Introduction to Algebra

Hartmut Laue

Mathematisches Seminar der Universität Kiel 2016
## Contents

1. Roots, radicals and zeros of polynomials 5
2. Algebraic structures and the extension principle 25
3. Cosets and homomorphisms 47
4. Field extensions 65
5. Solubility by radicals 89
6. Representations 101
7. Characters 117
8. Group actions 131
Prerequisites

It is assumed that the reader is familiar with basic concepts of Linear Algebra: vector spaces over fields, bases, linear dependence and independence, linear mappings and matrices. Furthermore, it will be helpful if the reader has taken notice of the notions of a ring and of a group, but no deeper results about these are expected. Recall, however, the so-called

**Subgroup Criterion:** A subset $H$ of a group $G$ is a subgroup of $G$ if and only if $H \neq \emptyset$ and $xy^{-1} \in H$ for all $x, y \in H$.

Basic terminology (terms like homomorphism, monomorphism, epimorphism, isomorphism, automorphism, endomorphism, equivalence relation, partition, operations, compositions of mappings) should be known although even these are briefly recalled in due course.

Some notational hints will be useful as these may differ from other authors's style:

- $\mathbb{N} = \{1, 2, 3, 4, \ldots\}$, $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$,
- $\mathfrak{n} = \{k | k \in \mathbb{N}, \ 1 \leq k \leq n\}$ for all $n \in \mathbb{N}_0$,
- $\check{K} := K \setminus \{0_K\}$ if $K$ is a ring.

The image of an element $x$ under a function $f$ is in this course generally denoted by $xf$ (with “inevitable” exceptions for historic reasons like, e.g., $f = \exp$). Other known choices of notation are the traditional $f(x)$ or $x^f$. The traditional left-hand notation of mappings certainly has the advantage of being widely spread, but as it seems only for purely historic reasons. If the choice of notation is oriented by principles like consistency with foundations and smoothness of presentation from a systematic point of view – in particular, regarding advanced theories in non-commutative algebra like Group Theory, Combinatorial Algebra –, a closer look shows that the right-hand notation is doubtlessly preferable. Compositions of functions are a fundamental tool, and it is here where those differences in notation have noticeable consequences. In accordance with the aforementioned choice of the right-hand notation, the notation $fg$ for the composition of mappings $f, g$ has the meaning to apply first $f$, then $g$ to an element $x$ of the domain of $f$. Hence $x(fg) = (xf)g$.

(Brackets are not needed here so that we may also just write $xfg$.)
1 Roots, radicals and zeros of polynomials

Let \( f = (f_n)_{n \in \mathbb{N}_0}, \ g = (g_n)_{n \in \mathbb{N}_0} \) be number sequences. We set

\[
(f_n)_{n \in \mathbb{N}_0} \uparrow (g_n)_{n \in \mathbb{N}_0} := (f_n + g_n)_{n \in \mathbb{N}_0},
\]

\[
(f_n)_{n \in \mathbb{N}_0} \bullet (g_n)_{n \in \mathbb{N}_0} := \left( \sum_{k=0}^{n} f_k g_{n-k} \right)_{n \in \mathbb{N}_0} \quad \text{(Cauchy product)}.
\]

This defines two operations on the set \( \Pi \) of all number sequences, and \((\Pi, \uparrow, \bullet)\) is a ring, i.e., \((\Pi, \uparrow)\) is an abelian group, \(\bullet\) is associative, and both distributive laws hold.

The sequence \( e \) where \( e_0 = 1, \ e_n = 0 \) for all \( n \in \mathbb{N}_0 \), is neutral with respect to \(\bullet\): If \((f_n)_{n \in \mathbb{N}_0} \in \Pi, \ n \in \mathbb{N}_0 \), then

\[
\sum_{k=0}^{n} e_k f_{n-k} = e_0 f_{n-0} + e_1 f_{n-1} + \cdots + e_n f_0 = f_n + 0 + \cdots + 0 = f_n,
\]

likewise \( \sum_{k=0}^{n} f_k e_{n-k} = f_n \), hence \( e \bullet (f_n)_{n \in \mathbb{N}_0} = (f_n)_{n \in \mathbb{N}_0} = (f_n)_{n \in \mathbb{N}_0} \bullet e \). A ring which contains a multiplicatively neutral element (also called unit element) is called unitary if this is different from the zero element.\(^1\) Thus \((\Pi, +, \bullet)\) is a unitary ring. Put

\[
\Pi_0 := \{ f | f \in \Pi, \ \exists n \in \mathbb{N}_0 \ \forall m \in \mathbb{N}_{>n} \ \ f_m = 0 \}.
\]

The defining condition means that almost all sequence values are zero. The structure \((\Pi_0, \uparrow)\) is an abelian group, and \(\Pi_0\) is closed with respect to \(\bullet\). Moreover, \(e \in \Pi_0\). Therefore \(\Pi_0\) is a subring of \(\Pi\) with a unit element, and the latter coincides with the unit element of \(\Pi\).\(^2\) A unitary subring of a unitary ring is called unital in this course if both rings have the same unit element.

The element \( t := (0, 1, 0, 0, \ldots) \) of \(\Pi\) plays a special role. We calculate its powers and obtain from the definition of \(\bullet\)

\[
t^2 = \ t \bullet t = (0, 0, 1, 0, 0, \ldots)
\]
\[
t^3 = \ t^2 \bullet t = (0, 0, 0, 1, 0, 0, \ldots)
\]
\[
\vdots
\]
\[
t^k = \ t^{k-1} \bullet t = (0, \ldots, 0, 1, 0, 0, \ldots)
\]

\(^1\)If the additively neutral element is at the same time multiplicatively neutral, it turns out easily that it is the only ring element, therefore the ring truly “trivial”.

\(^2\)This is by no means always the case. As a simple example consider the ring \(\mathbb{Z} \oplus \mathbb{Z}\) with componentwise addition and multiplication. The element \((1, 1)\) is multiplicatively neutral. The subring \(\mathbb{Z} \oplus \{0\}\) is unitary, but its unit element \((1, 0)\) is distinct from \((1, 1)\).
for all \( k \in \mathbb{N} \). The first power \( t^1 \) equals \( t \). The power \( t^0 \) finally is a special case of a so-called *empty product*, a product without any factor. In every structure with a neutral element, the empty product (in case of additive notation, the *empty sum*) is defined to be that element. Thus we complete our list by

\[
\begin{align*}
t^1 &= (0, 1, 0, 0, \ldots) \\
t^0 &= (1, 0, 0, 0, \ldots)
\end{align*}
\]

Let \( K \) be the domain of numbers over which the set \( \Pi \) of sequences was considered. Defining

\[
\forall c \in K \forall f \in \Pi \quad c \cdot f := (cf_n)_{n \in \mathbb{N}_0},
\]

we have for all nonzero \( f \in \Pi_0 \) the finite additive decomposition

\[
f = (f_0, f_1, f_2, \ldots) = f_0(1, 0, 0, 0, \ldots) + f_1(0, 1, 0, 0, \ldots) + \cdots + f_n(0, \ldots, 0, 1, 0, 0, \ldots)
\]

where \( n := \max \{ k \vert k \in \mathbb{N}_0, \ f_n \neq 0 \} \). Thus every \( f \in \Pi_0 \) is a \( K \)-linear combination of powers of \( t \). If \( c_0, c_1, \ldots, c_n \in K \), \( c_n \neq 0 \) and \( c_0 t^0 + c_1 t^1 + \cdots + c_n t^n = f \), then \( c_0 = f_0 \), \( c_1 = f_1 \), \ldots, \( c_n = f_n \), for the left hand sides \((c_0, \ldots, c_n, 0, 0, \ldots)\). It follows that the representation of \( f \in \Pi_0 \) as a \( K \)-linear combination of powers of \( t \) with scalars \( \neq 0 \) is unique. Note that the zero element of \( \Pi_0 \) is represented as the empty sum.

1.1 Definition. Let \( K \) be a commutative unitary ring, \( V \) an abelian group, \( \cdot : K \times V \rightarrow V \) an (external) operation such that the laws “known from vector spaces” hold, i.e., \((c + d) \cdot v = c \cdot v + d \cdot v, \ c \cdot (u + v) = c \cdot u + c \cdot v, \ (cd) \cdot v = c \cdot (d \cdot v), \ 1_K \cdot v = v\) for all \( c, d \in K, u, v \in V \). Then \( V \) is called a \( K \)-space.\(^3\) A subset \( T \) of \( V \) is called a \( K \)-generating system of \( V \) if every element of \( V \) is a \( K \)-linear combination of elements of \( T \). If for every finite subset \( S \) of \( T \) and for all choices of \( c_v, d_v \in K \) \((v \in S)\) the implication

\[
\sum_{v \in S} c_v v = \sum_{v \in S} d_v v \Rightarrow \forall v \in S \quad c_v = d_v
\]

holds true, then \( T \) is called \( K \)-linearly independent. A \( K \)-linearly independent \( K \)-generating system of \( V \) is called a \( K \)-basis of \( V \).

An inspection of all foregoing steps shows that we may start from an arbitrary commutative unitary ring \( K \) which need not necessarily be a domain of numbers and obtain our \( K \)-spaces \( \Pi, \Pi_0 \) in this more general setting. Furthermore, the set \( T := \{ t^0, t^1, t^2, \ldots \} \) is a \( K \)-basis of \( \Pi_0 \). Clearly, \( \Pi_0 \) is a \( K \)-subspace of \( \Pi S \), where we borrow the meaning of this term in the obvious way from vector space theory.

Now suppose that \( V \) a \( K \)-space. Let \( \bullet \) be an operation on \( V \) with the properties

\[
(i) \quad \forall u, v, w \in V \quad u \bullet (v + w) = u \bullet v + u \bullet w, \quad (v + w) \bullet u = v \bullet u + w \bullet u,
\]

\(^3\)Usually, the point \( \cdot \) for the multiplication between “scalars” from \( K \) and “vectors” from \( V \) is omitted. In this definition, however, we used it in order to make a distinction between that multiplication and the internal one within \( K \).
\(\forall c \in K \forall u, v \in V \quad c \cdot (u \bullet v) = (c \cdot u) \bullet v = u \bullet (c \cdot v).\)

Then \(V\) (more precisely, \((V, +, \bullet)\)) is called a \(K\)-algebra. It is called associative (commutative resp.) if \(\bullet\) is associative (commutative resp.). In every \(K\)-algebra \(V\) we write \(V\) for the set of all nonzero elements of \(V\).

**Examples.** (1) Let \(L\) be a unitary commutative ring containing \(K\) as a unital subring. Then \(L\) is an associative and commutative unitary \(K\)-algebra. The trivial case of \(L = K\) shows that \(K\) itself may be viewed as a \(K\)-algebra.

(2) For every \(n \in \mathbb{N}\), the set of all \(n \times n\) matrices over \(K\) (for which we shall write \(K^{n \times n}\) throughout) is an associative unitary \(K\)-algebra with respect to the usual matrix operations as addition and multiplication (within \(K^{n \times n}\)) and componentwise multiplication by a scalar as product between \(K\) and \(K^{n \times n}\). For \(n > 1\) it is not commutative.

(3) \(\Pi, \Pi_0\) (as constructed above) are associative and commutative unitary \(K\)-algebras.

Let \(V\) be an associative unitary \(K\)-algebra which contains an element \(t\) such that the elements \(t^j\) \((j \in \mathbb{N}_0)\) are mutually distinct and form a \(K\)-basis of \(V\). Then \(V\) is called a polynomial ring over \(K\) in the indeterminate (or variable) \(t\). A polynomial over \(K\) is an element of a polynomial ring over \(K\). A polynomial of the form \(ct^j\) where \(c \in K\), \(j \in \mathbb{N}_0\), is called a monomial in \(t\). It follows from the definition that polynomial rings are commutative.

We have shown that \(\Pi_0\) is a polynomial ring over \(K\). It does not contain \(K\) as a unital subring, but the special sequences \((c, 0, 0, \ldots) \in \Pi\) form a substructure in which calculations take place as in \(K\): For any \(c, d \in K\) we have

\[(c, 0, 0, \ldots) \dagger (d, 0, 0, \ldots) = (c + d, 0, 0, \ldots), \quad (c, 0, 0, \ldots) \bullet (d, 0, 0, \ldots) = (cd, 0, 0, \ldots).\]

In other words, the mapping \(\iota : K \to \Pi_0, \; c \mapsto (c, 0, 0, \ldots)\), is an injective homomorphism with respect to addition, internal multiplication, and external multiplication ("scalar times vector"). Note that when \(K\) is considered as a \(K\)-algebra, both scalars and vectors are the elements of \(K\) and there is no difference in calculating an internal or an external product. An injective homomorphism is called a monomorphism. A \(K\)-algebra homomorphism from a \(K\)-algebra into another structure with two internal operations ("addition" and "multiplication") and an external multiplication with elements of \(K\) is a mapping which is a homomorphism with respect to the two additions and with respect to the two multiplications, and which satisfies the homogeneity rule known from \(K\)-linear mappings in vector space theory.\(^4\) For \(\iota\), we have for all \(c, d \in K\)

\[c \cdot d = cd \mapsto (cd, 0, 0, \ldots) = (cd, c0, c0, \ldots) = c \cdot (d, 0, 0, \ldots),\]

\(^4\)The second algebraic structure need not be assumed to be a \(K\)-algebra to apply this definition. In many applications, however, this will be the case. As we shall see later (3.11), there is a general reason for the fact that the second structure necessarily at least involves (contains) a \(K\)-algebra which turns out to be closely related to the initially given one: the set of all images under \(\varphi\).
and indeed \((d, 0, 0, \ldots)\) is the image of \(d\) under \(\iota\).

In this course, usually the right notation for mappings will be used, i.e., the image of an element \(x\) under a mapping \(\varphi\) will be denoted by \(x\varphi\). Thus we may express the homogeneity of \(\iota\) as follows: \(\forall c, d \in K \quad (c \cdot d) \iota = c \cdot d \iota\).

The elements of \(\Pi\) are called power series over \(K\), \(\Pi\) a power series algebra over \(K\) in the indeterminate (or variable) \(t\). The set-theoretic situation which we are studying, including the \(K\)-algebra monomorphism \(\iota\), may be visualized by the following diagram:

\[\begin{array}{c}
\mathbb{K} \\
\downarrow \iota \\
\mathbb{K}_0 \\
\downarrow \iota \\
\mathbb{K}_1 \\
\downarrow \iota \\
\cdots \\
\downarrow \iota \\
\mathbb{K}_k \\
\downarrow \iota \\
\mathbb{K}_n \\
\end{array}\]

Now let \(V\) be an arbitrary polynomial ring over \(K\) in the indeterminate \(t\). The elements of \(V\) are uniquely given in the form \(\sum_{j=0}^{n} c_j t^j\) where \(n \in \mathbb{N}_0\), \(c_j \in K\) for all \(j\), \(c_n \neq 0\).\(^5\) Note that the zero element of \(V\) is represented by the empty sum. For every element \(f \in V\), the number \(n\) in the representation \(f = \sum_{j=0}^{n} c_j t^j\) with \(c_0, \ldots, c_n \in K\), \(c_n \in \hat{K}\), is called its degree with respect to \(t\). We use the notation \(\deg f\) for the degree of a nonzero element \(f \in V\) (\(\deg f\) if the reference to \(t\) is not obvious from the context). If \(c_n = 1_K\), \(f\) is called normed.

1.1.1. \(\deg fg \leq \deg f + \deg g\) for all \(f, g \in V\) with \(fg \neq 0_V\).

**Proof.** Let \(n, m \in \mathbb{N}_0\), \(c_0, \ldots, c_n, d_0, \ldots, d_m \in K\) such that \(c_n \neq 0_K\), \(d_m \neq 0_K\) and \(f = \sum_{j=0}^{n} c_j t^j\), \(g = \sum_{j=0}^{m} d_j t^j\). Then \(fg\) is a \(K\)-linear combination of powers \(t^k\) of \(t\) with \(k \leq n + m\). The claim follows. \(\Box\)

It should be noted that the exponent \(n+m\) of \(t\) arises only if \(c_n t^n\) and \(d_m t^m\) are multiplied: \(c_n t^n d_m t^m = c_n d_m t^{n+m}\). We know that \(c_n, d_m \neq 0_K\), but in general it may happen that still \(c_n d_m = 0_K\).\(^6\) An element \(a \in K\) is called a zero divisor if there exists an element \(b \in K\) such that \(ab = 0_K\). In many rings \(K\), however, \(0_K\) is the only zero divisor, i.e., a product of nonzero elements is never zero. For these rings, the inequality in 1.1.1 thus becomes an equality because the coefficient of the term with the highest possible

\(^5\)The elements of the power series algebra \(\Pi\) are similarly written in the form \(\sum_{n \in \mathbb{N}_0} c_k t^k\). But note that this is just a formal notation which has not the meaning of a sum while \(\sum_{k=0}^{n} c_k t^k\) is indeed an ordinary sum of \(n+1\) monomials. The notation for power series follows the pattern given by polynomials but should not give rise to the inadequate idea of a “sum of infinitely many monomials”!

