -2. Hence we can test $a = \pm 1, \pm 2$ as zeros of $X^3 + X^2 + X - 2$ and find that none of them is. So $X^3 + X^2 + X - 2$ is irreducible in $\mathbb{Q}[X]$.

4.2 Gauss Lemma

In the previous subsection we have seen that a rational zero of a monic polynomial in $\mathbb{Z}[X]$ is automatically integral. This is a general phenomenon which holds for any unique factorization domain R instead of $\mathbb{Z}$.

The key ingredient is the existence of a gcd-concept in unique factorisation domains. Let a, b be two elements of a unique factorisation domain and let $a_1 \cdots a_n$ and $b_1 \cdots b_m$ be their factorisation into irreducible elements. If none of the pairs a_i, b_j are associates we say that $\gcd(a, b) = 1$. If there are associate pairs a_i, b_j we rearrange the ordering in such a way that a_i, b_i are associate for $i = 1, \ldots, r$ and a_i, b_j are not associate for any $i, j > r$. For any $i = 1, 2, \ldots, r$ we can find a unit u_i such that $b_i = a_i u_i$. Hence $a = a_1 \cdots a_r \cdot a'$ and $b = a_1 \cdots a_r \cdot b'$ where $\gcd(a', b') = 1$. We call $d = a_1 \cdots a_r$ a greatest common divisor of a and b. Notation, $\gcd(a, b) = d$. Notice that $\gcd(a, b)$ is not uniquely determined, but only up to units.
Moreover we see that if $d = \gcd(a, b)$, then $\gcd(a/d, b/d) = 1$. In a similar way we can define the gcd of a finite number of elements $A_1, \ldots, A_s$. Let $d = \gcd(A_1, \ldots, A_s)$, then again $\gcd(A_1/d, \ldots, A_s/d) = 1$.
Letting R be a unique factorization domain we call a polynomial $P \in R[X]$ *primitive* if the gcd of the coefficients of P is a unit in R.

Theorem 4.2.1 *Let R be a unique factorization domain and K its quotient field. Then any polynomial $P \in K[X]$ can be written as $P = c(P)P^*$ where $c(P) \in K^*$ and $P^* \in R[X]$ is a primitive polynomial. Moreover, $c(P)$ is uniquely determined up to multiplication by elements of R^*.*

We call the element $c(P)$ the *content* of the polynomial P.

Proof: Find an element $r \in R$ such that $rP \in R[X]$. One can take for example the product of the denominators of the coefficients of P. Let d be a greatest common divisor of the coefficients of rP. The rP/d is a primitive polynomial. Call it P^* and we get $P = (d/r)P^*$.
It remains to show the uniqueness of $c(P)$ (up to units). Suppose that $P = \lambda A^*$ and $P = \mu B^*$ with $\lambda, \mu \in K$ and $A^*, B^* \in R[X]$ primitive. Then there exist elements $r, s \in R$ such that $rA^* = sB^*$. We can assume $\gcd(r, s) = 1$. Suppose r has an irreducible factor π. Then all coefficients of sB^* are divisible by π and hence the coefficients of B^* are divisible by π. But then B^* is not primitive. Therefore, r is a unit in R. Similarly we argue that $s \in R^*$.

□

Lemma 4.2.2 (Gauss Lemma) *Let R be a unique factorisation domain and $A, B \in R[X]$ primitive polynomials. Then AB is also primitive.*

Proof: Suppose AB is not primitive. Then there exists an irreducible element π such that $AB \equiv 0(\text{mod } \pi)$. Let $a_n X^n$ be the highest degree term in A for which π does not divide a_n and $b_m X^m$ the highest degree term in B for which b_m is not divisible by π. Such m, n exist because P, Q are primitive polynomials. Then $PQ \equiv 0(\text{mod } \pi)$ implies that $a_n b_m \equiv 0(\text{mod } \pi)$. In other words π divides $a_n b_m$. Since π is irreducible and R is a unique factorization domain, this implies that π divides a_n or b_m, which is impossible. Hence we conclude that AB is a primitive polynomial.

□

Corollary 4.2.3 *Let R be a unique factorization domain and $A, B \in R[x]$ nontrivial polynomials. Then $c(AB) = c(A)c(B)$ (up to unit factors). In others words, the content is a multiplicative function.*

Proof: Write $A(X) = c(A)A^*(X), B(X) = c(B)B^*(X)$, with $A^*, B^* \in R[X]$ primitive polynomials. Then $AB = c(A)c(B)A^*B^*$. According to Lemma 4.2.2 the product A^*B^* is a primitive polynomial. Therefore we conclude that $c(A)c(B)$ must coincide with $c(AB)$.

□

4.3 Factorization over a unique factorization domain

Let R be a unique factorization domain and K its quotient field. As a consequence of the previous section we show that factorization in $K[X]$ can be reduced to factorization in $R[X]$, which is often much easier.

Proposition 4.3.1 *Let R be a unique factorization domain and K its quotient field. Let $P(X) \in R[X]$ be a primitive polynomial. Then $P(X)$ is irreducible in $K[X]$ if and only if it is irreducible in $R[X]$.*

Proof: We will show that $P(X)$ is reducible in $K[X]$ if and only if it is reducible in $R[X]$.
Suppose that $P(X) = A(X)B(X)$ with $A(X), B(X) \in K[X]$. We rewrite $A(X)$ as $c(A)A^*(X)$ where $c(A) \in K^*$ and $A^*(X) \in R[X]$ primitive. Similarly $B(X) = c(B)B^*(X)$. So we get $P(X) = c(A)c(B)A^*(X)B^*(X)$. Since both P and A^*B^* are primitive (the latter by Gauss' Lemma) we find that Theorem 4.2.1 implies $c(A)c(B) \in R^*$. So P factors in $R[X]$.
Suppose $P(X)$ is reducible in $R[X]$. Since it is primitive it factors into two factors of lower degree than $P(X)$. Hence $P(X)$ is reducible in $K[X]$.

□

We apply this in the following examples.

Example 4.3.2. Suppose we are given a fourth degree polynomial $P(X) = \sum_{i=0}^4 a_iX^i \in \mathbb{Q}[X]$ with $a_4, a_0 \neq 0$ which we like to factor in $\mathbb{Q}[X]$. After division by its content $c(P)$, if necessary, we can assume that P is in $\mathbb{Z}[X]$ and primitive. According to the remarks above it suffices to factor in $\mathbb{Z}[X]$. First we look for linear factors, i.e. factorizations of the form $P(X) = Q(X)(qX - p)$ with $Q(X) \in \mathbb{Z}$ and $p, q \in \mathbb{Z}$. Then it follows that $q|a_4$ and $p|a_0$. Since $a_0, a_4 \neq 0$ this gives us a finite number of possibilities which we can test one by one.
Now suppose that there are no linear factors. Then we should try a factorisation of the form

$$P(X) = \sum_{i=0}^{4} a_iX^i = (b_2X^2 + b_1X + b_0) \cdot (c_2X^2 + c_1X + c_0)$$

with $b_i, c_j \in \mathbb{Z}$. Comparison of coefficients gives us

i. $b_2c_2 = a_4$

ii. $b_2c_1 + b_1c_2 = a_3$

iii. $b_2c_0 + b_1c_1 + b_0c_2 = a_2$

iv. $b_1c_0 + b_0c_1 = a_1$

v. $b_0c_0 = a_0$.

The numbers b_2, c_2 are integers dividing a_4. So there are finitely many possibilities. Similarly there are finitely many choices for b_0, c_0. For a fixed choice of b_0, c_0, b_2, c_2 we can solve b_1, c_1 via the solution of the remaining linear equations in b_1, c_1. Although this method may be cumbersome, it always leads to a factoristion in a finite number of steps.

$\diamondsuit$

Example 4.3.3. Here is an example which illustrates how to factor in $\mathbb{C}[X, Y]$. We take $R = \mathbb{C}[X]$ and consider $\mathbb{C}[X, Y]$ as the polynomial ring $R[Y]$. Suppose we are asked to factor $P(X, Y) = Y^3 + (X^2+1)Y - (X^2-X)$ in $\mathbb{C}[X, Y]$. Denote by $\deg_Y$ the degree in the variable Y. Suppose that $P(X, Y) = A(X, Y)B(X, Y)$ with $A, B \notin \mathbb{C}$ and $\deg_Y(A) < \deg_Y(B)$. Since the coefficients of P in Y have gcd one, $\deg_Y(A)$ cannot be zero. So we have $\deg_Y(A) = 1$. Suppose $A(X, Y) = A_1(X)Y - A_0(X)$. Then A_1 divides 1, the leading terms of P as polynomial in Y. Hence $A_1 \in \mathbb{C}$ and we might as well assume $A_1 = 1$. The term $A_0(X)$ divides X^2-X the constant term of P. Hence $A_0(X)$ equals a constant times $1, X, X-1$ or $X^2 - X$. Then substitute consecutively $Y = \lambda, \lambda X, \lambda(X-1), \lambda(X^2-X)$ in $P(X, Y)$. In all cases we get a contradiction (check this!). We thus conclude that $P(X, Y)$ is irreducible in $\mathbb{C}[X, Y]$.

$\diamondsuit$

4.4 Unique factorisation in polynomial rings

We have seen that $K[X]$ is a unique factorisation domain when K is a field. The following theorem is a generalization of this fact.

Theorem 4.4.1 *Suppose R is a unique factorisation domain. Then so is $R[X]$.*

Proof: Let $P \in R[X]$ be the polynomial to be factored. First we write it as $c(P)P^*$ with $c(P) \in R$ and $P^* \in R[X]$ primitive. The content $c(P)$ is uniquely determined (up to units) and can be factored uniquely into irreducible elements of R (R being a unique factorization domain). Any factorization of P^* in $K[X]$ is, up to units in K, the same as a factorization of the form $P = A_1A_2 \cdots A_r$ into irreducible elements $A_i \in R[X]$ which are also primitive. By unique factorization in $K[X]$ these factors A_i are uniquely determined up to units and ordering of the indices i.

$\square$

Corollary 4.4.2 *Let R be a unique factorisation domain. Then the polynomial ring $R[X_1, \dots, X_n]$ in n variables is a unique factorisation domain.*

Proof: Use induction on n and Theorem 4.4.1.

□

4.5 Eisenstein's criterion

Let R be a unique factorisation domain. There are not many general criteria to decide whether a polynomial in $R[X]$ is irreducible in $R[X]$. However, there is one important sufficient criterion known as the *criterion of Eisenstein* named after (Gotthold Eisenstein, German mathematician, 1823-1852).

Definition 4.5.1 *Let R be a unique factorisation domain and π an irreducible element of R. A polynomial $f = a_nX^n + \cdots + a_1X + a_0 \in R[X]$ is called an Eisenstein polynomial with respect to π if $\pi \nmid a_n$, $\pi | a_i$ for all $i < n$ and $\pi^2 \nmid a_0$.*

Theorem 4.5.2 *Let R be a unique factorisation domain with quotient field K. Then an Eisenstein polynomial in $R[X]$ is irreducible in $K[X]$.*

Proof: Let f be an Eisenstein polynomial with respect to the irreducible element π. Without loss of generality we assume it is primitive. Suppose f is reducible in $K[X]$. Then, as a consequence of Proposition 4.3.1 $f = AB$ where $A, B \in R[X]$ and $\deg(A), \deg(B) < \deg(f)$. Write $A = \sum_{i=0}^{n} a_iX^i$ and $B = \sum_{j=0}^{m} b_jX^j$ and let f_{n+m} be the leading coefficient of f. Consider the factorisation $f = AB$ modulo π. We get $f_{n+m}X^{n+m} \equiv AB(\text{mod } \pi)$. Since $f_{n+m} = a_nb_m$ and $f_{n+m} \not\equiv 0(\text{mod } \pi)$ we see that a_n, b_m are not divisible by π. We like to show that a_0 and b_0 are both divisible by π. Suppose this is not the case and that, for example π does not divide a_0. Let b_iX^i be the lowest degree term for which π does not divide b_i (such a term exists because the leading term of B is not divisible by π). Then $AB(\text{mod } \pi)$ has the non-zero term $a_0b_iX^i$ with $i \leq m$. Since $f_{n+m}X^{n+m} \equiv AB(\text{mod } \pi)$ we get a contradiction and conclude that a_0 is divisible by π. By the same argument b_0 is divisible by π. But then the constant term of f, which is a_0b_0, is divisible by π^2. This contradicts the assumption that f is Eisenstein. Therefore we conclude that f is irreducible.

□

Example 4.5.3. Consider $X^5 + 2X^3 - 6 \in \mathbb{Z}[X]$. This is an Eisenstein polynomial for $\pi = 2$, hence it is irreducible in $\mathbb{Q}[X]$. Since it is primitive it is also irreducible in $\mathbb{Z}[X]$.

♢

Example 4.5.4. Consider $X^3 + (Y^4 - 1)X - (Y^2 + 1) \in \mathbb{R}[Y][X]$ as polynomial in X. This is an Eisenstein polynomial for $\pi = Y^2 + 1$, hence irreducible in $R(Y)[X]$. It is also primitive and therefore irreducible in $\mathbb{R}[Y][X] = \mathbb{R}[X, Y]$. Similarly this works for the polynomial $X^2 + Y^2 - 1 \in \mathbb{C}(Y)[X]$ with $\pi = Y - 1$.

♢

Theorem 4.5.5 *Let p be a prime number in $\mathbb{Z}$. Then the polynomial*

$$\phi_p = \frac{X^p - 1}{X - 1} = X^{p-1} + X^{p-2} + \cdots + X + 1$$

is irreducible in $\mathbb{Q}[X]$.

Proof: Let us replace X by $Y + 1$ and consider ϕ_p as a polynomial in Y. Notice that by binomial expansion of the p-th power we get

$$\phi_p(Y+1) = \frac{(Y+1)^p - 1}{Y} = \frac{1}{Y}(Y^p + \cdots + \binom{p}{k} Y^k + \cdots + pY).$$

The latter polynomial considered modulo p reads Y^{p-1} because $\binom{p}{k}$ is divisible by p for all $k \in \{1, 2, \ldots, p-1\}$. Moreover one sees that the constant coefficient is p and not divisible by p^2. Therefore $\phi_p(Y+1)$ is Eisenstein and irreducible in $\mathbb{Q}[Y]$. A fortiori the same holds for $\phi_p(X) \in \mathbb{Q}[X]$.

□

4.6 Exercises

1. Consider the ring:

$$R = \mathbb{Z}[\sqrt{-13}] = \left\{ a + b\sqrt{-13} : \quad a, b \in \mathbb{Z} \right\}.$$

   a. Prove that 2, $7 \in R$ are irreducible.
   Hint: use the norm map

$$N : R \to \mathbb{Z}, \quad N(a + b\sqrt{-13}) = a^2 + 13b^2,$$

   see Example 1.2.5.

   b. Show that $14 = 2 \cdot 7$ en $14 = (1 + \sqrt{-13})(1 - \sqrt{-13})$ are two different factorisations of 6 into irreducible elements.

2. Let R be the ring of polynomial functions on the circle:

$$R = \mathbb{R}[X, Y]/I, \qquad I = (X^2 + Y^2 - 1),$$

   and let $x := X(\text{mod } I)$, $y := Y + (\text{mod } I) \in R$.

   a. Prove that every residue class modulo I has a unique representing element of the form $A(X) + B(X)Y$.
   We shall denote the corrsponding class by $A(x) + B(x)y$.

   b. On R we define the norm map $N : R \to \mathbb{R}[x]$ by

$$N : A(x) + yB(x) \mapsto (A(x) + yB(x))(A(x) - B(x)y) = A(x)^2 - (1 - x^2)B(x)^2.$$

   Prove that $N(\alpha\beta) = N(\alpha)N(\beta)$ for any two $\alpha, \beta \in R$.

c. Prove that $A(x) + yB(x)$ is a unit in R if and only if the norm is in $\mathbb{R}^*$.

d. (subtle) Prove that $x - 1$ and $y - 1$ are irreducible in R. Hint: use the norm map N.

e. Show that $a = (x + y - 1)^2 = 2(x - 1)(y - 1)$ are two distinct factorisations of a as product of irreducible elements (and the unit 2).

f. Restrict the polynomials $(X + Y - 1)^2$ and $(X - 1)(Y - 1)$ as functions on the circle. In other words, consider the functions $(\cos\phi + \sin\phi - 1)^2$ and $(\cos\phi - 1)(\sin\phi - 1)$ on the unit circle parametrized by ϕ. Determine the zeros of these functions (with multiplicities). Do you see the relation with the previous exercise?

g. Draw a picture of the unit circle and the lines $X + Y - 1 = 0$, $X - 1 = 0$ and $Y - 1 = 0$. Find other elements in R which allow two different factorisations in R.

h. Consider the ring $\mathbb{C}[X, Y]/(X^2 + Y^2 - 1)$ (so we take $\mathbb{C}$ instead of $\mathbb{R}$). Show that the element $x - 1$ is reducible in this ring.

3. Determine whether the following elements in $\mathbb{Z}[\sqrt{-3}]$ are irreducible:

$$\sqrt{-3},\ \ 1,\ \ 2,\ \ 1 + \sqrt{-3},\ \ 5.$$

4. Let $R = \{a/b \in \mathbb{Q} :\ a, b \in \mathbb{Z}, b \text{ odd}\}$. This is a subring of $\mathbb{Q}$.

a. Determine R^*.

b. Prove that every $x \in R,\ x \neq 0$, can be written uniquely in the form $x = 2^k \cdot u$, where $k \in \mathbb{Z}_{\geq 0}, u \in R^*$.

c. Show that 2 is, up to multiplication by units, the only irreducible element in R.

5. Let R be a unique factorisation domain and $d \in R$ the gcd of a, $b \in R$: $d = \gcd(a, b)$. Suppose $c \in R$ divides both a and b. Prove that c is a divisor of d.

6. Factor $X^8 - 16$ and $X^6 + 27$ into irreducible factors in $\mathbb{Q}[X]$.

7. Is $5X^4 + 10X + 10$ an Eisenstein polynomial in $\mathbb{Z}[X]$? Is it irreducible in $\mathbb{Z}[X]$? and in $\mathbb{Q}[X]$?

8. Prove that $X^n + 2$ is irreducible in $\mathbb{Z}[X]$ for all $n \in \mathbb{Z}_{>0}$.
Prove that $Y^n - X$ is irreducible in $K[X, Y]$ (K is a field) for all $n \in \mathbb{Z}_{>0}$.

9. a. Find an example of an irreducible polynomial $f \in \mathbb{Z}[X]$ with the property that $f(X^2)$ is *not* irreducible.

b. Let $f \in \mathbb{Z}[X]$ be a monic Eisenstein polynomial. Prove that $f(X^2)$ is irreducible in $\mathbb{Q}[X]$.

10. Let R be a unique factorisation domain. Prove that $\cup_{n\geq 0} R[X_1, X_2, ..., X_n]$ is a unique factorisation domain.

11. Factor the following polynomials into irreducible factors in $\mathbb{Z}[X]$, in $\mathbb{Q}[X]$ and in $(\mathbb{Z}/5\mathbb{Z})[X]$ (except for the fourth one):

$$\begin{aligned} 4X^2+4, \\ 2X^{10}+4X^5+3, \\ X^4-7X^2+5X-3, \\ X^{111}+9X^{74}+27X^{37}+27, \\ X^3+X+3. \end{aligned}$$

12. Factor the following polynomials into irreducible factors in $\mathbb{Z}[X]$, and in $\mathbb{Q}[X]$:

$$\begin{aligned} \tfrac{1}{7}((X+1)^7-X^7-1), \\ X^3+3X^2+6X+9, \\ X^4+2X^3+3X^2+9X+6, \\ X^{12}-1, \\ X^4-X^3+X^2-X+1. \end{aligned}$$

13. Factor the following polynomials into irreducible factors in $\mathbb{Q}[X,Y]$:

$$\begin{aligned} Y^4+X^2+1, \\ Y^3-(X+1)Y^2+Y+X(X-1), \\ X^n+Y^3+Y \quad (n\geq 1), \\ X^4+4Y^4, \\ X^4+2X^3+X^2-Y^2-2Y-1, \\ Y^n-13X^4 \qquad (n\geq 1). \end{aligned}$$

14. Let $f\in\mathbb{Z}[X]$ be a monic polynomial such that $f(0)$ is a prime number. Prove that f has at most *three* distinct zeros in $\mathbb{Q}$.

15. Determine all irreducible polynomials $f\in(\mathbb{Z}/2\mathbb{Z})[X]$ of degree ≤ 3.

Chapter 5

Prime and maximal ideals

There are many rings which are not principal ideal rings or unique factorisation rings but which, nevertheless, play an extremely important role in algebra and its applications. Notably we think of rings of algebraic numbers, such as $\mathbb{Z}[\zeta_n]$ where $\zeta_n = e^{2\pi i/n}$ which play a crucial role in number theory. Or polynomial rings in several variables which form the basis of algebraic geometry. In these extensions the concepts of prime ideals and maximal ideals are important.

5.1 Prime ideals

We have seen in the previous sections the importance of the concept primality. So for any prime number $p \in \mathbb{Z}$ we have $p|ab \Rightarrow p|a$ or $p|b$. Using ideals this can be reformulated as follows: $ab \in (p) \Rightarrow a \in (p)$ or $b \in (p)$. In general, ideals which share this property with (p) will be called prime ideals.

Definition 5.1.1 *Let R be a ring. A prime ideal in R is an ideal $I \subset R$ which satisfies the following properties:*

(P1) *$I \neq R$;*

(P2) *For all a, $b \in R$: $ab \in I \Rightarrow a \in I$ or $b \in I$.*

Example 5.1.2. We have seen that (p) is a prime ideal in $\mathbb{Z}$ whenever p is a prime number. An ideal (n) where $n \neq 0$ is not a prime is not a prime ideal. Namely, if $n = 1$ then condition (P1) does not hold. When $n > 1$ and $n = ab$ with $a, b < n$ we observe that $ab \in (n)$ but $a \notin (n)$ and $b \notin (n)$.
Finally, the ideal $(0) \subset \mathbb{Z}$ is considered a prime ideal.

$\diamondsuit$

An important question is how to recognize a prime ideal. Here is an instance which generalizes the situation in $\mathbb{Z}$.

Theorem 5.1.3 *Let R be a unique factorisation domain and $\pi \in R, \pi \neq 0$. Then (π) is a prime ideal if and only if π is irreducible.*

Proof: Suppose that (π) is a prime ideal and $\pi = ab$ where $a, b \in R$ are proper divisors. Then $ab \in (\pi)$ and $a \notin (\pi)$ and $b \notin (\pi)$. This contradicts the fact that (π) is a prime ideal. Hence π cannot be reducible.
Suppose now that π is irreducible. Then $ab \in (\pi)$ implies that π divides ab. By unique factorization the factor π must occur either in the factorization of a or the factorization of b (or both). At any rate, $a \in (\pi)$ or $b \in (\pi)$. Hence (π) is a prime ideal.

□

Example 5.1.4. The ideal $(X^2-1) \subset \mathbb{R}[X]$ is not a prime ideal since it contain $(X+1)(X-1)$, but neither $(X+1)$ nor $(X-1)$.
The ideal $(X^2+1) \subset \mathbb{R}[X]$ is a prime ideal, since X^2+1 is an irreducible element of $\mathbb{R}[X]$.

♢

In more advanced cases we can use another criterion.

Theorem 5.1.5 *Let R be a ring and $I \subset R$ with $I \neq R$ an ideal. Then:*

$$I \text{ is a prime ideal } \iff R/I \text{ is a domain.}$$

Proof: For $a \in R$ we denote $\bar{a} = (a+I) \in R/I$. Suppose I is a prime ideal. Consider the ring R/I. The property $ab \in I \Rightarrow a \in I$ or $b \in I$ can be translated into $\bar{a}\bar{b} = \bar{0} \Rightarrow \bar{a} = \bar{0}$ or $\bar{b} = \bar{0}$. Hence R/I has no zero divisors. So R/I is a domain.
Suppose conversely that R/I is a domain. Then $\bar{a}\bar{b} = \bar{0} \Rightarrow \bar{a} = \bar{0}$ or $\bar{b} = \bar{0}$. This can be translated directly into the property $ab \in I \Rightarrow a \in I$ or $b \in I$. Therefore I is a prime ideal.

□

Here we give a number of examples in a non-principal ideal domain.

Example 5.1.6. Let $R = \mathbb{R}[X, Y]$ and consider the ideal $(X, Y-1)$. Then $(X, Y-1)$ is the kernel of the homomorphism $\phi : \mathbb{R}[X, Y] \to \mathbb{R}$ given by $\phi : p(X, Y) \mapsto p(0, 1)$. To see this suppose that $p(X, Y)$ is in the kernel of ϕ. Expand p into powers of X and $Y-1$. Then $p = \sum_{i,j} a_{i,j} X^i (Y-1)^j$. Any non-constant term contains a factor X or a factor $Y-1$. Hence each non-constant term is in $(X, Y-1)$. Since $p(0, 1) = 0$ we get $a_{00} = 0$, so there is no constant term. We conclude that $p \in (X, Y-1)$ and $\ker \phi \subset (X, Y-1)$. Conversely, any polynomial linear combination of $X, Y-1$ vanishes at the point $(0, 1)$. So $(X, Y-1)$ is contained in $\ker(\phi)$, hence we have equality.
Since ϕ is surjective we conclude by the isomorphism theorem that $\mathbb{R}[X, Y]/(X, Y-1) \cong \mathbb{R}$. Since $\mathbb{R}$ is a domain we conclude that $(X, Y-1)$ is a prime ideal.

♢

Example 5.1.7. Let $R = \mathbb{Z}[\sqrt{-5}]$ and $I = (3, 1-\sqrt{-5})$. In Exercise 7 it has been shown that I is the kernel of the homomorphism $\phi : \mathbb{Z}[\sqrt{-5}] \to \mathbb{Z}/3\mathbb{Z}$ given by $\phi : a + b\sqrt{-5} \mapsto a + b(\text{mod } 3)$. Since ϕ is surjective we get $\mathbb{Z}[\sqrt{-5}]/(3, 1-\sqrt{-5}) \cong \mathbb{Z}/3\mathbb{Z}$. Since $\mathbb{Z}/3\mathbb{Z}$ is a domain we conclude that $(3, 1-\sqrt{-5})$ is a prime ideal.

♢

Example 5.1.8. Let $I = (X + Y,\, X^2 + 1) \subset R = \mathbb{R}[X, Y]$. Consider the homomorphism $\phi : \mathbb{R}[X, Y] \to \mathbb{C}$ given by $\phi : p(X, Y) \mapsto p(i, -i)$. We claim that $\ker(\phi) = (X+Y, X^2+1)$. Clearly $X+Y$ and X^2+1 are in the kernel of ϕ, hence $(X + Y, X^2 + 1) \subset \ker(\phi)$. Now suppose that $p(X, Y) \in \ker(\phi)$, in other words, $p(i, -i) = 0$. This implies that the polynomial $p(X, -X)$ evaluated at $X = i$ is zero and so $p(X, -X)$ is divisible by $X^2 + 1$. So $p(X, -X) = (X^2 + 1)Q(X)$ for some $Q \in \mathbb{R}[X]$. Furthermore, $p(X, Y) - p(X, -X)$ is divisible by $Y - (-X) = Y + X$ in $\mathbb{R}[X, Y]$. Hence $p(X, Y) = p(X, -X) + (Y + X)T(X, Y) = (X^2 + 1)Q(X) + (Y + X)T(X, Y)$. So $p(X, Y) \in (X + Y, X^2 + 1)$.
By the isomorphism we see that $\mathbb{R}[X, Y]/(X + Y, X^2 + 1) \cong \mathbb{C}$. Since $\mathbb{C}$ is a domain we conclude that $(X + Y, X^2 + 1)$ is a prime ideal.