\(^6\)For example, if \(K = \mathbb{Z} \oplus \mathbb{Z}\) with componentwise addition and multiplication, \(c_n = (1, 0), d_m = (0, 1)\). Clearly, this phenomenon does not occur if \(c_n = 1_K\) or \(d_m = 1_K\). Hence we have equality in 1.1.1 if \(f\) or \(g\) is normed.
exponent \( n + m \) cannot vanish. A commutative unitary ring \( K \) is called an integral domain if \( K \) is multiplicatively closed, which is obviously just a reformulation of the aforementioned property. We therefore have

**1.1.2.** If \( K \) is an integral domain, then so is the polynomial ring \( V \), and \( \deg fg = \deg f + \deg g \) for all \( f, g \in V \). \( \square \)

Considering \( \deg \) as a function from \( \hat{V} \) into \( \mathbb{N}_0 \), the contents of 1.1.2 may also be expressed in the following form, observing that every \( n \in \mathbb{N}_0 \) occurs as the degree of some polynomial (e.g., of \( t^n \)):

**1.1.2’** If \( K \) is an integral domain, then the function \( \deg \) is an epimorphism\(^7\) of \( (\hat{V}, \cdot) \) onto \((\mathbb{N}_0, +)\). \( \square \)

Every field, more generally every unital subring of a field, is an integral domain. A typical example of an integral domain which is not a field is the familiar ring \( \mathbb{Z} \). The ring \( \mathbb{Z} \oplus \mathbb{Z} \) is commutative and unitary but not an integral domain (cf. footnote 6).

Let \( f = \sum_{j=0}^n c_j t^j \in V \) (where \( c_0, \ldots, c_n \in K \), \( c_n \neq 0_K \)), and let \( b \) be an element of a unitary associative \( K \)-algebra \( A \). We set

\[
f(b) := \sum_{j=0}^n c_j b^j
\]

and call the mapping \( F_b : V \to A, \ f \mapsto f(b) \), the substitution (mapping) of \( b \). If \( f(b) = 0_A \), then \( b \) is called a zero (note, however, footnote 9) of \( f \) in \( A \). In other words, the zeros of \( f \) in \( A \) are the solutions in \( A \) of the equation \( \sum_{k=0}^n c_j x^j = 0_A \).

**1.1.3.** Let \( A \) be a unitary associative \( K \)-algebra, \( b \in A \). Then \( F_b \) is a unital \( K \)-algebra homomorphism.

*Proof.* If \( f, g \in V, \ c \in K, \ f = \sum_{j=0}^n c_j t^j, \ g = \sum_{i=0}^m d_i t^i \) (where \( n, m \in \mathbb{N}_0 \), \( c_j, d_i \in K \)) we trivially have \((f \cdot cg)(b) = f(b) + cf(b)\), and moreover \((fg)(b) = \sum_{k=0}^{m+n} (\sum_{j=0}^n c_j d_{k-j}) b^k = \sum_{j=0}^n c_j b^j \sum_{i=0}^m d_i b^i = f(b)g(b)\) (where \( c_j : = 0_K \) if \( j > n \), \( d_i = 0_K \) if \( i > m \)). \( \square \)

We write \( K[t] \) for a polynomial ring over \( K \) in the indeterminate \( t \) which contains \( K \) as a unital subring.\(^8\) An element \( f \in K[t] \) is called pure if there exist an \( n \in \mathbb{N} \) and an \( a_0 \in K \) such that \( f = t^n + a_0 \). Assume that \( K \) is a unital subalgebra of a unitary associative \( K \)-algebra \( A \). An element \( r \in A \) is called a radical over \( K \) if \( r \neq 0_K \) and there exists a pure polynomial \( f \in K[t] \) such that \( f(r) = 0_K \). By definition, there exists then a positive integer \( n \) such that \( r^n \in K \). Any such \( n \) will be called an exponent for \( r \) (with respect to \( K \)). The most important case in this course is that of a field \( A \) containing \( K \) as a subfield.

\(^7\)Recall that a surjective homomorphism is called an epimorphism.

\(^8\)Hence \( t^0 = 1_K \). The polynomial ring \( \Pi_0 \) constructed above does not have this property so that in this moment the existence of a polynomial ring containing \( K \) is unclear. But this is only a temporary gap as will be seen in due course (2.9.2).
For example, roots \( \neq 0_K \) are special radicals. Recall from basic Analysis what a root is: If \( c \in \mathbb{R}_{\geq 0}, n \in \mathbb{N} \), then there exists a unique non-negative real number the \( n \)-th power of which equals \( c \). This number is called the \( n \)-th root of \( c \) and denoted by \( \sqrt[n]{c} \).

**In this course, the notation** \( \sqrt[n]{c} \) **is used only if** \( c \) **is a non-negative real number.**

Then \( \sqrt[n]{c} \) is a zero of the pure polynomial \( t^n - c \), thus for \( c \neq 0 \) it is a radical over every subfield \( K \) of \( \mathbb{R} \) with \( c \in K \). (If \( c \notin K \), \( t^n - c \) is not a polynomial over \( K \!).

**Examples.**

(1) \( t^2 - 3 \) has two zeros in \( \mathbb{R} \), the square root of 3, \( \sqrt{3} \), and its additive inverse \(-\sqrt{3} \) which of course does not satisfy the definition of a root of a real number.

Both numbers are radicals over \( \mathbb{Q} \), and every even number is an exponent for them.

(2) The complex number \( i\sqrt{2} \) is a radical over \( \mathbb{Q} \) of exponent 6, because it is a zero of the pure polynomial \( t^6 + 4 \in \mathbb{Q}[t] \). Over \( \mathbb{R} \), 2 is an exponent for it as it is a zero of the pure polynomial \( t^2 + \sqrt{4} \in \mathbb{R}[t] \).

(3) \( 1 + \sqrt{2} \) is not a radical over \( \mathbb{Q} \): A trivial induction shows that for all \( n \in \mathbb{N} \) there exist \( a, b \in \mathbb{N} \) such that \((1 + \sqrt{2})^n = a + b\sqrt{2} \). Assuming \((1 + \sqrt{2})^n \in \mathbb{Q} \) for some \( n \in \mathbb{N} \) we thus would have \( a, b \in \mathbb{N} \), \( c \in \mathbb{Q} \) such that \( a + b\sqrt{2} = c \), hence \( \sqrt{2} = \frac{c-a}{b} \in \mathbb{Q} \), a contradiction.

(4) The roots of the polynomial \( t^2 + t + 1 \) in \( \mathbb{C} \) are radicals over \( \mathbb{Q} \), because in \( \mathbb{Q}[t] \) we have the equation

\[
(t^2 + t + 1)(t - 1) = t^3 - 1.
\]

We conclude that every zero of \( t^2 + t + 1 \) is also a zero of \( t^3 - 1 \), hence a radical of exponent 3 over \( \mathbb{Q} \).

The equation in (4) is of course the special case \( n = 3 \) of the general statement that, for all \( n \in \mathbb{N} \),

\[
(t^{n-1} + \cdots + t + 1_K)(t - 1_K) = t^n - 1_K,
\]

for any unitary commutative ring \( K \). The zeros of \( t^n - 1_K \) in \( K \) are called the \( n \)-th roots of unity in \( K \).\(^{10} \) If \( K \) is not specified, the term refers to the case \( K = \mathbb{C} \).

**1.1.4.** Let \( K \) be a unital subring of an integral domain \( L \). Let \( f, g, h \in K[t], f = gh, b \in L \) and \( f(b) = 0_K \). Then \( g(b) = 0_K \) or \( h(b) = 0_K \).

**Proof.** If \( f(b) = 0_K \), then \( g(b)h(b) = 0_K \), hence the claim as \( L \) has no zero divisors. \( \square \)

\(^9\) The terminology introduced here is not used in the same way by all authors. It is common use to talk about “roots of a polynomial” instead of zeros. In order to prevent any confusion with roots of real numbers we avoid this usage of the term “root” throughout this course. (Exception: footnote 10.)

\(^{10}\) Here we deviate from our convention about the usage of the term “root” in favour of the traditional terminology.
It follows that the \( n \)-th roots of unity in an integral domain \( L \) are \( 1_K \) and the zeros of \( t^{n-1} + \cdots + t + 1 \) in \( L \). The 3rd roots of unity \( \neq 1 \) in \( \mathbb{C} \) (see Example (4) on p. 10) are the complex solutions of the equation

\[
0 = x^2 + x + 1 = \left(x + \frac{1}{2}\right)^2 + \frac{3}{4}.
\]

This familiar transformation reduces the determination of the zeros in the form \( a + bi \) with \( a, b \in \mathbb{R} \) to the treatment of the simple case of a pure polynomial of degree 2 (namely, \( t^2 + \frac{3}{4} \)). Thus we obtain the two solutions

\[
x_1 = -\frac{1}{2} + \frac{1}{2}i\sqrt{3}, \quad x_2 = -\frac{1}{2} - \frac{1}{2}i\sqrt{3}.
\]

The case \( n = 4 \) is easier as we have \( t^4 - 1 = (t^2 + 1)(t + 1)(t - 1) \) so that the complex 4th roots of unity are \( 1, -1, i, -i \). But it is obvious that we will not always be lucky if the degree \( n \) increases and we want to describe the \( n \)-th roots of unity, as above, as an \( \mathbb{R} \)-linear combination of 1 and \( i \). There is, however, a different idea which quickly leads to this standard representation. Recall the fundamental equation of the complex exponential function: \( \exp(z + z') = \exp z \exp z' \) for all \( z, z' \in \mathbb{C} \), which means, in other words:

1.1.5. \( \exp \) is a homomorphism of \((\mathbb{C}, +)\) into \((\mathbb{C}, \cdot)\). \( \square \)

In particular, by a standard induction we obtain from 1.1.5

1.1.6. \( \exp(nz) = (\exp z)^n \) for all \( n \in \mathbb{N}_0, z \in \mathbb{C} \). \( \square \)

The replacement of \( z \) by \( iz \) leads to Moivre’s formula:

\[
\cos nz + i \sin nz = (\cos z + i \sin z)^n \quad \text{for all } n \in \mathbb{N}_0, z \in \mathbb{C},
\]

making use of Euler’s formula, the fundamental connection between \( \exp, \sin, \cos \):

\[
\exp(iz) = \cos z + i \sin z \quad \text{for all } z \in \mathbb{C}. \quad \text{(11)}
\]

We exploit this equation for real values of \( z \) and obtain for arbitrary \( a, b \in \mathbb{R} \) the chain of equivalences

\[
\exp(a + bi) = 1 \iff e^a(\cos b + i \sin b) = 1 \iff e^a \cos b = 1, e^a i \sin b = 0 \iff b \in \mathbb{Z} \cdot 2\pi, a = 0
\]

The kernel of a group homomorphism \( f \), denoted by \( \ker f \), is the set of all group elements which have the same image under \( f \) as the neutral element. In the case of the homomorphism \( \exp \) (see 1.1.5), we therefore have \( z \in \ker \exp \iff \exp z = \exp 0 \iff \exp z = 1 \)

which is equivalent to \( z \in \mathbb{Z} \cdot 2\pi i \) by what we have just shown. This proves:

(11) Substituting \( t \) by \( iz \) in the formal power series \( \sum_{n \in \mathbb{N}_0} \frac{(iz)^n}{n!} \) (which converges everywhere in the complex plane and defines the complex exponential function), we obtain

\[
\exp(iz) = \sum_{n=0}^{\infty} \frac{(iz)^n}{n!} = \sum_{k=0}^{\infty} \frac{(-1)^k z^{2k}}{(2k)!} + i \sum_{k=1}^{\infty} \frac{(-1)^{k-1} z^{2k-1}}{(2k-1)!} = \cos z + i \sin z.
\]

In particular, \( \exp(2\pi i) = 1 \), which is the famous Eulerian equation, combining in the form \( e^{2\pi i} = 1 \) “the five most important numbers of Mathematics” 1, 2, \( e, \pi, i \).
1.1.7. $\ker \exp = \mathbb{Z} \cdot 2\pi i$. \hfill \qed

It is easily seen that the mapping $\kappa : ]0, 2\pi[ \to K(0, 1), \ c \mapsto \exp(ci)$, is a bijection, commonly called the trigonometric parameterization of the unit circle. \hfill 12

Now put for every $n \in \mathbb{N}$

$$R_n(1) := \{w | w \in \mathbb{C}, \ w^n = 1\}, \quad R(1) := \bigcup_{n \in \mathbb{N}} R_n(1).$$

For all $w \in R_n(1)$ we have $1 = |w^n| = |w|^n$, hence $|w| = 1$. Thus $R_n(1) \subseteq K(0, 1)$. Now let $z \in K(0, 1)$ and $c \in ]0, 2\pi]$ such that $z = \exp(ci)$. Then, by 1.1.6 and 1.1.7,

$$z \in R_n(1) \iff \exp(nci) = 1 \iff nc \in \mathbb{Z} \cdot 2\pi \iff \exists k \in \mathbb{Z} \quad z = \exp \frac{k \cdot 2\pi i}{n}$$

The values $\exp \frac{k \cdot 2\pi i}{n}$ where $k > n$ or $k < 1$ are just repetitions of the values where $k \in \mathbb{N}$. Therefore,

1.1.8. $\forall n \in \mathbb{N} \quad R_n(1) = \{\exp \frac{k \cdot 2\pi i}{n} | k \in \mathbb{Z} \}$. \hfill \qed

From this description of $R_n(1)$ we obtain a beautiful geometric interpretation for the $n$-th roots of unity: Recall that each point $P$ of the unit circle determines an angle, given by the line segment connecting it with the origin and the positive real axis. Its radian measure is given by the length of the corresponding arc on the unit circle (between $P$ and the point representing 1). If $P$ represents an $n$-th root of unity $\exp \frac{k \cdot 2\pi i}{n}$, the corresponding angle corresponds to $\frac{k}{n}$-th of the complete periphery which is of length $2\pi$. Therefore the $n$-th roots of unity are represented by the vertices of the regular $n$-gon inscribed in the unit circle such that the point representing 1 is one of them.

We are going to give a specifically algebraic interpretation of $R_n(1)$ and start with the following definition of the most general of all algebraic structures:

12$\kappa$ is injective as $ck = c'k$ (where $0 < c < c' \leq 2\pi$) would imply that $\exp((c' - c)i) = 1$, hence $c' - c \in \mathbb{Z} \cdot 2\pi$ by 1.1.7, a contradiction because $0 < c' - c < 2\pi$. Furthermore, let $a, b \in \mathbb{R}$ such that $a + bi \in K(0, 1)$. Then $a^2 + b^2 = 1$ so that $|a|, |b| \leq 1$. There are exactly two values $c \in ]0, 2\pi]$ such that $\cos c = a$, and for both of them we have $b^2 = 1 - (\cos c)^2 = (\sin c)^2$, hence $b = \sin c$ or $b = -\sin c$. As $\kappa$ is injective, the sine values for those two elements $c$ cannot coincide. Hence we may choose $c \in ]0, 2\pi]$ such that $\cos c = a, \sin c = b$. It follows that $ck = a + bi$. Thus $\kappa$ is bijective.
1.2 Definition. A magma is a pair \((X, \circ)\) where \(X\) is any set and \(\circ\) is any operation on \(X\). A submagma of \((X, \circ)\) is then a subset of \(X\) which is closed with respect to \(\circ\). If \(S\) is a submagma of \((X, \circ)\) such that \((S, \circ)\) is a semigroup (i.e., such that the associative law holds in \((S, \circ)\)), then \(S\) is called a subsemigroup of \((X, \circ)\). If there exist neutral elements \(e_X\) of \((X, \circ)\), \(e_S\) of \((S, \circ)\), we say that \(S\) is a unital submagma if \(e_X = e_S\). A subgroup of \((X, \circ)\) is a submagma \(S\) such that \((S, \circ)\) is a group. A semigroup with a neutral element is called a monoid.