♢

Example 5.1.9. Let $I = (5, X^2 + Y + 1) \subset R = \mathbb{Z}[X, Y]$. Consider the homomorphism $\phi_1 : \mathbb{Z}[X, Y] \to (\mathbb{Z}/5\mathbb{Z})[X]$ given by $\phi : p(X, Y) \mapsto p(X, -X^2 - 1)(\mathrm{mod}\ 5)$. We claim that the ideal $I = (5, X^2+Y+1)$ is the kernel of ϕ. Clearly this ideal is in $\ker(\phi)$. Now suppose that $p(X, Y)$ is in the kernel. Denote by $\bar{p}(X, Y)$ the polynomial where the coefficients are reduced modulo 5. Then $\bar{p}(X, -X^2 - 1) = 0$. So $\bar{p}(X, Y)$ is divisible by $Y - (-X^2 - 1) = Y + X^2 + 1$. Let $Q \in \mathbb{Z}[X, Y]$ be such that $\bar{p}(X, Y) = (Y + X^2 + 1)\bar{Q}$. Then $p(X, Y)$ differs a multiple of 5 from $(Y + X^2 + 1)Q(X, Y)$. Hence $p(X, Y)$ is in $(5, Y + X^2 + 1)$. Thus we get via the isomorphism theorem $\mathbb{Z}[X, Y]/(5, X^2+Y+1) \cong (\mathbb{Z}/5\mathbb{Z})[X]$. The latter ring is a domain, hence $(5, X^2 + Y + 1)$ is a prime ideal.

♢

Example 5.1.10. Let $I = (YZ - X^2, X^2 - Z) \subset \mathbb{C}[X, Y, Z]$. First note that all polynomials in I become zero at the point $X = 1, Y = -1, Z = 1$.
We assert that I is not a prime ideal. Namely $Z(Y - 1) = YZ - Z = (YZ - X^2) + (X^2 - Z) \in I$ and the other hand, $Z, Y - 1 \notin I$ since these polynomials do not vanish at the point $(1, -1, 1)$.

♢

5.2 Maximal ideals

Definition 5.2.1 *Let R be a ring. An ideal M in R is called a maximal ideal if*

(M1) *$M \neq R$,*

(M2) *For every ideal J satisfying $M \subset J \subset R$ we have either $J = M$ or $J = R$.*

In other words, maximal ideals are ideals which are not properly contained in any ideal except R itself.
Examples of non-maximal ideals are $(9) \subset \mathbb{Z}$ since (3) lies in between; and $(2) \subset \mathbb{Z}[X]$, since $(2, X)$ lies in between.

Theorem 5.2.2 *Let R be a ring. Let M be a maximal ideal in R. Then M is a prime ideal.*

Proof: Let $a, b \in R$ be such that $ab \in M$. Suppose $a \notin M$. Then $M + (a)$ strictly contains M and since M is maximal, $M + (a) = R$. Hence there exist $m \in M$ and $r \in R$ such that $m + ra = 1$. Multiply both sides by b to get $bm + rab = b$. Both terms on the left are in M. Then their sum b must also be in M. Similarly, if $b \notin M$ then $a \in M$. So M is a prime ideal.

□

In principal ideal domains maximal ideals are precisely the ideals generated by irreducible elements.

Theorem 5.2.3 *Let R be a principal ideal domain and $I \subset R$ an ideal, $I \neq (0)$. Then the following statments are equivalent:*

(i) I is a prime ideal.

(ii) I is a maximal ideal.

(iii) $I = (\pi)$ where $\pi \in R$ is irreducible.

Proof: (i)⇒ (iii) follows from Theorem 5.1.3
(iii)⇒ (ii). Suppose π is irreducible. Suppose there is an ideal J with $(\pi) \subset J \subset R$. Since R is principal there exists $r \in R$ such that $J = (r)$. Hence r divides π. Since π is irreducible we have either $r \in R^*$, in which case $(r) = R$, or r is associate of π in which case $(r) = (\pi)$. Therefore (π) is maximal.
(ii)⇒ (i) follows from Theorem 5.2.2

□

For non-principal ideal domains there is another criterion to verify whether an ideal is maximal.

Theorem 5.2.4 *Let R be a ring, and $M \subset R$ an ideal with $M \neq R$. Then the following statements are equivalent:*

i) M is a maximal ideal.

ii) R/M is a field.

Proof: i) ⇒ ii). Let $a \in R$ and $a \notin M$. We will show that there exists $x \in R$ such that $ax \equiv 1(\text{mod } M)$. The ideal $M + (a)$ contains M and by maximality of M we have $M + (a) = M$ or $M + (a) = R$. The first case contradicts $a \notin M$, so we are left with $M + (a) = R$. In particular, there exist $m \in M$ and $x \in R$ such that $m + ax = 1$. So, $ax \equiv 1(\text{mod } M)$.
ii) ⇒ i). Let J be an ideal with $M \subset J \subset R$ and suppose that $J \neq M$. Choose $a \in J, a \notin M$. Since R/M is a field there exists $b \in R$ such that $ab \equiv 1(\text{mod } M)$. Hence $ab = 1(\text{mod } J)$. Together with $a \in J$ this implies that $1 \in J$. And we conclude that $J = R$.

□

As application we reconsider the examples given in the previous section.

- $\mathbb{R}[X,Y]/(X,Y-1) \cong \mathbb{R}$. Since $\mathbb{R}$ is a field we conclude that $(X,Y-1)$ is a maximal ideal.
- $\mathbb{Z}[\sqrt{-5}]/(3,1-\sqrt{-5}) \cong \mathbb{Z}/3\mathbb{Z}$. Since $\mathbb{Z}/3\mathbb{Z}$ is a field we conclude that $(3,1-\sqrt{-5})$ is a maximal ideal.
- $\mathbb{R}[X,Y]/(X+Y,X^2+1) \cong \mathbb{C}$. Since $\mathbb{C}$ is a field we conclude that $(X+Y,X^2+1)$ is a maximal ideal.
- $\mathbb{Z}[X,Y]/(5,X^2+Y+1) \cong (\mathbb{Z}/5\mathbb{Z})[X]$. The latter ring is a domain, not a field. Hence $(5,X^2+Y+1)$ is a prime ideal, but not a maximal ideal. The latter is also clear since $(5,X^2+Y+1,X)$ is a proper ideal which strictly contains $(5,X^2+Y+1)$.
- $I=(YZ-X^2,X^2-Z) \subset \mathbb{C}[X,Y,Z]$. Since I is not a prime ideal, it cannot be a maximal ideal. The ideal (X,Y,Z) properly contains I.

5.3 Existence of maximal ideals (optional)

In this section we deal with the following question which often arises in algebra questions.

Question 5.3.1 *Given an ideal I in a ring R, does there exist a maximal ideal M such that $I \subset M$?*

Naively we would proceed as follows. Choose an ideal $I_1 \neq R$ such that $I \subsetneq I_1$. If such an I_1 does not exist then obviously I is maximal. If I_1 exists we choose $I_2 \neq R$ such that $I_1 \subsetneq I_2$. Again if such an I_2 does not exist, I_1 is maximal. We continue this way and find a chain of ideals

$$I \subsetneq I_1 \subsetneq I_2 \subsetneq \cdots \subsetneq R.$$

If at some point we cannot find an ideal I_n, then clearly the previous ideal I_{n-1} is maximal. If the chain continues indefinitely our approach fails.
In the case when R is a principal ideal domain we can always find a maximal ideal between I and R. Namely, let $I=(r)$ (r is not a unit because $I=(r)\neq R$). By Theorem 3.3.2 r is divisible by an irreducible element π. Then (π) is a maximal ideal by Theorem 5.2.3 and $(r) \subset (\pi) \subset R$. In the proof of Theorem 3.3.2 have used Proposition 3.3.3 which states that every chain of ideals $J \subset J_1 \subset J_2 \subset \cdots$ stabilises. That is, there exists an index n_0 such that $J_n = J_{n_0}$ for every $n \geq n_0$. In other words, every strictly increasing chain $J \subsetneq J_1 \subsetneq J_2 \subsetneq \cdots$ has finite length.

Definition 5.3.2 *A ring R satisfies the ascending chain condition or is called a Noetherian ring if every strictly increasing chain of ideals $I \subsetneq I_1 \subsetneq I_2 \subsetneq \cdots$ has finite length.*

The name Noetherian ring refers to Emmy Noether (1882-1935), one of the cofounders of modern algebra. So we have seen that principal ideal domains are examples of Noetherian rings. A direct consequence is the following Theorem.

Theorem 5.3.3 *Let R be a Noetherian ring. Then, for every ideal $I \neq R$ there exists a maximal ideal M containing I.*

An important further property of Noetherian rings is the following.

Theorem 5.3.4 *Let R be a ring. Then the following statements are equivalent,*

a. R satisfies the ascending chain condition.

b. Every ideal is finitely generated. In other words to every ideal $I \subset R$ there exist $a_1, \dots, a_n$ such that $I = (a_1, a_2, \dots, a_n)$.

Proof: Suppose R satisfies the ascending chain condition. Let I be an ideal. Choose $a_1 \in I$ and consider $I_1 = (a_1)$. Suppose there exists $a_2 \in I$ and $a_2 \notin I_1$. Let $I_2 = (a_1, a_2)$. We continue in this way. So if there exists $a_{n+1} \in I$ and $a_{n+1} \notin (a_1, \dots, a_n)$ we define $I_{n+1} = (a_1, \dots, a_{n+1})$. In this way we get a strictly increasing chain of ideals $I_1 \subsetneq I_2 \subsetneq\subsetneq \cdots$. By assumption this chain is finite, so there must be an index n such that $I = (a_1, \dots, a_n)$.
Suppose conversely that every ideal is finitely generated. Consider an ascending chain of ideals $I_1 \subsetneq I_2 \subsetneq \cdots$. The union $I = \cup_{i \geq 1} I_i$ is again an ideal. So there exists a finite number of generators $a_1, \dots, a_m$ such that $I = (a_1, \dots, a_m)$. Let n be the smallest index such that $a_j \in I_n$ for $j = 1, \dots, m$. Then $\cap_{i \geq 1} I_i = I \subset I_n$. In other words, $I_i = I_n$ for all $i \geq n$. So the chain stabilises.

□

One of the most important results in the early stages of modern algebra is the following Theorem.

Theorem 5.3.5 (Hilbert) *Let K be a field. Then, for any integer $n \geq 1$ the polynomial ring $K[X_1, \dots, X_n]$ is Noetherian.*

This Theorem follows by induction on n from the following Theorem.

Theorem 5.3.6 *Let R be a Noetherian ring. Then the polynomial ring $R[X]$ is also a Noetherian ring.*

Proof: Let I be a non-zero ideal in $R[X]$. Let f_1 be a non-zero polynomial of minimal degree in I. Choose a polynomial $f_2 \in I \setminus (f_1)$ with minimal degree. In general, for any $k \geq 1$, we choose f_{k+1} to be a polynomial of minimal degree in $I \setminus (f_1, \dots, f_k)$. Of course this proces continues as long as $I \neq (f_1, \dots, f_k)$. For every $i \geq 1$ let d_i be the degree of f_i and a_i its leading coefficient. Consider the sequence of ideals

$$(a_1) \subset (a_1, a_2) \subset \cdots \subset (a_1, \dots, a_i) \subset \cdots$$

Since R is Noetherian there exists k such that $(a_1, \dots, a_k) = (a_1, \dots, a_{k+1})$. Hence there exist $r_j \in R$ such that $a_{k+1} = r_1 a_1 + \cdots + r_k a_k$. Notice that the polynomial

$$g = f_{k+1} - \sum_{j=1}^{k} r_j X^{d_{k+1} - d_j} f_j$$

has degree less than d_{k+1} and is not contained in $(f_1, \ldots, f_k)$. This contradicts the minimality of $\deg(f_{k+1})$.
Hence there does not exist $f_{k+1} \in I \setminus (f_1, \ldots, f_k)$, and we conclude that $I = (f_1, \ldots, f_k)$.

□

Another fundamental result in the theory of polynomials in several variables is the following.

Theorem 5.3.7 *The maximal ideals in* $\mathbb{C}[X_1, \ldots, X_n]$ *are precisely the ideals of the form* $(X_1 - a_1, \ldots, X_n - a_n)$ *where* $a_1, \ldots, a_n \in \mathbb{C}$.

It is not hard to see that the ideal $I = (X_1 - a_1, \ldots, X_n - a_n)$ is maximal in $\mathbb{C}[X_1, \ldots, X_n]$, Just consider the evaluation map $\mathbb{C}[X_1, \ldots, X_n] \to \mathbb{C}$ given by $p(X_1, \ldots, X_n) \mapsto p(a_1, \ldots, a_n)$. It has I as kernel and the field $\mathbb{C}$ as image.
The converse statement that any maximal ideal has this form is harder to prove and we shall not do it here. There is an important consequence though.

Corollary 5.3.8 (Hilbert Nullstellensatz) *Let* $f_1, \ldots, f_r \in \mathbb{C}[X_1, \ldots, X_n]$. *Then the set of equations*

$$f_1 = 0, \quad f_2 = 0, \quad , \ldots, \quad f_r = 0$$

has no common solution in $\mathbb{C}$ *if and only if there exist* $g_1, \ldots, g_r \in \mathbb{C}[X_1, \ldots, X_n]$ *such that*

$$g_1 f_1 + g_2 f_2 + \cdots + g_r f_r = 1.$$

Proof: Clearly if such g_i exist there cannot be a common solution of $f_1 = \cdots = f_r = 0$.
Suppose that there is no common solution. Let $I = (f_1, \ldots, f_r)$. Suppose $I \neq \mathbb{C}[X_1, \ldots, X_n]$. Then there exists a maximal ideal M such that $I \subset M \subset \mathbb{C}[X_1, \ldots, X_n]$. Suppose $M = (X_1 - a_1, \ldots, X_n - a_n)$. Then all polynomials in M, hence in I vanish at the point $a_1, \ldots, a_n$. So we have a common zero contradicting our assumption. Therefore $I = \mathbb{C}[X_1, \ldots, X_n]$ and hence $1 \in I$, as asserted.

□

5.4 Zorn's Lemma (optional)

In this section we prove the following general theorem.

Theorem 5.4.1 *Let* R *be a ring and* $I \subset R$ *an ideal,* $I \neq R$. *Then there exists a maximal ideal* M *such that* $I \subset M \subset R$.

However, for the proof we require the so-called Zorn Lemma which, in its turn, is equivalent to the Axiom of choice. Whether one is prepared to work with the Axiom of choice is very much a matter of taste. If one is not prepared to

so then Theorem 5.4.1 cannot be proven and we have to content ourselves with partial results, such as the results from the previous section.
The axiom of choice states that to any surjective map $f : A \to B$ there exists $g : B \to A$ such that $f \circ g$ is the identity on B. Here is the statement of Zorn's Lemma, although 'Zorn's Theorem' would be more appropriate.

Theorem 5.4.2 (Zorn's Lemma) *Let P be a partially ordered set. Then P contains at least one maximal chain.*

Some explanation of terminology is in order here. A *partially ordered set* is a set P with a binary relation $\leq$ satisfying two properties:

a. If $x \leq y$ and $y \leq z$ then $x \leq z$.

b. For all x, y: $x \leq y$ and $y \leq x \iff x = y$.

A chain is a subset K of a partially ordered set such that for any $x, y \in$ we have either $x \leq y$ or $y \leq x$. A *maximal chain* in a partially ordered set is a chain K which is not contained in a strictly larger chain.
Proof: of Theorem 5.4.1. As partially ordered set P we take all ideals J such that $I \subset J \subsetneq R$ with the inclusion of sets as order relation. According to Zorn's Lemma P contains a maximal chain, say $\mathcal{M}$. Let M be the union of all ideals in $\mathcal{M}$. So $M = \cup_{J \in \mathcal{M}} J$. Then M itself is an ideal. Namely, suppose $a, b \in M$. Then there exist $J, J' \in \mathcal{M}$ such that $a \in J$ and $b \in J'$. Then, since $\mathcal{M}$ is a chain, we have $J \subset J'$ or $J' \subset J$. Suppose the latter occurs. Then $a, b \in J$ and since J is an ideal, $a - b \in J \subset M$. If $a \in M$ then there exists an ideal $J \in \mathcal{M}$ such that $a \in J$. Consequently for any $r \in R$ we get $ra \in J \subset M$. Hence we see that M is an ideal. It is also a maximal ideal. First of all, $M = R$ would imply the existence of $J \in \mathcal{M}$ such that $1 \in J$, which contradicts $J \neq R$. Secondly, there cannot be an ideal strictly between M and R, since this would contradict the maximality of the chain $\mathcal{M}$.

□

5.5 Exercises

1. Let R be a domain. Prove: the ideal generated by X en Y in $R[X, Y]$ equals $\{f \in R[X, Y] : f(0, 0) = 0\}$ and is a prime ideal in $R[X, Y]$.

2. Let K be a field, $n \in \mathbb{Z}_{>0}$, and $\alpha_1, \alpha_2, ..., \alpha_n \in K$. Prove: $(X_1 - \alpha_1, X_2 - \alpha_2, ..., X_n - \alpha_n)$ is a maximal ideal.

3. Prove: $(5) \subset \mathbb{Z}[i]$ is not a prime ideal.

4. Let K be a field. Prove that the ideal $(X, Y) \subset K[X, Y, Z]$ is prime but not maximal.

5. Which of the following ideals in $\mathbb{Z}[X]$ are prime or maximal:

$$(X, 3); \quad (X^2 - 3); \quad (5, X^2 + 3).$$

6. Which of the following ideals in $\mathbb{Q}[X,Y]$ are prime or maximal:

$$(X^2+1); \quad (X-Y, Y^2+1); \quad (X^2+1, Y^2+1); \quad (X^2+1, Y^2-2).$$

7. Let R be a ring and $I \subset R$ an ideal. Prove: I is a prime ideal in R if and only if there is a field K and a ring homomorphism $f : R \to K$ with $I = \ker(f)$.

8. Let $R = \{\sum a_i X^i \in \mathbb{Q}[X] : a_1 = 0\}$, see example 3.3.8.

   a. Let
   $$\Phi_0 : R \longrightarrow \mathbb{R}, \qquad f \mapsto f(0)$$
   the evaluation homomorphism in 0. Prove that $\ker(\Phi_0) = (X^2,\, X^3)$.

   b. Prove that $ker(\Phi_0)$ is not a principal ideal, but it is a maximal ideal.

9. Let R be a ring, let $I \subset R$ be an ideal and $\phi : R \to R/I$ the natural homomorphsim. Let $J \subset R$ be a prime ideal with $I \subset J$.

   Prove that $\phi(J)$ is a prime ideal in R/I and, conversely, that any prime ideal in R/I is of this form. (Hint: combine 5.1.5 and 18).

10. Same as Exercise 9, but with 'prime ideal' replaced by 'maximal ideal'.

11. Let $f : R_1 \to R_2$ be a ring homomorphism, let $I_2 \subset R_2$ be an ideal, and $I_1 = f^{-1}(I_2) \subset R_1$.

    a. Prove: I_1 is an ideal in R_1, and R_1/I_1 is isomorphic to a subring of R_2/I_2.

    b. Prove: if I_2 is prime in R_2 then I_1 is prime in R_1.

    c. Show by an example that 'prime ideal' cannot be replaced by 'maximal ideal' in part b).

12. Let R be a Boolean ring (see Exercise 18 on page 19).

    a. Prove: R is a domain $\Leftrightarrow$ R is a field $\Leftrightarrow$ $R \cong \mathbb{Z}/2\mathbb{Z}$.

    b. Let $I \subset R$ be an ideal. Prove: I is a prime ideal $\Leftrightarrow$ I is a maximal ideal $\Leftrightarrow$ $R/I \cong \mathbb{Z}/2\mathbb{Z}$.

13. Let R be a ring,and $I \subset R$ an ideal, $I \neq R$. We are given that every $x \in R, x \notin I$, satisfies $x^2 - 1 \in I$.

    a. Prove: $R/I \cong \mathbb{Z}/2\mathbb{Z}$ or $R/I \cong \mathbb{Z}/3\mathbb{Z}$.

    b. Is I a prime ideal in R?

14. Let $I \subset \mathbb{Z}[X]$ be a prime ideal.

    a. Prove that $I \cap \mathbb{Z}$ is a prime ideal in $\mathbb{Z}$.

b. Prove that either $I = \{0\}$ or $I = (f)$, where $f \in \mathbb{Z}[X]$ is irreducible, or $I = (p)$ where $p \in \mathbb{Z}$ is a prime number, or $I = (p, f)$ where $f \in \mathbb{Z}[X]$ and f is irreducible modulo the prime number p.

c. Determine all maximal ideals in $\mathbb{Z}[X]$.

15. Let R be a ring, and $I \subset R$ an ideal of finite index (that is, R/I is a finite ring).

    Prove: I is a prime ideal $\Leftrightarrow$ I is a maximal ideal.

16. Let R be a ring in which every ideal $I \neq R$ is a prime ideal. Prove that R is a field.

17. Let R be a ring such that $I \cap J \neq \{0\}$ for any pair of ideals $I \neq \{0\}$, $J \neq \{0\}$ in R.
    Show that $\{a \in R : a \text{ is a zero divisor }\} \cup \{0\}$ is a prime ideal in R.

18. Let R be a ring without 1, with additive group $\mathbb{Q}$ and multiplication $xy = 0$ for all $x, y \in R$. Prove: R contains no ideal which satisfies axioms (M1) and (M2) for maximal idelas. Why doesn't this contradict Theorem 5.4.1?

19. Let $R = C([0,1])$ be the ring of real continuous functions on the interval $[0,1]$.

    a. Show that the units in $C([0,1])$ are given by the nowhere vanishing functions.

    b. Let $a \in [0,1]$ and define $M_a = \{f \in R | \ f(a) = 0\}$. Prove that M_a is a maximal ideal in R.

    c. Challenge: show that every prime ideal in $C([0,1])$ is of the form M_a for some $a \in [0,1]$.

20. Let $R = \mathbb{R}[X,Y]/I$ with $I = (X^2 + Y^2 - 1)$ be the ring of polynomial functions on the unit circle. Let $x := X + I$, $y = Y + I \in R$.

    a. Prove that $(x - a, \ y - b)$ with $a, b \in \mathbb{R}$ is a maximal ideal in R precisely when $a^2 + b^2 = 1$.

    b. For which $b \in \mathbb{R}$ is $(y - b)$ a maximal ideal in R ?

21. Let R be a ring, and $a \in R$ an element such that $a^n \neq 0$ for all positive integers n. Prove that R contains a prime ideal I such that $a \notin I$. (Hint: apply Zorn's Lemma to the set of ideals not containing any power of a.)

22. The *radical* $\sqrt{0}$ of a ring R is defined by the set of nilpotent elements in R. That is,
$$\sqrt{0} = \{a \in R : \exists n \in \mathbb{Z}_{>0} : a^n = 0\}.$$
    Prove that $\sqrt{0}$ is an ideal in R. Prove that $\sqrt{0} = \cap_I I$, where I runs over all prime ideals in R (hint: use Exercise 21).

23. The *Jacobson-radical* $J(R)$ of a ring R is defined by

$$J(R) = \{x \in R : \forall r \in R : 1 + rx \in R^*\}.$$

    a. Let $x \in J(R)$, and let $M \subset R$ be a maximal ideal. Prove that $x \in M$.

    b. Let M be a maximal ideal in R and let $x \in M$. Prove that $1 + x \notin M$.

    c. Prove that $J(R) = \cap_M M$, where M runs over all maximal ideals in R.

    d. Prove that $J(R)$ is an ideal in R.

24. Let R be a ring. Let $S \subset R$ be a non-empty subset with the property that $0 \notin S$ and $\forall s, t \in S$: $st \in S$.

    Show that there is a prime ideal I in R such that $I \cap S = \emptyset$. (Hint: use the ring $S^{-1}R$ from Exercise 16 on page 19, and apply 5.4.1 and Exercise 11(b)). What is the relationship with Exercise 21?

25. Let $R = \{a/b \in \mathbb{Q} : a, b \in \mathbb{Z}, b \not\equiv 0 \bmod 5\}$. Prove the following:

    (a) Determine the irreducible elements of R.

    (b) Prove that R is a unique factorization domain.

    (c) Prove that R contains exactly one maximal ideal M.

    (d) Prove that $R/M \cong \mathbb{Z}/5\mathbb{Z}$.

26. Let R be a ring. We call R a *local ring* when $R - R^*$ is an ideal in R.

    a. Prove: R is a local ring $\Leftrightarrow$ R has precisely one maximal ideal.

    b. Let R be a local ring. Let $x \in R$ be such that $x^2 = x$. Prove that $x = 0$ or $x = 1$.

27. Let R be a ring and $I \subset R$ a prime ideal. Let $S = R - I$.

    a. Prove: $\forall s, t \in S : st \in S$.

    b. Prove that the ring $S^{-1}R$ from Exercise 16 on page 19 is a local ring (see Exercise 26).

Chapter 6

Fields

6.1 Prime fields and characteristic

Let L be a field. A subset $K \subset L$ is called a *subfield* if:

a. $1 \in K$,

b. $a, b \in K \quad \Longrightarrow \quad a - b \in K$,

c. $a, b \in K,\ b \neq 0 \quad \Longrightarrow \quad ab^{-1} \in K$.

A subfield K of L is again a field with the addition and multiplication inherited from L.
One easily verifies that the intersection of any set of subfields of L is again a subfield. The intersection of all subfields of L is called the the *prime field* of K. In a sense one can say that the prime field of L is the smallest subfield of L.

Theorem 6.1.1 *Let K be a field. Then the prime field of K is isomorphic to either $\mathbb{Z}/p\mathbb{Z}$ for some prime p, or $\mathbb{Q}$.*

Proof: Denote the prime field of K by K_0. Of course $1 \in K_0$. Consider the homomorphism $\phi : \mathbb{Z} \to K$ obtained by sending 1 to 1. The image, being a subring of K, is a domain. Hence ϕ has a prime ideal as kernel. The only prime ideals in $\mathbb{Z}$ are (0) and (p) with p prime. When the kernel is (0) the map ϕ is injective and embeds $\mathbb{Z}$ into K. Of course K should then also contain the quotient field of $\mathbb{Z}$, that is: $\mathbb{Q}$. When the kernel of ϕ is (p) for some prime p, the image is then $\mathbb{Z}/p\mathbb{Z}$ which is indeed a field.

□

Definition 6.1.2 *Let K be a field. Suppose its prime field is $\mathbb{Q}$. Then we say that the characteristic of K is 0. When the prime field is $\mathbb{Z}/p\mathbb{Z}$ we say that the characteristic is p.*
Notation $\mathrm{char}(K) = 0$ *or* $\mathrm{char}(K) = p$.

Examples of characteristic zero fields are $\mathbb{Q}$ itself, $\mathbb{R}, \mathbb{C}$ and $\mathbb{Q}(T)$ the field of rational functions with coefficients in $\mathbb{Q}$.
Examples of characteristic p fields are $\mathbb{Z}/p\mathbb{Z}$ and $(\mathbb{Z}/p\mathbb{Z})(T)$ and finite fields (to be treated later).
A very interesting property of characteristic p fields is the following.

Proposition 6.1.3 *Let K be a field of characteristic p. Then, for any $a, b \in K$ we have*

$$(a+b)^p = a^p + b^p.$$

(High school student's dream)

Proof: Simply develop $(a+b)^p$ by Newton's binomial law to get

$$(a+b)^p = a^p + \binom{p}{1} a^{p-1} b + \binom{p}{2} a^{p-2} b^2 + \cdots + b^p.$$

Since $\binom{p}{k}$ is divisible by p, hence zero mod p, whenever $0 < k < p$ our assertion follows.

□

6.2 Algebraic and transcendental elements

Let L be a field and $K \subset L$ a subfield. In its turn the field L will be called a *field extension* of K.
Let $\alpha \in L$. We distinguish two cases,

1. There exists a non-trivial polynomial $P(X) \in K[X]$ such that $P(\alpha) = 0$.

2. There exists no such polynomial.

In the first case we say that α is *algebraic* over K, in the second case we say that α is *transcendental* over K. When α is algebraic the set $I = \{P \in K[X] | P(\alpha) = 0\}$ is a non-trivial ideal in $K[X]$, hence a principal ideal of the form (f) where $f \in K[X]$. We can take f to be monic, which uniquely determines f as the monic polynomial of minimal degree in the ideal I. We summarize,

Definition 6.2.1 *Let L and K be as above and $\alpha \in L$ an element algebraic over K, The monic polynomial $f(X) \in K[X]$ of minimal degree such that $f(\alpha) = 0$ is called the minimal polynomial of α. The degree of f is called the degree of α over K. Notation* $\deg(\alpha)$.

Example 6.2.2. Let K be a field and consider $L = K(X)$, the quotient field of $K[X]$. Take $\alpha = X \in K(X)$ then it is clear that X is transcendental over K. It is more difficult to find examples of real numbers which are transcendental over $\mathbb{Q}$. In fact, it was only fairly recently in math history that the existence of transcendental numbers was established. Around 1840 Liouville (Joseph Liouville, French mathematician, 1809-1882) showed that the number $\sum_{k=1}^{\infty} 10^{-k!}$ is transcendental (over $\mathbb{Q}$). This is an artificially constructed number. Later, the transcendence of e was proven in 1873 by Charles Hermite (French mathematician, 1822-1901) and in 1882 Lindemann (Carl Louis Ferdinand von Lindemann, German mathematician) proved transsendence of π, thereby solving the ancient problem of *squaring the circle* by construction (see Chapter 7).

♢

Example 6.2.3. The complex number $i = \sqrt{-1}$ is algebraic over $\mathbb{R}$ with minimal polynomial $X^2 + 1$. It is also algebraic over $\mathbb{Q}$ with the minimal polynomial $X^2 + 1$.
For any $k, n \in \mathbb{Z}_{>0}$ the number $\alpha = \sqrt[n]{k} \in \mathbb{R}$ is algebraic over $\mathbb{Q}$, since it is a zero of the polynomial $X^n - k$. However, this need not always be the minimal polynomial of α.
Also the complex numbers

$$e^{\frac{2\pi ik}{n}} := \cos\frac{2\pi k}{n} + i\sin\frac{2\pi k}{n} \qquad (k \in \mathbb{Z}$$

are algebraic over $\mathbb{Q}$, being zeros of $X^n - 1$. We call these numbers *roots of unity.* Their minimal polynomial is in general less easy to determine, see Section 6.5.

♢

Definition 6.2.4 *Let L be a field extension of the field K. Let $\alpha \in L$. The field generated by α over K is the smallest subfield of L which contains both K and α. Notation: $K(\alpha)$*
Let $\alpha_1, \ldots, \alpha_n \in L$. The field generated over K by $\alpha_1, \ldots, \alpha_n$ is the smallest subfield of L containing K and α_i for $i = 1, \ldots, n$. Notation $K(\alpha_1, \ldots, \alpha_n)$.
An extension of the form $K(\alpha)$ is called a simple extension.

Example 6.2.5. When α is transcendental over K the field $K(\alpha)$ simply consists of the elements

$$\frac{a_n\alpha^n + \cdots + a_1\alpha + a_0}{b_m\alpha^m + \cdots + b_1\alpha + b_0}$$

with $b_m \neq 0$.
When α is algebraic this becomes more subtle because of the algebraic relations that exist for α. For example, the smallest subfield of $\mathbb{C}$ containing both $\mathbb{R}$ and i is $\mathbb{C}$ itself, but $\mathbb{C} = \{a + bi | \; a, b \in \mathbb{R}\}$. I.e. there are no quotients to be seen.

♢

Theorem 6.2.6 *Let L be a field extension of the field K. Let $\alpha \in L$ be algebraic over K and let f be the minimal polynomial of α over K.*
Then f is irreducible and $K(\alpha) \cong K[X]/(f)$.

Proof: Consider the ringhomomorphism $\phi : K[X] \to K(\alpha)$ given by $\phi : X \mapsto \alpha$ and $\phi(a) = a$ for any $a \in K$. The kernel of ϕ is (f), hence the image of ϕ is $K[X]/(f)$. Since L is a field the image is a domain and therefore (f) is a prime ideal. Since $K[X]$ is a principal ideal domain, Theorem 5.2.3 tells us that we have the equivalences

$$f \text{ irreducible} \iff (f) \text{ prime ideal} \iff (f) \text{ maximal ideal.}$$

So f is irreducible, and since (f) is a maximal ideal the image $K[X]/(f)$ is a subfield of $K(\alpha)$. Since the latter is the minimal subfield of L containing α and K we conclude that $K(\alpha) = K[X]/(f)$.

□

Remark 6.2.7 *Suppose K is a field and α an element of an extension of K which is algebraic. Let $f(X)$ be its minimal polynomial given by $X^n + f_{n-1}X^{n-1} + \cdots + f_1X + f_0$. Then it follows from the isomorphism $K(\alpha) = K[X]/(f)$ that any element can be uniquely written in the form*

$$a_0 + a_1\alpha + a_2\alpha^2 + \cdots + a_{n-1}\alpha^{n-1}$$

with $a_0, a_1, \ldots, a_{n-1} \in K$. In calculations with these expressions we simply keep track of the relation $\alpha^n = -f_{n-1}\alpha^{n-1} - \cdots - f_0$, just as we do with $i^2 = -1$ in the case of complex numbers.

When working in $K(\alpha)$ for some algebraic α, multiplication and addition are more or less straightforward. How to take inverses is perhaps less obvious. We shall illustrate this in a number of examples.

Example 6.2.8. Let $d \in \mathbb{Q}$ with $\alpha := \sqrt{d} \notin \mathbb{Q}$. The minimal polynomial of $\sqrt{d}$ over $\mathbb{Q}$ is $X^2 - d$. Clearly it is irreducible, for $\sqrt{d} \notin \mathbb{Q}$. So

$$\mathbb{Q}(\sqrt{d}) = \{a + b\sqrt{d} \in \mathbb{C} : a,\, b \in \mathbb{Q}\}.$$

The inverse of $a + b\sqrt{d}$ can be computed in an analogous way as with complex numbers,

$$\frac{1}{a + b\sqrt{d}} = \frac{1}{a + b\sqrt{d}}\frac{a - b\sqrt{d}}{a - b\sqrt{d}} = \frac{a - b\sqrt{d}}{a^2 - db^2}.$$

The latter number equals $\frac{a}{a^2-db^2} - \frac{b}{a^2-db^2}\sqrt{d}$ which is of the desired form. Notice that $a^2 - db^2$ cannot be zero unless $a = b = 0$. This is because d is not the square of a rational number.

♢

Example 6.2.9. For extensions of higher degree computation of an inverse becomes more involved. As example we take $\alpha \in \mathbb{R}$ which is a zero of the polynomial $f = X^3 + X^2 - 1$. Note that f is irreducible in $\mathbb{Q}[X]$ because f has no zero in $\mathbb{Z}$ and thus no zero in $\mathbb{Q}$. So α has degree 3 and

$$\mathbb{Q}(\alpha) = \mathbb{Q}[\alpha] = \{a_0 + a_1\alpha + a_2\alpha^2 |\ a_0, a_1, a_2 \in \mathbb{Q}\}.$$

During computations we use the fact that $\alpha^3 = -\alpha^2 + 1$. We also need,

$$\alpha^4 = \alpha(\alpha^3) = \alpha(-\alpha^2 + 1) = -\alpha^3 + \alpha = \alpha^2 + \alpha - 1.$$

Suppose we want to compute the inverse of $\beta = \alpha^2 + 3\alpha + 1$. Denote its inverse by $\xi = x_2\alpha^2 + x_1\alpha + x_0$ where $x_2, x_1, x_0 \in \mathbb{Q}$. Then

$$\begin{aligned} 1 &= \beta\xi \\ &= (\alpha^2 + 3\alpha + 1)(x_2\alpha^2 + x_1\alpha + x_0) \\ &= x_2\alpha^4 + (x_1 + 3x_2)\alpha^3 + (x_0 + 3x_1 + x_2)\alpha^2 + (3x_0 + x_1)\alpha + x_0 \\ &= (x_0 + 2x_1 - x_2)\alpha^2 + (3x_0 + x_1 + x_2)\alpha + x_0 + x_1 + 2x_2 \end{aligned}$$

In the last step we have used the relations $\alpha^3 = -\alpha^2 + 1$ and $\alpha^4 = \alpha^2 + \alpha - 1$ (please verify!). Equating the left and right hand side of our equation we get the system of linear equations

$$\begin{aligned} x_0 + x_1 + 2x_2 &= 1 \\ 3x_0 + x_1 + x_2 &= 0 \\ x_0 + 2x_1 - x_2 &= 0 \end{aligned}$$

The solution of this system is $x_2 = 5/11, x_1 = 4/11, x_0 = -3/11$. Hence the inverse of $\alpha^2 + 3\alpha + 1$ equals

$$\frac{5}{11}\,\alpha^2 + \frac{4}{11}\,\alpha - \frac{3}{11}.$$

$\diamondsuit$

Example 6.2.10. There is also a shortcut to the example above which reflects the gcd algorithm applied to $X^3 + X^2 - 1$ and $X^2 + 3X + 1$.
We need to solve $(\alpha^2+3\alpha+1)\xi = 1$. We know for certain that $(\alpha^3+\alpha^2-1)\xi = 0$. Substract $\alpha-1$ times the first from the second to get $(5\alpha+1)\xi = 2-\alpha$. Substract this $(\alpha/5 + 14/25)$ times from the first to get $\frac{11}{25}\xi = 1 - (2-\alpha)(\alpha/5 + 14/25) = (5\alpha^2 + 4\alpha - 3)/25$. We conclude that $\xi = (5\alpha^2 + 4\alpha - 3)/11$.

$\diamondsuit$

6.3 Finite and algebraic extensions

Let L be a field extension of K. Then L can be considered as a linear vector space over K. That is, L forms a space of vectors with the usual addition group and the field K plays the role of scalar elements of the vector space.
For example, the complex numbers $\mathbb{C}$ can be seen as a two dimensional vector space over $\mathbb{R}$. In fact $\mathbb{C}$ is always identified with $\mathbb{R}^2$ by speaking of the 'complex plane'. Of course there is also a difference between $\mathbb{R}^2$ and $\mathbb{C}$, which is the multiplication of elements in $\mathbb{C}$. In $\mathbb{R}^2$ we do not have a priori such a multiplication.
In general, in viewing a field extension of K as a vector space of K we ignore the multiplication structure. For example, the non-isomorphic field extensions $\mathbb{Q}(\sqrt{2})$ and $\mathbb{Q}(\sqrt{5})$ of $\mathbb{Q}$ yield a two-dimensional vector space over $\mathbb{Q}$ (with basis $1, \sqrt{d}$ where $d = 2, 5$) but the vector spaces are both isomorphic to $\mathbb{Q}^2$.

Definition 6.3.1 *Let L be a field extension of K. We say that L is a finite extension of K when L, considered as K-vector space, is finite dimensional.*
The dimension of L as K-vector space is called the degree of L over K. Notation: $[L : K]$.
We call L an algebraic extension of K if every element of L is algebraic over K.

Example 6.3.2. Consider the extension $\mathbb{Q}(\sqrt[3]{2})$ of $\mathbb{Q}$. The minimal polynomial of $\sqrt[3]{2}$ over $\mathbb{Q}$ is $X^3 - 2$. So from Remark 6.2.7 we know that

$$\mathbb{Q}(\sqrt[3]{2}) = \{a_0 + a_1\sqrt[3]{2} + a_2\sqrt[3]{4} |\ a_0, a_1, a_2 \in \mathbb{Q}\}.$$

This representation is exactly the representation of a $\mathbb{Q}$-vector space with basis $1, \sqrt[3]{2}, \sqrt[3]{4}$ (note $\sqrt[3]{4} = (\sqrt[3]{2})^2$. So we see that the dimension is 3, in other words: $[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = 3$.
In Theorem 6.3.3 we generalise this argument.

♢

Theorem 6.3.3 *Let L be a finite field extension over K. Then every element of L is algebraic over K (in other words: L is an algebraic extension of K).*
Let $\alpha \in L$. Then α is algebraic over K if and only if $K(\alpha)$ is finite over K. Moreover, in that case we have $[K(\alpha) : K] = \deg(\alpha)$ and a K-basis of $K(\alpha)$ is given by $1, \alpha, \ldots, \alpha^{n-1}$ where $n = \deg(\alpha)$.

Remark 6.3.4 *The converse of the first part of Theorem 6.3.3 need not hold, an algebraic extension need not be a finite extension.*
As a counter example, the smallest subfield of $\mathbb{R}$ which contains all numbers of the form $\sqrt[n]{2}$ with $n \in \mathbb{Z}_{>0}$ is an algebraic, but infinite extension of $\mathbb{Q}$. See also Exercise 9.

Proof: Suppose $[L : K] = n < \infty$ and $\alpha \in L$. Since any $n + 1$-tuple of vectors in n-dimensional space is dependent, there is a non-trivial relation with coefficients in K between the elements $1, \alpha, \ldots, \alpha^{n-1}$. Let us say $a_0 \cdot 1 + a_1 \cdot \alpha + ... + a_n \cdot \alpha^n = 0$ with $a_0, a_1, ..., a_n \in K$, not all zero. Hence α is a zero of the polynomial $a_0 + a_1X + \ldots + a_nX^n \in K[X]$. Therefore α is algebraic over K.
As a particular case we see that if $K(\alpha)$ is a finite extension of K, then α is algebraic.
Now suppose that α is algebraic over K of degree n. We have seen in Remark 6.2.7 that any element in $K(\alpha)$ can be written uniquely in the form. $a_0 + a_1\alpha + \ldots + a_{n-1}\alpha^{n-1}$ with $a_i \in K$ en $n := \deg(\alpha)$. Therefore $[K(\alpha) : K] = n$ and $K(\alpha)$ is a finite extension.

□

6.4 Composite extensions

In practice we often encounter the following situation. We are given a field extension L over K and $\alpha_1, \ldots, \alpha_n \in L$ algebraic elements over K. We then consider $K(\alpha_1, \ldots, \alpha_n)$, the smallest subfield of L containing K and $\alpha_1, \ldots, \alpha_n$. From Theorem 6.4.2 it follows that $K(\alpha_1, \ldots, \alpha_n)$ is again finite over K and we would like to compute its degree. For example, what is the degree of $\mathbb{Q}(\sqrt{2}, \sqrt{5})$ over $\mathbb{Q}$, or $[\mathbb{Q}(\sqrt{2}, \sqrt[3]{2}) : \mathbb{Q}]$, or $\mathbb{Q}(e^{\pi i/4}, \sqrt{2}) : \mathbb{Q}]$. The most important tool in these calculations will be the following Theorem.

Theorem 6.4.1 (Tower relation for field degrees) *Let K be a field, L a field extension of K and let M be a field extension of L (so $K \subset L \subset M$). Then M is finite over K if and only if M is finite over L and L is finite over K.*
Moreover, if M is finite over K, we have $[M : K] = [M : L] \cdot [L : K]$.

Proof: Suppose that M is finite over K. Since L is a sub-K-vectorspace of M, L is also finite over K. Let $\alpha_1, ..., \alpha_n$ be a basis of the K-vectorspace M. That is, every element of M is a K-linear combination of $\alpha_1, \ldots, \alpha_n$. Since any K-linear combination is also an L-linear combination, the L-span of $\alpha_1, \ldots, \alpha_n$ certainly equals M. Hence $[M : L] \le n$ is finite.
Suppose that $[M : L] = n$ and $[L : K] = m$ are both finite. Let $\alpha_1, \alpha_2, ..., \alpha_m$ be a K-basis of L and $\beta_1, \beta_2, ..., \beta_n$ an L-basis of M. We will show that $\{\alpha_i\beta_j : 1 \le i \le m, 1 \le j \le n\}$ is a K-basis of M.
Let $x \in M$, then there exist $y_1, \ldots, y_n \in L$ such that

$$x = \sum_{j=1}^{n} y_j \beta_j.$$

Since $\alpha_1, ..., \alpha_m$ is a K-basis of L each y_j can be written as

$$y_j = \sum_{i=1}^{m} a_{ij} \alpha_i$$

where $a_{ij} \in K$ for all i, j. Hence

$$x = \sum_{1 \le i \le m, 1 \le j \le n} a_{ij} \alpha_i \beta_j.$$

So every element $x \in M$ is in the K-linear span of $\alpha_i\beta_j$ with $1 \le i \le m, 1 \le j \le n$.
It remains to show linear independence of the elements $\alpha_i\beta_j$. Suppose there exists a linear relation

$$\sum_{1 \le i \le m, 1 \le j \le n} c_{ij} \alpha_i \beta_j = 0,$$

with $c_{ij} \in K$. After regrouping the terms this is the same as

$$\sum_{j=1}^{n} (\sum_{i=1}^{m} c_{ij} \alpha_i) \beta_j = 0$$

where $\sum_{i=1}^m c_{ij}\alpha_i \in L$. From the L-linear independence of the β_j it follows that $\sum_{i=1}^m c_{ij}\alpha_i = 0$ for $j = 1, \ldots, n$. Since the α_i are K-linear independent these relations imply that $c_{ij} = 0$ for all i and j. Hence the elements $\alpha_i\beta_j$ are indeed linearly independent over K.

So we conclude that $[M : K] = mn = [M : L] \cdot [L : K]$.

□

Corollary 6.4.2 *Let L be a field extension of K and let $\alpha_1, \alpha_2, ..., \alpha_n \in L$. If $\alpha_1, \ldots, \alpha_n$ are algebraic over K then $K(\alpha_1, \ldots, \alpha_n)$ is finite over K.*

Proof: Consider the sequence of simple extensions

$$K \subset K(\alpha_1) \subset K(\alpha_1, \alpha_2) \subset \cdots \subset K(\alpha_1, \ldots, \alpha_n).$$

By Theorem 6.3.3 each simple extension is finite. Hence, by Theorem 6.4.1 and induction on n it follows that $K(\alpha_1, \ldots, \alpha_n)$ is a finite extension of K.

□

Corollary 6.4.3 *Let L be a finite field extension of K. Then there exists a finite number of $\alpha_1, \ldots, \alpha_n \in L$ such that $L = K(\alpha_1, \ldots, \alpha_n)$.*

Moreover, let d_i be the degree of α_i for $i = 1, \ldots, n$, then $[L : K] \le d_1 d_2 \cdots d_n$.

In particular, when $\gcd(d_i, d_j) = 1$ for all $i \ne j$ we have $[L : K] = d_1 d_2 \cdots d_n$.

Proof: We choose elements $\alpha_1, \alpha_2, \ldots$ in L as follows. Let $K_0 = K$ and choose $\alpha_1 \notin K_0$. Define $K_1 = K_0(\alpha_1)$. Choose $\alpha_2 \notin K_1$ and define $K_2 = K_1(\alpha_2)$. We obtain a sequence of fields $K_0 \subsetneq K_1 \subsetneq K_2 \subsetneq \cdots$ of strictly increasing degrees, all less than or equal to $[L : K]$. Since the latter is finite there must exist an index n such that the degree of K_n over K equals $[L : K]$. Hence $L = K_n = K(\alpha_1, \ldots, \alpha_n)$.

Also note that, by the tower relation for degrees, $[L : K] = [K_n : K_{n-1}] \cdots [K_2 : K_1] \cdot [K_1 : K_0] \le d_n \cdots d_2 d_1$.

In particular, since $K(\alpha_i) \subset L$, we have that d_i divides $[L : K]$. If all d_i are pairwise relatively prime we obtain that $d_1 \cdots d_n$ divides $[L : K]$. Hence $[L : K] = d_1 \cdots d_n$.

□

Example 6.4.4. Consider the field $L = \mathbb{Q}(\sqrt{2}, \sqrt[3]{2})$. The elements $\sqrt{2}$ and $\sqrt[3]{2}$ have degrees 2 and 3, which are relatively prime. Hence $[\mathbb{Q}(\sqrt{2}, \sqrt[3]{2}) : \mathbb{Q}] = 6$.

When the degrees of α_1 and α_2 are not relatively prime, it is more difficult to say something about the degree of $K(\alpha_1, \alpha_2)$. For example $\sqrt[3]{2}$ and $\omega\sqrt[3]{2}$ (where $\omega = e^{2\pi i/3}$ is a cube root of unity) are both zeros of $X^3 - 2$ and hence have degree 3. But, $[\mathbb{Q}(\sqrt[3]{2}, \omega\sqrt[3]{2}) : \mathbb{Q}] = [\mathbb{Q}(\sqrt[3]{2}, \omega) : \mathbb{Q}]$ and the latter degree is 6 because $[\mathbb{Q}(\sqrt[3]{2} : \mathbb{Q}] = 3$ and $[\mathbb{Q}(\omega) : \mathbb{Q}] = 2$. For the latter equality recall that $\omega = -1/2 + \sqrt{-3}/2$ is a quadratic number.

As another example consider the composite extension $L = \mathbb{Q}(\sqrt{5}, \sqrt{3})$ of $\mathbb{Q}$. We claim that $[\mathbb{Q}(\sqrt{5}, \sqrt{3}) : \mathbb{Q}] = 4$. By the tower rule

$$[L : \mathbb{Q}] = [L : \mathbb{Q}(\sqrt{5})] \cdot [\mathbb{Q}(\sqrt{5} : \mathbb{Q}]$$

we need to show that $[L:\mathbb{Q}(\sqrt{5})]=2$. The field L is an extension of $\mathbb{Q}(\sqrt{5})$ with the element $\sqrt{3}$, which is a zero of the polynomial X^2-3. We need to show that this is irreducible in $\mathbb{Q}(\sqrt{5})[X]$. In other words, we need to show that it has no zero in $\mathbb{Q}(\sqrt{5})$. Let us suppose that there is a zero, say $a+b\sqrt{5}$ with $a,b\in\mathbb{Q}$. Then $(a+b\sqrt{5})^2-3=0$, hence $a^2+5b^2-3+2ab\sqrt{5}=0$. Hence $a^2+5b^2-3=0$ and $ab=0$. So either $a=0$, in which case the first equation reads $5b^2-3=0$, or $b=0$, in which case the first equation reads $a^2-3=0$. Note that $5b^2-3=0$ doesn't have a rational solution b and that $a^2-3=0$ doesn't have a rational solution a. So we get a contradiction and we see that X^2-3 is irreducible in $\mathbb{Q}(\sqrt{5})[X]$. Hence $[L:\mathbb{Q}]=4$.

A more complicated example is the field $L=\mathbb{Q}(\sqrt[3]{2},\sqrt[3]{5})$. It has degree 9. To see this we use the tower rule $[L:\mathbb{Q}]=[L:\mathbb{Q}(\sqrt[3]{2}]\cdot[\mathbb{Q}(\sqrt[3]{2}:\mathbb{Q}]$. The second factor is 3. It remains to show that $[L:\sqrt[3]{2}]=3$. Note that L is an extension of $\mathbb{Q}(\sqrt[3]{2})$ by $\sqrt[3]{5}$. So we need to show irreducibility of X^3-5 in $\mathbb{Q}(\sqrt[3]{2})[X]$. Reducibility would imply the existence of an element in $\mathbb{Q}(\sqrt[3]{2})$ as zero of X^3-5. Say $a+b\sqrt[3]{2}+c\sqrt[3]{4}$ with $a,b,c\in\mathbb{Q}$ is such a zero. Elaboration of $(a+b\sqrt[3]{2}+c\sqrt[3]{4})^3-5=0$ yields after some calculation,

$$\begin{aligned}
a^3+2b^3+4c^3+12abc &= 5\\
3a^2b+6cb^2+6ac^2 &= 0\\
3a^2c+3ab^2+6bc^2 &= 0
\end{aligned}$$

Multiply the second equation with c, the third by b and subtract. We obtain $3a(2c^3-b^3)=0$. Hence either $a=0$ or $2c^3-b^3=0$. When $a=0$, substitution in the second equation yields $cb=0$ hence $b=0$ or $c=0$. When $2c^3-b^3=0$ we conclude $b=c=0$. So two of the three numbers a,b,c are zero. But then the first equation gives an equation for the remaining variable which is impossible to solve. For example, when $b=c=0$ the first equation yields $a^3=5$ which is impossible to solve in rational a. Therefore we get a contradiction and conclude that X^3-5 is irreducible in $\mathbb{Q}(\sqrt[3]{2})[X]$.

$\diamondsuit$

The last two examples indicate that the actual computation in composite extension can be rather cumbersome. In a number of cases, such as the cases just discussed, things simplify considerably if we use some Galois theory. We shall come back to this point in the chapters on Galois theory.

6.5 Determination of minimal polynomials

Given a finite composite extension $K(\alpha,\beta)$ of a field K. How can we compute the degree of elements of elements such as $\alpha\beta$ or $\alpha+\beta$ given the minimal polynomials of α,β itself.

We shall illustrate two methods by way of the example $\sqrt{2}+\sqrt{3}\in\mathbb{Q}(\sqrt{2},\sqrt{3})$.

a. As first method we use the fact that every element in $\mathbb{Q}(\sqrt{2},\sqrt{3})$ is a $\mathbb{Q}$-linear combination of $1,\sqrt{2},\sqrt{3},\sqrt{2}\sqrt{3}=\sqrt{6}$. We express the first 4 powers of $\sqrt{2}+\sqrt{3}$ in terms of this basis.

$$\begin{array}{rcl} 1 &=& 1 \\ \sqrt{2}+\sqrt{3} &=& \sqrt{2}+\sqrt{3} \\ (\sqrt{2}+\sqrt{3})^2 &=& 5+2\sqrt{6} \\ (\sqrt{2}+\sqrt{3})^3 &=& 11\sqrt{2}+9\sqrt{3} \\ (\sqrt{2}+\sqrt{3})^4 &=& 49+20\sqrt{6} \end{array}$$

Using linear algebra we see that there is a $\mathbb{Q}$-linear relation between the right hand sides, namely $49+20\sqrt{6}-10(5+2\sqrt{6})+1=0$. Hence $\gamma^4-10\gamma^2+1=0$ where $\gamma=\sqrt{2}+\sqrt{3}$.

b. The second method uses the so-called conjugates of $\sqrt{2}$ and $\sqrt{3}$. Notice that $\sqrt{2}+\sqrt{3}$ is a zero of the polynomial

$$(X-(\sqrt{2}+\sqrt{3}))\times(X-(\sqrt{2}-\sqrt{3}))\times(X-(-\sqrt{2}+\sqrt{3}))\times(X-(-\sqrt{2}-\sqrt{3})).$$

One might hope that by taking this symmetric combination the product is a polynomial with coefficients in $\mathbb{Q}$. And indeed, after some calculation one obtains the product X^4-10X^2+1.

6.6 Exercises

1. Prove that every $\alpha\in\mathbb{Q}(\sqrt{2})$ is algebraic over $\mathbb{Q}$.

2. Prove that the set of complex numbers which are algebraic over $\mathbb{Q}$ is a countable set. (Hint: show that $\mathbb{Z}[X]$ is countable.)

   Prove that there exist complex numbers, and also real numbers, which are transcendental over $\mathbb{Q}$.

3. Does there exist $\alpha\in\mathbb{R}$ such that $\mathbb{Q}(\alpha)=\mathbb{R}$? (Hint: what is the cardinality of $\mathbb{Q}(\alpha)$?)

4. Prove that for every $n\in\mathbb{Z}_{>0}$ the polynomial X^n-2 is the minimal polynomial of $\sqrt[n]{2}$ over $\mathbb{Q}$.

5. Let α be an algebraic element over a field K and let $f(X)=\sum_{i=0}^n a_iX^i$ be its minimal polynomial with $a_n=1$.

   Prove: if $\alpha\neq 0$ then $a_0\neq 0$, and $\alpha^{-1}=-\sum_{i=1}^n a_0^{-1}a_i\alpha^{i-1}$.

6. Compute the minimal polynomial and the degree over $\mathbb{Q}$ for each of the following α's:
   $2-\sqrt{3}$, $\sqrt[3]{2}+\sqrt[3]{4}$, $\sqrt{3+2\sqrt{2}}$; β^{-1}, $\beta+1$ where $\beta^3+3\beta-3=0$.

7. a. Prove: $\mathbb{Q}(\sqrt{2})(\sqrt{7})=\mathbb{Q}(\sqrt{2}+\sqrt{7})$, and $\dim_{\mathbb{Q}}\mathbb{Q}(\sqrt{2}+\sqrt{7})=4$.

   b. Compute the minimal polynomial of $\sqrt{2}+\sqrt{7}$ over $\mathbb{Q}$.

8. Let $\alpha \in \mathbb{R}$, $\alpha^3 - \alpha - 1 = 0$. Write each of the following elements in the form $a + b\alpha + c\alpha^2$, with $a, b, c \in \mathbb{Q}$:
$$\alpha^{10}, \quad \alpha^{-10}, \quad (\alpha^2 + \alpha + 1)^2, \quad (\alpha^2 + 1)^{-1}.$$

9. Let $L = \cup_{n=1}^{\infty} \mathbb{Q}(\sqrt[n]{2})$. Prove:
   a. L is a field (hint: $\mathbb{Q}(\sqrt[n]{2}) \cup \mathbb{Q}(\sqrt[m]{2}) \subset \mathbb{Q}(\sqrt[nm]{2})$);
   b. L is algebraic over $\mathbb{Q}$;
   c. For each $n \in \mathbb{N}$ the field L contains a field of degree n over $\mathbb{Q}$ (so L is not finite over $\mathbb{Q}$).