The group \((\mathbb{Z}, +)\) has many submags which are no subgroups, for example, any subset \(\mathbb{Z}_{\geq y}\) where \(y \in \mathbb{Z}\). For every integer \(a\), the set of its multiples \(a\mathbb{Z}\) is a subgroup of \((\mathbb{Z}, +)\) and, as is easily seen, the smallest subgroup containing \(a\). If \(a \neq 0\), the mapping \(\mathbb{Z} \to a\mathbb{Z}, \, z \mapsto az\), is an isomorphism. If \(b \in \mathbb{Z}\), then \(a\mathbb{Z} = b\mathbb{Z}\) if and only if \(a|b\) and \(b|a\), i.e., if and only if \(b \in \{a, -a\}\). It is not hard to prove that every subgroup of \((\mathbb{Z}, +)\) has the form \(n\mathbb{Z}\) for a uniquely determined non-negative integer \(n\), where the uniqueness is obvious from what we just observed:13

1.2.1. Let \(H\) be a subgroup of \((\mathbb{Z}, +)\). Then there exists a unique non-negative integer \(n\) such that \(H = n\mathbb{Z}\).

Proof. If \(H = \{0\}\) the claim is trivial. Otherwise \(H\) contains a positive integer. Let \(n := \min H \cap \mathbb{N}\). Then \(n\mathbb{Z} \subseteq H\). Assuming that \(n\mathbb{Z} \neq H\), let \(h \in H \setminus n\mathbb{Z}\). Considering the largest multiple of \(n\) which is smaller than \(h\) exhibits an integer \(z\) such that \(0 < h - nz < n\), a contradiction because \(h - nz \in H \cap \mathbb{N}\). Hence \(n\mathbb{Z} = H\).

We define for all \(U, V \subseteq X\) the product set of \(U\) and \(V\) by

\[U \circ V := \{u \circ v | u \in U, v \in V\}.\]

Then \((\mathfrak{P}(X), \circ)\) is a magma and \(f : X \to \mathfrak{P}(X), x \mapsto \{x\}\), is a monomorphism of \((X, \circ)\) into \((\mathfrak{P}(X), \circ)\) as \(f\) is injective and \((x \circ x')f = \{x \circ x'\} = \{x\} \circ \{x'\} = xf \circ x'f\) for all \(x, x' \in X\). If \((X, \circ)\) has a neutral element \(e\), then \(X \circ X = X\) because \(X \circ X \subseteq X = X \circ \{e\} \subseteq X \circ X\).

---

13 Note that it is not required in this definition that \((X, \circ)\) be a group. For example, our definition allows us to say that, for a field \(K\), \(K\) is a subgroup of \((K, \cdot)\), etc. It may happen that \((X, \circ)\) has no neutral element, and if it has one, it may happen that it does not coincide with that of a subgroup \(S\). On the other hand, if \((X, \circ)\) is a group, then it is easily seen (and should be well known) that all of its subgroups have the same neutral element. Every magma has at most one neutral element. Thus, in a magma with neutral element \(e\), no explicit reference to \(e\) is needed when one talks about an inverse element \(y\) of an element \(x\), i.e., an element with the property that \(x \circ y = e = y \circ x\). One should bear in mind that, in general, such an element need not exist, and if it does, there may be more than one inverse of an invertible element. Again, if \((X, \cdot)\) is a group, then every element has a uniquely determined inverse. Therefore, given a group, it is reasonable to choose a notation for the (!) inverse of \(x\) which directly refers to \(x\). Multiplicatively, \(x^{-1}\) is the common notation for it, additively \(-x\). But note that this notation generally makes no sense if applied outside a group theoretic context where neither existence nor uniqueness of an inverse may be assumed.

14 Thus the group \((\mathbb{Z}, +)\) has the remarkable property that all of its subgroups \(\neq \{0\}\) are isomorphic to \((\mathbb{Z}, +)\).
1.3 Proposition. Let \((X, \circ), (Y, \cdot)\) be magmas and \(f\) a homomorphism of \((X, \circ)\) into \((Y, \cdot)\).

1. (1) \(Xf\) is a submagma of \((Y, \cdot)\).

2. If \((X, \circ)\) is a semigroup, then so is \((Xf, \cdot)\).

3. If \(e\) is neutral in \((X, \circ)\), then \(ef\) is neutral in \((Xf, \cdot)\). If \(x, x' \in X\) such that \(x \circ x' = x' \circ x\), then \(xf \cdot x'f = x'f \cdot xf\). If \(x'\) is an inverse of \(x\) in \((X, \circ)\), then \(x'f\) is an inverse of \(xf\) in \((Xf, \cdot)\).

4. If \((X, \circ)\) is a group, then \(Xf\) is a subgroup of \((Y, \cdot)\), and \(\ker f := \{x | x \in X, xf = ef\}\) is a subgroup of \((X, \circ)\) (where \(e\) denotes the neutral element of \((X, \circ)\)).

Proof. (1) Let \(y, y' \in Xf\). Then there exist \(x, x' \in X\) such that \(y = xf, y' = x'f\), and \(y \cdot y' = xf \cdot x'f = (x \circ x')f \in Xf\).

(2) Let \(y, y', y'' \in Xf\). Then there exist \(x, x', x'' \in X\) such that \(y = xf, y' = x'f, y'' = x''f\), and \((y \cdot y') \cdot y'' = (xf \cdot x'f) \cdot x''f = (x \circ x')f \cdot x''f = ((x \circ x') \circ x'')f = (x \circ (x' \circ x''))f = xf \cdot (x' \circ x'')f = xf \cdot (x'f \cdot x''f) = y \cdot (y' \cdot y'').\)

(3) Let \(y \in Xf\). Then there exists an \(x \in X\) such that \(y = xf\), and \(y \cdot ef = xf \cdot ef = (x \circ e)f = xf = y\), analogously \(ef \cdot y = y\). Clearly \(x \circ x' = x' \circ x\) implies \(xf \cdot x'f = (x \circ x')f = (x' \circ x)f = x'f \cdot xf\). This holds, in particular, if \(x'\) is an inverse of \(x\). In this case the mid terms equal \(ef\).

(4) The first claim follows from (1), (2), (3). Since \((X, \circ)\) is a group, we may apply the subgroup criterion for the second claim: Clearly, \(e \in \ker f\). If \(x, x' \in \ker f\), then \((xf)^{-1} = (ef)^{-1} = ef\), the latter being neutral, hence the inverse of itself. It follows by (3) that \((x^{-1} \circ x')f = x^{-1}f \cdot x'f = (xf)^{-1} \cdot x'f = (ef)^{-1} \cdot ef = ef\), thus \(x^{-1} \circ x' \in \ker f\).

Let us recall one of the most important examples of a group, the symmetric group on a set \(X\), consisting of all permutations of \(X\), endowed with the composition of mappings as operation. It is usually denoted by \(S_X\). If \((X, \circ)\) is a magma, the set Aut \((X, \circ)\) (commonly written simply as Aut \(X\)) of all automorphisms of \((X, \circ)\) clearly is a subset of \(S_X\). In a sense, the following simple remark is of fundamental importance, as it shows that studying an arbitrary algebraic structure by means of their automorphisms always means that a group comes into play:

1.3.1. Let \((X, \circ)\) be any magma. Then Aut \(X\) is a subgroup of \(S_X\).

Proof. We just have to observe that \(\text{id}_X \in \text{Aut} X\) (which is trivial) and \(\alpha, \beta \in \text{Aut} X\) implies that \(\alpha \beta \in \text{Aut} X\) and \(\beta^{-1} \in \text{Aut} X\). Clearly, \(\alpha \beta, \beta^{-1} \in S_X\). Compositions of homomorphisms are homomorphisms, and for all \(x, y \in X\) we have \(((x \beta^{-1}) \circ (y \beta^{-1})) \beta = x \circ y\), hence \((x \beta^{-1}) \circ (y \beta^{-1}) = (x \circ y) \beta^{-1}\). □

\(^{15}\)Clearly, if \((Y, \cdot)\) is a group and \(e_Y\) its neutral element, then \(\ker f = \{x | x \in X, xf = e_Y\}\) as then \(ef = e_Y\).
There is a second way to look at $\text{Aut} \, X$, observing first that $\text{End} \, X$, the set of all endomorphisms of $(X, \circ)$, is a monoid where the operation is given by composition, and then that $\text{Aut} \, X$ is the set of invertible elements of the monoid $\text{End} \, X$. It is easily seen that in every monoid, the set of invertible elements is a subgroup. Thus $\text{Aut} \, X$ arises as an example of this general remark.

If $(G, \circ)$ is a group, we write $H \leq G$ to say that $H$ is a subgroup of $(G, \circ)$ if due to the context there is no doubt about the operation $\circ$ to which this statement refers. We want to illustrate in which way 1.3(4) can be instrumentalized, by showing that, with respect to multiplication,

1.3.2. $R_d(1) \leq R_n(1) \leq R(1) \leq K(0, 1) \leq \mathbb{C}$ where $d, n \in \mathbb{N}$ such that $d|n$.

Proof. (a) $f : \hat{\mathbb{C}} \to \mathbb{R}$, $z \mapsto |z|$, is a multiplicative homomorphism and $\ker f = K(0, 1)$. Hence $K(0, 1) \subseteq \hat{\mathbb{C}}$. (b) $f : \mathbb{Q} \to \hat{\mathbb{C}}$, $x \mapsto \exp(x \cdot 2\pi i)$, is a homomorphism of $(\mathbb{Q}, +)$ into $(\hat{\mathbb{C}}, \cdot)$ and $\mathbb{Q}f = R(1)$. Hence $R(1) \leq \hat{\mathbb{C}}$. (c) $f : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$, $z \mapsto z^n$, is a multiplicative homomorphism and $\ker f = R_n(1)$. Hence $R_n(1) \leq \hat{\mathbb{C}}$. If $d|n$, then again of course $R_d(1) \leq \hat{\mathbb{C}}$, and additionally $R_d(1) \subseteq R_n(1)$ as $z^d = 1$ implies $z^n = 1$.

For a second proof of (c), observe that $f_n : \mathbb{Z} \to \hat{\mathbb{C}}$, $k \mapsto \exp\left(\frac{k2\pi i}{n}\right)$, is a homomorphism of $(\mathbb{Z}, +)$ into $(\hat{\mathbb{C}}, \cdot)$ and $\mathbb{Z}f_n = R_n(1)$. Hence $f_n$ is an epimorphism of $(\mathbb{Z}, +)$ onto $(R_n(1), \cdot)$, implying (c). Furthermore, 1.1.7 immediately allows to describe the kernel of $f_n$ as the set of all multiples of $n$. Summarizing, we have

1.3.3. $f_n$ is an epimorphism of $(\mathbb{Z}, +)$ onto $(R_n(1), \cdot)$, and $\ker f_n = n\mathbb{Z}$.

The subgroups $R_n(1)$ where $n \in \mathbb{N}$ are in fact the only finite subgroups of $(\hat{\mathbb{C}}, \cdot)$:

1.3.4. Let $U$ be a finite subgroup of $(\hat{\mathbb{C}}, \cdot)$. Then $U = R_n(1)$ where $n = |U|$.

Proof. 16 As $|U| = n$ it suffices to prove that $U \subseteq R_n(1)$. Let $x \in U$. Then $U = xU$, implying $\prod U = \prod_{u \in U}(xu) = x^n(\prod U)$. Hence $x^n = 1$, i.e., $x \in R_n(1)$.

A multiplication of $n$-th roots of unity may be visualized (thanks to 1.3.3) by adding their defining angles (given counter-clockwise between the connecting line with the origin and the positive real axis). Therefore the special $n$-th root of unity $w_1 = \exp\frac{2\pi i}{n}$, visualized as the “first upper neighbour” of the trivial root of unity 1, has the property that its powers range over all of $R_n(1)$, equivalently, that all powers $w_1^k$ where $k \in \mathbb{N}$ are distinct. The same holds for the “first lower neighbour” of 1 but certainly not for all $n$-th roots of unity. For example, 1 never has this property unless $n = 1$. For $n = 4$, the powers of $i$ and $-i$ range over all 4-th roots of unity while the powers of 1, −1 do not.

1.4 Definition. Let $n \in \mathbb{N}$, $w \in R_n(1)$. Then $w$ is called \textit{primitive} if $w, w^2, \ldots, w^n$ are mutually distinct.

1.4.1. $w$ is primitive if and only if $w^l \neq 1$ for all $l \in n - 1$.

16 The proof is valid for arbitrary fields in place of $\mathbb{C}$ if we adjust the meaning of $R_n(1)$ accordingly.
Proof. We have $w^n = 1$, thus trivially $w^l \neq 1$ for all $l \in \mathbb{Z}$ if $w$ is primitive. On the other hand, if there are $j, k \in \mathbb{Z}$ such that $j < k$ and $w^j = w^k$, it follows that $w^{k-j} = 1$ and $k - j \in \mathbb{Z}$.

We write $\mathcal{P}_n$ for the set of all primitive $n$-th roots of unity and put $\varphi(n) := |\mathcal{P}_n|$. Then $\varphi$ is a mapping of $\mathbb{N}$ into $\mathbb{N}$, called the Eulerian function. For example, $\mathcal{P}_4 = \{i, -i\}$, $\varphi(4) = 2$. It is easily seen that $\mathcal{P}_3 = R_3(1) \setminus \{1\}$, $\mathcal{P}_5 = R_5(1) \setminus \{1\}$, thus $\varphi(3) = 2$, $\varphi(5) = 4$.

1.4.2. Let $k \in \mathbb{Z}$, $w = \exp \frac{k\cdot2\pi i}{n}$. Then $w \in \mathcal{P}_n$ if and only if $\gcd(k,n) = 1$.

Proof. Suppose first that $\gcd(k,n) = 1$. Let $l \in \mathbb{N}$ such that $1 = w^l = \exp \frac{lk\cdot2\pi i}{n}$. Then, by 1.1.7, $\frac{lk}{n} \in \mathbb{Z}$. But $k$ and $n$ have no common prime divisor, hence $\gcd(l, n)$ which means that $n = l$. By 1.4.1, it follows that $w \in \mathcal{P}_n$.

Now let $k \in \mathbb{Z}$ arbitrary, put $c := \gcd(k,n)$, $d := \frac{n}{c}$, $k' := \frac{k}{c}$. Clearly then $\gcd(k',d) = 1$, hence

$$\exp \frac{k\cdot2\pi i}{n} = \exp \frac{k'\cdot2\pi i}{d} \in \mathcal{P}_d$$

by our first part. In particular, if $c \neq 1$, then $d$ is a proper divisor of $n$, hence, making use of the first (and here strict) inclusion in 1.3.2, $(\exp \frac{k\cdot2\pi i}{n})^j \leq R_d(1) < R_n(1)$.

Let $\mathcal{T}_n := \{d|d \in \mathbb{N}, d|n\}$. We have proved:

1.4.3. Every $n$-th root of unity is a primitive $d$-th root of unity for a (unique) $d \in \mathcal{T}_n$.

As $R_d(1) \subseteq R_n(1)$ if $d \in \mathcal{T}_n$ we conclude

1.4.4. $R_n(1) = \bigcup_{d \in \mathcal{T}_n} \mathcal{P}_d = \sum_{d \in \mathcal{T}_n} \varphi(d)$.

Here the dot above the sign for the union means that the union is disjoint. The second part of 1.4.4 follows from the first as $n = |R_n(1)| = \sum_{d \in \mathcal{T}_n} |\mathcal{P}_d| = \sum_{d \in \mathcal{T}_n} \varphi(d)$.

We have seen that 1.4.3 and its converse hold. Hence $\varphi(n) = |\mathcal{P}_n|$ is the number of all $k \in \mathbb{Z}$ such that $\gcd(k,n) = 1$ which may therefore be viewed as a purely number-theoretic definition of $\varphi(n)$. The second part of 1.4.4 is in fact a classical result in elementary number theory (see [HW, Theorem 63]) for which we have given an algebraic proof. It happens frequently and is thus a pleasant general component of abstract algebra that a structural line of reasoning has an interesting consequence in number theory which would not suggest an algebraic background at first glance.\textsuperscript{18}

\textsuperscript{17}It should not be overlooked, however, that with this conclusion we have made use of a basic result of elementary number theory which has not been dealt with as a topic in this chapter (see Example (1) on p. 28 and the condition (\textasteriskcentered) on p. 29).

\textsuperscript{18}"Count, for every divisor of a positive integer $n$, how many numbers between 1 and the divisor are prime to the divisor. If you sum up these numbers, you get : $n$." The charm is that you do not need more than a chapter from the teaching curriculum for 12-year-olds to understand and – empirically – enjoy this, while the general proof makes use of properties of algebraic structures!
The discussion of radicals has opened up a subtopic of independent interest, that of roots of unity. These play an important role in many branches of algebra and mathematics in general. We won’t leave this subject without mentioning a famous problem concerning their geometric visualisation in the complex plane which leads to regular $n$-gons as we have seen on p. 12:

**Problem.** Let $n \in \mathbb{N}$. Starting from a line segment of length 1, is it possible to construct the $n$-th roots of unity in the complex plane by ruler and compass?

The complete answer makes use of the theory to be developed in this course, hence of abstract field theory! It is easily seen that the $n$-th roots of unity are constructible for $n = 1, 2, 3, 4, 6, 8$, but why is the answer “yes” for $n = 5$, why “no” for $n = 7, 9$?

The problem is famous because C. F. Gauss (1777-1855) settled the case $n = 17$ (in the affirmative) ingeniously at the age of 19: not by giving an explicit construction, but by developing a line of reasoning that such a construction exists. Clearly, a quality like this gives far more evidence of an outstanding mathematical talent than any technical ability of concrete constructing. Why should Gauss execute for real a complicate construction when there would be no further gain in insight by doing so?

Having discussed complex roots of unity, it is a very simple step to treat general pure equations over $\mathbb{C}$:

**1.5 Proposition.** Let $a \in \mathbb{C}$, $n \in \mathbb{N}$. Write $R_n(a)$ for the set of all complex solutions of the equation $x^n = a$. Then $|R_n(a)| = n$. If $r \in \mathbb{R}$ such that $a = |a| \exp(ri)$, $w \in \mathbb{P}_n$ and $c := \sqrt[n]{|a|} \exp \frac{2\pi}{n}$, then $R_n(a) = \{ c, cw, \ldots, cw^{n-1} \}$.

**Proof.** We have $c \neq 0$ as $a \neq 0$, and $|\{c, cw, \ldots, cw^{n-1}\}| = n$ as $w \in \mathbb{P}_n$. Furthermore, $(cw^k)^n = c^n w^{kn} = |a| \exp(ri) = a$ for every $k \in \mathbb{N}_0$. This proves the inclusion “⊇”. To obtain equality it is sufficient to observe that the polynomial $t^n - a$ can have at most $n$ zeros in $\mathbb{C}$.

Instead of giving a special argument for the last statement in this proof we show the following proposition the final part of which is much more general than what is needed for 1.5.

**1.6 Proposition.** Let $K$ be a commutative unitary ring, $f \in K[t]$.

1. Let $g \in K[t]$ be normed. Then there exist $h, r \in K[t]$ such that

   (i) $f = gh + r$, (ii) $r = 0_K$ or $\deg r < \deg g$.

---

19 The first explicit construction was published, almost 30 years after Gauss’s discovery, by someone else.

20 Slightly more generally, it suffices to assume that the coefficient $b$ of the term of highest degree in $g$ be invertible in $K$. If this is the case, we may apply 1.6(1) to the polynomial $b^{-1}g$ which is normed. Then it suffices to observe that $f = (b^{-1}g)h + r = g(b^{-1}h) + r$, $\deg g = \deg(b^{-1}g)$. In particular, the claim in 1.6(1) holds for every nonzero $g \in K[t]$ if $K$ is a field. – The commutativity of $K$ is essential for 1.6. Even if $K$ satisfies all axioms of a field apart from the commutative law of multiplication, a polynomial over such a structure (called skew field) can have infinitely many zeros in $K$. 

---

17
(2) If \( b \in K \) is a zero of \( f \), there exists a polynomial \( h \in K[t] \) such that \( f = (t - b)h \).

(3) If \( K \) is an integral domain, \( f \neq 0_K \), \( n := \deg f \), then there exist at most \( n \) zeros of \( f \) in \( K \).

Proof. (1) Put \( m := \deg g \). If \( f = 0_K \), the claim is trivial (\( h := 0_K =: r \)). For \( f \neq 0_K \) we proceed by induction on \( \deg f \), observing first that the claim is trivial whenever \( m > \deg f \) as then we may set \( h := 0_K, r := f \). Thus for \( f \in K \) we must only consider the case \( m = 0 \) which means that \( g = 1_K \) as \( g \) is normed. Then again the claim is trivial as it suffices to set \( h := f, r := 0_K \).

For the inductive step suppose that \( f \) is of degree \( n > 0 \) and the claim holds for all nonzero polynomials of degree \( < n \) in place of \( f \). Let \( c_0, \ldots, c_n \in K \) and \( f = \sum_{j=0}^n c_j t^j \), \( c_n \neq 0_K \). Then the polynomial \( \hat{f} := f - g c_n t^{n-m} \) is either \( 0_K \) or of degree \( < n \). If \( \hat{f} = 0_K \), the claim is clear as then \( f = g c_n t^{n-m} \) and we may put \( r := 0_K \). But if \( \hat{f} \in K[t] \) we know inductively that \( f = g h + r \) for polynomials \( h, r \in K[t] \) where \( r = 0_K \) or \( \deg r < m \). Then \( f = g c_n t^{n-m} + g h + r \), hence the claim, putting \( h = c_n t^{n-m} + \hat{h} \).

(2) We apply (1) with \( g := t - b \). It follows that there exist \( h \in K[t], r \in K \) such that \( f = (t - b)h + r \). As \( b \) is a zero of \( f \) this implies that \( r = 0_K \), hence \( f = (t - b)h \).

(3) is again proved by induction on \( n \): The case where there is no zero of \( f \) in \( K \) is trivial, in particular the case \( n = 0 \). Now let \( n > 0 \) and suppose there exists a zero \( b \in K \) of \( f \). Then (2) implies that \( f = (t - b)h \) where \( h \in K[t], \deg h < n \) by 1.1.2. Assuming inductively that \( h \) has at most \( n - 1 \) zeros in \( K \), our claim follows: By 1.1.4, the zeros of \( f \) in \( K \) are \( b \) and the zeros of \( h \), hence their number is at most \( n \).

1.7 Definition. Let \( K \) be an integral domain, \( b \in K, f \in K[t], j \in N_0 \). Then \( b \) is called a \( j \)-fold zero of \( f \) if there exists a polynomial \( h \in K[t] \) such that \( f = (t - b)^j h \), \( h(b) \neq 0_K \). The number \( j \), called the multiplicity of \( b \) with respect to \( f \), is uniquely determined: The assumption \( (t - b)^j h = f = (t - b)^i \tilde{h} \) where \( 0 \leq i < j, h, \tilde{h} \in K[t], h(b) \neq 0_K \neq \tilde{h}(b) \) would imply that \( (t - b)^{j-i} (h - \tilde{h}) = 0_K \), hence \( (t - b)^{j-i} h - \tilde{h} = 0_K \), i.e., \( (t - b)^{j-i} h = \tilde{h} \). Now the substitution of \( b \) gives a contradiction. The element \( b \) is called a multiple zero of \( f \) if \( j \geq 2 \). We want to derive a criterion for an element to be a multiple zero and introduce the following formal differential calculus for this purpose: For all \( n \in N_0, c_0, \ldots, c_n \in K \) we set

\[
\left( \sum_{j=0}^n c_j t^j \right)^{'} := \sum_{j=1}^n j c_j t^{j-1}.
\]

The mapping \( d : K[t] \to K[t], f \mapsto f' \), is obviously \( K \)-linear. If \( m, n \in N \), we have

\[
(t^m \cdot t^n)' = (t^{m+n})' = (m + n) t^{m+n-1} = mt^{(m-1)+n} + nt^{m+(n-1)} = (t^m)'t^n + t^m(t^n)',
\]

and the rule \( (t^m \cdot t^n)' = (t^m)'t^n + t^m(t^n)' \) holds trivially if \( m = 0 \) or \( n = 0 \). Exploiting the \( K \)-linearity of \( d \) we conclude easily the following product rule, otherwise known from classical differential calculus:
1.8 Proposition. Let $K$ be an integral domain, $f \in K[t]$, $b \in K$. The following are equivalent:

(i) $b$ is a multiple zero of $f$,

(ii) $b$ is a zero of $f$ and a zero of $f'$.

Proof. Let $j \in \mathbb{N}_0$, $h \in K[t]$ such that $f = (t - b)^j h$, $h(b) \neq 0_K$. If (i) or (ii) holds we have $f(b) = 0_K$, hence $j \geq 1$ by 1.6(2). It follows from 1.7.1 that

$$f' = j(t - b)^{j-1}h + (t - b)^{j-1}(jh + (t - b)h'),$$

and we see that for $j = 1$ the element $b$ is not a zero of this polynomial because $h(b) \neq 0_K$. Therefore $f'(b) = 0_K$ if and only if $j \geq 2$. 

As an example, we consider the case of a pure polynomial (cf. 1.5) over an arbitrary integral domain $K$. Let $a \in K$, $f = t^n - a$ for some $n \in \mathbb{N}$. We define $n1_K := 1_K + \cdots + 1_K$. Then $f' = nt^{n-1} = t^{n-1} + \cdots + t^{n-1} = (n1_K)t^{n-1}$, by the distributive law. Hence, by 1.8, either $n1_K = 0_K$ or $f$ has no multiple zero in $K$. As a consequence we note

1.8.1. Let $K$ be a field containing $\mathbb{Q}$ as a subfield, $a \in K$, $n \in \mathbb{N}$. Then $t^n - a$ has no multiple zero in $K$. 

Having considered pure polynomials in some detail, we now turn to polynomials of a more complicated nature over a field $K$. Under the hypothesis that $k, n \in \mathbb{N}, n1_K \neq 0_K$, we set $\frac{k}{n} := k(n1_K)^{-1}$. The case of a polynomial of degree 2 over a field $K$ should be well known if $K = \mathbb{R}$, and the solution in this case is easily generalized to arbitrary fields provided that $1_K + 1_K \neq 0_K$: Let $a_0, a_1 \in K$. In $K[t]$ we have the equation

$$t^2 + a_1t + a_0 = \left(t + \frac{a_1}{2}\right)^2 - \left(\frac{a_1^2}{4} - a_0\right).$$

Instead of seeking immediately a solution of the equation $x^2 + a_1x + a_0 = 0_K$ we first approach the pure equation $x^2 - a = 0$ where $a = \frac{a_1^2}{4} - a_0$. If $K$ contains a solution $r$ of this equation, it suffices to put $b := r - \frac{a}{2}$ to obtain a zero of the polynomial originally given. Explicitly, the zero is then given in the form

$$-\frac{a_1}{2} + r \quad \text{where} \quad r^2 = \frac{a_1^2}{4} - a_0,$$

hence expressed by means of (a simple additive shift of) a radical of exponent 2 over $K$.

\begin{quote}
The tricky transformation given by $(*)$ is thus to be understood as a reduction of the problem of solving an arbitrary equation of degree 2 to the case of a pure equation of degree 2. Finding a solution in the general case is reduced to finding a radical: This is the message of $(*)$.
\end{quote}
Although likewise true, it should not simply be considered as a step in deriving “a formula for the solutions of a quadratic equation” – which generations of students had to memorize as if it were of major importance for everybody’s lifelong happiness. (For the majority of them, the opposite holds.)

The fundamental question which emanates from this success in treating the quadratic case is now of course the following: Is it possible to reduce the problem of solving an arbitrary equation of any degree to the problem of solving pure equations? Is it always possible to express general solutions in terms of radicals?

We observe that (*) is a special case of a simplifying step for a polynomial of arbitrary degree \( n \geq 2 \) which, however, will not immediately lead to a pure polynomial if \( n > 2 \):

1.9 Proposition. Let \( n \in \mathbb{N}_{>1} \), \( K \) a field such that \( n1_K \neq 0_K \). Let \( a_0, \ldots, a_{n-1} \in K \) and put \( s := t + \frac{a_{n-1}}{n} \) in \( K[t] \). Then there exist \( b_0, \ldots, b_{n-2} \in K \) such that

\[
t^n + a_{n-1}t^{n-1} + a_{n-2}t^{n-2} + \cdots + a_1t + a_0 = s^n + b_{n-2}s^{n-2} + \cdots + b_1s + b_0.
\]

Proof. We just have to make standard use of the binomial theorem to find

\[
t^n + a_{n-1}t^{n-1} + \sum_{j=0}^{n-2} a_j t^j = \left(s - \frac{a_{n-1}}{n}\right)^n + a_{n-1}\left(s - \frac{a_{n-1}}{n}\right)^{n-1} + \sum_{j=0}^{n-2} a_j \left(s - \frac{a_{n-1}}{n}\right)^j
\]

for some \( u \in \langle 1_K, s, \ldots, s^{n-2} \rangle_K \). Hence \( t^n + \sum_{j=0}^{n-1} a_j t^j = s^n + u \), proving the claim.

Thus, given any equation, we can always in a first step get rid of the term of second highest degree if the hypothesis on \( K \) in 1.9 is satisfied. For \( n = 3 \) this means: We can solve an arbitrary equation of degree 3 over a field \( K \) with \( 1_K + 1_K + 1_K \neq 0_K \) if we can solve an arbitrary equation of the form

(**) \( x^3 + b_1x + b_0 = 0_K \) (where \( b_0, b_1 \in K \)).

We shall now show that the latter always allows indeed a reduction to finding appropriate radicals if \( K \) is a field where \( 1_K + 1_K + 1_K \neq 0_K \neq 1_K + 1_K \). This hypothesis will be assumed in the sequel. The solubility by radicals of cubic equations may be viewed as a cornerstone in the development of modern Algebra. The result dates back to the first half of the 16th century and and was not found, but finally published by Gerolamo Cardano in 1545. We present it in a form which emphasizes the character of a reduction to radicals in the same spirit as in the case of an equation of degree 2.

Case 1: \( b_1 = 0_K \). Then the equation is pure, its solutions are radicals by definition.

Case 2: \( b_1 \neq 0_K \). Let \( z \in K \) be a solution of the pure quadratic equation

\[
x^2 - \left(\frac{b_1^2}{27} + \frac{b_0^2}{4}\right) = 0_K.
\]
Let $y \in K$ be a solution of the pure cubic equation

$$x^3 + \left(\frac{b_0}{2} - z\right) = 0_K.$$

Then $y \neq 0_K$ and $y - \frac{b_0}{3y}$ is a solution of (**).

(Note that the coefficient of the pure cubic equation involves the solution $z$ of the foregoing quadratic equation!) In Case 2, we first show that $y \neq 0_K$. Otherwise $\frac{b_0}{2} - z = 0_K$, hence $z = \frac{b_0}{2}$, implying that $\left(\frac{b_0}{2}\right)^2 - \left(\frac{b_1}{27} + \frac{b_2}{4}\right) = 0_K$. This means that $b_1 = 0_K$, contrary to the hypothesis. Thus we have $y \neq 0_K$, $z^2 = \frac{b_1^3}{27} + \frac{b_2^2}{4}$, $y^3 = z - \frac{b_0}{2}$. Hence

$$(y - \frac{b_1}{3y})^3 + b_1\left(y - \frac{b_1}{3y}\right) + b_0 = y^3 - b_1y + \frac{b_1^2}{3y} - \frac{b_1^3}{27y^3} + b_1y - \frac{b_2^2}{3y} + b_0$$

$$= z - \frac{b_0}{2} - \frac{b_1^3}{27(z - \frac{b_0}{2})} + b_0 = z + \frac{b_0}{2} - \frac{z^2 - \frac{b_2^2}{4}}{z - \frac{b_0}{2}} = 0_K.$$

In the case of a quadratic equation, the field $K$ obviously contains a solution if and only if the radical used in its description belongs to $K$. In the case of a cubic equation, however, $K$ may contain a solution even if the pure equations of degree 2 and 3 (see Case 2 above) are not soluble in $K$. The solubility of these is only sufficient but not necessary in order to find a solution in $K$. At this point it is not even clear if, in the case where those required radicals do not exist in $K$, there exists a solution by radicals which is different from Cardano’s. Thus we can see that the case of a cubic equation quickly leads into rather intricate problems. We may summarize these by formulating the question:

**Problem.** What has to be done if $K$ does not contain the required radicals $z, y$?

Furthermore, there is another problem to be noted with respect to Cardano’s solution by radicals: In the classical case of $K = \mathbb{C}$, the solutions are still far from the standard representation in the form $a + bi$ with $a, b \in \mathbb{R}$. In the example $x^3 - 2x + 1 = 0$, the attempt to transform the solutions obtained by radicals into the standard representation will turn out to be a hard piece of work. But 1 is clearly a zero of the polynomial $t^3 - 2t + 1 = (t - 1)(t^2 + t - 1)$, and the other two zeros turn out to be $-\frac{1}{2} + \frac{1}{2}\sqrt{5}$, $-\frac{1}{2} - \frac{1}{2}\sqrt{5}$. In particular, all zeros are real numbers which is not evident from the Cardano solution. The polynomial $t^3 - 2t + 1$ has rational coefficients, but in $\mathbb{Q}$ there is just one zero. Still it makes sense to talk about zeros in $\mathbb{R}$ or $\mathbb{C}$. While we know these fields, what would we have to do if an arbitrary field $K$ is given? May we expect that, even if $K$ does not contain zeros of a polynomial, there will exist a larger field in which we find the requested zeros?