10. a. Prove that there exist no $a, b \in \mathbb{Q}$ such that $(a + b\sqrt{2})^2 = 3$. Conclude from this that $X^2 - 3$ is irreducible in $\mathbb{Q}(\sqrt{2})[X]$.
   b. Prove: $[\mathbb{Q}(\sqrt{2}, \sqrt{3}) : \mathbb{Q}] = 4$.

11. Let L be a *finite* extension of a field K, and $\alpha \in L$. Prove that the degree of the minimal polynomial of α over $\mathbb{Q}$ divides $[L : K]$.

12. Let $f = X^4 - 4X^3 - 4X^2 + 16X - 8$. Prove that $\frac{1}{8} \cdot X^4 f(2/X)$ is an Eisenstein polynomial for 2. Conclude that f is irreducible in $\mathbb{Q}[X]$.

13. Let $\beta = 1+\sqrt{2}+\sqrt{3}$. Express $\sqrt{2}$, $\sqrt{3}$ and β^{-1} with respect to the $\mathbb{Q}$-basis $1, \beta, \beta^2, \beta^3$ of $\mathbb{Q}(\beta)$.

14. a. Prove: $\mathbb{Q}(\sqrt{2}, \sqrt[3]{5}) = \mathbb{Q}(\sqrt{2} \cdot \sqrt[3]{5}) = \mathbb{Q}(\sqrt{2} + \sqrt[3]{5})$.
   b. Determine the minimalpolynomial over $\mathbb{Q}$ of $\alpha = \sqrt{2} \cdot \sqrt[3]{5}$ and $\alpha = \sqrt{2} + \sqrt[3]{5}$.

15. a. Verify that $X^5 - 1 = (X-1)(X^4+X^3+X^2+X+1) =: (X-1)\Phi_5$ and that Φ_5 is irreducible in $\mathbb{Q}[X]$ (hint: substitute X:=X+1 in Φ_5).
   b. Let
$$M := \mathbb{Q}[X]/(\Phi_5), \qquad \zeta := X + (\Phi_5),$$
$$\text{en zij} \qquad \beta := X + X^4 + (\Phi_5) \in M, \qquad L := \mathbb{Q}[\beta] \subset M.$$
   Determine $a, b \in \mathbb{Q}$ such that $\beta^2 = a\beta + b$ and determine the minimal polynomial of β over $\mathbb{Q}$.
   c. Determine $[M : L]$ and the minimal polynomial of ζ over L.
   d. Give a formula for cos $\frac{2\pi}{5}$ in terms of rational numbers and their square roots.

16. Let $\alpha \in \mathbb{R}$, $\alpha^3 - \alpha - 1 = 0$. Determine the minimal polynomial over $\mathbb{Q}$ of the following numbers:
$$\alpha - 1, \quad \alpha^2 + \alpha + 1, \quad (\alpha^2 + 1)^{-1}.$$

17. Let α be algebraic over a field K and suppose that $[K(\alpha) : K]$ is *odd*. Prove: $K(\alpha) = K(\alpha^2)$.

18. Let L be a field extension of K and let en K_0 be the algebraic closure of K in L.

    Prove: every $\alpha \in L, \alpha \notin K_0$ is transcendental over K_0.

19. Let α be transcendental over a field K, and $\beta \in K(\alpha)$, $\beta \notin K$. Prove:

    a. α is algebraic over $K(\beta)$ (hint: let $\beta = f(\alpha)/g(\alpha)$, and consider the polynomial $f(X) - \beta g(X)$).

    b. β is transcendental over K.

20. Let K be a field.

    a. ('partial fractions'). Prove that the following set forms a K-basis of $K(X)$:

    $$\{X^n : n \in \mathbb{Z}_{\geq 0}\} \cup \{X^i \cdot f^{-m} : f \in K[X]\},$$

    the $f \in K[X]$ are monic and irreducible and $m \in \mathbb{Z}_{>0}, 0 \leq i < gr(f)\}$.

    b. Let α be transcendental over K. Prove that $[K(\alpha) : K]$ equals the cadinality of K when K is infinite and that $[K(\alpha) : K]$ is countable if K is finite.

21. Let $K = \mathbb{F}_2(X, Y) = Q(\mathbb{F}_2[X, Y])$, (the quotient field of $\mathbb{F}_2[X, Y]$).

    a. Let $f = T^2 + X \in K[T]$. Prove that f is irreducible and let

    $$L := K[T]/(f), \qquad t := T + (f) \quad \in L.$$

    b. Let $g = S^2 + Y \in L[S]$. Prove that g is irreducible and let

    $$M := L[S]/(g), \qquad s := S + (g) \quad \in M.$$

    c. Notice that $K \subset L \subset M$ and prove that $1,\ t,\ s,\ st$ form a K-basis of M.

    d. Prove that for every $\alpha \in M$, $\alpha \notin K$: the degree over K is 2. Conclude that the extension M of K is not simple.

22. Let $f = X^4 + aX^3 + bX^2 + cX + d \quad \in K[X]$ where K is a field with $kar(K) \neq 2$, 3 and let $\alpha_1, \ldots, \alpha_4$ be the zeros of f (in an extension of K).

    a. Define:

    $$\begin{array}{rcl} C_1 & = & (\alpha_1 + \alpha_2 - \alpha_3 - \alpha_4)^2 \\ C_2 & = & (\alpha_1 - \alpha_2 + \alpha_3 - \alpha_4)^2 \\ C_3 & = & (\alpha_1 - \alpha_2 - \alpha_3 + \alpha_4)^2. \end{array}$$

    Express α_1 in terms of $\sqrt{C_i}$ and the coefficients a of f.

    b. Verify that the S_4 action (permutation of the α_i) permutes the C_i. Also verify that the subgroup $H = \{(1), (12)(34), (13)(24), (14), (23)\}$ fixes the C_i.

c. Show that:

$$\begin{aligned} C_1+C_2+C_3 &= 3a^2-8b \\ C_1C_2+C_1C_3+C_2C_3 &= 3a^4-16a^2b+16b^2+16ac-64d \\ C_1C_2C_3 &= (a^3-4ab+8c)^2. \end{aligned}$$

d. Verify that, given this information, one can solve the general fourth degree equation.

23. Let $f = X^3 + aX^2 + bX + c \in \mathbb{Q}[X]$ be an irreducible polynomial with α_1, α_2, $\alpha_3 \in \mathbb{C}$. Define $\Omega^f_{\mathbb{Q}} \cong \mathbb{Q}(\alpha_1, \alpha_2, \alpha_3)$. Then :

$$\sqrt{\triangle} := (\alpha_1-\alpha_2)(\alpha_1-\alpha_3)(\alpha_2-\alpha_3) \in \Omega^f_{\mathbb{Q}}, \quad \text{en} \quad \triangle \in \mathbb{Q},$$

where $\triangle$ is called the discriminant of f.

a. Prove that $[\Omega^f_{\mathbb{Q}} : \mathbb{Q}] = 3$ or 6.

b. Prove that $\sqrt{\triangle} \notin \mathbb{Q} \;\Rightarrow\; [\Omega^f_{\mathbb{Q}} : \mathbb{Q}] = 6$.

c. Suppose that $\sqrt{\triangle} \in \mathbb{Q}$. Write $f = (X-\alpha_1)(X^2+rX+s) \in \mathbb{Q}(\alpha_1)[X]$. Prove that $\alpha_2 \in \mathbb{Q}(\alpha_1)$ by expressing α_2 in terms of $\sqrt{\triangle}, a, b, c, r, s, \alpha_1 \in \mathbb{Q}(\alpha_1)$. Conclude $[\Omega^f_{\mathbb{Q}} : \mathbb{Q}] = 3 \Leftrightarrow \sqrt{\triangle} \in \mathbb{Q}$.

Chapter 7

Motivatie Galoistheorie (in Dutch)

Galoistheorie bestudeert de abstracte structuur van de verzameling wortels van een veelterm of, in moderne bewoordingen, de automorfismen van lichaamsuitbreidingen (in casu het splijtlichaam van de veelterm). In deze inleiding wordt óók geprobeerd de overgang naar die abstractie te duiden in historische-wiskundig kader, en via talrijke rekenopgaven waarvan sommige uitwerkingen voorhanden zijn.

De gebruikelijke twee mogelijkheden voor de presentatie van het materiaal worden allebei besproken, nl. diegene waarin (à la Galois) de stelling van het primitieve element eerst wordt bewezen, en vervolgens de hoofstelling eruit wordt afgeleid; en diegene waarbij de hoofdstelling wordt bewezen op basis van het lemma van Dedekind, waaruit dan de stelling van het primitieve element volgt.

Gunther Cornelissen

Figure 7.1: Enige helden der Galoistheorie: N.-H. Abel (1802-1829), E. Artin (1898-1962), E. Galois (1811-1832) en P. Ruffini (1765-1822)

7.1 Korte geschiedenis van het oplossen van veeltermvergelijkingen

Het is welbekend hoe de wortels van een vierkantsvergelijking $X^2 + bX = c$ bepaald worden. De wortels worden gegeven door de formules

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 + 4c}}{2}.$$

Een "bewijs" vinden we in het boek *Al-jabr w'al muqabala*, van Musa al-Khowarizmi (± 830 A.D.). Mohammed ibn Musa al-Khowarizmi geboortig uit Khiwa schreef verschillende boeken. Twee ervan hebben, door een Latijnse vertaling ervan, aanzienlijke invloed uitgeoefend. Zijn elementaire rekenkunde, bewaard gebleven in een latijnse vertaling van de twaalfde eeuw, heeft bijgedragen tot de verspreiding van het decimale positiestelsel in de Arabische en later in de Latijnse landen. Dezelfde vertaling met als aanhef "Algorismi de numero Indorum" heeft het woord *algoritme* blijvend aan de wiskundige taal toegevoegd. De latijnse vertaling van het boek *Al-jabr w'al muqabala*, wat staat voor "Leer der vergelijkingen", heeft tot het woord *algebra* gevoerd.
Euclides geeft een rigoureus meetkundig bewijs van de formules in zijn *Elementen*, boek II, stelling 5 *Als een rechte lijn in gelijke (C) en ongelijke delen wordt verdeeld (D) dan is de rechthoek begrensd door de ongelijke segmenten van het geheel (AH) samen met het vierkant op de rechte lijn tussen de snijpunten* (Z^2) *gelijk aan het vierkant (CF) op de helft.*

Figure 7.2: Het bewijs van Euclides

Als we aannemen dat de lengte van het lijnstuk $AB = b$ en de oppervlakte van de rechthoek $(AH) = c$ gekend zijn en dat de ongelijke segmenten $AD = x$ en $DB = y$ onbekend zijn, dan geeft de stelling de oplossing van het stelsel

$$\begin{cases} x + y = b \\ xy = c \end{cases}$$

Dus van de vergelijking

$$X^2 - bX + c.$$

Stellen we namelijk $Z = x - \frac{b}{2}$ dan zegt de stelling dat $c + Z^2 = (\frac{b}{2})^2$, waaruit

$$Z = \sqrt{(\frac{b}{2})^2 - c}$$

en

$$x = \frac{b}{2} + \sqrt{(\frac{b}{2})^2 - c}, y = \frac{b}{2} - \sqrt{(\frac{b}{2})^2 - c}$$

volgt.

De methode van al-Khowarizmi staat dichter bij de manipulaties van vergelijkingen waaraan wij gewend zijn, namelijk: maak van het linkerlid van de vergelijking

$$X^2 + bX = c$$

een volledig kwadraat door er $\frac{b^2}{4}$ bij te tellen, dit geeft

$$X^2 + bX + \frac{b^2}{4} = (X + \frac{b}{2})^2 = c + \frac{b^2}{4}.$$

De methode van al-Khowarizmi lag uiteindelijk ook aan de basis van de veralgemeningen naar hogeregraadsvergelijkingen. Deze veralgemeningen waren niet voor de hand liggend. In 1494 schrijft Luca Pacioli in zijn *Summa de Arithmetica, Geometria, Proportione et Proportionalita* dat de oplossingen van $X^3 + mX = n$ en $X^3 + n = mX$ (in moderne notatie) even onmogelijk zijn als de kwadratuur van de cirkel. Rond 1515 kon Scipione del Ferro Pacioli tegenspreken, hij vond de oplossing van het probleem: *Een kubus plus enkele van zijn ribben is gelijk aan een getal. Bepaal de ribbe.*
Scipione del Ferro maakte zijn oplossing niet bekend maar gaf ze door aan zijn leerlingen. Rond 1535 vond Nicolo Fontana bijgenaamd Tartaglia, door de methode van al-Khowarizmi naar drie dimensies te veralgemenen, eveneens de oplossing van dit probleem:

$$X^3 + bX = X^3 + 3uvX = c$$

Zoek een u, v zodat

$$\begin{cases} 3uv = b \\ u - v = X \end{cases}$$

dan is

$$(u - v)^3 + 3uv(u - v) = c$$
$$(u - \frac{b}{3u})^3 + 3u\frac{b}{3u}(u - \frac{b}{3u}) = c$$

En u^3 kan nu bekomen worden door een vierkantsvergelijking op te lossen:

$$27(u^3)^2 - 27(u^3)c - b^3 = 0.$$

We laten het als oefening aan de lezer om hieruit de formules voor de oplossing van de derdegraadsvergelijking af te leiden.
Tartaglia's methode werd verder uitgebreid en gepubliceerd door Cardano (1501-1576). De formules voor de wortels van de algemene derdegraadsvergelijking noemt men de formules van Cardano. Ludovico Ferrari (1522-1565), een leerling van Cardano, vond analoge formules voor de vierdegraadsvergelijking.
De oplossingen voor de derde- en vierdegraadsvergelijking staan enerzijds symbool voor de snelle vooruitgang in de theorie van de (algebraïsche) vergelijkingen in het midden van de 16de eeuw maar vormen anderzijds een "eindpunt" voor die vooruitgang. Het duurt enkele eeuwen, tot het werk van Abel en Galois, voor resultaten in verband met vijfdegraadsvergelijkingen (en met vergelijkingen van hogere graad) gevonden worden. Er was bv. nog geen goede notatie om

met vergelijkingen om te gaan. Een belangrijk concept ontbrak nog, *de veelterm*. In zijn *L'Arithmetique (1585)* combineert Simon Stevin nieuwe notaties en theoretische vooruitgang van vroegere auteurs (o.a. Bombelli en Nunes) tot een eerste samenvattend werk over “veeltermen”. (Stevin noemt ze gehele algebraïsche getallen[1]). Stevin noteert de variabele als $\boxed{1}$, het kwadraat van de variabele als $\boxed{2}$ enz.,

$$3\boxed{3}+5\boxed{2}-4\boxed{1}+6\boxed{0}$$

staat dus voor

$$3X^3+5X^2-4X+6$$

in onze notatie. Belangrijker is echter dat Stevin een rekenkunde van veeltermen ontwikkelde, in moderne taal betekent dit dat hij vaststelde dat de veeltermen een *ring* vormen. Hij bewees eveneens dat er voor veeltermen een Euclidisch delingsalgoritme (deling van veeltermen met rest) bestaat (dit speelt een belangrijke rol in de theorie).

Naast het werk van Simon Stevin vormt het in 1591 verschenen boek *In Artem Analyticem Isagoge* van François Viète een tweede mijlpaal voor de theorie van de veeltermen. Het is Viète's idee om veeltermen in onbepaalde coëfficiënten te beschouwen (hij stelde zowel de coëfficiënten als de variabelen door letters voor). Hij is dus de ontdekker van wat we nu de *algemene* (of *generieke*) *veelterm* noemen.

Om Viète's inzicht in “vergelijkingen” te illustreren vermelden we zijn oplossing van een probleem dat Adriaan van Roomen in 1593 stelde aan “alle wiskundigen van de hele wereld”. Zoek een oplossing van de vergelijking[2]:

$$\begin{gathered}45X-3795X^3+95634X^5-1138500X^7+7811375X^9-34512075X^{11}+105306075X^{13}\\ -232672680X^{15}+384942375X^{17}-488494125X^{19}+483841800X^{21}-378658800X^{23}\\ +236030652X^{25}-117679100X^{27}+46955700X^{29}-14945040X^{31}+3764565X^{33}\\ -740259X^{35}+111150X^{37}-12300X^{39}+945X^{41}-45X^{43}+X^{45}=A\end{gathered}$$

van Roomen gaf een paar voorbeelden:

- $A=\sqrt{2+\sqrt{2+\sqrt{2+\sqrt{2}}}}$ dan is $X=\sqrt{2-\sqrt{2+\sqrt{2+\sqrt{2+\sqrt{3}}}}}$.
- $A=\sqrt{2+\sqrt{2}}$ dan is $X=\sqrt{2-\sqrt{2+\sqrt{\frac{3}{16}}+\sqrt{\frac{15}{16}}+\sqrt{\frac{5}{8}-\sqrt{\frac{5}{64}}}}}$.

en vroeg de oplossing voor

$$A=\sqrt{1+\frac{3}{4}-\sqrt{\frac{5}{16}}-\sqrt{1+\frac{7}{8}-\sqrt{\frac{45}{64}}}}.$$

[1] De terminologie “gehele algebraïsche getallen” heeft nu een andere betekenis

[2] van Roomen gebruikte de notatie van Simon Stevin

Viète herkende de vergelijking als deze die $2\sin(45\alpha)$ uitdrukt als een functie van $2\sin\alpha$. De oplossing wordt dus bepaald door een hoek α te zoeken zodat $2\sin(45\alpha) = A$, $X = 2\sin\alpha$. Hij bepaalde de numerieke waarde van de oplossing $2\sin(\frac{\pi}{3^3.5^2})$ tot op 9 decimalen maar stelde tevens dat er 23 positieve oplossingen waren en 22 negatieve oplossingen. (Als $2\sin(45\alpha) = A$ dan ook $2\sin(45(\alpha + k\frac{2\pi}{45})) = A$, $k = 0, \ldots, 44$. Voor $k = 0, \ldots, 22$ is $2\sin(\alpha_k) \geq 0$ en voor $k = 23, \ldots, 44$ is $2\sin(\alpha_k) \leq 0$.)
Ook de methode die Viète gebruikte is opmerkelijk, hij liet zien dat aangezien $45 = 3^2 \cdot 5$ de vergelijking kan terug gebracht worden tot twee vergelijkingen van graad 3 en één vergelijking van graad 5. Dit idee zal later een centrale rol spelen in het werk van Euler, Laplace en Gauss. De Galoistheorie geeft inzicht in de structuur van vergelijkingen zodat we kunnen begrijpen waarom sommige vergelijkingen terug te brengen zijn tot vergelijkingen van lagere graad en sommige niet.
Viète stelde in een later werk *De Recognitione Aequationum* ("Over het begrijpen van vergelijkingen", dat postuum gepubliceerd werd in 1615), de vraag naar de *relatie tussen de wortels en de coëfficiënten van een vergelijking.* De oplossing van dit probleem vinden we in het werk van Girard (1629). Girard formuleerde ook het eerst *dat een vergelijking van graad n ook "n wortels moet hebben".* Een inzicht waaraan Viète's resultaten zeker veel hebben bijgedragen. De moderne vorm van deze uitspraak "elke veelterm van graad n over de complexe getallen heeft n wortels in de complexe getallen" staat bekend als de **Hoofdstelling van de Algebra** (alhoewel ironisch eigenlijk een stelling uit de analyse). Het eerste rigoreuze bewijs hiervan werd gegeven door Gauss, maar daarvoor hadden belangrijke wiskundigen (Euler, Laplace, Lagrange ...) al pogingen ondernomen.
Een getal dat bekomen wordt door het herhalen van de gebruikelijke bewerkingen van de rekenkunde en worteltrekken uitgaande van getallen in een lichaam K heet "uitgedrukt is in radicalen over K"; bijvoorbeeld

$$\sqrt[n]{\sqrt[m]{a} - b\sqrt{c}} + \sqrt[l]{d - \sqrt[k]{e}}$$

met a, b, c, d, e elementen van een lichaam K en n, m, l, k gehele getallen. We weten al dat de wortels van eerste-, tweede-, derde- en vierdegraadspolynomen allen kunnen uitgedrukt worden in radicalen over hun coëfficiëntenlichaam (als dat lichaam niet karakteristiek 2 of 3 heeft). Abel bewees in 1823 dat de wortels van een algemene vijfdegraadsvergelijking niet kunnen uitgedrukt worden in radicalen. Jaren daarvoor in 1799 publiceerde Ruffini twee volumes: *Teoria Generale delle Equazioni*, waarin hij eveneens beweerde een bewijs te geven voor dit feit. Alhoewel men niet expliciet een fout of een tekortkoming in het bewijs kon aangeven werd Ruffini's bewijs op heel veel scepticisme onthaald. Abels bewijs schiep meer duidelijkheid en werd gepubliceerd in het eerste nummer van het tijdschrift *Journal für die reine und angewandte Mathematik*. De theorie die ontwikkeld werd door Evarist Galois in 1830 geeft een definitieve oplossing van het probleem. Galois' werk geeft meer inzicht in het bewijs van Abel en levert tevens een criterium dat aangeeft of de wortels van een gegeven polynoom al dan niet uitgedrukt kunnen worden in radicalen. Het zal verschillende jaren duren

voor Galois' resultaten begrepen worden door de wiskundige gemeenschap maar als het zover is wordt de Galoistheorie één van de belangrijkste hoekstenen in de verdere ontwikkeling van de algebra en de getaltheorie.
Volgens Galois is het centrale object dat de theorie van een vergelijking $f(X) = 0$ beheerst de groep van permutaties van de wortels van f (als complexe getallen) die "alle polynoomrelaties tussen die wortels invariant laten". Het probleem ermee is dat er oneindig veel "polynoomrelaties tussen de wortels" in voorkomen. Door het moderne concept van splijtlichaam wordt dit weer in een *eindig* dimensionale vectorruimte ondergebracht, en via de hoofdstelling van de Galoistheorie wordt de studie van deze eindig-dimensionale vectorruimte dan weer herleid tot de studie van een eindige groep.

7.2 Construeerbaarheid

In de oude Griekse wiskunde zijn veel constructieproblemen te vinden, o.a. de de zgn. *drie Delische problemen*:

1. (een cirkel kwadrateren) Gegeven een cirkel, construeer een vierkant met dezelfde oppervlakte.

2. (een kubus verdubbelen) Gegeven een kubus, construeer een kubus met het dubbele volume.

3. (een hoek driedelen) Gegeven een hoek, construeer een hoek waarvan er precies drie in de gegeven hoek passen.

Dat de theorie van lichaamsuitbreidingen leidt tot de oplossing van deze problemen kan vreemd lijken, toch is de reduktie van het meetkundige naar het algebraïsche probleem niet zo moeilijk. Als deze vertaling is geschied, kunnen de "negatieve" resultaten, d.w.z. het feit dat bepaalde constructies niet met passer en liniaal alleen uitgevoerd kunnen worden, bewezen worden door ze in verband te brengen met de graad van een lichaamsuitbreiding. Positieve resultaten, zoals het bewijs dat een regelmatige 17-hoek kan geconstrueerd worden met passer en liniaal, kunnen verkrijgen worden als toepassing van de Galoistheorie.
We moeten eerst heel precies beschrijven wat we bedoelen met een *constructie*. We beginnen met de toegestane constructieregels.

1. Twee punten in het vlak zijn gegeven. Ze worden als *construeerbaar* beschouwd.

2. Zijn twee punten geconstrueerd, dan kunnen we hun verbindingsrechte trekken of om één van de punten een cirkel construeren die door het andere punt gaat. Zulke cirkels en rechten noemen we *construeerbaar*.

3. De snijpunten van geconstrueerde cirkels en rechten noemen we *construeerbare punten*.

Opmerking 7.2.1 *1. Opgelet: niet alle punten van een construeerbare rechte of een construeerbare cirkel zijn construeerbaar.*
2. Wanneer men deze regels volgt kan men het liniaal alleen gebruiken om een rechte te trekken door twee (construeerbare) punten. Men mag bijvoorbeeld het liniaal niet gebruiken om "lengten" af te passen.

We beschrijven een aantal welbekende basisconstructies. De nummering in de tekeningen geeft de volgorde van de uitvoering aan. We gebruiken hulppunten in de verschillende constructies. Alhoewel deze hulppunten totaal willekeurig zijn, nemen we steeds aan dat we deze punten (met de 3 regels) eerst geconstrueerd hebben. Dit is geen beperking van de algemeenheid aangezien het resultaat, in dit geval de geconstrueerde loodlijn, onafhankelijk is van de gekozen hulppunten. Door deze afspraak te maken hoeven we in de redenering niet bij te houden welke punten hulppunten zijn en welke construeerbare punten zijn.

Constructie 7.2.2 *Men kan een loodlijn construeren door een (construeerbaar) punt p op een (construeerbare) rechte l, cf. figuur 7.3.*

Figure 7.3: Constructie: loodlijn uit een punt p op een rechte l, links $p \notin l$ en rechts $p \in l$.

Constructie 7.2.3 *Men kan een evenwijdige door een (construeerbaar) punt aan een (construeerbare) rechte construeren, cf. figuur 7.4 (we gebruiken constructie 7.2.2).*

Constructie 7.2.4 *Op een gegeven (construeerbare) rechte kan men een lijnstuk afpassen met een gegeven (construeerbaar) beginpunt en met een lengte gelijk aan de afstand tussen twee gegeven (construeerbare) punten, cf. figuur 7.4 (we gebruiken hierbij constructie 7.2.3).*

Figure 7.4: Constructie van evenwijdige

Definitie 7.2.5 *Een reëel getal $a \in \mathbf{R}$ heet* construeerbaar *als de absolute waarde van a als afstand tussen twee construeerbare punten voorkomt, waarbij de afstand tussen de twee gegeven basispunten bij definitie lengte 1 heeft.*

Met de constructies 7.2.2, 7.2.3, 7.2.4, kunnen we in het vlak een assenstelsel invoeren, waarin de twee basispunten coördinaten $(0,0)$ en $(0,1)$ hebben.

Stelling 7.2.6 *Een punt $p = (a,b) \in \mathbf{R}^2$ is juist dan construeerbaar als zijn coördinaten a en b, (in een assenstelsel zoals hiervoor bepaald werd), construeerbaar zijn.*

Bewijs: Gegeven een construeerbaar punt $p = (a, b)$ dan kunnen we de coördinaten construeren door uit p de loodlijn op de assen neer te laten.
Omgekeerd: als a en b construeerbare reële getallen zijn, dragen we de lengtes $|a|$ en $|b|$ op de assen over (in de goede richting) en richten in de verkrijgen punten de loodlijnen op. Het snijpunt van beide loodlijnen is het punt p. □

Stelling 7.2.7 *De construeerbare reële getallen vormen een deellichaam van* $\mathbf{R}$ *dat* $\mathbf{Q}$ *bevat.*

Bewijs: We tonen aan dat voor elk koppel construeerbare getallen a, b met $a, b \geq 0$ de som $a + b$, het verschil $a - b$, het product ab en als $a \neq 0$ ook a^{-1} construeerbaar zijn. Hieruit volgt het gestelde vanwege definitie 7.2.5.
De optelling en aftrekking kunnen geconstrueerd worden met constructie 7.2.4.
Voor de vermenigvuldiging gebruiken we gelijkvormige rechthoekige driehoeken, zie figuur 7.5:

Figure 7.5: Constructie van het product

Als de eerste driehoek en een zijde van de tweede driehoek gegeven zijn, kan men de tweede driehoek construeren door evenwijdigen te trekken (constructie 7.2.3).
Om ab te construeren, kiezen we $r = 1, s = a$ en $r' = b$. Uit $\dfrac{r}{s} = \dfrac{r'}{s'}$ volgt dan $s' = ab$. Om a^{-1} te construeren, kiezen we $r = a, s = 1$ en $r' = 1$. Dan is $s' = a^{-1}$. □

Stelling 7.2.8 *Als* $a \in \mathbf{R}$, $a > 0$ *een construeerbaar getal is, dan is ook* $\sqrt{a}$ *construeerbaar.*

Bewijs: Hiervoor gebruiken we de constructie in figuur 7.6. De linkse driehoek heeft zijde $r = a$ en de rechtse (gelijkvormige) driehoek heeft zijden $r' = s$ en $s' = 1$. Dus geldt $a = rs' = r's = s^2$, dus $s = \sqrt{a}$.