**Problem.** What has to be done if $K$ does not contain any solution of a given polynomial equation?

Cardano knew already that not only the equations of degree 3, but also those of degree 4
are soluble by radicals. A suitable reduction to a cubic polynomial was found so that the step from degree 3 to degree 4 is much less exciting than that from degree 2 to degree 3. It took about 300 years after the discovery of the solution by radicals of the cubic equation until the cases of a degree > 4 could be settled – in the negative: It was proved (Abel 1826) that a general solution by radicals does not exist for equations of degree > 4. Even worse, there exist equations of degree 5 with rational coefficients which are in no way soluble by radicals (in \( \mathbb{C} \)). Examples of this kind are \( x^5 - 4x + 2 = 0 \), \( 2x^5 - 10x + 5 = 0 \).

In 1829, Galois characterized the equations which are soluble by radicals. The main instrument to obtain this characterization is the famous Galois Theory which will be developed in this course. It is of utmost importance for countless developments and directions of modern Algebra which may be devoted to questions which are not at all connected with its historical origin.

1.10 Definition. Let \( K \) be a subfield of a field \( L \), \( f \in K[t] \) normed. \( L \) is called a splitting field of \( f \) over \( K \) if there exist \( b_1, \ldots, b_n \in L \) (not necessarily distinct) such that

(i) \( f = (t - b_1) \cdots (t - b_n) \),

(ii) If \( \tilde{L} \) is a subfield of \( L \) such that \( K \cup \{b_1, \ldots, b_n\} \subseteq \tilde{L} \), then \( \tilde{L} = L \).

Then \( b_1, \ldots, b_n \) are the zeros of \( f \) in \( K \) as \( f(b) = 0_K \) for some \( b \in K \) if and only if at least one of the differences \( b - b_j \) vanishes, i.e., if and only if \( b \in \{b_1, \ldots, b_n\} \). As a simple example, consider \( f = t^2 - 2 \in \mathbb{Q}[t] \). The subring \( \mathbb{Q}[\sqrt{2}] \) of \( \mathbb{R} \) is a field \(^{21}\) and clearly does not have a proper subfield containing \( \mathbb{Q} \cup \{\sqrt{2}, -\sqrt{2}\} \). (In fact there is no proper subfield \( \neq \mathbb{Q} \) in \( \mathbb{Q}[\sqrt{2}] \) at all!).

Problem. Given a normed polynomial \( f \) over a field \( K \), does there exist a splitting field of \( f \) over \( K \)?

We shall need some more preparation to see that the answer is positive. For a subfield \( K \) of \( \mathbb{C} \), this may be concluded from the following famous result:

Fundamental Theorem of Algebra (Gauss, doctoral dissertation 1799) Every polynomial in \( \mathbb{C}[t] \setminus \mathbb{C} \) splits in \( \mathbb{C}[t] \) into linear factors.

A field \( L \) such that every polynomial in \( L[t] \setminus L \) splits in \( L[t] \) into linear factors is called algebraically closed. Thus \( \mathbb{C} \) is closed. The following important result will not be proved in this course:

Theorem For every field \( K \) there exists an algebraically closed field \( L \) which contains \( K \) as a subfield.

Proofs may be found in numerous textbooks, see e.g. [St, 3.24]. The same book contains two proofs of the fundamental theorem of algebra, in its appendix §.A1. If \( f \) is a normed polynomial over a subfield \( K \) of \( \mathbb{C} \), \( n \) its degree, then the fundamental theorem of algebra ensures that \( f = (t - b_1) \cdots (t - b_n) \) for suitable complex numbers \( b_1, \ldots, b_n \). Thus \( \mathbb{C} \)

\(^{21}\)If \( a, b \in \mathbb{Q} \) not both 0, then \( a^2 - 2b^2 \neq 0 \) as 2 is not a square in \( \mathbb{Q} \), and

\[
(a + b\sqrt{2})^{-1} = \frac{a}{a^2 - 2b^2} - \frac{b}{a^2 - 2b^2} \sqrt{2} \in \mathbb{Q}[\sqrt{2}].
\]
satisfies condition (i) in the definition of a splitting field. The other condition, however, will not hold in general for \( \mathbb{C} \) itself but for some subfield of \( \mathbb{C} \), by an application of the following remark:

1.10.1. Let \( M \) be a field and \( \mathcal{X} \) a nonempty set of subfields of \( M \). Then \( \bigcap \mathcal{X} \) is a subfield of \( M \).

**Proof.** Clearly, \( \bigcap \mathcal{X} \) is an additive subgroup of \( M \), and \( \bigcap \mathcal{X} \setminus \{0_M\} \) is a multiplicative subgroup of \( M \). The claim follows. \( \square \)

If \( f = (t - b_1) \cdots (t - b_n) \) (where \( b_1, \ldots, b_n \in \mathbb{C} \)), let \( \mathcal{X} \) be the set of all subfields of \( \mathbb{C} \) containing \( K \cup \{b_1, \ldots, b_n\} \). Then \( \mathbb{C} \in \mathcal{X} \), and the field \( L := \bigcap \mathcal{X} \) is a splitting field of \( f \) over \( K \). Since Gauss's first proof of the fundamental theorem of algebra several others have been found, also by Gauss himself. To some extent, they all make use of some analytical properties of the individual field \( \mathbb{C} \). The general theorem on the existence of a splitting field of a given polynomial over an arbitrary field \( K \) would follow in the same way by making use of the quoted theorem on the existence of an embedding of \( K \) into an algebraically closed field. It will, however, be obtained in a much simpler way in due course (4.3).

Let us now assume that \( f \) splits completely into linear factors \( t - b_j \) in \( K[t] \). We have

\[
\begin{align*}
f &= (t - b_1) \cdots (t - b_n) \\
&= t^n - b_1 t^{n-1} - b_2 t^{n-1} - \cdots - b_n t^{n-1} \\
&\quad + b_1 b_2 t^{n-2} + b_1 b_3 t^{n-2} + \cdots + b_{n-1} b_n t^{n-2} \\
&\quad - b_1 b_2 b_3 t^{n-3} - b_1 b_2 b_4 t^{n-3} - \cdots - b_{n-2} b_{n-1} b_n t^{n-3} \\
&\quad \vdots \\
&\quad + (-1)^n b_1 \cdots b_n \\
&= t^n - \cdots + (-1)^k (b_{i_1} b_{i_2} \cdots b_{i_k} + \cdots) t^{n-k} + \cdots + (-1)^n b_1 \cdots b_n
\end{align*}
\]

where the sum in brackets extends over all products of \( k \) factors, \( i_1 < i_2 < \cdots < i_k \). We write \( s_k(b_1, \ldots, b_n) \) for this sum, i.e., we set

\[
s_k(b_1, \ldots, b_n) := \sum_{i_1 < i_2 < \cdots < i_k} b_{i_1} b_{i_2} \cdots b_{i_k} \quad \text{for all } k \in \mathbb{N},
\]

and additionally \( s_0(b_1, \ldots, b_n) := 1_K \). The result of our calculation may then be expressed as follows:

1.11 Theorem. (Viète) Let \( K \) be a field, \( n \in \mathbb{N}, b_1, \ldots, b_n \in K \). Then

\[
(t - b_1) \cdots (t - b_n) = \sum_{k=0}^{n} (-1)^k s_k(b_1, \ldots, b_n) t^{n-k}.
\]

\( \square \)
As an example, let \( b_1, \ldots, b_n \) be the \( n \)-th roots of unity. Then \((t - b_1) \cdots (t - b_n) = t^n - 1\) so that 1.11 implies

\[
s_k(b_1, \ldots, b_n) = 0 \quad \text{for all } k \in \{0, 1, \ldots, n - 1\} \quad \text{and} \quad s_n(b_1, \ldots, b_n) = b_1 \cdots b_n = (-1)^n.
\]

Now let \( m := \varphi(n) \) and \( b_1, \ldots, b_m \) be the primitive \( n \)-th roots of unity. The polynomial \( \Phi_n := \prod_{i \in m}(t - b_i) \) is called the \( n \)-th cyclotomic polynomial, being related to the problem of dividing the unit circle into \( n \) equal parts (cf. p.12). Clearly, \( \deg \Phi_n = \varphi(n) \). Apart from the case \( n = 1 \) we have \( \Phi_n(0) = 1 \) because \( \mathcal{P}_n \) is closed under taking multiplicative inverses (and, for \( n = 2 \), the assertion is obvious). We will prove in 4.21.1 that \( \Phi_n \in \mathbb{Z}[t] \).

For example, for \( n = 1, 2, 3, 4, 5 \) we obtain

\[
\varphi(1) = 1, \quad \varphi(2) = 1, \quad \varphi(3) = 2, \quad \varphi(4) = 2, \quad \varphi(5) = 4.
\]

The value \( s_k(b_1, \ldots, b_n) \) does not depend on the order of the elements \( b_i \), i.e.,

1.11.1. \( \forall \pi \in S_n \forall b_1, \ldots, b_n \in K \quad s_k(b_1, \ldots, b_n) = s_k(b_{1\pi}, \ldots, b_{n\pi}) \)

where \( S_n := S_n \).

1.12 Definition. Let \( B, A \) be sets, \( n \in \mathbb{N} \). A mapping \( g \) of \( B^n \) to \( A \) is called symmetric if \((b_1, \ldots, b_n)g = (b_{1\pi}, \ldots, b_{n\pi})g \) for all \( b_1, \ldots, b_n \in B \). Then the contents of 1.11.1 may be expressed as follows: For all \( k \in \mathbb{N} \), the mapping

\[
K^n \to K, \quad (b_1, \ldots, b_n) \mapsto s_k(b_1, \ldots, b_n),
\]

is symmetric. It is called the \( k \)-th elementary symmetric function. Viète’s theorem 1.11 essentially means that, up to the sign, the coefficients of a normed polynomial \( f \) over a field are the values of the elementary functions for the tuple of its zeros in a splitting field. In an intricate but most regular way, the coefficients of \( f \) are composed by them:

1.11’ Let \( K \) be a field, \( n \in \mathbb{N}, a_0, \ldots, a_{n-1} \in K, f = t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0 \). Suppose that we are given a splitting field of \( f \) so that \( f \) splits into linear factors \( t - b_1, \ldots, t - b_n \). Then

\[
a_{n-k} = (-1)^k s_k(b_1, \ldots, b_n) \quad \text{for all } k \in \mathbb{N}.
\]

Although we know, by Viète’s theorem, in which way the zeros of a polynomial (in a splitting field) are transformed into its coefficients by means of the elementary symmetric functions, there is no visible general way back to “decipher this code” and reveal, vice versa, the zeros from those.
2 Algebraic structures and the extension principle

For the sake of a clear orientation let us briefly recall some fundamental algebraic terminology:

(I) Algebraic structures with one operation
   a) Magma: Any set with an operation (see 1.2).
      Examples: \((\mathbb{N}, \wedge)\) where \(a \wedge b := a^b\); \((\mathbb{Z}, -)\).
   b) Semigroup: Associative magma.
      Examples: \((X, \dashv)\) where \(X\) is any set, \(a \dashv b := b\); \((\mathbb{N}, +)\).
   c) Monoid: Semigroup with neutral element.
      Examples: \((\mathbb{N}_0, +)\); \((\mathbb{N}, \cdot)\); \((X^X, \circ)\) where \(X\) is any set and \(\circ\) denotes the composition of mappings, i.e., \(f \circ g\) maps \(x \in X\) to \((xf)g\). Usually we just write \(fg\) and use this notation in its natural sense, i.e., the mapping written at the first place is the mapping which is applied first.
   d) Group: Monoid in which every element is invertible.
      Examples: \((\text{End} A, \dagger, \circ)\) where \(A\) is an abelian group.

(II) Algebraic structures with two operations
   a) Double magma\(^22\): Any set with two operations.
   b) Ring: Double magma \((R, +, \cdot)\) where \((R, +)\) is an abelian\(^23\) group, \((R, \cdot)\) is a semigroup, and both distributive laws hold.
      Example: \((\text{End} A, \dagger, \circ)\) where \(A\) is an abelian group.
   c) Integral domain: Commutative unitary ring \((R, +, \cdot)\) such that \(R\) is multiplicatively closed.
      Example: \((\mathbb{Z}, +, \cdot)\).

---

\(^22\)This term is not commonly used in the literature.
\(^23\)for groups, synonymous with “commutative”
\(^24\)With regard to the notion of a ring, confusion of terminology has been created and spread by a certain influential book on Algebra. Therefore we have reason to emphasize that a ring does not necessarily contain a unit element by definition. If there is one and distinct from the additively neutral element, the ring is called unitary in this text.
d) **Field**: Integral domain \((R, +, \cdot)\) such that \(R\) is multiplicatively a group.

Examples: \((\mathbb{Q}, +, \cdot), (\mathbb{C}, +, \cdot), (\mathbb{Q}[\sqrt{2}], +, \cdot)\) (see 1.10); \(\{0, 1\}, +, \cdot\) where +, \cdot are defined as follows:

<table>
<thead>
<tr>
<th>(0)</th>
<th>(1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0)</td>
<td>(0)</td>
</tr>
<tr>
<td>(1)</td>
<td>(1)</td>
</tr>
</tbody>
</table>

(III) Algebraic structures with a set of operators (which we assume here to be a commutative unitary ring \(K\))

a) **\(K\)-magma, \(K\)-double magma** resp. : Magma, double magma resp., \(M\) for which there is given some “outer product” \(K \times M \to M\).

b) **\(K\)-space** : see 1.1. A \(K\)-space where \(K\) is a field is called a vector space.

c) **\(K\)-algebra** : see 1.1.

Examples: \(K^{n \times n}, \text{End}_K V\) for a \(K\)-space \(V\); any polynomial ring over \(K\).

We have constructed a polynomial ring over \(K\) at the beginning of chapter 1 and introduced the notation \(K[t]\) for a polynomial ring (in the indeterminate \(t\)) over \(K\) which contains \(K\) as a unital subring (see p. 9), but we have not yet established its existence. This will be an immediate consequence of the following general proposition:

**2.1 Proposition** (Extension Principle). Let \(B, M\) be sets and \(\varphi\) an injective mapping of \(B\) into \(M\). Then there exists a set \(\hat{B}\) containing \(B\) and a bijection \(\hat{\varphi}\) of \(\hat{B}\) onto \(M\) such that \(\hat{\varphi}|_B = \varphi\).

![Diagram of Extension Principle]

**Supplement.** Let \(\circ\) be an operation on \(B\), \(\cdot\) an operation on \(M\) such that \(\varphi\) is a homomorphism. Then there exists an extension \(\hat{\circ}\) of \(\circ\) to an operation on \(\hat{B}\) such that \(\hat{\varphi}\) is an isomorphism of \((\hat{B}, \hat{\circ})\) onto \((M, \cdot)\).\(^{25}\)

Given a commutative unitary ring \(K\), we have a monomorphism of \(K\) into the polynomial ring \(\Pi_0\) over \(K\) (see p. 1.1). Applying 2.1 and its supplement to both operations on \(K\), we obtain that \(\iota\) extends to an isomorphism of a double magma containing \(K\) onto \(\Pi_0\). This double magma is therefore a polynomial ring over \(K\), containing \(K\) as a unital subring. We have thus justified our usage of the notation \(K[t]\). A considerable generalization will be considered in 2.9.

**Proof of 2.1.** We will make use of the following purely set theoretic (not completely

\(^{25}\) Similarly, if \(B\) and \(M\) are \(K\)-magmas, \(\hat{B}\) may be made into a \(K\)-magma by extending the definition of the product between \(K\) and \(B\) such that \(\hat{\varphi}\) is an isomorphism of the \(K\)-magmas \(\hat{B}\) and \(M\).
trivial) statement:

Given any two sets \( A, B \), there exists a set \( A' \) which is equipotent with \( A \) and disjoint from \( B \).\(^{26}\)

We apply this to \( A := M \setminus B \varphi \), set \( \hat{B} := B \cup A' \) and choose a bijection \( \psi \) of \( A' \) onto \( A \). Then the mapping \( \hat{\varphi} := \varphi \cup \psi \) is a bijection\(^{27}\), and \( \hat{\varphi}|_B = \varphi \).