Figure 7.6: Constructie van de vierkantswortel

□

Stelling 7.2.9 *Gegeven vier punten* $p_i = (a_i, b_i)$, $i = 1, 2, 3, 4$, *met coördinaten in een deellichaam* K *van* $\mathbf{R}$. *Zij* A, B *rechten of cirkels die uit de gegeven punten geconstrueerd werden (met passer en liniaal volgens de toegestane methoden). Dan liggen de coördinaten van de snijpunten van* A *en* B *of in* K *of in een kwadratische uitbreiding* $K(\sqrt{r})$ *van* K, *met* $r \in K$, $r > 0$.

Bewijs: De rechte door twee punten $p = (a, b)$ en $q = (c, d)$ is bepaald door een lineaire vergelijking

$$(c - a)(y - b) = (d - b)(x - a).$$

De cirkel rond p en door het punt q is bepaald door een kwadratische vergelijking

$$(x-a)^2 + (y-b)^2 = (c-a)^2 + (d-b)^2.$$

Het snijpunt van twee rechten door de gegeven punten wordt dus bepaald door een stelsel van 2 lineaire vergelijkingen met coëfficiënten in K op te lossen. De coördinaten van het snijpunt liggen dus in K.
De coördinaten van de snijpunten van een rechte en een cirkel geconstrueerd vanuit de gegeven punten bekomt men door in de vergelijking van de cirkel één variabele te elimineren gebruikmakend van de lineaire relatie gegeven door de vergelijking van de rechte. Men moet dan een kwadratische vergelijking over K oplossen.
De oplossingen liggen in $K(\sqrt{D})$ met D de discriminant van de kwadratische vergelijking. Is $D = 0$ dan ligt de oplossing (en dus de coördinaten van de snijpunten) in K. Is $D > 0$ dan ligt de oplossing in een kwadratische uitbreiding van de gezochte vorm. $D < 0$ komt niet voor, omdat dit betekent dat de rechte en de cirkel elkaar niet snijden.
Tenslotte om de coördinaten te verkrijgen van de snijpunten van twee cirkels moeten we een stelsel kwadratische vergelijkingen oplossen. We hebben echter te maken met een bijzonder geval, het verschil van de twee kwadratische vergelijkingen is een lineaire vergelijking (in beide vergelijkingen is het deel bepaald door de kwadratische termen gelijk aan x^2+y^2). Dus dit geval herleidt zich tot het voorgaande. □

Stelling 7.2.10 *Zij $a_1, \ldots, a_m$ een stel construeerbare reële getallen. Dan bestaat er een toren van uitbreidingen*

$$\mathbf{Q} = K_0 \subset K_1 \subset \cdots \subset K_n = K$$

met de volgende eigenschappen:

1. *K is een deellichaam van $\mathbf{R}$,*

2. *$a_1, \ldots, a_m \in K$,*

3. *voor $i = 0, \ldots, n-1$ geldt*

$$K_{i+1} = K_i(\sqrt{r_i})$$

met $0 < r_i \in K_i$ en $\sqrt{r_i} \notin K_i$.

Omgekeerd: zij

$$\mathbf{Q} = K_0 \subset K_1 \subset \cdots \subset K_n = K$$

een keten die aan bovenstaande eigenschappen 1 en 3 voldoet, dan is elk element van K construeerbaar.

Bewijs:
Kies het coördinatensysteem zoals hiervoor, de 2 basispunten hebben dus coördinaten in $\mathbf{Q}$. De constructie van de getallen a_i bestaat uit verschillende stappen. In elke stap bepaalt men snijpunten van (construeerbare) rechten en (construeerbare) cirkels. Als men de coördinaten van de verkrijgen punten aan $\mathbf{Q}$

toevoegt, bouwen deze vanwege stelling 7.2.9 een toren van lichaamsuitbreidingen die voldoet aan de drie eigenschappen.
Zij omgekeerd een toren van lichaamsuitbreidigen met de twee eigenschappen gegeven, dan volgt uit stelling 7.2.8 en 7.2.7 dat elk element in K construeerbaar is. □

Gevolg 7.2.11 *Zij a een construeerbaar reëel getal, dan is a algebraïsch over* $\mathbf{Q}$ *en*

$$[\mathbf{Q}(a):\mathbf{Q}] = 2^n.$$

BEWIJS: Zij K het laatste lichaam in een toren met de eigenschappen uit stelling 7.2.10, en zodat $a \in K$. Dan geldt $[K:\mathbf{Q}] = 2^m$ vanwege de torenformule. De torenformule impliceert eveneens dat $[\mathbf{Q}(a):\mathbf{Q}]|2^m$, aangezien $\mathbf{Q}(a) \subset K$.

□

Gevolg 7.2.11 impliceert de “onmogelijkheid” van de constructies waarnaar in de Delische problemen wordt gevraagd. Daarvoor moeten we wel nog definiëren wat we bedoelen met een “hoek”: het bestaat uit twee halve rechten met gemeenschappelijk beginpunt. *Een hoek θ is construeerbaar precies als $\cos\theta$ een construeerbaar getal is.* Dit is een natuurlijke definitie, omdat het geven van de hoek hetzelfde is als het aangeven van het beginpunt van de twee halfrechten en de twee halfrechten. Als we één ervan als coördinatenas aannemen, dan kunnen we de andere vinden als we $\cos\theta$ kennen.

Stelling 7.2.12 *a) Er bestaat een cirkel die met passer en lineaal niet kan worden gekwadrateerd.*
b) Er bestaat een kubus die met paseer en lineaal niet kan worden verdubbeld.
c) Er bestaat een hoek die met passer en lineaal niet kan worden gedriedeeld.

BEWIJS:
a) Neem een cirkel met straal 1, het vierkant met dezelfde oppervlakte zou zijde π moeten hebben, en aangezien π transcendent is[3] is het zeker niet algebraïsch en dus niet construeerbaar.
b) Neem een kubus met zijde 1, de dubbele kubus heeft dan zijde $\sqrt[3]{2}$, maar dat getal is niet construeerbaar want de minimaalvergelijking van $\sqrt[3]{2}$ is X^3-2 en $\sqrt[3]{2}$ is dus een algebraïsch getal van graad 3.
c) De hoek $\theta = 60^0$ is construeerbaar maar de hoek $\frac{1}{3}\theta = 20^0$ is niet construeerbaar. Omdat $\cos 60^0 = \dfrac{1}{2}$ is θ per definitie construeerbaar. We beweren dat $\cos 20^0$ een algebraïsch getal is van graad 3. Gevolg 7.2.11 impliceert dan dat de hoek van 20^0 niet construeerbaar is. De formules voor cosinus en sinus van sommen van hoeken geven $\cos 3\alpha = 4\cos^3\alpha - 3\cos\alpha$. Voor $x = \cos\frac{1}{3}\theta = \cos 20^0$ verkrijgen we dan $\dfrac{1}{2} = 4x^3 - 3x$ of $x = \cos 20^0$ is een wortel van de veelterm $8X^3 - 6X - 1$. Dit impliceert dat $2x$ voldoet aan $X^3 - 3X - 1 = 0$, maar dit

[3] voor een bewijs, zie bv. L. Berggren; J. Borwein, P. Borwein, Pi: a source book. Springer-Verlag, New York, 2000.

polynoom is irreducibel: een wortel zou wegens Gauss' lemma geheel moeten zijn en een deler van -1, maar zulke wortels zijn er niet. Omdat de graad drie is is het polynoom dan irreducibel. □

Het omgekeerde van gevolg 7.2.11 geldt niet. Bijvoorbeeld zijn er uitbreidingen $\mathbb{Q}(\alpha)$ van graad 4 met α niet construeerbaar. Het is moeilijker om stelling 7.2.10 te gebruiken om positieve resultaten te verkrijgen, daarvoor moeten we informatie hebben over de deellichamen van een lichaamsuitbreiding. De hoofdstelling van de Galoistheorie geeft ons zulke informatie en kan dan ook gebruikt worden om "construeerbaarheid" te bewijzen, bv. om de vraag te beantwoorden welke regelmatige n-hoeken met passer en lineaal kunnen worden geconstrueerd, en dan een constructie aan te geven (dit probleem werd voor het eerst door Gauss opgelost).

Opgaven

7.2.1. Stel dat $\zeta = e^{\frac{2\pi i}{5}}$. Toon aan dat ζ minimaalpolynoom $X^4 + X^3 + X^2 + X + 1$ heeft over $\mathbb{Q}$, dat $\xi := \zeta + \zeta^{-1} = \cos 72^0$, dat $\xi^2 + \xi - 1 = 0$, en gebruik dit om een constructie te geven van een regelmatige vijfhoek.

7.2.2. Zijn de hoeken 90^0 en 120^0 te driedelen met passer en lineaal? Bewijs of weerleg.

7.2.3. Gegeven zijn twee construeerbare punten P, Q op afstand 1. Is het mogelijk met passer en lineaar een punt te construeren op de rechte PQ zodat de afstand van dat punt tot P het inverse kwadraat is van de afstand tot Q?

Chapter 8

Splitting fields and Galois groups

8.1 Splitting fields

In previous chapters we have considered fields extensions L/K and studied elements in L, algebraic over K, and their minimal polynomials. In this section we shall reverse the approach. Given a field K and a polynomial $f \in K[x]$, does there exist a field extension of K which contains a zero of f? More generally, does there exist a field extension which contains all zeros of f?
We begin by answering the first question.

Theorem 8.1.1 *Let K be a field and $f(X) \in K[X]$ a non-constant polynomial. Then there exists a finite field extension L/K such that L contains a zero of $f(X)$.*

Proof: When f is irreducible in $K[X]$ we know that $K[X]/(f)$ is a field which contains the element $X(\text{mod } f)$ as obvious zero.
Suppose f is reducible in $K[X]$. Let g be an irreducible factor of f. Then clearly $K[X]/(g)$ is again a finite extension of K which contains a zero of g, hence it contains a zero of f.

□

We now address the question if there is a field extension which contains all zeros of a given polynomial f. To put the question more precisely we use the term splitting field.

Definition 8.1.2 *Let K be a field and $f \in K[X]$ a monic polynomial of degree $n > 0$. A field extension L of K is called a splitting field of f over K if*

1. *There exist $\alpha_1, \ldots, \alpha_n \in L$ such that $f = \prod_{i=1}^{n}(X - \alpha_i)$ and*
2. *$L = K(\alpha_1, \ldots, \alpha_n)$.*

Theorem 8.1.3 *Let K be a field and $f \in K[X]$ a non-constant polynomial. Then there exists a splitting field of f over K.*

Proof: We argue by induction on the degree of f, which we assume to be monic. When $\deg(f) = 1$ we see that K itself is a splitting field.
Suppose $n > 1$ and assume that we proved our statement for all polynomials of degree $n-1$. By Theorem 8.1.1 we know that there exists a finite extension K_1/K and $\alpha \in K_1$ such that $f(\alpha) = 0$. Hence there exists $g \in K_1[X]$ such that $f(X) = (X - \alpha)g(X)$. The degree of g is $n-1$, hence by our induction hypothesis there is a splitting field L of g over K_1. A fortiori f factors into linear factors $X - \alpha_i$ in $L[X]$. For the splitting field of f over K we simply take the field extension $K(\alpha_1, \ldots, \alpha_n) \subset L$.

□

A question that arises immediately is whether the splitting field of a polynomial is uniquely determined. In answering this question it is important to state when two extensions are considered the same. Before dealing with the question of the uniqueness of a splitting field in Corollary 8.3.9 we first introduce the concept of field isomorphisms.

8.2 The Galois group

Definition 8.2.1 *Let K, K' be two fields. Consider a ring homomorphism $\phi : K \to K'$. The kernel of ϕ is an ideal in K, hence either (0) or K itself. Since $\phi(1) = 1$ we get that $\ker(\phi) = (0)$ and so a field morphism is always injective. This is the reason we speak of a field embedding rather than a field homomorphism.*
A field embedding is called a field isomorphism if it is a bijection.
A field isomorphism of a field K to itself is called a field automorphism.
The group of automorphisms of a field K is denoted by $\mathrm{Aut}(K)$.

Example 8.2.2. $\mathrm{Aut}(\mathbb{Q}) = \mathrm{id}$.

♢

Example 8.2.3. Complex conjugation $z \mapsto \bar{z}$ is a field automorphism of $\mathbb{C}$.

♢

Example 8.2.4. Let K be a finite field. Since K cannot contain $\mathbb{Q}$ it must have positive characteristic p. The map $F_p : K \to K$ given by $F_p(x) = x^p$ is a field embedding. Because a field embedding is injective and K finite, F_p is automatically a bijection and thus an element of $\mathrm{Aut}(K)$. We call F_p the *Frobenius automorphism* of K.
An example of a finite field other than $\mathbb{F}_p$ is $\mathbb{F}_3[X]/(X^2 + 1)$.

♢

Example 8.2.5. Consider the subfields $\mathbb{Q}(\sqrt[3]{2})$, $\mathbb{Q}(\omega\sqrt[3]{2})$ and $\mathbb{Q}(\omega^2\sqrt[3]{2})$ of $\mathbb{C}$. Here $\omega = e^{2\pi i/3}$ is a primitive cube root of unity. Notice that all three numbers $\sqrt[3]{2}, \omega\sqrt[3]{2}, \omega^2\sqrt[3]{2}$ are zeros of the polynomial $X^3 - 2$ which is irreducible in $\mathbb{Q}[X]$. By Theorem 6.2.6 all three fields are isomorphic to $Q[X]/(X^3 - 2)$. Notice that $\mathbb{Q}(\sqrt[3]{2})$ is a subfield of $\mathbb{R}$, whereas the other two are not. In spite of this all three fields are isomorphic. For example $\mathbb{Q}(\sqrt[3]{2})$ is isomorphic to $\mathbb{Q}(\omega\sqrt[3]{2})$ via the isomorphism σ determined by $\sigma(x) = x$ for all $x \in \mathbb{Q}$ and $\sigma(\sqrt[3]{2}) = \omega\sqrt[3]{2}$.

$\diamondsuit$

The above example is an example of an isomorphism between two extensions L, L' of $\mathbb{Q}$ which fixes the elements of $\mathbb{Q}$. Such isomorphisms, where $\mathbb{Q}$ can be any ground field K, are going to be the important maps in Galois theory.

Definition 8.2.6 *Let K be a field and L, L' two finite extensions. An isomorphism $\sigma : L \to L'$ is called a K-isomorphism if it is a field isomorphism with the additional property that $\sigma(x) = x$ for all $x \in K$.*
Let K be a field and L a finite extension. An automorphism $\sigma : L \to L$ is called a K-automorphism if it is a field isomorphism with the additional property that $\sigma(x) = x$ for all $x \in K$.
The group of K-automorphisms of a finite extension L is called the Galois group of the extension L/K. Notation: $\mathrm{Gal}(L/K)$.

Remark 8.2.7 *Let K be a field, $f \in K[X]$ a non-constant polynomial and L a splitting field of f over K. Let $\alpha \in L$ be a zero of f and $\sigma \in \mathrm{Gal}(L/K)$. Since σ is a K-automorphism it fixes the elements of K. In particular, it follows from $\sigma(f(\alpha)) = 0$ that $f(\sigma(\alpha)) = 0$. This means that an element σ of $\mathrm{Gal}(L/K)$ permutes the zeros of f. Moreover, the action of σ is uniquely determined by the way in which the zeros of f are permuted.*
However, not necessarily all permutations of the zeros of f occur as elements of $\mathrm{Gal}(L/K)$.

The following Proposition allows us to determine the Galois group in the case of simple extensions.

Proposition 8.2.8 *Let K be a field, and L a finite extension of K generated by one element α. So $L = K(\alpha)$. Let f be the minimal polynomial of α. Suppose that L contains another zero β of f. The field embedding σ determined by $\sigma(\alpha) = \beta$ and $\sigma(x) = x$ for all $x \in K$ is an element of $\mathrm{Gal}(L/K)$.*
Conversely, every element of $\mathrm{Gal}(L/K)$ is given in this way. In particular $|\mathrm{Gal}(L/K)|$ equals the number of distinct zeros of f which lie in L.

Proof: From Lemma 6.2.6 we know that we have

$$K(\alpha) \cong K[X]/(f) \cong K(\beta)$$

via the K-isomorphisms given by $\alpha \mapsto X(\mathrm{mod}\ f) \mapsto \beta$. Hence the isomorphism σ determined by $\sigma(\alpha) = \beta$ and $\sigma(x) = x$ for all $x \in K$ is in $\mathrm{Gal}(L/K)$.
Conversely, every element $\sigma \in \mathrm{Gal}(L/K)$ maps α to another zero of f. If the number of distinct zeros of f in L is m, the number of K-automorphisms of L is also m.

$\square$

Example 8.2.9. $|\mathrm{Gal}(\mathbb{Q}(\sqrt{2})/\mathbb{Q})| = 2$. The automorphisms are given by $\sqrt{2} \mapsto \sqrt{2}$, which is the identity map, and $\sqrt{2} \mapsto -\sqrt{2}$.

$\diamondsuit$

Example 8.2.10. $|\mathrm{Gal}(\mathbb{C}/\mathbb{R})| = 2$ and the group is generated by complex conjugation, which is an $\mathbb{R}$-automorphism of $\mathbb{C}$.

♢

Example 8.2.11. Consider the fifth root of unity $\zeta = e^{2\pi i/5}$. Its minimal polynomial over $\mathbb{Q}$ is $(X^5-1)/(X-1) = X^4+X^3+X^2+X+1$. The other zeros are given by $\zeta^k = e^{2\pi k i/5}$ for $k = 1,2,3,4$ which clearly belong to $\mathbb{Q}(\zeta)$. Consider the K-automorphism σ given by $\sigma : \zeta \mapsto \zeta^3$. Notice that

$$\sigma^2(\zeta) = (\zeta^3)^3 = \zeta^4, \quad \sigma^3(\zeta) = (\zeta^3)^4 = \zeta^2$$

and finally $\sigma^4(\zeta) = (\zeta^3)^2 = \zeta$. So $\mathrm{Gal}(\mathbb{Q}(\zeta)/\mathbb{Q}) = \{\mathrm{id}, \sigma, \sigma^2, \sigma^3\}$. This is an example of a field over $\mathbb{Q}$ with a cyclic Galois group.

♢

Example 8.2.12. Let $K = \mathbb{F}_p(t)$ be the function field in one variable t and coefficients in $\mathbb{F}_p$. The polynomial $X^p - t$ is irreducible in $K[X]$ because it is an Eisenstein polynomial with respect to the irreducible element $t \in \mathbb{F}_p[t]$. Denote a zero of $X^p - t$ by $t^{1/p}$ and define $L = K(t^{1/p})$. Then, in L we have the factorisation $X^p - t = (X - t^{1/p})^p$. In other words, $X^p - t$ has precisely one distinct zero. Therefore the Galois group $\mathrm{Gal}(L/K)$ can map $t^{1/p}$ only to $t^{1/p}$. Hence $\mathrm{Gal}(L/K)$ consists of precisely one element, which is the identity map.

♢

Example 8.2.13. Let $L = \mathbb{Q}(\sqrt[3]{2})$. Any element $\sigma \in \mathrm{Gal}(L/\mathbb{Q})$ maps $\sqrt[3]{2}$ to a zero of $X^3 - 2$. But L contains only one zero of $X^3 - 2$ namely $\sqrt[3]{2}$ itself. Hence $\mathrm{Gal}(L/\mathbb{Q})$ consists of only one element, namely the identity map.

♢

Example 8.2.14. Let $L = \mathbb{Q}(\sqrt{2}, \sqrt{3})$. This is an extension of degree 4 over $\mathbb{Q}$, but it is not written as a simple extension. Of course an element in $\mathrm{Gal}(L/\mathbb{Q})$ maps $\sqrt{2}$ to $\pm\sqrt{2}$ and $\sqrt{3}$ to $\pm\sqrt{3}$. But it is not clear if all choices of $\pm$-signs are possible.
To answer this question we note that $\sqrt{2} + \sqrt{3}$ is an element of L of degree 4. Hence $\mathbb{Q}(\sqrt{2}, \sqrt{3}) = \mathbb{Q}(\sqrt{2} + \sqrt{3})$. The minimal polynomial of $\sqrt{2} + \sqrt{3}$ reads $f = X^4 - 10X^2 + 1$ and the complete set of zeros reads $\pm\sqrt{2} \pm \sqrt{3}$, each of which is contained in $\mathbb{Q}(\sqrt{2}, \sqrt{3})$. According to Theorem 8.2.8 the Galois group consists of 4 elements. Hence for every choice of signs $\epsilon_2, \epsilon_3 \in \{\pm 1\}$ there exists $\sigma \in \mathrm{Gal}(L/\mathbb{Q})$ such that $\sigma(\sqrt{2}) = \epsilon_2\sqrt{2}$ and $\sigma(\sqrt{3}) = \epsilon_3\sqrt{3}$.
Let us define in particular $\sigma \in \mathrm{Gal}(\mathbb{Q}(\sqrt{2}, \sqrt{3})/\mathbb{Q})$ by $\sigma(\sqrt{2}) = -\sqrt{2}$ and $\sigma(\sqrt{3}) = \sqrt{3}$. We define τ by $\tau(\sqrt{2}) = \sqrt{2}$ and $\tau(\sqrt{3}) = -\sqrt{3}$. We easily check that $\sigma^2 = \tau^2 = \mathrm{id}$ and $\sigma\tau = \tau\sigma$. Thus we see that $\mathrm{Gal}(\mathbb{Q}(\sqrt{2}, \sqrt{3})/\mathbb{Q})$ is isomorphic to the fourgroup of F.Klein.

♢

Motivated by the last example one might have the idea to determine the Galois group of a finite extension L/K in the following way. Determine an element $\alpha \in L$ such that $L = K(\alpha)$. Let $f \in K[X]$ be the minimal polynomial of α over K. Then determine the number of distinct zeros of f that are contained in L. This is precisely the order of $\mathrm{Gal}(L/K)$.

Although this is certainly a feasible approach, we shall not follow it here. Instead we prefer to get some more insight into general properties of Galois groups.

8.3 Galois extensions

We first show the following Theorem.

Theorem 8.3.1 *Let K be a field and L a finite extension. Then $|\mathrm{Gal}(L/K)| \le [L:K]$.*
Moreover, if $|\mathrm{Gal}(L/K)| = [L:K]$ then every irreducible polynomial $f \in K[X]$ with a zero in L contains precisely $\deg(f)$ distinct zeros in L.

Proof: We proceed by induction on $[L:K]$. When $[L:K]=1$ we have $L=K$ and the theorem is obvious. Let now $n>1$ and assume our Theorem is proved for all extensions L/K with $[L:K]<n$. Now assume that $[L:K]=n$. Let α be any element in L and not in K. Let $G=\mathrm{Gal}(L/K)$ and $H=\mathrm{Gal}(L/K(\alpha))$. Clearly H is a subgroup of G. Consider the right coset decomposition $G = \cup_{i=1}^r g_iH$ where $g_i \in G$ for $i=1,\ldots,r$ and the sets g_iH are pairwise disjoint. Let f be the minimal polynomial of α over K and suppose it has degree d. Now note that $g_i(\alpha)$ are zeros of f for $i=1,\ldots,r$. Moreover, they are distinct. For if $g_i(\alpha)=g_j(\alpha)$ for some $i\ne j$ we would get $g_i^{-1}g_j(\alpha)=\alpha$ and hence $g_i^{-1}g_j \in H$. This contradicts the definition in which g_i, g_j belong to different cosets. Since the $g_i(\alpha)$ are distinct elements in a set of at most d elements we find that $r \le d$. By the induction hypothesis $|H| = |\mathrm{Gal}(L/K(\alpha))| \le [L:K(\alpha)]$, so we get

$$|\mathrm{Gal}(L/K)| = |G| = r|H| \le d|H| \le [K(\alpha):K][L:K(\alpha)] = [L:K].$$

This completes the induction step.
Moroever, if $|\mathrm{Gal}(L/K)| = [L:K]$ we infer that $r=d$. In other words, f has precisely d distinct zeros, as asserted.

□

We shall now be interested in those finite extensions L/K for which $|\mathrm{Gal}(L/K)| = [L:K]$. We call these extensions *Galois extensions*. One of the reasons to be interested in them is the following property.

Proposition 8.3.2 *Let L/K be a finite extension and suppose that $|\mathrm{Gal}(L/K)| = [L:K]$. Let $\alpha \in L$ be such that $\sigma(\alpha)=\alpha$ for all $\sigma \in \mathrm{Gal}(L/K)$. Then $\alpha \in K$.*
Roughly speaking, an element of a Galois extension which is fixed under the Galois group, belongs to the ground field.

Proof: It is given that $\sigma(\alpha)=\alpha$ for all $\sigma \in \mathrm{Gal}(L/K)$. Hence $\mathrm{Gal}(L/K) = \mathrm{Gal}(L/K(\alpha))$ and we see that

$$[L:K] = |\mathrm{Gal}(L/K)| = |\mathrm{Gal}(L/K(\alpha))| \le [L:K(\alpha)].$$

Since also $[L:K]=[L:K(\alpha)][K(\alpha):K]$ we conclude that $[K(\alpha):K]=1$, hence $\alpha \in K$.

□

Computation of the Galois group of Galois extensions will be our main concern. Of course we need some easy criteria by which we can recognize Galois extensions.
Before we continue we need two more definitions.

Definition 8.3.3 *Let L/K be a finite extension. Then L is called normal over K if every irreducible polynomial in $K[X]$ with a zero in L has all of its zeros in L.*

Definition 8.3.4 *Let L/K be a finite extension. An element $\alpha \in L$ is called separable over K if its minimal polynomial f has no multiple zeros in a splitting field of f. The element α is called inseparable over K if f has α as a multiple zero.*
The extension L/K is called separable if all of its elements are separable over K.

An example of an inseparable extension is given by Example 8.2.12. There the element $t^{1/p}$ is inseparable over $\mathbb{F}_p(t)$. Notice that in $\mathbb{F}_p(t)$ the derivative of $X^p - t$ with respect to X is identically 0. This is the way to recognize inseparable elements.

Proposition 8.3.5 *Let L/K be a finite extension. Let $\alpha \in L$ and let f be its minimal polynomial over K. Then α is inseparable over K if and only if K has positive characteristic p and the derivative of f is identically zero. The latter condition is equivalent to $f \in K[X^p]$.*

Proof: Suppose f has a multiple zero in some splitting field of f over K. Then f and its derivative f' have a common divisor. A fortiori they should have a common divisor in $K[X]$. Since f is irreducible in $K[X]$ this implies that f divides f'. Since $\deg(f') < \deg(f)$ this is only possible if $f' = 0$. Let a_iX^i be any non-trivial term in f. Its derivative is ia_iX^{i-1}. This can only be zero if i is divisible by p. Hence $X^i = (X^p)^{i/p}$. We conclude that $f \in K[X^p]$.
Suppose conversely that $f(X) = g(X^p) \in K[X^p]$. Then $f(X) = f(X) - f(\alpha) = g(X^p) - g(\alpha^p) = (X^p - \alpha^p)h(X^p)$ for some $h \in L[X]$. Hence $f(X) = (X - \alpha)^p h(X^p)$ and f has a multiple zero α.

□

We see that separability questions only play a role if the ground field has positive characteristic. In characteristic zero, for example when $K = \mathbb{Q}$ we need not worry about (in)separability.
The main result of this section will be the following Theorem.

Theorem 8.3.6 *Let L/K be a finite extension. Then the following statements are equivalent,*

1. *$|\mathrm{Gal}(L/K)| = [L : K]$ (i.e. L/K is a Galois extension).*

2. *L/K is a normal and separable extension.*

3. *L is a splitting field over K of a polynomial $f \in K[X]$ with distinct zeros. (We call such f separable).*

Note that the implication 8.3.6(1)⇒ 8.3.6(2) follows immediately from the second part of Theorem 8.3.1.
The implication 8.3.6(2)⇒ 8.3.6(3) is also straightforward. Suppose that $L = K(\alpha_1, \ldots, \alpha_r)$. Let $g_i \in K[X]$ be the minimal polynomial of α_i for $i = 1, \ldots, r$. By the normality assumption all zeros of every g_i are in L and moreover, by the separability condition, these zeros are distinct for each g_i. Let now f be the product over the distinct elements in the set $\{g_1, g_2, \ldots, g_r\}$. Then f has distinct zeros and L is its splitting field.
The implication 8.3.6(3)⇒ 8.3.6(1) is the hardest part of the proof of Theorem 8.3.6 and the remainder of this Section will be devoted to it.
The fundamental tool will be the following Proposition.