As for the supplement, we define the operation \( \hat{o} \) on \( \hat{B} \) by \( x \hat{\hat{o}} x' := (x \hat{\varphi} \cdot x' \hat{\varphi}) \hat{\varphi}^{-1} \) for all \( x, x' \in \hat{B} \). For all \( x, x' \in B \) then \( x \hat{\hat{o}} x' = (x \circ x') \hat{\varphi}^{-1} = (x \circ x') \hat{\varphi}^{-1} = x \circ x' \), i.e., \( \hat{o} \) is an extension of \( o \). By definition of \( \hat{o} \), \( (x \hat{\hat{o}} x') \hat{\varphi} = (x \hat{\varphi} \cdot x' \hat{\varphi}) \hat{\varphi}^{-1} \hat{\varphi} = x \hat{\varphi} \cdot x' \hat{\varphi} \) for all \( x, x' \in \hat{B} \).

Thus \( \hat{\varphi} \) is a homomorphism. Being bijective, it is an isomorphism. □

2.2 Definition. A submonoid of a magma \((M, \cdot)\) is a subset \( S \) of \( M \) such that \((S, \cdot)\) is a monoid. Assume in the following that \((M, \cdot)\) is a monoid. As usual, we call a submonoid \( S \) unital if \( 1_M \in S \). Frequently, only unital submonoids are of interest in a given context.\(^{28}\) The following neat rule plays an important role in parts of group theory but hold under more general hypotheses as follows:

2.2.1. (Generalized “Dedekind’s law”) Let \( S \) be a unital submonoid of \((M, \cdot)\), \( U \) a set of invertible elements of \((S, \cdot)\). Then for all \( T \subseteq M \)

\[
U(T \cap S) = (UT) \cap S, \quad (S \cap T)U = S \cap (TU)
\]

(The iuxtaposition here stands for the subset product induced by \( \cdot \) as introduced in 1.2.)

Proof. We prove the first equation, the second is obtained similarly. Obviously \( U(T \cap S) \) is a subset of both \( UT \) and \( S \). Let \( s \in (UT) \cap S \). Then there exist \( u \in U, t \in T \) such that \( s = u \cdot t \). It follows that \( t = (u^{-1} \cdot u) \cdot t = u^{-1} \cdot (u \cdot t) \in S \). □

Clearly, we have

2.2.2. Let \( X \) be a nonempty set of unital submonoids of \((M, \cdot)\). Then \( \bigcap X \) is a unital submonoid of \((M, \cdot)\).

In particular, for any \( T \subseteq M \) the intersection of all unital submonoids containing \( T \) is a unital submonoid of \((M, \cdot)\). It is obviously uniquely determined as the smallest unital submonoid containing \( T \). It is called the unital submonoid generated by \( T \). If it equals \( M, T \) is called a generating system of \((M, \cdot)\). The multiplicative closure of \( T \),

\[
\{m|m \in M, \exists l \in \mathbb{N}_0 \exists t_1, \ldots, t_l \in T \quad m = t_1 \cdots t_l\}
\]

\(^{26}\)Sketch of a proof: Let \( X := \{x|x \in \bigcup B, x \notin x\} \), \( A' := \{\{a, X\}|a \in A\} \), \( f : A \to A', a \mapsto \{a, X\} \).

Then \( f \) is a bijection and \( A' \cap B = \emptyset \). (Observe first that \( x|x \in M, x \notin x \notin M \) for every set \( M \).)

\(^{27}\)explicitly, \( \hat{\varphi} : \hat{B} \to M, x \mapsto x, x \in \hat{B} \).

\(^{28}\)Therefore, many authors find it convenient to define a submonoid as unital, thus exclude non-unital submonoid structures right from the beginning. If the neutral element of \((M, \cdot)\) is the only idempotent element of \( M \), i.e., satisfying the equation \( x^2 = x \), then clearly every submonoid is automatically unital. This is the case if \((M, \cdot)\) is a group, but also in other important monoids like \((\mathbb{N}, \cdot)\), \((\mathbb{N}_0, +)\).
is a unital submonoid and contained in every unital submonoid containing \( T \). Hence it is the unital submonoid generated by \( T \).

Now let \( \cdot \) be commutative. Then every element of the unital submonoid generated by \( T \) may be written in the form

\[
(*) \quad t_{i_1}^{j_1} \cdots t_{i_n}^{j_n} \quad \text{where } n \in \mathbb{N}_0, \ j_1, \ldots, j_n \in \mathbb{N}, \ t_1, \ldots, t_n \text{ mutually distinct.}
\]

The order of the elements \( t_1, \ldots, t_n \) is arbitrary as \( \cdot \) is commutative. The set \( T \) is called independent if changing the order of the factors in (*) is the only possibility to represent an element of the unital submonoid generated by \( T \) as a product of powers \( \neq 1_M \) of mutually distinct elements of \( T \), i.e., if for all \( n, n' \in \mathbb{N}_0, \ t_1, \ldots, t_n, s_1, \ldots, s_{n'} \in T, \ j_1, \ldots, j_n, i_1, \ldots, i_{n'} \in \mathbb{N} \) the following holds: If \( t_1, \ldots, t_n \) are mutually distinct, \( s_1, \ldots, s_{n'} \) are mutually distinct, and \( t_{i_1}^{j_1} \cdots t_{i_n}^{j_n} = s_{i_1}^{j_1} \cdots s_{i_{n'}}^{j_{n'}} \), then \( n = n' \), and there exists a permutation \( \pi \) of \( \mathbb{N}_0 \) such that \( t_k = s_{k\pi}, \ j_k = i_{k\pi} \) for all \( k \in \mathbb{N}_0 \). An independent generating system of \((M, \cdot)\) is called a basis of the commutative monoid \((M, \cdot)\). If \( T \) is a basis of \((M, \cdot)\) and \( t \in T \), obviously the powers \( t, t^2, t^3, \ldots \) must be mutually distinct. In particular, no basis of \((M, \cdot)\) contains \( 1_M \), and furthermore

2.2.3. If a commutative monoid \((M, \cdot)\) with \(|M| \neq 1\) has a basis, then \( M \) is infinite. \( \square \)

A commutative monoid is called free if it has a basis. If \( T \) is a basis, then the monoid is called free over \( T \) or freely generated by \( T \). In contrast with the notion of a basis in vector space theory, we have

2.2.4. A commutative monoid has at most one basis. If \( T \) is a basis, \( U \) any generating system, then \( T \subseteq U \).

Proof. It suffices to prove the last assertion because it implies that any two bases contain each other, hence coincide. Let \( t \in T \). There exist \( u_1, \ldots, u_k \in U \) such that \( t = u_1 \cdots u_k \). Every \( u_j \) has a representation as a term in (*) so that \( t \) turns out to be a product of such terms. Now the definition of independence implies that \( k = 1, \ t = u_1, \) hence \( t \in U \). \( \square \)

For example, the commutative monoid \((\mathbb{N}_0, +)\) is free over \( \{1\} \), \((\mathbb{N}_0 \times \mathbb{N}_0, +)\) is free over \( \{(1, 0), (0, 1)\} \), but the commutative monoid \((\mathbb{Z}, +)\) is not free: Any generating system \( U \) of the monoid \((\mathbb{Z}, +)\) must contain at least one positive integer \( n \) and one negative integer \( z \). But for all \( k \in \mathbb{N} \) we have \( (k|z)|n + (kn)z = kn(-z + z) = 0 \). Hence \( U \) is not independent. We mention without proof two important examples of free commutative monoids:

Examples. (1) \((\mathbb{N}, \cdot)\) is a free commutative monoid.

(2) Let \( K \) be a field, \( V \) a polynomial ring over \( K \) and \( N \) the set of all normed elements of \( V \). Then \((N, \cdot)\) is a free commutative monoid.

\[\text{We have chosen the multiplicative notation here. Of course, in an additively written context, products must be replaced by sums, powers by multiples, so } "u_1^{j_1} \cdots u_n^{j_n}" \text{ becomes } "j_1t_1 + \cdots + j_nt_n", \text{ etc.} \]
In fact, (1) is just a different way of formulating the so-called fundamental theorem of arithmetic about existence and uniqueness of a prime decomposition of positive integers. The (by 2.2.4 unique) basis of \((\mathbb{N}, \cdot)\) is the set of primes. This example shows how closely some algebraic and number-theoretic notions are related. The familiar notion of divisibility of integers has a natural generalization for arbitrary commutative monoids which provides a unified approach to examples like (1), (2): Let \((M, \cdot)\) be a commutative monoid, \(a, b \in M\). We write \(b \mid a\) ("\(b\) divides \(a\) in \(M\)") if there exists an element \(x \in M\) such that \(bx = a\), in which case we call \(b\) a divisor of \(a\) in \(M\). The relation \(\mid\) is reflexive and transitive. The divisors of \(1_M\) are clearly the invertible elements (also called units) of \(M\) and form a multiplicative (unital) subgroup of \((M, \cdot)\) (cf. the comments on 1.3.1).  

2.2.5. \(u \mid b\) for all \(b \in M\) and units \(u\) of \(M\), 

because \(ux = 1_R\) implies, for all \(b \in M\), that \(u(xb) = b\). \(\square\)

An element \(q \in M\) is called indecomposable if \(q\) is not a unit of \(M\) and \(q = ab\) with \(a, b \in M\) implies that \(a\) is a unit or \(b\) is a unit of \(M\). In other words, indecomposable elements are those non-units which admit a product decomposition only by means of a unit. Indecomposable elements are also called irreducible. In particular, this is the standard term used in the case of the multiplicative structure of a polynomial ring over a commutative unitary ring. In the examples (1), (2), it is readily proved by a standard induction that the considered monoid \((M, \cdot)\) is generated by the set \(T\) of indecomposable elements. The independence of \(T\) is a consequence of the following nontrivial property for arbitrary elements \(q \in T\):

\[
(\star) \quad \forall a, b \in M \quad (q \mid ab \Rightarrow q \mid a \text{ or } q \mid b).
\]

If \(t_1, \ldots, t_n,\) likewise \(s_1, \ldots, s_{n'}\), are mutually distinct elements of \(T\), and \(t_1^{n_1} \cdots t_n^{n_n} = s_1^{j_1} \cdots s_{n'}^{j_{n'}}\), \(n \geq 1\), then (\(\star\)) implies (after a standard inductive extension to finitely many factors) that \(t_n \mid s_j\) for some \(j \in \mathbb{N}'\). As \(t_n\) is a non-unit and \(s_j\) is indecomposable, we obtain – not in general, but for the monoids considered in (1), (2) – that \(t_n = s_j\), hence \(t_1^{n_1} \cdots t_{n-1}^{n_{n-1}} t_n^{-1} = s_1^{j_1} \cdots s_{j-1}^{j_{j-1}} \cdots s_{n'}^{j_{n'}}\). Now the number of factors on both sides is less than before which shows in which way (\(\star\)) opens the way for an induction argument for a proof of the independence of \(T\). Its further details are standard and left to the reader. The main argument for a proof of (\(\star\)) may be found on p. 63.

\(^3\)!\(^3\)See [HW, Theorem 2].

\(^3\)!\(^1\)Note, however, that a zero divisor \(b\) in a commutative ring \(K\) (recall 1.1) satisfies an equation \(bx = 0_K\) for some \(x \in K\). The condition that \(x \neq 0_K\) is of course essential as \(b \mid 0_K\) holds for every \(b \in K\): \(b0_K = 0_K\). Furthermore, the reader should be aware of the difference between the term “unit element” as a synonym for “neutral element” (see p. 5) and the term “unit” as a synonym for “invertible element”.

\(^3\)!\(^2\)Clearly, for every unit \(u\) we have \(u(u^{-1}b) = b\) for all \(b \in M\) so that product decompositions where one of the factors is a unit always.
We introduce a notation for an “essentially finite” product (sum resp.) in a commutative monoid \((M, \cdot)\) \((\langle M, + \rangle\) resp.): Given a set \(T\) and a mapping \(T \rightarrow M, t \mapsto x_t\), such that \(x_t = 1_M\) \((x_t = 0_M\) resp.) for almost all \(t \in T\), we set

\[
\prod_{t \in T} x_t := \prod_{t \in T, x_t \neq 1_M} x_t, \quad \left( \sum_{t \in T} x_t := \sum_{t \in T, x_t \neq 0_M} x_t \right) \text{ resp.}
\]

Using this notation, we may express the basis property of a subset \(T\) of a (multiplicatively written) commutative monoid \((M, \cdot)\) as follows: \(T\) is a monoid basis if and only if for every \(m \in M\) there exists a unique mapping \(T \rightarrow \mathbb{N}_0, j \mapsto j_t\), such that \(j_t = 0\) for almost all \(t \in T\) and \(m = \prod_{t \in T}^t j_t\). In the sequel, we simply write \(M\) instead of \((M, \cdot)\) if the operation \(\cdot\) is either generic or well-known from the context. In general, we use the multiplicative notation unless otherwise specified.

2.3 Proposition (The universal property). Let \(M\) be a free commutative monoid, \(T\) its basis. Let \(N\) be any commutative monoid and \(f_0\) a mapping of \(T\) into \(N\). Then there exists a unique extension of \(f_0\) to a unital homomorphism of \(M\) into \(N\).

Proof. \(^34\) Uniqueness: Let \(f, g\) be homomorphisms of \(M\) into \(N\) such that \(f|_T = f_0 = g|_T\). Let \(m \in M\), represented in the form \(\prod_{t \in T} t^j\) as discussed above. Then \(mf = \prod_{t \in T}^t (tf_0)^j = mg\). Hence \(f = g\).

Existence: If \(m = \prod_{t \in T}^t t^j \in M\), the independence of \(T\) allows us to define without ambiguity \(mf := \prod_{t \in T}^t (tf_0)^j\). Thus we obtain a mapping \(f\) of \(M\) into \(N\) which is an extension of \(f_0\). As \(1_M\) is the empty product in \(M\), \(1_N\) the empty product in \(N\), we have \(1_Mf = 1_N\). Let \(m' = \prod_{t \in T}^t t^j \in M\). Then

\[
(mm')f = \left( \prod_{t \in T}^t (t^{j_t+j'_t}) \right) f = \prod_{t \in T}^t (tf_0)^{j_t} \prod_{t \in T} (tf_0)^{j'_t} = mf \cdot m'f
\]

\(\Box\)

We make a first application, to be seen as a model case, of the Extension Principle 2.1:

2.4 Proposition. Let \(T\) be a set.

(1) There is a commutative monoid which is freely generated by \(T\).

\(^{33}\) A mapping with this property is commonly called “of finite support”. If the condition holds for all \(t \in T\), \(\prod_{t \in T} x_t\) becomes the empty product in \((M, \cdot)\) \((\text{the empty sum in } (M, + )\) resp.) which equals \(1_M\) \((0_M\) resp.)

\(^{34}\) A proof of an assertion of existence and uniqueness of a certain object may frequently be approached comfortably by considering first the uniqueness part and afterwards the existence part: Proving uniqueness means to show that there is just one single candidate for the object in question. In many cases, this line of reasoning sheds enough light on that unique candidate to define it then as is needed for the existence part.
(2) Let $\beta$ be a bijection of $T$ onto a set $T'$. Suppose that $M$, $M'$ are commutative monoids which are freely generated by $T$, $T'$ resp. Then there exists a unique extension of $\beta$ to an isomorphism of $M$ onto $M'$.

In particular, a commutative monoid which is freely generated by $T$ is unique up to isomorphism.

Proof. (1) For all $f, g \in \mathbb{N}_0^T$ we put

$$f \uparrow g : T \to \mathbb{N}_0, \ t \mapsto tf + tg.$$  

Then $(\mathbb{N}_0^T, \uparrow)$ is a commutative monoid, $o : T \to \mathbb{N}_0, \ t \mapsto 0$, its neutral element. For every $s \in T$ define

$$\delta_s : T \to \mathbb{N}_0, \ t \mapsto \begin{cases} 1 & \text{if } t = s \\ 0 & \text{otherwise} \end{cases}$$

and set $S := \{\delta_s | s \in T\}$. We claim:

(*): $S$ is an independent subset of $\mathbb{N}_0^T$

Let $n, n' \in \mathbb{N}_0$, $t_1, \ldots, t_n \in T$ mutually distinct, likewise $s_1, \ldots, s_{n'}$, moreover $j_1, \ldots, j_n$, $i_1, \ldots, i_{n'} \in \mathbb{N}$ such that $j_1 f_{t_1} \uparrow \cdots \uparrow j_n f_{t_n} = i_1 f_{s_1} \uparrow \cdots \uparrow i_{n'} f_{s_{n'}}$. For every $k \in \mathbb{N}$ we have $t_k(j_1 f_{t_1} \uparrow \cdots \uparrow j_n f_{t_n}) = j_k \neq 0$, hence there exists $l \in \mathbb{N}'$ such that $t_k = s_l$. As the elements $s_l$ are mutually distinct this implies $t_k(i_1 f_{s_1} \uparrow \cdots \uparrow i_{n'} f_{s_{n'}}) = i_l$, hence $j_k = i_l$. Therefore, for each $k \in \mathbb{N}$ there exists a unique $l \in \mathbb{N}'$ such that $t_k = s_l$, $j_k = i_l$. This also holds reversely. We conclude that $n = n'$ and there is a permutation $\pi$ of $\mathbb{N}$ such that $t_k = s_{k\pi}$, $j_k = i_{k\pi}$ for all $k \in \mathbb{N}$. This shows (*).

Thus the submonoid $\hat{S}$ of $\mathbb{N}_0^T$ generated by $S$ is free over $S$. This puts us in a position to apply 2.1: The mapping $\beta : T \to S$, $t \mapsto \delta_t$, is a bijection. Hence there exists a set $\hat{T}$ containing $T$, an extension of $\beta$ to a bijection $\hat{\beta}$ of $\hat{T}$ onto $\hat{S}$, and an operation on $\hat{T}$ such that $\hat{\beta}$ is an isomorphism. It follows that $\hat{T}$ is a commutative monoid and $T$ is a basis of $\hat{T}$ because $S$ is a basis of $\hat{S}$.

(2) By the universal property 2.3, there exists an extension of $\beta$ to a homomorphism $\hat{\beta}$ of $M$ into $M'$. On the other hand, applying 2.3 to $\beta^{-1}$, we obtain an extension of $\beta^{-1}$ to a homomorphism $\hat{\beta}^{-1}$ of $M'$ into $M$. We have $\hat{\beta} \beta^{-1}|_T = id_T$, hence $\hat{\beta} \beta^{-1} = id_M$, likewise

31
\(\widetilde{\beta}^{-1}\beta = \text{id}_{M'}\). It follows that \(\widetilde{\beta}\) is bijective, hence an isomorphism. Homomorphisms which coincide on \(T\) are equal as \(M\) is the multiplicative closure of \(T\). It follows that the homomorphic extension \(\widetilde{\beta}\) of \(\beta\) is unique. \(\square\)

We write \(T^*\) for the (up to isomorphism) unique free commutative monoid over the set \(T\). Its elements may be assumed in the form

\[t_1^{j_1} \cdots t_k^{j_k}\] where \(k \in \mathbb{N}_0, t_1, \ldots, t_k \in T\) mutually distinct, \(j_1, \ldots, j_k \in \mathbb{N}\).

Let \(K\) be a commutative unitary ring, \(V\) a \(K\)-space, \(\bullet\) an operation on \(V\). Observe that the two conditions in 1.1 which make \(V\) into a \(K\)-algebra may be replaced by the requirement that the operation \(\bullet\) be a \(K\)-bilinear mapping \(V \times V \to V\). Thus a \(K\)-algebra is a special case of the notion defined in the following:

2.5 Definition. Let \(K\) be a commutative unitary ring, \(A, V\) \(K\)-spaces, \(\circ\) a bilinear mapping of \(A \times V\) into \(V\). Then \(V\)\(^{35}\) is called an \(A\)-left module over \(K\). If \(\circ\) is a bilinear mapping of \(V \times A\) into \(V\), \(V\) is called an \(A\)-right module over \(K\).

We write \(T \leq A\) if \(T\) is a \(K\)-subspace of \(V\) such that \(a \circ t \in T\) \((t \circ a \in T\) resp.) for all \(a \in A, t \in T\), called an \(A\)-submodule of \(V\), also an \(A\)-invariant \(K\)-subspace.

Now let \((A, +, \bullet)\) be an associative unitary \(K\)-algebra. Let \(\circ\) be a bilinear mapping of \(A \times V\) into \(V\). Then \(V\) is called an \(A\)-algebra left module over \(K\) if

\[\forall a, b \in A \forall v \in V \quad (a \bullet b) \circ v = a \circ (b \circ v).\]

If, moreover, \(1_A \circ v = v\) for all \(v \in V\), the \(A\)-module \(V\) is called unital. (Analogously, if \(\circ\) is a bilinear mapping of \(V \times A\) into \(V\) such that \(v \circ (a \bullet b) = (v \circ a) \circ b\) for all \(a, b \in A, v \in V\), we call \(V\) an \(A\)-algebra right module over \(K\), and unital if \(v \circ 1_A = v\) for all \(v \in V\).)

The investigations of unital \(A\)-algebra modules for a given associative unitary \(K\)-algebra \(A\) is one of the topics of major interest in current algebraic research, the area of Representation Theory. A systematic approach to the fundamentals on modules and associative algebras may be found in [Pi]. The simplest example of an \(A\)-module is given by \(V = A, \circ = \bullet\). Then we have two actions of \(A\) on itself: left multiplication, thus making \(A\) into an \(A\)-algebra left module, and right multiplication, making \(A\) into an \(A\)-algebra right module. A submodule with respect to left multiplication is called a left ideal, with respect to right multiplication a right ideal. A one-sided ideal of \(A\) is a subspace of \(A\) which is a left ideal or a right ideal.

Examples. (1) Let \(n \in \mathbb{N}, A := K^{n \times n}\). Then

\[T := \left\{ \begin{pmatrix} c_1 & \cdots & c_n \\ 0_K & \cdots & 0_K \\ \vdots & \ddots & \vdots \\ 0_K & \cdots & 0_K \end{pmatrix} \mid c_i \in K \right\}\]

\(^{35}\)More precisely: the pair \((V, \delta)\) where \(\delta\) is the mapping of \(A\) into \(\text{End}_K V\) which assigns to every \(a \in A\) the endomorphism of \(V\) induced by \(a\) via \(\circ\). This \(K\)-linear mapping \(\delta\) is called the action of \(A\) on \(V\).
is a right ideal of $A$ while the set of corresponding transposes,

$$^tT := \left\{ \begin{pmatrix} c_1 & 0_K & \ldots & 0_K \\ \vdots & \vdots & \ddots & \vdots \\ c_n & 0_K & \ldots & 0_K \end{pmatrix} \mid c_i \in K \right\}$$

is a left ideal of $A$.

---

(2) Let $A, B$ be commutative unitary associative $K$-algebras $(B, +, \cdot)$. Then $B$ becomes a unital $A$-algebra if we define

$$\forall a \in A \forall b \in B \quad a \circ b := a \cdot b \quad (b \circ a := b \cdot a \text{ resp.})$$

(i.e., $\circ$ is the restriction of $\cdot$ to $A \times B$). In particular, if the algebra $A$ is a field, the $A$-module properties of $B$ amount to saying that $B$ is an $A$-vector space.

(3) Consider the case $A = K$. Given a $K$-space $V$, which actions $\circ$ of $K$ exist to make $V$ a unital $K$-module? For any such action we must have $c \circ v = (c \cdot 1_K) \circ v = 1_K \circ (cv) = cv$, exploiting (very weakly) that $\circ$ is bilinear and the module is unital. Hence the only possible definition of $\circ$ is given by the $K$-space structure of $V$ we started from: $K$-modules (over $K$) are nothing else than $K$-spaces.

(4) Let $(A, +, \cdot)$ be a unital associative unitary $K$-algebra, $X$ a set, $V := A^X$. We make $V$ into a $K$-space by defining, for all $f, g \in V$,

$$f \uparrow g : X \to A, \quad x \mapsto xf + xg, \quad c \cdot f : X \to A, \quad x \mapsto c(f(x)) \quad \text{for all } c \in K.$$ 

Now we set $a \circ f : X \to A, x \mapsto a \cdot (f(x))$, for all $a \in A, f \in V$. Then the distributive laws and the $K$-space property of $A$ imply that $\circ$ is bilinear. Furthermore, exploiting that $A$ is associative we obtain for all $a, b \in A, f \in V, x \in X$

$$x((a \cdot b) \circ f) = (a \cdot b)xf = a \cdot (b \cdot (f(x))) = a \cdot (x(b \circ f)) = x((a \circ (b \circ f))),$$

hence $(a \cdot b) \circ f = a \circ (b \circ f)$. Finally, $x(1_A \circ f) = 1_A \cdot (f(x)) = xf$, i.e., $1_A \circ f = f$. Thus $V$ is a unital $A$-algebra left module.

---

36 Instead of the first, any other row (column resp.) could have been chosen in this definition which thus supplies $n$ right ideals and $n$ left ideals of $A$. In the case of a field $K$ it is an easy exercise to show that these one-sided ideals are minimal in the sense that they are $\neq \{0_A\}$ and do not contain any proper one-sided ideal of $A$ apart from $\{0_A\}$. This shows that the $K$-algebra $K^{n \times n}$ is a direct sum of $n$ minimal left ideals and a direct sum of $n$ minimal right ideals.
Moreover, a finite support: \( f \)

Hence \( f \) is uniquely determined by \( f_0 \).

Existence: As \( Y \) is an \( A \)-basis of \( V \), for every \( v \in V \) there is a unique mapping of finite support \( Y \to A, y \mapsto a_y^{(v)} \), such that \( v = \sum_{y \in Y} a_y^{(v)} \circ y \). Therefore we may define \( f : V \to W \) by \( vf := \sum_{y \in Y} a_y^{(v)} \circ (yf_0) \) which obviously is a \( K \)-linear extension of \( f_0 \). Moreover, \( (a \circ v)f = \sum_{y \in Y} (a \bullet a_y^{(v)})(yf_0) = \sum_{y \in Y} a \circ' (a_y^{(v)} \circ yf_0) = a \circ' (vf) \) for all \( a \in A, v \in V \).

\(^{37}\)From this it should also be read off what an \( A \)-mono/-epi/-iso/-endo/-automorphism is: an \( A \)-homomorphism which is injective/surjective/bijection/from \( V \) into \( V \)/from \( V \) into \( V \) and bijective.
2.6 Proposition. Let $K$ be a commutative unitary ring, $B$ a unitary associative $K$-algebra, $A$ a unital subalgebra of $B$, $V$ a unitary $B$-algebra left module, $Z \subseteq B$, $Y \subseteq V$. Set $Z \circ Y := \{ z \circ y \mid z \in Z, y \in Y \}$.

(1) If $\langle Z \rangle_A = B$, $\langle Y \rangle_B = V$, then $\langle Z \circ Y \rangle_A = V$.

(2) If $Z$ is $A$-independent, $Y$ is $B$-independent, then the mapping $Z \times Y \to Z \circ Y$, $(z, y) \mapsto z \circ y$ is bijective and $Z \circ Y$ is $A$-independent.

(3) If $Z$ is an $A$-basis of $B$, $Y$ is a $B$-basis of $V$, then $Z \circ Y$ is an $A$-basis of $V$.

Special case Suppose that $A$, $B$ are fields and $\dim_A B$, $\dim_B V$ are finite. Then $\dim_A V$ is finite and $\dim_A V = \dim_A B \cdot \dim_B V$.

Proof. (1) Let $v \in V$. By hypothesis, there exist elements $y_1, \ldots, y_n \in Y$, $b_1, \ldots, b_n \in B$ such that $v = \sum_{j \in \mathbb{N}} b_j \circ y_j$, and, for all $j \in \mathbb{N}$, elements $z^{(j)}_1, \ldots, z^{(j)}_k \in Z$, $a^{(j)}_1, \ldots, a^{(j)}_k \in A$ such that $b_j = \sum_{i \in \mathbb{N}} a^{(j)}_i \cdot z^{(j)}_i$. It follows that

$$v = \sum_{j \in \mathbb{N}} \left( \sum_{i \in \mathbb{N}} a^{(j)}_i \cdot z^{(j)}_i \right) \circ y_j = \sum_{j \in \mathbb{N}} \sum_{i \in \mathbb{N}} a^{(j)}_i \circ \left( z^{(j)}_i \circ y_j \right) \in \langle Z \circ Y \rangle_A .$$

(2) The mapping $Z \times Y \to Z \circ Y$, $(z, y) \mapsto z \circ y$ is surjective by definition of $Z \circ Y$. If $z, z' \in Z$, $y, y' \in Y$ such that $z \circ y = z' \circ y'$, then $z \circ y - z' \circ y' = 0_V$, hence $y = y'$, $z = z'$ as $Y$ is $B$-independent and $z, z' \neq 0_B$. Now let $U$ be a finite subset of $Z \circ Y$. For each $y \in Y$ set $Z_y := \{ z \in Z, z \circ y \in U \}$, and for every pair $(z, y) \in Z \times Y$ such that $z \circ y \in U$ let $a_{z, y} \in A$. Then

$$0_V = \sum_{z \in Z, y \in Y} a_{z, y} \circ (z \circ y) = \sum_{y \in Y} \left( \sum_{z \in Z_y} a_{z, y} \cdot z \right) \circ y$$

implies that $\sum_{z \in Z_y} a_{z, y} \cdot z = 0_B$ for every $y \in Y$ as $Y$ is $B$-independent. Furthermore, the $A$-independence of $Z$ implies that for each $y \in Y$ we have $a_{z, y} = 0_B$ for all $z \in Z_y$. Hence $Z \circ Y$ is $A$-independent. (3) is an immediate consequence of (1), (2), and trivially implies the special case. □

2.7 Proposition (Generalized Independence Lemma (Dedekind)). Let $R$ be an integral domain, $X$ a magma, $\mathcal{H}$ the set of nonzero homomorphisms of $X$ into $(R, \cdot)$. Then $\mathcal{H}$ is an $R$-independent subset of the $R$-module $R^X$ (see Example (4) on p. 33, $A := K := R$).

Proof. Otherwise let $n \in \mathbb{N}$ be minimal such that $\sum_{i \in \mathbb{N}} c_i f_i = 0_{R^X}$ for some $c_i \in R$ and mutually distinct $f_1, \ldots, f_n \in \mathcal{H}$. Then $c_i \neq 0_R$ for all $i \in \mathbb{N}$. Exploiting our $R$-linear relation for the $f_i$ in two different ways we obtain, for arbitrary $x, x' \in X$, the equations

$$\sum_{i \in \mathbb{N}} c_i (x' f_i)(x f_i) = \sum_{i \in \mathbb{N}} c_i (x' x) f_i = 0_R,$$

$$\sum_{i \in \mathbb{N}} c_i (x' f_n)(x f_i) = x' f_n \sum_{i \in \mathbb{N}} c_i (x f_i) = 0_R,$$
hence $\sum_{i=n-1}^n c_i(x'f_i - x'f_n)(xf_i) = 0_R$ by subtraction. For arbitrary $x' \in X$ we conclude $\sum_{i=n-1}^n (c_i x'(f_i - f_n)) f_i = 0_{RX}$, implying $c_i x'(f_i - f_n) = 0_R$ for every $i \in \mathbb{N}$ by the choice of $n$. For all $i \in \mathbb{N}$ we have $c_i \neq 0_R$, hence $f_i - f_n = 0_{RX}$, i.e., $f_i = f_n$, as $R$ is an integral domain. Since the functions $f_i$ are mutually distinct, we obtain $n = 1$. Hence $c_1 f_1 = 0_{RX}$ and finally $f_1 = 0_{RX}$ as $c_1 \neq 0_R$ and $R$ is an integral domain, a contradiction. \hfill \Box

2.8 Proposition. Let $K$ be a commutative unitary ring, $A$ an associative unitary $K$-algebra, $X$ a set.

(1) There exists a unital $A$-algebra module with $A$-basis $X$.

(2) Let $\beta$ be a bijection of $X$ onto a set $X'$. Suppose that $V, V'$ are unital $A$-algebra modules with $A$-bases $X, X'$ resp. Then there exists a unique extension of $\beta$ to an $A$-isomorphism of $V$ onto $V'$.

In particular, a unital $A$-algebra module with basis $X$ is unique up to $A$-isomorphism.

The proof of this 2nd application of the Extension Principle 2.1 follows the pattern of 2.4:

(1) For every $y \in X$ define

$$\delta_y : X \to A, \ x \mapsto \begin{cases} 1_A & \text{if } x = y \\ 0_A & \text{otherwise} \end{cases}$$

and $Y := \{\delta_y \mid y \in X\}$. With reference to Example (4) on p.33 we claim:

(*) $Y$ is an independent subset of the $A$-module $A^X$

Clearly, the mapping $X \to Y, y \mapsto \delta_y$, is a bijection. Let $n \in \mathbb{N}, y_1, \ldots, y_n \in X$ be mutually distinct, $a_1, \ldots, a_n \in A$ such that $\sum_{i=1}^n a_i \circ \delta_{y_i} = 0_{AX}$. By applying the left hand side to the elements $y_i$ we see that $a_i = 0_A$ for all $i \in \mathbb{N}$. This proves (*).