Proposition 8.3.7 *Let K and K' be fields which are isomorphic via an isomorphism $\sigma : K \to K'$. Let L/K and L'/K' be two finite extensions. To any $p \in K[X]$ given by $p(X) = \sum_i p_i X^i$ we associate the polynomial $p^\sigma \in K'[X]$ given by $p^\sigma(X) = \sum_i \sigma(p_i)X^i$.*
Let $\alpha \in L$ and let f be the minimal polynomial of α over K. Suppose that L' contains a zero α' of f^σ. Then there is an isomorphism $\tau : K(\alpha) \to K'(\alpha')$ given by $\tau(\alpha) = \alpha'$ and $\tau(x) = \sigma(x)$ for all $x \in K$. In other words, τ is an extension of the isomorphism $\sigma : K \to K'$ to $K(\alpha)$.

Proof: By Theorem 6.2.6 we know that $K(\alpha) \cong K[X]/(f)$ and $K(\alpha') \cong K'[X]/(f^\sigma)$. We now show that $K[X]/(f) \cong K'[X]/(f^\sigma)$. Consider the homomorphism $\phi : K[X] \to K'[X]/(f^\sigma)$ given by $p \mapsto p^\sigma (\text{mod } f^\sigma)$. Notice that

$$p \in \ker(\phi) \iff f^\sigma | p^\sigma \iff f|p \iff p \in (f).$$

Hence $\ker(\phi) = (f)$ and via the isomorphism theorem we get $K[X]/(f) \cong K'[X]/(f^\sigma)$. Notice also that ϕ restricted to K is simply the embedding σ.
The resulting isomorphism is now given by $x \mapsto \sigma(x)$ for all $x \in K$ and $\alpha \mapsto X(\text{mod } f) \mapsto X(\text{mod } f^\sigma) \mapsto \alpha'$. This is precisely our desired map τ.

□

An important application is the following Proposition.

Proposition 8.3.8 *Let K, K' be two fields isomorphic via $\sigma : K \to K'$. Let $f \in K[X]$. Let L be a splitting field of f over K and L' a splitting field of f^σ over K'. Then the isomorphism σ can be extended to an isomorphism $\tau : L \to L'$ (in other words $\tau|_K = \sigma$).*

Proof: We proceed by induction on $[L : K]$. When $[L : K] = 1$ our statement is clear, we then have $L = K$ and $L' = K'$. Let $n > 1$ and suppose our Proposition is proven for all L/K with $[L : K] < n$. Suppose $[L : K] = n$.
Let g be an irreducible factor of f of degree > 1 and $\alpha \in L$ a zero of g. Let $\alpha' \in L'$ be a zero of g^σ. By Proposition 8.3.7 there exists an isomorphism $\rho : K(\alpha) \to K'(\alpha')$ such that $\rho|_K = \sigma$.

The degree of $[L : K(\alpha)]$ is less than n. Furthermore, L is a splitting field of f over $K(\alpha)$ and L' is a splitting field of f^σ over $K'(\alpha')$. Hence our induction hypothesis applies and we can extend ρ to an isomorphism $\tau : L \to L'$ such that $\tau|_L = \rho$ and, a fortiori, $\tau|_K = \sigma$.

□

When we take $K' = K$ and for σ the identity map we immediately get the following Corollary.

Corollary 8.3.9 *Let K be a field, $f \in K[X]$ a non-constant polynomial and L, L' two splitting fields of f over K. Then L, L' are K-isomorphic.*

Hence the splitting field of a polynomial over K is uniquely determined up to K-isomorphism. From now on we could also speak of *the* splitting field of a polynomial if one wants.

Proof: (of 8.3.6(3)⇒ 8.3.6(1)). We again proceed by induction on $[L : K]$. Again the case $[L : K] = 1$ is clear. Let $n > 1$ and suppose our statement is proven for all extensions of degree $< n$. Now assume that $[L : K] = n$. Let g be an irreducible factor of f of degree $d > 1$. Then by separability of f the polynomial g has d distinct zeros in L which we denote by $\alpha^{(1)}, \ldots, \alpha^{(d)}$. We abbreviate $\alpha^{(1)} = \alpha$. According to Proposition 8.3.7 with $K' = K$ there is a K-isomorphism $\sigma_i : K(\alpha) \to K(\alpha^{(i)})$ such that $\sigma_i(\alpha) = \alpha^{(i)}$ for $i = 1, \ldots, r$. According to Proposition 8.3.8 with $L' = L$ we can extend these isomorphisms to isomorphisms $\tau_i : L \to L$ such that $\tau_i|_K(\alpha) = \sigma_i$ for $i = 1, \ldots, d$. And a fortiori, since the σ_i are K-isomorphisms, the τ_i are K-isomorphisms. Hence $\tau_1, \ldots, \tau_d \in \mathrm{Gal}(L/K)$.

Note that L is a splitting field of f over $K(\alpha)$. Hence our induction hypothesis applies and we find that $|\mathrm{Gal}(L/K(\alpha))| = [L : K(\alpha)]$. Denote $H = \mathrm{Gal}(L/K(\alpha))$. Of course H is a subgroup of $\mathrm{Gal}(L/K)$. We assert that all elements $\tau_i h$ with $i = 1, \ldots, d$ and $h \in H$ are distinct. Suppose two of them are equal, say $\tau_i h = \tau_j h'$. Apply these two elements to α. We then get $\tau_i h(\alpha) = \tau_i(\alpha) = \alpha^{(i)}$. Similarly $\tau_j h'(\alpha) = \alpha^{(j)}$. Hence $\alpha^{(i)} = \alpha^{(j)}$ and so $i = j$. Consequently $h = h'$.

Hence we have found that

$$|\mathrm{Gal}(L/K)| \geq d|H| = [K(\alpha) : K][L : K(\alpha)] = [L : K].$$

Since also $|\mathrm{Gal}(L/K)| \leq [L : K]$ our desired equality follows.

□

An interesting Corollary of Theorem 8.3.6 is the equivalence of 8.3.6(2) and 8.3.6(3) which we explicitly state here.

Corollary 8.3.10 *Let K be a field. An extension of K is normal and separable if and only if it is a splitting field of a separable polynomial.*

It turns out, but we will not prove it here, that we can toss out the word separable.

Theorem 8.3.11 *Let K be a field. An extension of K is normal if and only if it is a splitting field of a polynomial.*

In particular, in a splitting field L/K every polynomial in $K[X]$, which has a zero in L, factors completely into linear factors in $L[X]$.
In a similar vein we can characterize separable extensions.

Theorem 8.3.12 *Let K be a field. A finite extension L of K is separable if and only if $L = K(\alpha_1, \ldots, \alpha_n)$ with α_i separable over K for $i = 1, 2, \ldots, n$.*

8.4 Exercises

1. Determine a splitting field of X^3-7 in $\mathbb{C}$, determine all of its $\mathbb{Q}$-automorphisms and all $\mathbb{Q}$-embeddings of the root extension in the splitting field.

2. Let L be a splitting field of the polynomial f over K and $f = \prod_{i=1}^{n} (X - \alpha_i)$. Prove: $L = K(\alpha_1, ..., \alpha_{n-1})$ (so one alpha less!).

3. Let ζ be a zero of $f = X^4+X^3+X^2+X+1$. Prove that $\zeta^5 = 1$ and that ζ^2, ζ^3 and ζ^4 are the other zeros of f in $\mathbb{C}$. Prove that $\mathbb{Q}(\zeta)$ is a splitting field of f over $\mathbb{Q}$.

4. Show that $\mathbb{Q}(\sqrt[4]{2})$ and $\mathbb{Q}(i\sqrt[4]{2})$ are two $\mathbb{Q}$-isomorphic fields.

5. Indicate whether the following fields are splitting fields or not (here t is transcendental over $\mathbb{F}_3$):

   (a) $\mathbb{F}_3(t)(\sqrt[3]{t})$ over $\mathbb{F}_3(t)$?

   (b) $\mathbb{F}_3(t)(\sqrt[4]{t})$ over $\mathbb{F}_3(t)$?

6. Give an example of a tower of field extensions $F \subset K \subseteq E$ where K/F and E/K are normal, but E/F is not normal.

7. Are the following polynomials separable or not?

   (a) $x^3 + x^2 - x - 1$ over $\mathbb{Q}$?

   (b) $x^4 + x^2 + 1$ over $\mathbb{F}_2$?

   (c) $x^{10} + x^5 + 3$ over $\mathbb{F}_3$?

   (d) $x^{10} + 4t$ over $\mathbb{F}_5(t)$ where t is transcendental over $\mathbb{F}_5$?

Chapter 9

The Main Theorem of Galois theory

9.1 Fixed fields

The Main Theorem (9.2.1) of Galois theory establishes a one to one correspondence between the subgroups of $\mathrm{Gal}(L/K)$ of a Galois extension L/K and the intermediate fields $K \subset M \subset L$. For its proof we need two important ingredients, namely the theorems of Dedekind and Artin.

Theorem 9.1.1 (Dedekind) *Let L be a field and $\sigma_1, \ldots, \sigma_n$ distinct elements of $\mathrm{Aut}(L)$. Suppose that there exist $a_1, \ldots, a_n \in L$ such that*

$$a_1\sigma_1(x) + a_2\sigma_2(x) + \cdots + a_n\sigma_n(x) = 0$$

for all $x \in L$. Then $a_1 = a_2 = \cdots = a_n = 0$.

Proof: We proceed by induction on n. When $n = 1$ we see that $a_1\sigma_1(x) = 0$ for all $x \in L$, in particular for $x = 1$. Hence $a_1 = 0$.
Let $n > 1$ and suppose our theorem is proved for every $n-1$-tuple of elements in $\mathrm{Aut}(L)$. Choose an index i with $1 \le i < n$. Since σ_i and σ_n are distinct functions, there exists $\xi \in L$ such that $\sigma_i(\xi) \ne \sigma_n(\xi)$. From $a_1\sigma_1(x) + \cdots + a_n\sigma_n(x) = 0$ for all $x \in L$ it follows that $a_1\sigma_1(\xi x) + \cdots + a_n\sigma_n(\xi x) = 0$ for all $x \in L$. We get $a_1\sigma_1(\xi)\sigma_1(x) + \cdots + a_n\sigma_n(\xi)\sigma_n(x) = 0$ for all $x \in L$. Subtract from this $\sigma_n(\xi)$ times the original relation. We get

$$a_1(\sigma_1(\xi) - \sigma_n(\xi))\sigma_1(x) + \cdots + a_{n-1}(\sigma_{n-1}(\xi) - \sigma_n(\xi))\sigma_{n-1}(x) = 0$$

for all $x \in L$. According to our induction hypothesis all coefficients are zero, in particular $a_i(\sigma_i(\xi) - \sigma_n(\xi)) = 0$. Since $\sigma_i(\xi) \ne \sigma_n(\xi)$ this implies $a_i = 0$. This argument holds for all i with $i < n$ and we are left with $a_n\sigma_n(x) = 0$ for all $x \in L$ which in its turn implies that $a_n = 0$.

□

The reason we have proved Dedekind's theorem is the following Corollary, which is the only thing we will need from Dedekind's theorem.

Corollary 9.1.2 *Let L be a field and $\sigma_1, \ldots, \sigma_n$ distinct automorphisms of L. Then there exists $\xi \in L$ such that $\sigma_1(\xi) + \cdots + \sigma_n(\xi) \neq 0$.*

Notice that if L has characteristic zero, the proof of Corollary 9.1.2 is trivial, without need for Dedekind's Theorem. One simply takes $\xi = 1$ and notes that $\sigma_1(1) + \cdots + \sigma_n(1) = n \neq 0$. When L has a characteristic dividing n however, we get $\sigma_1(1) + \cdots + \sigma_n(1) = n = 0$. So we have to look for another element ξ and we need to take recourse to Dedekind's Theorem for that.
Corollary 9.1.2 plays an important role in the proof of the following crucial Theorem on fixed fields. Let L be a field and H a subgroup of the automorphism group of L, then we denote the *fixed field* under H by

$$L^H = \{x \in L |\ \sigma(x) = x \text{ for all } \sigma \in H\}.$$

Theorem 9.1.3 (E.Artin) *Let L be a field and H a finite subgroup of the automorphism group of L. Let L^H be its fixed field. Then L/L^H is a finite Galois extension with Galois group H. In particular $|H| = [L : L^H]$.*

Proof: Let $H = \{\sigma_1, \ldots, \sigma_h\}$. For any $x \in L$ we define $tr(x) = \sigma_1(x) + \cdots + \sigma_h(x)$. Note that $tr(x)$ is fixed under the action of each $\sigma_j \in H$ and therefore $tr(x) \in L^H$ for every $x \in L$.
Now suppose that L contains $h + 1$ elements $\alpha_1, \ldots, \alpha_{h+1}$ which are L^H-linear independent. We derive a contradiction as follows. Consider the homogeneous system of h equations

$$\begin{aligned} x_1\sigma_1(\alpha_1) + \cdots + x_{h+1}\sigma_1(\alpha_{h+1}) &= 0 \\ &\vdots \\ x_1\sigma_h(\alpha_1) + \cdots + x_{h+1}\sigma_h(\alpha_{h+1}) &= 0 \end{aligned}$$

in the $h + 1$ unknowns $x_i \in L$. According to linear algebra there must be a non-trivial solution. Choose such a solution. Without loss of generality we can assume that $x_1 \neq 0$ and after multiplication by a suitable factor if necessary we may assume that $\sigma_1(x_1) + \cdots + \sigma_h(x_1) \neq 0$. The possibility of this follows from Corollary 9.1.2. To the i-th equation above we apply σ_i^{-1} for $i = 1, 2, \ldots, h$. After renumbering we obtain the system of equalities

$$\begin{aligned} \sigma_1(x_1)\alpha_1 + \cdots + \sigma_1(x_{h+1})\alpha_{h+1} &= 0 \\ &\vdots \\ \sigma_h(x_1)\alpha_1 + \cdots + \sigma_h(x_{h+1})\alpha_{h+1} &= 0 \end{aligned}$$

Addition of these equalities yields

$$tr(x_1)\alpha_1 + \cdots + tr(x_{h+1})\alpha_{h+1} = 0.$$

We know that $tr(x_i) \in L^H$ for every i and $\alpha_1, \ldots, \alpha_{h+1}$ are L^H-linearly independent. Hence $tr(x_i) = 0$ for all i. But this contradicts our carefully doctored provision $tr(x_1) \neq 0$.

Hence $[L : L^H] \leq h = |H|$. On the other hand $|H| \leq [L : L^H]$ because $H \subset \mathrm{Gal}(L/L^H)$. Thus we conclude that $|H| = [L : L^H]$. In other words, L/L^H is a finite Galois extension with Galois group H.

□

9.2 Main Theorem

Let L/K be a Galois extension. The main theorem of Galois theory gives us a bijection between the subgroups of $\mathrm{Gal}(L/K)$ and the intermediate fields M such that $K \subset M \subset L$.

Note that for such an intermediate field M the extension L/M is again a Galois extension. This is clear, by Theorem 8.3.6 L is the splitting field of a separable polynomial f over K. But f can also be considered as polynomial in $M[X]$, so L is also a splitting field of f over M. The group $\mathrm{Gal}(L/M)$ is a subgroup of $\mathrm{Gal}(L/K)$. By Theorem 8.3.6 we have $|\mathrm{Gal}(L/M)| = [L : M]$.

Note that in general M is *not* a Galois extension of K. For example, the splitting field of $X^3 - 2$ over $\mathbb{Q}$ is given by $\mathbb{Q}(\sqrt[3]{2}, \omega)$ with $\omega = e^{2\pi i/3}$. But the subfield $\mathbb{Q}(\sqrt[3]{2})$ is not Galois over $\mathbb{Q}$ since it contains only one zero of $X^3 - 2$.

Theorem 9.2.1 *Let L/K be a Galois extension. Then the map from intermediate fields M of L/K and subgroups of* $\mathrm{Gal}(L/K)$ *given by $M \mapsto$* $\mathrm{Gal}(L/M)$ *is a bijection.*

The inverse map is given by $H \mapsto L^H$.

Proof: Denote the map $M \mapsto \mathrm{Gal}(L/M)$ by α and denote $H \mapsto L^H$ by β. Notice that

$$\alpha \circ \beta : \; H \mapsto L^H \mapsto \mathrm{Gal}(L/L^H).$$

By Artin's Theorem 9.1.3 the latter group is isomorphic to H. Hence $\alpha \circ \beta$ is the identity map from the set of subgroups of $\mathrm{Gal}(L/K)$ to itself. Notice also that

$$\beta \circ \alpha : \; M \mapsto \mathrm{Gal}(L/M) \mapsto L^{\mathrm{Gal}(L/M)}.$$

Since L/M is a Galois extension, the latter field must be M. So $\beta \circ \alpha$ is the identity map from the set of intermediate fields to itself. In general, if we have two maps $\alpha : A \to B$ and $\beta : B \to A$ between finite sets A, B such that both $\alpha \circ \beta$ and $\beta \circ \alpha$ are the identity map, it follows that α and β are injective. Hence $|A| \leq |B|$ and $|B| \leq |A|$. So, $|A| = |B|$ and our theorem is proved.

□

We call the bijection in Theorem 9.2.1 the *Galois correspondence*. It turns out that under the Galois correspondence normal extensions of K correspond to normal subgroups of L. We can state this slightly more general as follows.

Theorem 9.2.2 *Let L/K be a finite Galois extension. Then an intermediate extension M with $K \subset M \subset L$ is normal over K if and only if* $\mathrm{Gal}(L/M)$ *is a normal subgroup of* $\mathrm{Gal}(L/K)$. *Moreover, we have the isomorphism*

$$\mathrm{Gal}(M/K) \cong \mathrm{Gal}(L/K)/\mathrm{Gal}(L/M).$$

To prove this theorem we need a preliminary observation.

Lemma 9.2.3 *Let L/K be a finite Galois extension and M an intermediate field, in other words $K \subset M \subset L$. Then M is normal over K if and only if $\sigma(M) = M$ for every $\sigma \in \mathrm{Gal}(L/K)$.*

Proof: Suppose that the intermediate field M is normal over K and let $\alpha \in M$. Let g be the minimal polynomial of α over K. Then $\sigma(\alpha)$ is also a zero of g for any $\sigma \in \mathrm{Gal}(L/K)$. By the normality of M this zero is also contained in M. Hence σ maps elements of M to itself.
Conversely suppose that $\sigma(M) = M$ for every $\sigma \in \mathrm{Gal}(L/K)$. Then, if an irreducible polynomial g has a zero $\alpha \in M$, all of its zeros are images of α under $\mathrm{Gal}(L/K)$ and hence all zeros are contained in M.

□

We are now ready to prove Theorem 9.2.2.
Proof: Suppose that M is normal over K. We have just seen that this implies that $\sigma(M) = M$ for every $\sigma \in \mathrm{Gal}(L/K)$. So we can consider the restriction of σ to M, denoted by $\sigma|_M$, for every $\sigma \in \mathrm{Gal}(L/K)$. Clearly the map $\pi : \mathrm{Gal}(L/K) \to \mathrm{Gal}(M/K)$ given by $\sigma \mapsto \sigma|_M$ is a group homomorphism. It is surjective since any element of $\mathrm{Gal}(M/K)$ can be extended to an element of $\mathrm{Gal}(L/K)$ by Proposition 8.3.8. Its kernel is $\mathrm{Gal}(L/M)$ and so this is a normal subgroup, as asserted. Moreover, by the isomorphsim theorem for groups we get

$$\mathrm{Gal}(M/K) \cong \mathrm{Gal}(L/K)/\mathrm{Gal}(L/M).$$

Suppose conversely that H is a normal subgroup of $\mathrm{Gal}(L/K)$. Let $\alpha \in L^H$ and $\sigma \in \mathrm{Gal}(L/K)$. Then, for any $h \in H$ there is $h' \in H$ such that $h\sigma = \sigma h'$. Hence $h(\sigma(\alpha)) = \sigma(h'(\alpha) = \sigma(\alpha)$. So $\sigma(\alpha)$ is fixed under all $h \in H$ and hence $\sigma(\alpha) \in L^H$. We conclude that $\sigma(L^H) = L^H$ and by the above lemma L^H is a normal extension of K.

□

9.3 Examples of Galois correspondences

Example 9.3.1. Let $L/\mathbb{Q}$ be the splitting field of $X^3 - 2$ over $\mathbb{Q}$. We are asked to determine $\mathrm{Gal}(L/\mathbb{Q})$. The first step is to determine the degree $[L : \mathbb{Q}]$. We have seen at earlier occasions that $L = \mathbb{Q}(\sqrt[3]{2}, \omega)$ where $\omega = e^{2\pi i/3}$ is a cube root of unity. The degrees over $\mathbb{Q}$ of $\sqrt[3]{2}$ and ω are 3 and 2 respectively. Hence $[L : \mathbb{Q}] = 2 \cdot 3 = 6$. By Theorem 8.3.6 we now know that $|\mathrm{Gal}(L/\mathbb{Q})| = 6$. An element of the Galois group maps $\sqrt[3]{2}$ to a zero of $X^3 - 2$, so $\sqrt[3]{2}, \omega\sqrt[3]{2}$ or $\omega^2\sqrt[3]{2}$. The element ω is mapped to ω or ω^2. In principle there are 6 possibilities. Since we know that the Galois group has precisely 6 elements, every possibility should occur. Let us consider two special elements $\sigma, \tau \in \mathrm{Gal}(L/\mathbb{Q})$ defined by

$$\sigma(\sqrt[3]{2}) = \omega\sqrt[3]{2}, \quad \sigma(\omega) = \omega, \quad \tau(\sqrt[3]{2}) = \sqrt[3]{2}, \quad \tau(\omega) = \omega^2.$$

Notice that

$$\begin{aligned}\sigma\tau(\sqrt[3]{2}) &= \sigma(\sqrt[3]{2}) = \omega\sqrt[3]{2}\\ \sigma\tau(\omega) &= \sigma(\omega^2) = \omega^2.\end{aligned}$$

Similarly,

$$\begin{aligned}\tau\sigma^2(\sqrt[3]{2}) &= \tau\sigma(\omega\sqrt[3]{2}) = \tau(\omega^2\sqrt[3]{2}) = \omega\sqrt[3]{2}\\ \tau\sigma^2(\omega) &= \tau\sigma(\omega) = \tau(\omega) = \omega^2.\end{aligned}$$

From these equalities we conclude that $\sigma\circ\tau = \tau\circ\sigma^2$. Similarly we can check that $\sigma^3 = \mathrm{id}$ and $\tau^2 = \mathrm{id}$ (please verify). These relations are precisely the defining relations for the symmetric group S_3. Hence $\mathrm{Gal}(L/\mathbb{Q}) \cong S_3$. The element σ plays the role of a cyclic permutation of order 3. In fact it permutes the zeros of $X^3 - 2$ in a cyclic way (check!). The element τ fixes $\sqrt[3]{2}$ and it interchanges the other two zeros of $X^3 - 2$. It coincides with complex conjugation on L if L is considered subfield of $\mathbb{C}$.

The group S_3 has 6 subgroups which we tabulate together with their invariant fields:

Group H	Invariant field L^H
$\{\mathrm{id}\}$	L
S_3	$\mathbb{Q}$
$A_3 = \{\mathrm{id}, \sigma, \sigma^2\}$	$\mathbb{Q}(\omega) = \mathbb{Q}(\sqrt{-3})$
$\{\mathrm{id}, \tau\}$	$\mathbb{Q}(\sqrt[3]{2})$
$\{\mathrm{id}, \sigma\tau\}$	$\mathbb{Q}(\omega^2\sqrt[3]{2})$
$\{\mathrm{id}, \sigma^2\tau\}$	$\mathbb{Q}(\omega\sqrt[3]{2})$

The subgroup A_3 is a normal subgroup of S_3 and its fixed field $\mathbb{Q}(\sqrt{-3})$ is normal over $\mathbb{Q}$.

$\diamondsuit$

Example 9.3.2. We compute the Galois group of the splitting field L of $X^4 - 4X^2 + 1$ over $\mathbb{Q}$ and determine the subgroups and the Galois correspondence.

First we need to find the degree of L. Notice that $X^4 - 4X^2 + 1 = (X^2 - 2)^2 - 3$. Hence the four zeros are given by $\pm\sqrt{2 \pm \sqrt{3}}$. The field L contains $\sqrt{2 + \sqrt{3}}$. Write $\alpha = \sqrt{2 + \sqrt{3}}$. From the very special form of the equation we easily see that $-\alpha, 1/\alpha, -1/\alpha$ are also zeros. Hence we conclude that $L = \mathbb{Q}(\sqrt{2 + \sqrt{3}}) = \mathbb{Q}(\alpha)$. So L is a simple extension and we can apply Proposition 8.2.8. Since $[L : \mathbb{Q}] = 4$ the Galois group has four elements. So $\mathrm{Gal}(L/\mathbb{Q})$ is either cyclic or it is Klein's fourgroup. Let us define the element σ by $\sigma(\alpha) = -\alpha$ and the element τ by $\tau(\alpha) = 1/\alpha$. Then note that $\sigma\tau(\alpha) = -1/\alpha$. In addition $\sigma\tau = \tau\sigma$ and $\sigma^2 = \tau^2 = \mathrm{id}$. So $\mathrm{Gal}(L/\mathbb{Q})$ is Klein's fourgroup, denoted by V_4.

Let us determine the fixed field under the subgroups of order 2. For example, the element $\alpha + 1/\alpha$ is fixed under τ. Notice that

$$(\alpha + 1/\alpha)^2 = \alpha^2 + 2 + 1/\alpha^2 = 2 + \sqrt{3} + 2 + 2 - \sqrt{3} = 6.$$

So, perhaps surprisingly, we discovered that $\sqrt{6} \in \mathbb{Q}(\alpha)$. Similarly $\alpha - 1/\alpha$ is fixed under $\sigma\tau$. Notice that

$$(\alpha - 1/\alpha)^2 = \alpha^2 - 2 + 1/\alpha^2 = 2\sqrt{3} - 2 + 2 - \sqrt{3} = 2.$$

Hence $\sqrt{2} \in \mathbb{Q}(\alpha)$. We now tabulate the subgroups of $\mathrm{Gal}(L/\mathbb{Q})$ and their fixed fields

Group H	Invariant field L^H
$\{\mathrm{id}\}$	L
V_4	$\mathbb{Q}$
$\{\mathrm{id}, \sigma\}$	$\mathbb{Q}(\sqrt{3})$
$\{\mathrm{id}, \tau\}$	$\mathbb{Q}(\sqrt{6})$
$\{\mathrm{id}, \sigma\tau\}$	$\mathbb{Q}(\sqrt{2})$

Note that all subgroups and all fixed fields are normal.

♢

Example 9.3.3. We compute the Galois group of the splitting field L of $X^4 - 2X^2 - 2$ over $\mathbb{Q}$ and determine the subgroups and the Galois correspondence. The polynomial $X^4 - 2X^2 - 2$ is an Eisenstein polynomial with respect to $p = 2$, hence irreducible in $\mathbb{Q}[X]$. The zeros are given by $\pm\sqrt{1 \pm \sqrt{3}}$. The element $\sqrt{1 + \sqrt{3}}$ has degree 4 over $\mathbb{Q}$ but this time $\mathbb{Q}(\sqrt{1 + \sqrt{3}})$ is not a splitting field. For example the element $\sqrt{1 + \sqrt{3}}\sqrt{1 - \sqrt{3}} = \sqrt{1^2 - 3} = \sqrt{-2}$ should be in L. Clearly $\sqrt{-2} \notin \mathbb{Q}(\sqrt{1 + \sqrt{3}})$ since the latter field is a subfield of $\mathbb{R}$ and $\sqrt{-2}$ is not a real number.

Write $\alpha = \sqrt{1 + \sqrt{3}}$. The field $\mathbb{Q}(\sqrt{-2}, \alpha)$ has degree 8 over $\mathbb{Q}$. The zeros of $X^4 - 2X^2 - 2$ are $\alpha, -\alpha, \sqrt{-2}/\alpha, -\sqrt{-2}/\alpha$. So $L = \mathbb{Q}(\sqrt{-2}, \alpha)$ and $\mathrm{Gal}(L/\mathbb{Q})$ has order 8. Any element of the Galois group maps α to one of the four zeros of $X^4 - 2X^2 - 2$ and $\sqrt{-2}$ to $\pm\sqrt{-2}$. These are 8 possibilities and each actually occurs. Define the elements σ, τ of the Galois group by

$$\sigma(\alpha) = \sqrt{-2}/\alpha, \quad \sigma(\sqrt{-2}) = -\sqrt{-2}$$

and

$$\tau(\alpha) = \alpha, \quad \tau(\sqrt{-2}) = -\sqrt{-2}.$$

We can check that

$$\sigma(\alpha) = \sqrt{-2}/\alpha, \quad \sigma^2(\alpha) = -\alpha, \quad \sigma^3(\alpha) = -\sqrt{-2}/\alpha, \quad \sigma^4(\alpha) = \alpha.$$

Hence σ has order 4. It is easily seen that $\tau^2 = \mathrm{id}$. Moreover, we verify that $\sigma\tau = \tau\sigma^3 = \tau\sigma^{-1}$. The group generated by σ, τ, together with the given relations is precisely the symmetry group of the square, which is a group of order 8. We denote it by D_4 (dihedral group of order 8). The element σ permutes the zeros of $X^4 - 2X^2 - 2$ cyclically and τ interchanges the zeros $\sqrt{-2}/\alpha$ and $-\sqrt{-2}/\alpha$ and leaves the other two fixed.

To determine the Galois correspondence we first observe that L contains three quadratic fields, namely $\mathbb{Q}(\sqrt{3})$ (since $\alpha^2 = 1 + \sqrt{3}$), $\mathbb{Q}(\sqrt{-2})$ and $\mathbb{Q}(\sqrt{-6})$. We now tabulate the subgroups of the Galois group together with the fixed fields.

Group H	Invariant field L^H
$\{\mathrm{id}\}$	L
D_4	$\mathbb{Q}$
$C_4 = \{\mathrm{id}, \sigma, \sigma^2, \sigma^3\}$	$\mathbb{Q}(\sqrt{-6})$
$\{\mathrm{id}, \tau, \sigma^2, \tau\sigma^2\}$	$\mathbb{Q}(\sqrt{3})$
$\{\mathrm{id}, \tau\sigma, \sigma^2, \tau\sigma^3\}$	$\mathbb{Q}(\sqrt{-2})$
$\{\mathrm{id}, \sigma^2\}$	$\mathbb{Q}(\sqrt{-2}, \sqrt{3})$
$\{\mathrm{id}, \tau\}$	$\mathbb{Q}(\alpha)$
$\{\mathrm{id}, \tau\sigma\}$	$\mathbb{Q}(\alpha - \sqrt{-2}/\alpha)$
$\{\mathrm{id}, \tau\sigma^2\}$	$\mathbb{Q}(\sqrt{-2}/\alpha)$
$\{\mathrm{id}, \tau\sigma^3\}$	$\mathbb{Q}(\alpha + \sqrt{-2}/\alpha)$

To check the fixed fields notice that $\sigma(\alpha^2) = -2/\alpha^2$. Hence $\sigma(1 + \sqrt{3}) = -2/(1+\sqrt{3}) = 1-\sqrt{3}$. And so $\sigma(\sqrt{3}) = -\sqrt{3}$. By definition $\sigma(\sqrt{-2}) = -\sqrt{-2}$. Also by definition $\tau(\sqrt{-2}) = -\sqrt{-2}$ but $\tau(\sqrt{3}) = \sqrt{3}$. From this we conclude that $\sqrt{3}$ is fixed by σ^2 and τ, $\sqrt{-2}$ is fixed by $\tau\sigma$ and σ^2 and $\sqrt{-6}$ is fixed by σ. This explains the Galois correspondence for the subgroups of order four. In the case of subgroups of order two, we have indicated fourth degree extensions of $\mathbb{Q}$ whose generating elements are fixed under the corresponding subgroup, which is straightforward to verify.

◊

Example 9.3.4. Let $p \in \mathbb{Z}_{>0}$ be a prime and L the splitting field of $X^p - 1$ over $\mathbb{Q}$. Let $\zeta = e^{2\pi i/p}$. Then the zeros of $X^p - 1$ are given by ζ^k for $k = 0, 1, \ldots, p-1$. So $\mathbb{Q}(\zeta)$ is the splitting field of $X^p - 1$. We have seen earlier that ζ is a zero of the irreducible polynomial $\phi_p(X) = X^{p-1} + \cdots + X^2 + X + 1$. Hence ζ has degree $p - 1$ over $\mathbb{Q}$. The other zeros of $\phi_p(X)$ are ζ^k with $k = 1, 2, \ldots, p - 1$. The Galois group is given by $\zeta \mapsto \zeta^k$ for any $k \in (\mathbb{Z}/p\mathbb{Z})^*$. Denote the latter element by σ_k. Then it is easily verified that $\sigma_{kl} = \sigma_k\sigma_l = \sigma_l\sigma_k$. Hence $\mathrm{Gal}(L/\mathbb{Q})$ is an abelian group of order $p - 1$ which is isomorphic to $(\mathbb{Z}/p\mathbb{Z})^*$. A Theorem of Gauss states that such groups are cyclic, but we will not prove it here.

◊

Example 9.3.5. We specialise the previous example to the case $p = 17$. So let $\zeta = e^{2\pi i/17}$. Consider the element $\sigma \in \mathrm{Gal}(\mathbb{Q}(\zeta)/\mathbb{Q})$ given by $\sigma(\zeta) = \zeta^3$. Notice that $\sigma^k(\zeta) = \zeta^{3^k}$. A simple computation shows that modulo 17 we have

$$3 \equiv 3,\ 3^2 \equiv 9,\ 3^3 \equiv 10,\ 3^4 \equiv 13,\ 3^5 \equiv 5,\ 3^6 \equiv 15,\ 3^7 \equiv 11,\ 3^8 \equiv 16$$

$$3^9 \equiv 14,\ 3^{10} \equiv 8,\ 3^{11} \equiv 7,\ 3^{12} \equiv 4,\ 3^{13} \equiv 12,\ 3^{14} \equiv 2,\ 3^{15} \equiv 6,\ 3^{16} \equiv 1.$$

Thus every element of $\mathrm{Gal}(\mathbb{Q}(\zeta)/\mathbb{Q})$ can be written in the form σ^k, so we have a cyclic group of order 16. For every divisor d of the order of a cyclic group there is exactly one subgroup of that order. So in our case we have subgroups of orders $16, 8, 4, 2, 1$ and they are generated by $\sigma, \sigma^2, \sigma^4, \sigma^8, \sigma^{16} = \mathrm{id}$. Via the Galois-correspondence there exist fields $K_1 = \mathbb{Q}, K_2, K_4, K_8, K_{16} = \mathbb{Q}(\zeta)$ such that $K_i \subset K_{2i}$ and $[K_{2i} : K_i] = 2$ for $i = 1, 2, 4, 8$. Hence the number ζ is contructible by Theorem 7.2.10 and so the regular 17-gon is construcible. This was the striking discovery made by Gauss and the first significant progress in classic ruler and straight-edge constructions in about 2000 years!

♢

9.4 Simple extensions

In Proposition 8.2.8 we showed how to compute Gal(L/K) for simple extensions L/K. We noted that this might be an approach to computing the Galois group for every extension, simply by writing the extension as a simple extension. This was the approach which Galois originally took. Although we did not adopt this approach, we complete these remarks here by showing that finite separable extensions are indeed simple.

Theorem 9.4.1 *Let L/K be a finite separable extension. Then*

1. *there are finitely many fields M such that $K \subset M \subset L$.*

2. *there exists $\alpha \in L$ such that $L = K(\alpha)$ (in other words, L is a simple extension of K).*

Proof: Suppose $L = K(\alpha_1, \ldots, \alpha_r)$. Let $g_i \in K[X]$ be the minimal polynomial of α_i for $i = 1, \ldots, r$. Let $f(X)$ be the product of the distinct elements in the set $\{g_1, \ldots, g_r\}$. Let N be the splitting field of f over K. Then N/K is a Galois extension. The number of subgroups of Gal(N/K) is finite and so, by Galois correspondence, there are at most finitely many intermediate fields M with $K \subset M \subset N$. Hence the first part of our Theorem follows.
The proof of the second part only applies when $|K| = \infty$. When K is finite it follows from the theory of finite fields. The proof is by induction on r where $L = K(\alpha_1, \ldots, \alpha_r)$. When $r = 1$ the statement is trivial. Let $r > 1$ and suppose that every finite separable extension generated by $r - 1$ elements is simple. By the induction hypothesis there exists $\beta \in L$ be such that $K(\alpha_1, \ldots, \alpha_{r-1}) = K(\beta)$. We now prove that $L = K(\beta, \alpha_r)$ is simple. Put $\alpha = \alpha_r$. Consider the fields $K(\alpha + c\beta)$ where c runs over K. Since K is infinite and there are at most finitely many subfields of L containing K there exist distinct c, c' such that $K(\alpha + c\beta) = K(\alpha + c'\beta)$. Clearly $\alpha + c\beta \in K(\alpha, \beta)$. On the other hand both $\alpha + c\beta - \alpha - c'\beta = (c - c')\beta \in K(\alpha + c\beta)$ and $c'(\alpha + c\beta) - c'(\alpha + c\beta) = (c' - c)\alpha \in K(\alpha + c\beta)$. Since $c - c' \neq 0$ this implies that $\alpha, \beta \in K(\alpha + c\beta)$. Hence $L = K(\alpha, \beta) = K(\alpha + c\beta)$, so L is a simple extension of K.

□

Remark 9.4.2 *Theorem 9.4.1 need not be true for inseparable extensions. Consider for example the extension $L = \mathbb{F}_p(s^{1/p}, t^{1/p})$ of $K = \mathbb{F}_p(s, t)$ which has degree p^2. Every element of the extension L has degree p, which follows from the observat that $\alpha^p \in K$ for every $\alpha \in L$. So L cannot be a simple extension. Furthermore the elements $as^{1/p} + bt^{1/p}$ with $a, b \in K$ generate an infinite number of distinct extensions. So the number of intermediate fields of L/K is infinite.*

Here we present a second proof of Theorem 9.4.1(2) which does not use the Galois correspondence. Again we assume that K is infinite. As we have seen in the proof of Theorem 9.4.1(2) it suffices to show that any separable extension of the form $K(\alpha,\beta)$ over K is simple. Let $f,g \in K[X]$ be the minimal polynomials of α respectively β over K. We work in a splitting field L of $f(X)g(X)$. Let $\alpha_1,\ldots,\alpha_n$ be the zeros of f and $\beta_1,\ldots,\beta_m$ the zeros of g. Since the extension is separable, the β_i are distinct. Take $\alpha=\alpha_1, \beta=\beta_1$. Since K is infinite there exists $c \in K$ such that c is different from $(\alpha-\alpha_i)/(\beta_k-\beta)$ for all $1 \le i \le n, 1 < k \le m$. Hence $\alpha + c\beta \ne \alpha_i + c\beta_k$ for all $1 \le i \le n, 1 < k \le m$. Let $\delta = \alpha + c\beta$. We assert that $K(\delta) = K(\alpha,\beta)$. To that end we note that $f(\delta - cx)$ and $g(x)$ have the common zero β. By construction it is also the only zero. So the gcd of $f(\delta - cx)$ and $g(x)$ is $x-\beta$. Both polynomials are in $K(\delta)[x]$, hence $\beta \in K(\delta)$. Then also $\alpha = \delta - c\beta \in K(\delta)$. Hence $K(\alpha,\beta) \subset K(\delta)$. Since we already know that $\delta \in K(\alpha,\beta)$ the desired equality $K(\alpha,\beta) = K(\delta)$ follows.

9.5 Exercises

The majority of the problems below have occurred in one examination or another.

1. Let $f(X) = X^6 - 25$ and $G = \mathrm{Gal}(f/\mathbb{Q})$.

   (a) Factor f into irreducible factors over $\mathbb{Q}$.

   (b) Show that the splitting field of f over $\mathbb{Q}$ is given by $L = \mathbb{Q}(\alpha,\omega)$ where $\alpha = \sqrt[3]{5}$ (the positive real root) and $\omega^3 = 1$ with $\omega \ne 1$.

   (c) Prove that $[L:\mathbb{Q}] = 6$.

   (d) Show that there exist elements $\sigma, \tau \in G$ such that

   $$\begin{aligned} \sigma(\alpha) &= \alpha\omega,\ \sigma(\omega) = \omega; \\ \tau(\alpha) &= \alpha,\ \tau(\omega) = \omega^2. \end{aligned}$$

   Show also that there is no element $\nu \in G$ such that $\nu(\alpha) = \omega$?

   (e) Prove that $G \cong D_3$.

   (f) Show that the fixed field of $\langle \sigma\tau \rangle$ equals $M = \mathbb{Q}(\alpha^2\omega)$.

   (g) Determine a primitive element of $L/\mathbb{Q}$.

2. Let α a root of $X^3 + X^2 + 1 \in \mathbb{F}_2[X]$ and $K = \mathbb{F}_2(\alpha)$.

   (a) Show that K is a field with 8 elements. Complete the multiplication table below:

$\cdot$	$\alpha+1$	α^2+1
$\alpha+1$		
α^2+1		

   (b) Suppose that f is an irreducible polynomial in $K[X]$ of degree 4. Let β be a root of f, and L a splitting field of f over K.

i. What is the number of elements of L?
ii. How many intermediate field does the extension L/K have?
iii. Why are all zeros of f of the form β^k for some $k \in \mathbb{Z}_{\geq 0}$?

3. We are given a Galois extension L/K of degree 2002 $(= 2 \cdot 7 \cdot 11 \cdot 13)$. Let n be the number of 13-Sylow subgroups of $\mathrm{Gal}(L/K)$.

   (a) Show that L contains a subfield of degree 154 $(= 2 \cdot 7 \cdot 11)$ over K.

   (b) Let M be a subfield of L of degree 154 over K. Then

   $$M \text{ normaal over } K \iff n = \ldots$$

   Complete the statement and then prove it.

4. Let $f(X) = X^8 - 4 \in \mathbb{Q}[X]$, and let L be a splitting field of f over $\mathbb{Q}$.

   (a) Factor f into irreducible factors over $\mathbb{Q}$.

   (b) Prove that $[L : \mathbb{Q}] = 8$.

   (c) Determine $\mathrm{Gal}(L/\mathbb{Q})$.

   (d) Determine a primitive element for $L/\mathbb{Q}$.

5. Provide an example, or show that it does not exist of each of the following

   (a) A Galois extension $K/\mathbb{Q}$ with cyclic Galois group and two distinct intermediate fields K_1 and K_2 $(\neq K, \neq \mathbb{Q})$ such that $[K_1 : \mathbb{Q}] = [K_2 : \mathbb{Q}]$.

   (b) An irreducible polynomial of degree 6 over $\mathbb{Q}$ with solvable Galois group.

   (c) A polynomial in $\mathbb{F}_7[X]$ which is not separable and which has exactly 3 non-zero coefficients.

   (d) A construction with ruler and compasses of a 10° angle.

6. Let L be a splitting field over $\mathbb{Q}$ of the polynomial $f(X) = (X^3 - 4)(X^2 + 12)$.

   (a) Determine $[L : \mathbb{Q}]$ and $\mathrm{Gal}(L/\mathbb{Q})$.

   (b) How many subfields $(\neq L, \neq \mathbb{Q})$ does L have? Which subfields are normal over $\mathbb{Q}$? Determine a primitive element for each normal extension (this can be done without calculation!).

7. Let K be a splitting field of the polynomial $(X^3 + X - 1)(X^4 + X - 1)$ over $\mathbb{F}_3$.

   (a) How many elements does K contain?

   (b) How many subfields does K have? Give all possible subfields.

   (c) Let f be an irreducible polynomial of degree 2004 over $\mathbb{F}_3$. Into how many factors does f factor in $K[X]$? Prove this.

8. Let $\zeta = e^{\frac{2\pi i}{221}}$. Why does $\mathrm{Gal}(\mathbb{Q}(\zeta)/\mathbb{Q})$ contain an element σ with $\sigma(\zeta) = \zeta^{11}$ but no τ for which $\tau(\zeta) = \zeta^{13}$? You may use that $221 = 13 \cdot 17$ and

$$X^{221} - 1 = (X-1) \cdot \frac{X^{13}-1}{X-1} \cdot \frac{X^{17}-1}{X-1} \cdot g(X)$$

with $g(X)$ irreducible of degree 192 over $\mathbb{Q}$.

9. Are the following statements false or true? Motivate your answer.

   (a) If α is algebraic over $\mathbb{Q}$, then so is $\sqrt[3]{1+\sqrt{\alpha}}$.

   (b) Given two circles in the plane (e.g. center and radius), then it is possible to contruct with ruler and compass a third circle whose surface area is the sum of the areas of the two given circles.

   (c) There exists a field automorphism σ of $\mathbb{C}$ such that $\sigma(\sqrt{2}) = \sqrt[3]{2}$.

   (d) There exists a field K and an irreducible polynomial over K with multiple roots in an extension of K.

   (e) There exists a solvable equation over $\mathbb{Q}$ of which no root is constructible over $\mathbb{Q}$.

10. Let $K = \mathbb{F}_4(t)$ and $f(X) = X^9 - t \in K[X]$ with t transcendental over the field $\mathbb{F}_4$ with 4 elements.

    (a) Show that f is irreducible over K.

    (b) Determine the degree of the splitting field L of f over K.

    (c) Show that $G := \mathrm{Gal}(L/K)$ contains a normal subgroup H with $H \cong \mathbb{Z}/9$ an such that $G/H \cong \mathbb{Z}/3$, but G not isomorphic to $\mathbb{Z}/9 \times \mathbb{Z}/3$.

11. Let $f = X^4 + X^2 - 1$.

    (a) How many elements does the Galoisgroup G of f over $\mathbb{Q}$ have?

    (b) Is G abelian or not?

    (c) How many elements does the Galoisgroep van f have over a field with 9 elements?

12. Are the following statements true or false? Explain.

    (a) Let t be transcendental over $\mathbb{Q}$. The polynomial $X^{2007} - t^3$ is irreducible over $\mathbb{Q}[t]$.

    (b) According to a theorem of Feit and Thompson from 1962 every finite group of odd order is solvable. Hence every polynomial of odd degree is solvable.

    (c) Let t be transcendental over $\mathbb{F}_3$. There exists a field automorphism σ of the field $\mathbb{F}_3(t, \sqrt{t}, \sqrt[3]{t})$ over $\mathbb{F}_3(t)$ such that $\sigma(\sqrt{t}) = \sqrt[3]{t}$.

13. We are given three collinear points $(0,0),(1,0)$ en $(x_0,0)$ in the plane, with $x_0 \in \mathbb{Q}$. Show that a fourth point $(x,0)$, collinear with the given points, and the property that the product of its distances to the three given points equals 1, is constructable with ruler and compass if and only if x_0 has the form
$$x_0 = t \pm \frac{1}{t(t-1)}$$
with $t \in \mathbb{Q} - \{0,1\}$.

14. Let $f(X) = (X^3-3)(X^2+3)(X^2+X+1)$, and L the splitting field of f over $\mathbb{Q}$.
    (a) Determine the degree of L over $\mathbb{Q}$.
    (b) Determine a primitive element for L over $\mathbb{Q}$.

15. Let $K = \mathbb{Q}(t)$ and $f(X) = X^5 - t \in K[t]$ with t een transcendental over $\mathbb{Q}$.
    (a) Is f irreducible over K?
    (b) Determine the degree of the splitting field of f over K.
    (c) Determine the Galoisgroup of f over K, i.e. give generators and their relations.

16. Let $K = \mathbf{F}_9(t)$, with t transcendental over $\mathbf{F}_9$, and let $f(X) = X^4 - t \in K[X]$.
    (a) Prove that f is irreducible over K.
    (b) Prove that X^2+1 is reducible over K.
    (c) Let α be a zero of f in a splitting field. Prove that $L = K(\alpha)$ is the splitting field of f.
    (d) Determine the Galois group $G = \mathrm{Gal}(L/K)$.
    (e) Give all subgroups of G and the corresponding intermediate fields.
    (f) For each intermediate field E between K and L, determine a minimal polynomial for the extensions L/E and E/K.

17. (Ribet) Let L/K be a finite Galois extension, and M an intermediate field between L and K. Suppose that no intermediate field between L and M is Galois over K, except L itself. Prove: if N is a subfield of L which contains all fields $\sigma(M)$ for $\sigma \in \mathrm{Gal}(L/K)$, then $N = L$.

Chapter 10

Solving equations

10.1 Symmetric polynomials

Let K be a field and consider the polynomial ring $K[x_1,\ldots,x_n]$ in n x_i. A polynomial $g(x_1,\ldots,x_n)$ is called *symmetric* if $g(x_{\sigma(1)},\ldots,x_{\sigma(n)}) = g(x_1,\ldots,x_n)$ for every permutation σ of $\{1,2,\ldots,n\}$. Simple examples are $x_1+x_2+\cdots+x_n$ or $\sum_{i\ne j} x_ix_j$ or $x_1x_2\cdots x_n$. Let us write

$$(T-x_1)(T-x_2)\cdots(T-x_n) = T^n - s_1T^{n-1} + s_2T^{n-2} - \cdots + (-1)^n s_n$$

where $s_1 = x_1+\cdots+x_n$, $s_2 = \sum_{i<j} x_ix_j$ and $s_n = x_1x_2\cdots x_n$. The coefficients s_i are all symmetric functions of $x_1,\ldots,x_n$. They are called the *elementary symmetric functions*. We have the following theorem.

Theorem 10.1.1 *Let K be a field. Then any symmetric polynomial in $K[x_1,\ldots,x_n]$ can be written as a polynomial in the elementary symmetric functions $s_1,s_2,\ldots,s_n$.*

For example, $x_1^2+x_2^2+\cdots+x_n^2$ is a symmetric polynomial. It can be written as $s_1^2-2s_2$ (please verify). Another example, $x_1^3+x_2^3+\cdots+x_n^3 = s_1^3-3s_2s_1+3s_3$. We use this theorem to derive a solution of the third and fourth degree equation.

10.2 Solution of the cubic equation

Consider the polynomial X^3-aX^2+bX-c and let x_1,x_2,x_3 be its zeros. We assume that a,b,c lie in a field which contains the cube root of unity ω. Consider the expression $u_1 = x_1+\omega x_2+\omega^2x_3$. Apply the cyclic substition $x_1 \to x_2 \to x_3 \to x_1$ represented by the cycle (123). Then u_1 changes into ω^2u_1. So u_1^3 is fixed under the cyclic substitution. Let similarly $u_2 = x_1+\omega^2x_2+\omega x_3$ Then after the cyclic substitution u_2 is replaced by ωu_2. Under the substitution $x_2 \to x_3 \to x_2$ (represented by (23) the form u_1 changes into u_2 and vice versa. Hence u_1u_2 and $u_1^3+u_2^3$ are fixed under the group of permutations generated by (123) and (23), which is the full permutation group S_3. Hence they are symmetric polynomials in x_1,x_2,x_3. By Theorem 10.1.1 these polynomials can be expressed in the symmetric expressions a,b,c in the x_i. Straightforward calculation gives us

$$\begin{aligned} u_1^3+u_2^3 &= 2(x_1^3+x_2^3+x_3^3)-3(x_1^2x_2+x_1x_2^2+x_1^2x_3+x_1x_3^2+x_2^2x_3+x_2x_3^2) \\ &= 2a^3-9ab+27c \end{aligned}$$

and

$$u_1u_2 = x_1^2+x_2^2+x_3^2-x_1x_2-x_1x_3-x_2x_3 = a^2-3b.$$

Without loss of generality we can assume that $a=0$. So we see that u_1^3 and u_2^3 are solutions of the equation

$$(U-u_1^3)(U-u_2^3) = U^2-27cU+(-3b)^3 = 0.$$

Solution gives us

$$u_1^3 = (27c+\sqrt{27^2c^2+4\cdot 27b^3})/2 = 27(c/2)+27\sqrt{(c/2)^2+(b/3)^3}.$$

Hence

$$u_1 = 3\sqrt[3]{\left(c/2-\sqrt{(c/2)^2+(b/3)^3}\right)}, \quad u_2 = -3b/u_1.$$

Now note that

$$3x_1 = (x_1+x_2+x_3)+u_1+u_2 = a+u_1+u_2 = u_1+u_2$$

where u_1, u_2 are given above. These formulas are known as *Cardano's formulas* for the solution of the cubic equation.

The linear forms u_1, u_2 which made the solution possible are examples of *Lagrange resolvents.* We will see more of them later.

10.3 Solution of the quartic equation

Suppose we are given the equation

$$X^4-aX^3+bX^2-cX+d=0.$$

Let x_1, x_2, x_3, x_4 be the solutions. Consider the elements

$$\begin{aligned} y_1 &= (x_1+x_2-x_3-x_4)^2 \\ y_2 &= (x_1-x_2-x_3+x_4)^2 \\ y_3 &= (x_1-x_2+x_3-x_4)^2 \end{aligned}$$

Observe that after any permutation of x_1, x_2, x_3, x_4 the expressions y_1, y_2, y_3 are also permuted. Hence the elementary symmetric functions in y_1, y_2, y_3 are symmetric in x_1, x_2, x_3, x_4. Straightforward calculation gives us

$$\begin{aligned} y_1+y_2+y_3 &= 3a^2-8b \\ y_1y_2+y_1y_3+y_2y_3 &= 3a^4-16a^2b+16b^2+16ac-64d \\ y_1y_2y_3 &= a^6-8a^4b+16a^3c+16a^2b^2-64abc+64c^2 \end{aligned}$$

Let us assume for simplicity (and without loss of generality) that $a = 0$. Then y_1, y_2, y_3 are solutions of the equation

$$(Y - y_1)(Y - y_2)(Y - y_3) = Y^3 + 8bY^2 + (16b^2 - 64d)Y - 64c^2 = 0.$$

We know how to solve this equation by Cardano's formula. Having found y_1, y_2, y_3 we observe that

$$4x_1 = a + \sqrt{y_1} + \sqrt{y_2} + \sqrt{y_3}$$

and similarly for x_2, x_3, x_4. Thus we have recovered Ferrari's formula for the solution of the quartic equation.
The functions y_1, y_2, y_3 are fixed under the substitutions $(1), (12)(34), (13)(24), (14)(23)$ that is, Klein's fourgroup V_4.Note that V_4 is a normal subgroup of S_4 and that $S_3 \cong S_4/V_4$, which is the group by which the y_1, y_2, y_3 are permuted. In its turn S_3 has the alternating subgroup A_3 as normal subgroup and $S_3/A_3 \cong S_2$. The solution of the cubic equation is based on the construction of an invariant under A_3. In this way we see how group theoretic considerations arise in the ideas to solve polynomial equations. The step to Galois theory is a natural continuation.

10.4 Radical extensions

Of course similar considerations have been attempted to solve the equation of degree 5 (quintic equation). The goal of these attempts was to describe the zeros of a quintic polynomial as a result of repeated application of addition/subtraction, multiplication/division and taking n-th roots, starting with the coefficients of the polynomial equation. We call such a procedure *solution by radicals* where "radical" refers to taking n-th roots. In the case of cubic and quartic equations we have seen how this is done. As is known since the beginning of the 19-th century, the quintic equation cannot be solved by radicals. In order to prove this we have a closer look at field extensions by n-th roots (radicals).

Remark 10.4.1 *From now on all our fields have characteristic zero.*

First we consider the zeros of $X^n - 1$. When working in the complex numbers we know that they are given by $e^{2\pi ik/}$ for $k = 0, 1, \ldots, n-1$ and that they form a cyclic multiplicative group generated by $e^{2\pi i/n}$. In general we do not have the exponential function available. However, the following is still true.

Theorem 10.4.2 *Let $n \in \mathbb{N}$ and let K be a field. Suppose that $X^n - 1$ has n zeros in K. These zeros form a cyclic group. We call these zeros the n-th roots of unity in K.*

Let now K be a field of characteristic zero and $n \in \mathbb{N}$. Let $a \in K$ and let L be splitting field of $X^n - a$. It contains n distinct zeros and also their quotients, which are the n-th roots of unity. Let $\zeta \in L$ be a generator of the n-th roots of unity. In particular we see that the splitting field of $X^n - a$ contains the n-th roots of unity. We consider n-th roots as the simplest examples of radicals (of 1) and we will usually assume that they lie in the ground field we are considering.

Definition 10.4.3 *Let K be a field. A finite extension L of K is called a radical extension if there exist $n \in \mathbb{Z}_{n>1}$ and $a \in K^*$ such that $L = K(\sqrt[n]{a})$.*

Theorem 10.4.4 *Let $n \in \mathbb{Z}_{>1}$ and let K be a field. Suppose K contains a generator ζ of the zeros of $X^n - 1$.*
Then, for any $a \in K^*$ the splitting field of $X^n - a$ is a Galois extension of K with cyclic Galois group whose order divides n.
Conversely, any Galois extension L/K of degree n with a cyclic Galois group is a radical extension of the form $K(\sqrt[n]{a})$.

Proof: Suppose L is the splitting field of $X^n - a$ over K. Denote one zero of $X^n - a$ by $\sqrt[n]{a}$. Then the other zeros are given by $\zeta^k \sqrt[n]{a}$. Furthermore, any Galois element sends $\sqrt[n]{a}$ to $\zeta^k \sqrt[n]{a}$ for some $k \in \mathbb{Z}/n\mathbb{Z}$. Denote this element by σ_k. Note also that σ sends ζ to itself since it is contained in K. One easily verifies that $\sigma_k \sigma_l(\sqrt[n]{a}) = \zeta^{k+l} \sqrt[n]{a} = \sigma_{k+l}(\sqrt[n]{a})$. Hence $\mathrm{Gal}(L/K)$ is isomorphic to an additive subgroup of $\mathbb{Z}/n\mathbb{Z}$. Hence it is cyclic.
Suppose $\mathrm{Gal}(L/K)$ is cyclic of order n. By Theorem 9.4.1 there exists an element $\alpha \in L$ such that $L = K(\alpha)$. Let σ be a generator of $\mathrm{Gal}(L/K)$. Define for every $i = 0, 1, 2, \ldots, n-1$ the sum

$$\begin{aligned} s_i &= \alpha^i + \zeta\sigma(\alpha^i) + \cdots + \zeta^{n-1}\sigma^{n-1}(\alpha^i) \\ &= \alpha^i + \zeta\sigma(\alpha)^i + \cdots + \zeta^{n-1}\sigma^{n-1}(\alpha)^i \end{aligned}$$

Since the determinant of $(\sigma^j(\alpha)^i)_{i,j=0,\ldots,n-1}$ is non-zero at least one of the s_i is non-zero. Since, trivially, $s_0 = 1 + \zeta + \cdots + \zeta^{n-1} = 0$ we see that $s_i \neq 0 \Rightarrow i > 0$. Choose such an i. Notice that $\sigma(s_i) = \zeta^{-1} s_i$. Hence s_i has n distinct images under the Galois group and thus $K(s_i) = K(\alpha)$. Furthermore, $\sigma(s_i^n) = s_i^n$ and therefore $s_i^n \in K$. Let us now denote $a = s_i^n$. Then L is the splitting field of $X^n - a$ over K.

□

In order to be able to adjoin roots of unity to our extensions the following proposition describes their effect on the corresponding Galois groups.

Proposition 10.4.5 *Let L/K be a finite Galois extension. Let α be an element in a finite extension of L and consider the extension $L(\alpha)/K(\alpha)$. This is a Galois extension whose Galois group is isomorphic to a subgroup of* $\mathrm{Gal}(L/K)$.

Remark 10.4.6 *We can be even more precise and show that*

$$\mathrm{Gal}(L(\alpha)/K(\alpha)) \cong \mathrm{Gal}(L/L \cap K(\alpha)) \subset \mathrm{Gal}(L/K),$$

but we will not need it here.

Proof: There is a natural map $\mathrm{Gal}(L(\alpha)/K(\alpha)) \to \mathrm{Gal}(L/K)$ by restriction of a element $\sigma \in \mathrm{Gal}(L(\alpha)/K(\alpha))$ to L. Clearly σ sends L to itself since L is a splitting field. It suffices to show that our map is injective. Suppose that σ fixes L. By definition it also fixes α hence it fixes $L(\alpha)$. So σ is the identity in $\mathrm{Gal}(L(\alpha)/K(\alpha))$. This shows the required injectivity.

□

10.5 Solvability

In talking about solvability it is convenient to speak of the Galois group of a polynomial.

Remark 10.5.1 *In order to avoid intricacies we shall assume that the ground field K has characteristic zero in this and the following sections.*

Definition 10.5.2 *Let $f \in K[X]$ be a separable polynomial. Then by the Galois group of f we mean the Galois group of the splitting field of f. Notation* $\mathrm{Gal}(f/K)$.

We now formalise our concept of solvability by radicals.

Definition 10.5.3 *Let K be a field and $f \in K[X]$. Denote its splitting field over K by L. Then we say that f is solvable by radicals if and only if there exists a tower of extensions*

$$K = K_0 \subset K_1 \subset K_2 \subset \cdots \subset K_m$$

such that K_{i+1} is a radical extension of K_i for $i = 0, 1, 2, \ldots, m-1$ and such that $L \subset K_m$.

The sequence of fields $K_0 \subset K_1 \subset \cdots \subset K_m$ is called a *radical tower.* We will call the last field a *radical tower extension.* So a radical tower extension of a field K is a field obtained by consecutively adjoining roots of elements of the previous field.

Proposition 10.5.4 *Suppose we have radical tower extension K_m of a field K. Then K_m is contained in a radical tower extension L of K which is also normal.*

Proof: Suppose that $K_m = K(\alpha_1, \ldots, \alpha_r)$. Let $g_i \in K[X]$ be the minimal polynomial of α_i over K for $i = 1, \ldots, r$. Let L be the splitting field of $g_1 \cdots g_r$. We call L the *normal closure* of K_m. Note that L is the smallest field which contains all fields $\sigma(K_m)$ where σ runs through $\mathrm{Gal}(L/K)$. Each field $\sigma(K_m)$ can be obtained by consecutively adjoining roots of elements from the previous fields. Hence the same holds for L and therefore L is a radical tower extension.

□

We like to translate solvability of a polynomial f by a property of the Galois group $\mathrm{Gal}(f/K)$.

Definition 10.5.5 *A group G is called solvable If there is a finite tower of subgroups*

$$\mathrm{id} = G_0 \subset G_1 \subset \cdots \subset G_r = G$$

such that

- *For every $i = 0, 1, \ldots, r-1$ G_i is a normal subgroup of G_{i+1}.*
- *The quotients G_{i+1}/G_i are all cyclic.*

We call the sequence $G_0, G_1, \ldots, G_r$ a resolving sequence of subgroups

A nice observation is that the solvability property of a group is inherited by its subgroups and quotients by normal subgroups.

Proposition 10.5.6 *Let G be a solvable group and H a subgroup. Then H is again solvable. Moreover, if H is a normal subgroup then G/H is again solvable.*

Proof: Let $\mathrm{id} = G_0 \subset G_1 \subset \cdots \subset G_r = G$ be the resolving sequence of subgroups. Then the sequence consisting of $G_i \cap H$ forms the resolving sequence of subgroups for H. To this end we need to show that $H \cap G_i$ is a normal subgroup of $H \cap G_{i+1}$ and $(H \cap G_{i+1})/(H \cap G_i)$ is cyclic. To this end end consider the group homomorphism $G_{i+1} \to G_{i+1}/G_i$ restricted to $H \cap G_{i+1}$. Clearly the kernel is $H \cap G_i$, which is normal, and the image, being a subgroup of the cyclic group G_{i+1}/G_i is cyclic. So $(G_{i+1} \cap H)/(G_i \cap H)$ is a cyclic group. When H is normal in G consider the natural quotient map $\phi : G \to G/H$. For a resolving sequence $G_0, G_1, \ldots, G_r$ of G we define $H_i = \phi(\langle G_i, H\rangle)$ for $i = 0, \ldots, r$. Here $\langle G_i, H\rangle$ denotes the subgroup of G generated by the elements of G_i and H. We assert that $H_0, H_1, \ldots, H_r$ is a resolving sequence for G/H. First note that $\langle G_i, H\rangle$ is a normal subgroup of $\langle G_{i+1}, H\rangle$. We leave this as an exercise. Then we have the following group theoretic isomorphisms

$$H_{i+1}/H_i \cong \langle G_{i+1}, H\rangle / \langle G_i, H\rangle \cong G_{i+1}/\langle G_i, G_{i+1} \cap H\rangle .$$

Hence H_{i+1}/H_i is isomorphic to a quotient group of G_{i+1}/G_i, which is cyclic. Hence H_{i+1}/H_i is also cyclic.

□

There exist groups which are not solvable, as is shown by the following Theorem.

Theorem 10.5.7 *The symmetric group S_n with $n \geq 5$ is not solvable.*

Proof: We assert that if H is a subgroup of S_n containing all 3-cycles, and N is a normal subgroup in H such that H/N is abelian, then N also contains all 3-cycles.

Indeed, suppose that $\sigma = (ijk)$ and $\tau = (krs)$ are 3-cycles for any five distinct i, j, k, r, s. By assumption, $\sigma \in H$ en $\tau \in H$. One easily verifies that $\sigma\tau\sigma^{-1}\tau^{-1} = (rki)$.

Consider the group homomorphsim $\phi : H \to H/N$ with kernel N. Since H/N is abelian, we have $\phi(\sigma\tau\sigma^{-1}\tau^{-1}) = 1$, so $\sigma\tau\sigma^{-1}\tau^{-1} \in \ker(\phi) = N$. As a consequence N contains the 3-cycle (rki) for all triples of distinct r, k, i.
Suppose S_n is solvable, then there is resolving sequence

$$S_n = H_0 \supseteq H_1 \supseteq \ldots \supseteq H_r = \{1\}$$

where the consecutive quotients are abelian groups. So by induction we find that H_r contains all 3-cycles. This is clearly not possible.

□

Remark 10.5.8 *The argument only works if we can choose five distinct indices, so $n \geq 5$ in S_n. Furthermore S_n is sovable when $n \leq 4$: S_1 and S_2 are abelian. The group S_3 has the resolving sequence* $\mathrm{id} \subset A_3 \subset S_3$. *The group S_4 has the resolving sequence* $\mathrm{id} \subset V_4 \subset A_4 \subset S_4$ *with quotients $S_4/A_4 = \mathbb{Z}/2\mathbb{Z}$ en $A_4/V_4 = \mathbb{Z}/3\mathbb{Z}$. Here V_4 denotes Klein's fourgroup consisting of* $\mathrm{id}, (12)(34), (13)(24), (14)(23)$.

10.6 Main Theorem

We now come to our main point.

Theorem 10.6.1 *A separable polynomial $f \in K[X]$ is solvable by radicals if and only if* $\mathrm{Gal}(f/K)$ *is solvable.*

Proof: We first set some notation. Let L/K be the splitting field of f and G its Galois group.
Suppose f is solvable. Then there exists a tower of fields

$$K = K_0 \subset K_1 \subset \cdots \subset K_m$$

such that K_{i+1} is a radical extension of K_i for $i = 0, 1, \ldots, m-1$ and $L \subset K_m$. First of all, by using Proposition 10.5.4 we might as well assume that K_m/K is normal. Secondly, suppose the order of the radical extensions involved are $n_1, n_2, \ldots, n_m$ and $n = \mathrm{lcm}(n_1, \ldots, n_m)$. Let ζ be a primitive n-th root of unity. Then we can replace K_i by $K_i(\zeta)$ so that we get the chain $K \subset K(\zeta) \subset K_1(\zeta) \subset \cdots$ which is 1 element longer than the original chain.
Let $G_0 = \mathrm{Gal}(K_m(\zeta)/K(\zeta)) \supset G_1 \supset \cdots \supset G_m = \mathrm{id}$ be the corresponding sequence of subgroups of the Galois group, where $G_i = \mathrm{Gal}(K_m(\zeta)/K_i(\zeta))$. From Theorem 10.4.4 it follows that $K_{i+1}(\zeta)/K_i(\zeta)$ is a radical Galois extension with a cyclic Galois group. So G_{i+1} is a normal subgroup of G_i and G_i/G_{i+1} is cyclic. Hence the sequence of groups is solvable. We now extend the resolving sequence by $G_0 \subset \mathrm{Gal}(K_m(\zeta)/K)$. Note that $K(\zeta)/K$ is an extension by roots of unity, hence abelian and normal. So $\mathrm{Gal}(K_m(\zeta)/K(\zeta)$ is normal in $\mathrm{Gal}(K_m(\zeta)/K)$ with abelian quotient. Therefore $\mathrm{Gal}(K_m(\zeta)/K)$ is also solvable.
The field L is a normal extension of K. So $\mathrm{Gal}(K_m(\zeta)/L)$ is a normal subgroup of $\mathrm{Gal}(K_m(\zeta)/K)$ and

$$\mathrm{Gal}(L/K) \cong \mathrm{Gal}(K_m(\zeta)/K)/\mathrm{Gal}(K_m(\zeta)/L).$$

Since $\mathrm{Gal}(K_m(\zeta)/K)$ is solvable, the same holds for the quotient group $\mathrm{Gal}(L/K)$. Suppose conversely that $\mathrm{Gal}(L/K)$ is a solvable group. Suppose it is of order n. Let ζ be a generator of the zeros of $X^n - 1$ in some splitting field. Consider a resolving sequence $\mathrm{id} = G_0 \subset G_1 \subset \cdots \subset G_r = \mathrm{Gal}(L/K)$. By Galois correspondence there is a sequence of subfields $L = K_0 \supset K_1 \supset \cdots \supset K_r = K$ such that $\mathrm{Gal}(L/K_i) = G_i$ for $i = 0, 1, \ldots, r$. Since G_i is a normal subgroup of G_{i+1} the extension $K_i \supset K_{i+1}$ is normal. Moreover, G_{i+1}/G_i is cyclic, so $\mathrm{Gal}(K_i/K_{i+1})$ is cyclic. Hence, by Theorem 10.4.4 the extension $K_i(\zeta)/K_{i+1}(\zeta)$ is a radical extension. So we get the tower of radical extensions

$$K \subset K(\zeta) = K_r(\zeta) \subset K_{r-1}(\zeta) \subset \cdots \subset K_0(\zeta) \subset L(\zeta)$$

and $L \subset L(\zeta)$. Hence f is solvable by radicals.

□

Of course it is of interest to know if there are polynomial equations which are unsolvable by radicals. Here is a straightforward example.

Theorem 10.6.2 *Let k be a field and $L = k(t_1, \ldots, t_n)$ the function field in n variables. Let $K = k(s_1, s_2, \ldots, s_n)$ be the subfield of L generated by the elementary symmetric functions in $t_1, \ldots, t_n$. Then L/K is a finite extension of degree $n!$ with Galois group S_n.*

Proof: Note that L is the splitting field of $X^n - s_1X^{n-1} + s_2X^{n-2} - \cdots + (-1)^n s_n$ with zeros $t_1, \ldots, t_n$. The Galois group of L/K is a subgroup of S_n, the permutation group of $t_1, \ldots, t_n$. On the other hand, every permutation of $t_1, \ldots, t_n$ is a K-automorphism of L. Hence $\mathrm{Gal}(L/K) = S_n$.

□

Since S_n is unsolvable for $n \geq 5$, Theorem 10.6.1 implies that there cannot exist general formulas for the solution of the n-th degree equation with $n \geq 5$ of Cardano and Ferrari type. A fortiori, we can also provide individual equations which cannot be solved by radicals.

Theorem 10.6.3 *The zeros of $f(x) = x^5 - 6x + 3 \in \mathbb{Q}[x]$ cannot be written as radical expressions in $\mathbb{Q}$.*

Proof: The polynomial f is irreducible by Eisenstein's criterion for $p = 3$. From the graph of f we see that f has at least 3 real zeros. If there would exist four real zeros, then by Rolle's Theorem f' would have at least three and f'' at least two distinct zeros. This contradicts $f'' = 20x^3$. So there must be a pair complex conjugate zeros as well. The next Lemma implies that $\mathrm{Gal} f/\mathbb{Q} \cong S_5$, which is not solvable.

□

Lemma 10.6.4 *Suppose that p is prime and $f \in \mathbb{Q}[x]$ a polynomial of degree p, irreducible in $\mathbb{Q}[X]$ with exactly $p - 2$ real zeros. Then $G := \mathrm{Gal} f/\mathbb{Q} \cong S_p$.*

Proof: Let α be a zero of f in in a splitting field L. Then $[\mathbb{Q}(\alpha) : \mathbb{Q}] = \deg f = p$. So p divides $[L : \mathbb{Q}] = |G|$ where $G = \mathrm{Gal}(L/\mathbb{Q})$. Cauchy's theorem states that G contains an element σ of order p. G is a subgroup of S_p, so σ is a p-cycle in G.
Complex conjugation is an automorphism of $L/\mathbb{Q}$ which swaps the two complex zeros, but fixes the real zeros, so we get a transposition in G. In S_p any p-cycle and a transposition generate S_p and therefore $G = S_p$.

□

10.7 Exercises

1. (Galois group of equations of degree 3) Let $x^3 - px - q$ be a cubic polynomial with coefficients in a field K (characteristic not 2,3) and $G = \mathrm{Gal}(f/K)$. Let $\{\alpha_i\}_{i=1}^3$ be the three roots of f in a splitting field The *discriminant* of f is defined as

$$D = \prod_{i<j} (\alpha_i - \alpha_j)^2 = -\prod_{i \neq j} (\alpha_i - \alpha_j).$$

   (a) Show that D is invariant under G and hence $D \in K$.

   (b) Let f' be the derivative of f. Prove that D is up to a sign change equal to $\prod_{i=1}^{3} f'(\alpha_i)$. Then prove that $D = 4p^3 - 27q^2$.

   (c) Suppose that f is irreducible in $K[x]$ and D is a square in K. Prove that $G \cong A_3$.

   (d) Suppose that f is irreducible in $K[x]$ and D is not a square in K. Prove that $G \cong S_3$. (hint: $[K(\sqrt{D}) : K] = 2$ and $deg(f) = 3$ both divide $|G|$.)

2. Give an example of an algebraic number of degree 4 over $\mathbb{Q}$ which is *not constructible* over $\mathbb{Q}$.

3. Give an example of a polynomial of degree 7, irreducible over $\mathbb{Q}$ which is not solvable in radicals. Give an example of an irreducible polynomial of degree 7 over $\mathbb{Q}$ which is solvable in radicals.

4. In "Discours de la méthode pour bien conduire sa raison et chercher la verité dans les sciences; plus la dioptrique les meteores et la geometrie qui sont des essais de cete methode (Leiden, 1637) contains the section "La Geometrie" pp. 297-413. There we find on p. 323: " *Dont la raison est qu'il y a reigle generale pour reduire au cube toutes les difficultés qui vont au quarré de quarré, et au sursolide toutes celles qui vont au quarré de cube, de façon qu'on ne les doit point estimer plus composées.*"

   Henk Bos interpretes this (in "Redefining geometrical exactness. Descartes' transformation of the early modern concept of construction", Springer-Verlag 2001, pp. 356 e.v.) as follows: Descartes asserted that the general equation of degree 6 can be solved by additional radicals and solutions of fifth degree equations (by analogy with Ferrari who reduced the fourth degree equation to the cubic equation). Show that Descartes' statement, in this interpretation, is erroneous.

5. Prove: to every finite group G there exists a finite extension L/K such that $\mathrm{Gal}(L/K) \cong G$ (you are allowed to choose L and K dependent on G).

   Remark. The question whether this is always possible with $K = \mathbb{Q}$ is unanswered yet and belongs to one of the big open problems in Galois theory.

Chapter 11

Finite fields

11.1 Existence, unicity

Definition 11.1.1 *A finite field is a field with finite cardinality.*

Theorem 11.1.2 *a) A finite field has p^n elements where p is the characteristic. b) For every prime p and $n \geq 1$ there exists precisely one field (up to isomorphism) with $q = p^n$ elements. We denote it by $\mathbb{F}_q$. There are no other finite fields.*
c) Let $q = p^n$ be a power of the prime p. Then $\mathbb{F}_q/\mathbb{F}_p$ is a Galois extension with Galoisgroep $\mathbb{Z}/n$. It is generated by the Frobenius element $\phi : x \mapsto x^p$.
d) The subfields of $\mathbb{F}_{p^n}$ are precisely the fields $\mathbb{F}_{p^m}$ with $m|n$. Moreover,

$$\mathrm{Gal}\mathbb{F}_{p^n}/\mathbb{F}_{p^m} \cong \mathbb{Z}/d = \langle \phi^m \rangle$$

where $n = md$.

Proof: a) Let L be a finite field. Since L is finite, it has positive characteristic p, where p is a prime.
So L is a finite extension of $\mathbb{F}_p$ and therefore a finite dimensional vector space over $\mathbb{F}_p$. Suppose the dimension n, then $|L| = p^n$.
b) We first prove uniqueness. Let $q = p^n$ and let L be a field with q elements. Then L^* has $q-1$ elements, and hence $x^{q-1} = 1$ for all $x \in L^*$. So all elements of L satisfy $x^q - x = 0$ and L is the splitting field of $x^q - x$. Since a splitting field of a polynomial is uniquely determined (up to isomorphism), the same holds for L.
To show existence take any prime power $q = p^n$ and let L be the splitting field of $x^q - x \in \mathbb{F}_p[x]$. For any two zeros α, β we have $(\alpha+\beta)^q = \alpha^q + \beta^q = \alpha + \beta$ and $(\alpha\beta)^q = \alpha^q\beta^q = \alpha\beta$. Hence the zeros of $x^q - x$ form a field. Since $x^q - x$ is separable, the zeros are distinct and L is a field with q elements.
c) The extension $\mathbb{F}_q/\mathbb{F}_p$ is a Galois extension since $\mathbb{F}_q$ is the splitting field of the separable polynomial $x^q - x$. Hence G has order $[\mathbb{F}_q : \mathbb{F}_p] = n$. The Frobenius map ϕ is indeed an element of $\mathrm{Gal}(\mathbb{F}_q/\mathbb{F}_p)$. Let r be the order of ϕ. Then $r \leq n$. On the other hand suppose that $\phi^r(x) = x$ for all $x \in \mathbb{F}_q$. Then $x^{p^r} - x$ has at least p^n solutions and so $p^r \geq p^n$, hence $r \geq n$. We conclude that $r = n$ and the Galois group is cyclic.

d) Subgroups of $\mathbb{Z}/n$ are isomorphic to cyclic groups $\mathbb{Z}/d$ with $d|n$, so the first statement follows from Galois correspondence. The fixed field of $\mathbb{Z}/d$ (generated by ϕ^m) is precisely $\mathbb{F}_{p^m}$.

□

Remark 11.1.3 *In practice one constructs a finite field with q elements as follows. Suppose $q = p^n$ and let $f \in \mathbb{F}_p[X]$ be an irreducible polynomial of degree n. Then $\mathbb{F}_p[X]/(f)$ is a finite field of p^n elements and thus isomorphic to $\mathbb{F}_q$.*
Since $\mathbb{F}_q/\mathbb{F}_p$ is Galois, f has all its roots in $\mathbb{F}_q$. A fortiori, the above construction of $\mathbb{F}_q$ is independent of the choice of f. So every irreducible polynomial of degree n has all its zeros in $\mathbb{F}_q$. Since all elements $x \in \mathbb{F}_q$ satisfy $x^q - x = 0$ we also conclude that every irreducible $f(x) \in \mathbb{F}_p[x]$ of degree n divides $x^q - x$.

11.2 Exercises

Exercise 11.2.1. Show that $f(x) = x^3 - x + 1$ and $g(x) = x^3 - x - 1$ are irreducible over $\mathbb{F}_3$. Is there an isomorphism between $\mathbb{F}_3[x]/(f)$ and $\mathbb{F}_3[x]/(g)$?

Exercise 11.2.2. Show that $f(x) = x^4+x+1$ and $g(x) = x^4+x^3+x^2+x+1$ are irreducible over $\mathbb{F}_2$. Hence there is an isomorphism $\phi : \mathbb{F}_2[x]/(f) \to \mathbb{F}_2[x]/(g)$. Show that $x(\text{mod } f)$ has multiplicative order 15 and $x(\text{mod } g)$ has multiplicative order 5. Conclude that $\phi(x(\text{mod } f)) \neq x(\text{mod } g)$.

Exercise 11.2.3. Determine all subfields of the field of 1024 elements (up to isomorphism).

Exercise 11.2.4. Let j be a primitive element for the extension $\mathbb{F}_4/\mathbb{F}_2$, i.e. $\mathbb{F}_4 = \mathbb{F}_2(j)$, and let $f(X) = X^4 + X^2 + jX + 1 \in \mathbb{F}_4[X]$.
(a) Give the minimal equation of j over $\mathbb{F}_2$.
(b) Prove that $\mathbb{F}_4 = \{0, 1, j, j+1\}$ and construct the table of multiplication of this field.
(c) Prove that f is irreducible over $\mathbb{F}_4$.
(d) Let α be a zero of f in a splitting field. Use the theory of finite fields to answer the following questions :

(d1) $L = \mathbb{F}_4(\alpha)$ is the splitting field of f.
(d2) L is a finite field of characteristic 2, which field is it?
(d3) Gal$L/\mathbb{F}_4$ is cyclic with 4 elements.

(e) Factor f in L and write its zeros with respect to the basis $\{1, \alpha, \alpha^2, \alpha^3\}$ of $L/\mathbb{F}_4$.
(f) Prove that there is a unique field E strictly between L and $\mathbb{F}_4$, and determine it.
(g) Give minimal polynomials for the extensiuon L/E and $E/\mathbb{F}_4$.

Exercise 11.2.5. Let j be a zero of $X^2 + X - 1 \in \mathbb{F}_3[X]$, in a splitting field. Let $f(X) = X^4 + jX - j \in \mathbb{F}_3(j)[X]$.
(a) Prove that j is a primitive element for the extension $\mathbb{F}_9/\mathbb{F}_3$, i.e. $\mathbb{F}_9 = \mathbb{F}_3(j)$.

(b) Write down the multiplication table for $\mathbb{F}_9$.
(c) Prove that f is irreducible over $\mathbb{F}_9$.
(d) Let α be a zero of f in the splitting field of f.

(d1) Prove that $L = \mathbb{F}_9(\alpha)$ is the splitting field of f.
(d2) Determine the number of elements in L.
(d3) Determine the Galois group of f over $\mathbb{F}_9$.

(e) Prove that there is a unique field E strictly in between $\mathbb{F}_9$ and L, and determine a primitive element of the extension $E/\mathbb{F}_9$. Use equalities in the calculation:

$$\begin{array}{rcl}\alpha^{81} & = & (-j+1)\alpha^3 + \alpha^2 + (j-1)\alpha \\ \alpha^{162} & = & \alpha^3 + (j+1)\alpha^2 + (j+1)\alpha \\ \alpha^{243} & = & (j-1)\alpha^3 + (j+1)\alpha^2 + \alpha\end{array}$$

(f) Determine the minimal polynomial of the extension $E/\mathbb{F}_9$.

Exercise 11.2.6. Determine the Galoisgroup van of the polynomial $X^4 + 2X^2 - 2$

1. over a field of of three elements;

2. over a field with nine elements;

3. over $\mathbb{Q}$.

Exercise 11.2.7. Let j be a zero of $X^3 + X + 1 \in \mathbb{F}_2[X]$, and let $f(X) = X^4 + jX^3 + (j+1)X^2 + X + j^2 \in \mathbb{F}_2(j)[X]$.
(a) Prove that $\mathbb{F}_8 = \mathbb{F}_2(j)$.
(b) Give the table of multiplication for $\mathbb{F}_8$.
(c) Prove that f is irreducible over $\mathbb{F}_8$.
(d) Let α be a zero of f in a splitting field.

(d1) Prove that $L = \mathbb{F}_8(\alpha)$ is the splitting field of f.
(d2) Determine the number of elements in L.
(d3) Determine the Galois group of f over $\mathbb{F}_8$.

(e) Prove that there is a unique field E strictly between $\mathbb{F}_8$ and L, and determine a primitive element for $E/\mathbb{F}_8$. Useful equalities in the calculation :

$$\begin{array}{rcl}\alpha^{64} & = & \alpha^3 + (j+1)\alpha + j^2 + 1 \\ \alpha^{128} & = & j\alpha^3 + \alpha^2 + j^2\alpha + j^2 + j \\ \alpha^{196} & = & (j+1)\alpha^3 + j^2\alpha + 1\end{array}$$

(f) Determine the minimal polynomial of the extension $E/\mathbb{F}_8$.