It follows that the $A$-submodule $\widehat{Y}$ of $A^X$ generated by $Y$ is free over $Y$. By 2.1 there exists a set $\widehat{X}$ containing $X$, an extension of $\beta$ to a bijection $\widehat{\beta}$ of $\widehat{X}$ onto $\widehat{Y}$, and operations that make $\widehat{X}$ into an $A$-module which is isomorphic to $\widehat{Y}$ via $\widehat{\beta}$. The claim in (1) follows. The universal property 2.5.2 allows us to show (2) in complete analogy to the proof of 2.4(2).

We write $AX$ for the free $A$-algebra module with basis $X$. For every element $v$ of $AX$ there is a unique mapping of finite support $X \to A, x \mapsto a_x$, such that $v = \sum'_{x \in X} a_x \circ x$. In the following, we choose $A = K$ and suppose that some operation $\cdot$ on the set $X$ is given. In other words, we take the underlying set of a magma $(X, \cdot)$ as a basis of a $K$-space. The $K$-space $KKX$ is then equipped with a natural multiplicative operation $\bullet$, given by

$$\sum'_{x \in X} c_x x \bullet \sum'_{x \in X} d_x x := \sum_{x,y \in X} c_x d_y (x \cdot y) \quad (c_x, d_x \in K),$$

36
obviously an extension of the operation \( \cdot \) given on \( X \) to \( KX \). We distinguish here carefully between different kinds of multiplications for the sake of clarity in the basic definitions. Nevertheless we will later adopt the convention to not write any multiplication symbol at all in formulas if the nature of the factors (for example, “scalar times vector”) allows a unique interpretation as to which type of product is being referred to.

The two distributive laws

\[
\forall u, v, w \in KX \quad u \cdot (v + w) = u \cdot w + u \cdot v, \quad (v + w) \cdot u = v \cdot u + w \cdot u
\]

are readily verified. Thus \((AX, +, \cdot)\) is a \(K\)-algebra.

2.8.1. Let \( K \) be a commutative unitary ring, \( X \) a magma, \( B \) a \( K \)-algebra, \( \varphi_0 \) a multiplicative homomorphism of \( X \) into \( B \). Then

\[
\varphi : KX \to B, \quad \sum_{x \in X} c_x x \mapsto \sum_{x \in X} c_x (x \varphi_0)
\]

is an extension of \( \varphi_0 \) to an \( K \)-algebra homomorphism.

For a proof, it suffices to observe that \( \varphi \) is clearly \( K \)-linear and, for any choice of scalars \( c_x, d_x \in K \) (where only a finite number is different from 0),

\[
(\sum_{x \in X} c_x x) \cdot (\sum_{y \in X} d_y y) = \sum_{x, y \in X} c_x d_y (x \cdot y) = (\sum_{x \in X} c_x x) (\sum_{y \in X} d_y y) \quad \square
\]

Clearly, \( \cdot \) is associative (commutative resp.) if and only if \( \cdot \) is associative (commutative resp.). The algebra \( KX \) is called the semigroup algebra (or semigroup ring) of \( X \) over \( K \) if \( X \) is a semigroup, similarly the group algebra (or group ring) of \( X \) over \( K \) if \( X \) is a group. The latter plays a fundamental role in the Theory of Representations and Characters of Finite Groups.

If \( X \) is a monoid, the \( K \)-algebra \( KX \) is associative and unitary: \( 1_X \) is its multiplicatively neutral element. In the sequel we consider the case where \( X \) is a free commutative monoid over some set \( T \) (see p.32). If \(|T| = 1\), say \( T = \{t\} \), then \( X = T^* = \{1_X, t, t^2, t^3, \ldots\} \). In this case \( KX \) is a polynomial ring over \( K \) in the variable \( t \) (see 1.1).

2.9 Definition. Let \( K \) be a commutative unitary ring, \( (P, +, \cdot) \) an associative commutative unitary \( K \)-algebra. Suppose that there exists an independent subset \( T \) of the commutative monoid \( (P, \cdot) \) the multiplicative closure of which is a \( K \)-basis of \( P \). Then \( P \) is called a polynomial ring over \( K \) in the set of variables (or indeterminates) \( T \), and \( T \) is called a \( K \)-algebra basis of \( P \).

Clearly, the independence of \( T \) means that the multiplicative closure of \( T \) is a free commutative monoid over \( T \), hence the \( K \)-algebra isomorphism \( P \cong KT^* \) holds. Vice versa, we have

2.9.1. For every set \( T \) the \( K \)-algebra \( KT^* \) is a polynomial ring over \( K \) in the set of variables \( T \). \( \square \)
A routine third application of the Extension Principle 2.1 now gives

2.9.2. If $K \cap T = \emptyset$ there exists a polynomial ring over $K$ in the set of variables $T$ containing $K$ as a unital subring.

Proof. Define $\varphi : K \cup T \to KT^*$, $x \mapsto \begin{cases} x1_A & \text{if } x \in K \\ x & \text{if } x \in T \end{cases}$. Then $\varphi$ is injective and $\varphi|_K$ is a monomorphism. Applying 2.1 with $M = KT^*$, $B = K \cup T$, to both operations on $K$, $M$ resp., we obtain the claim.

We write $K[T]$ for a polynomial ring over $K$ in the set variables $T$ containing $K$ as a unital subring.

The elements of a polynomial ring $P$ over $T$ have the form

$$\sum_{j_1, \ldots, j_k \in \mathbb{N}_0} c_{j_1, \ldots, j_k} t_1^{j_1} \cdots t_k^{j_k}$$

where $t_1, \ldots, t_k$ are mutually distinct elements of $T$ and $\mathbb{N}_0^k \to K$, $(j_1, \ldots, j_k) \mapsto c_{j_1, \ldots, j_k}$, is a mapping of finite support.

2.9.3 (The universal property). Let $P$ be a polynomial ring over $K$ in the set of variables $T$, $B$ be an arbitrary associative commutative $K$-algebra. Then every mapping of $T$ into $B$ extends uniquely to a unital $K$-algebra homomorphisms of $P$ into $B$.

Proof. Let $F_0 \in B^T$. Any extension $F$ of $F_0$ to a unital $K$-algebra homomorphism of $P$ into $B$ must satisfy the following condition, for any choice of elements $t_i \in T$ and $c_{j_1, \ldots, j_k} \in K$:

$$(\sum_{j_1, \ldots, j_k \in \mathbb{N}_0} c_{j_1, \ldots, j_k} t_1^{j_1} \cdots t_k^{j_k}) F = \sum_{j_1, \ldots, j_k \in \mathbb{N}_0} c_{j_1, \ldots, j_k} (t_1 F_0)^{j_1} \cdots (t_k F_0)^{j_k}$$

Hence there is at most one such extension. On the other hand, denoting by $M$ the multiplicative closure of $T$ in $P$, $F_0$ extends to a unital multiplicative homomorphism $F_1$ of $M$ into $B$, by 2.3. We know that $P \cong KM$ so that 2.8.1 shows how to extend $F_1$ to a $K$-algebra homomorphism from $P$ into $B$. 38

As a consequence, we obtain by the standard argument as given in 2.4(2),

2.9.4. Let $P, P'$ be polynomial rings over $K$ in the set of variables $T, T'$ resp. Let $F_0$ be a bijection of $T$ onto $T'$. Then the extension of $F_0$ from 2.9.3 is a $K$-algebra isomorphism of $P$ onto $P'$.

In particular, a polynomial ring over $K$ in the set of variables $T$ is unique up to isomorphism. 38

38It turns out that the extension is indeed given by the rule $(\ast)$ which, with appropriate comments, could have been used as its definition. But then we would have had to prove the desired properties in detail which we obtained for free by applying 2.3 and 2.8.1.
Making further use of the notation introduced in the proof of 2.9.3, the unique extension $F$ of $F_0$ as described in $(\ast)$ is called the replacement homomorphism\textsuperscript{39} with respect to $F_0$: In every polynomial over $T$, each variable $t \in T$ is replaced by $tF_0$. If $T$ is finite and consists of the mutually distinct elements $t_1, \ldots, t_n$, we write $F_{b_1, \ldots, b_n}$ for the replacement homomorphism which maps $t_i$ to $b_i$ for all $i \in \mathbb{2}$. Thus we obtain a function
\[
P \times B^n \to B
\]
\[
(f, (b_1, \ldots, b_n)) \mapsto f F_{b_1, \ldots, b_n} =: f(b_1, \ldots, b_n).
\]
Its left component functions $(., (b_1, \ldots, b_n)) : f \mapsto f(b_1, \ldots, b_n)$, are the replacement homomorphisms of $P$ into $B$. Its right component functions $(f, .) : (b_1, \ldots, b_n) \mapsto f(b_1, \ldots, b_n)$, are the polynomial functions of $B^n$ into $B$ which, of course, are no homomorphisms in general. Thus each polynomial $f \in P$ induces a polynomial function. Easy examples show that there may exist distinct polynomials which induce the same polynomial function.\textsuperscript{40} Therefore one should bear in mind:

**Polynomials and polynomial functions are distinct mathematical objects which may not be (in whatever sense) “identified”.**\textsuperscript{41}

Let $P$ be a polynomial ring over $K$ in $n$ variables, $b_1, \ldots, b_n$ (not necessarily distinct) elements of a commutative associative unitary $K$-algebra $B$. We put
\[
K[b_1, \ldots, b_n] := P F_{b_1, \ldots, b_n}
\]
which, by applying various parts of 1.3, is easily seen to be a unital $K$-subalgebra of $B$. Likewise it is clear that $\ker F_{b_1, \ldots, b_n} = \{f \in P, f(b_1, \ldots, b_n) = 0_B\}$ is a subalgebra\textsuperscript{42} of $P$. The $n$-tuple $(b_1, \ldots, b_n) \in B^n$ is called algebraically independent if $\ker F_{b_1, \ldots, b_n} = \{0_P\}$, i.e., if $F_{b_1, \ldots, b_n}$ is an isomorphism of $P$ onto $K[b_1, \ldots, b_n]$. Otherwise $(b_1, \ldots, b_n)$ is called algebraically dependent. For every $f \in \ker F_{b_1, \ldots, b_n}$ we have $f(b_1, \ldots, b_n) = 0_B$.

\textsuperscript{39}This is the general form of the special case of one variable considered on p. 9.

\textsuperscript{40}For example, if $K$ is the field with just two elements (p. 26), then all polynomials $t^j$ where $j \in \mathbb{N}$ induce the same polynomial function on $K$: the identity mapping, but they are clearly mutually distinct as polynomials (in one variable).

\textsuperscript{41}The reader will have observed that we avoid throughout this by numerous authors frequently used term. The reason is simple: Either two objects are identical, making any “identifying process” superfluous, or they are not identical and “identifying” them means to make a mistake. The intention in using the expression is usually to refer to some 1-1 correspondence between objects which is considered natural in the context, and to indicate that the smart author is not willing to meddle with notational distinction. As a typical example – here mentioned only for convenience as we cited this otherwise recommendable book anyway – see [St, p 19]. Note that the author emphasizes the distinction between polynomials and polynomial functions on p. 20. But be aware that, confusingly, a symmetric polynomial is later called a “symmetric function” (17.1), despite the author’s own warning in 1.30. Thus, in the terminology of [St], a symmetric function is not a function. Terminological inconsistencies of this kind often occur as a result of an attempted compromise between conceptual clearness and certain traditions of mathematical language.

\textsuperscript{42}With respect to multiplication, $\ker F_{b_1, \ldots, b_n}$ is not only closed but has the much stronger property of being an ideal of $P$ (cf. 2.5). Clearly, in a commutative algebra there is no distinction between left and right ideals.

39
Any \( n \)-tuple \((b_1, \ldots, b_n) \in B^n\) with this property is called a zero of \( f\), generalizing the definition given on p. 9. The fundamental topic of Algebraic Geometry is to study the set of common zeros in \( B^n\) for a subset of \( P\). For example, if \( n = 2, B = \mathbb{R}, f = t_1^2 + t_2^2 - 1\), the set of zeros is \(\{(b_1, b_2) | b_1, b_2 \in \mathbb{R}, b_1^2 + b_2^2 = 1\}\) – the unit circle (in the real plane).

If \( T\) consists of exactly \( n \) elements \( t_1, \ldots, t_n\), we write \( K[t_1, \ldots, t_n]\) instead of \( K[T]\). We will adopt the convention that, for a given commutative associative unitary ring \( K\), the notation \( K[t_1, \ldots, t_n]\) – coming out of the blue – always stands for the polynomial ring over \( K\) in \( n\) variables \((t_1, \ldots, t_n)\) if the elements \( t_1, \ldots, t_n\) have not been introduced with a different meaning in the context. Then every \( \pi \in S_n\) gives rise to the special replacement homomorphism

\[
\bar{\pi} := F_{t_1, \ldots, t_n} : K[t_1, \ldots, t_n] \to K[t_1, \ldots, t_n],
\]

the canonical extension of the mapping \( F_0 : T \to K[t_1, \ldots, t_n], t_i \mapsto t_i\pi\) for all \( i \in \mathbb{N}\).

Clearly, \( \bar{\pi}\) is an automorphism of \( K[t_1, \ldots, t_n]\). For example, if \( n = 3, \pi = \left(\begin{array}{ccc} 1 & 2 & 3 \\ 3 & 2 & 1 \end{array}\right)\), then

\(\begin{align*}
(a) & \quad (1_K + t_1 + t_2t_3 + 2t_1^2t_2t_3)\bar{\pi} = 1_K + t_3 + t_2t_1 + 2t_1^2t_2t_3 = 1_K + t_3 + t_1t_2 + 2t_1t_2t_3^2, \\
(b) & \quad (t_1t_3 - t_2^2)\bar{\pi} = t_3t_1 - t_2^2 = t_1t_3 - t_2^2, \\
(c) & \quad (t_1^3 + t_2^3 + t_3^3)\bar{\pi} = t_3^3 + t_2^3 + t_1^3 = t_1^3 + t_2^3 + t_3^3.
\end{align*}\)

Part (a) illustrates the typical behaviour of \( \bar{\pi}\): It modifies a polynomial, maintaining, however, certain characteristics of it. The action of \( \bar{\pi}\) in the case of the polynomial in (b) turns out to be trivial. The intermediate feeling that “something is happening” vanishes when we compare the given polynomial with its image under \( \bar{\pi}\): \( t_1t_3 - t_2^2\) is a fixed point for \( \bar{\pi}\). While this observation clearly is due to the particular choice of \( \pi\), the polynomial in (c) is not only a fixed point for \( \bar{\pi}\) but in regard to any chosen permutation of \( \{1, 2, 3\}\).

A polynomial \( f \in K[t_1, \ldots, t_n]\) is called symmetric if \( f\bar{\pi} = f\) for all \( \pi \in S_n\). We write \( K_{sym}[t_1, \ldots, t_n]\) for the set of all symmetric polynomials in \( K[t_1, \ldots, t_n]\). It is easily seen to be a subalgebra of \( K[t_1, \ldots, t_n]\).

2.9.5. Let \( n \in \mathbb{N}, f \in K_{sym}[t_1, \ldots, t_n]\), \( B\) a commutative associative unitary \( K\)-algebra. Then the polynomial function \( B^n \to B, (b_1, \ldots, b_n) \mapsto f(b_1, \ldots, b_n)\), is symmetric. \(\square\)

The elementary symmetric functions, introduced in 1.12 in connection with Viète’s theorem, arise (in the sense of 2.9.5) from the so-called elementary symmetric polynomials in \( K[t_1, \ldots, t_n]\), defined by

\[
s_k := \sum_{t_{i_1} < t_{i_2} < \cdots < t_{i_k}} t_{i_1}t_{i_2}\cdots t_{i_k} \quad \text{for all } k \in \mathbb{N}.
\]

\(43\)Later we will see a simple general reason for this fact.
$s_0 := 1_K$, $s_k = 0$ for $k > n$. Their importance lies in the following result which will not be proved here; for a proof see, e.g., [Bo, 4.4]:

**Main Theorem on symmetric polynomials.** Let $K$ be a commutative unitary ring, $n \in \mathbb{N}$. Then $\{s_1, \ldots, s_n\}$ is a $K$-algebra basis of $K_{\text{sym}}[t_1, \ldots, t_n]$.

This means that $K[s_1, \ldots, s_n]$ is a polynomial ring over $K$ in the set of variables $\{s_1, \ldots, s_n\}$. Thus the replacement homomorphism of $K[t_1, \ldots, t_n]$ into $K_{\text{sym}}[t_1, \ldots, t_n]$ given by $t_i \mapsto s_i$ for all $i \in \mathfrak{u}$ (in other words: $f \mapsto f(s_1, \ldots, s_n)$) is a $K$-algebra isomorphism: $\mathbb{K}_{\text{sym}}[t_1, \ldots, t_n] = K[s_1, \ldots, s_n] \cong K[t_1, \ldots, t_n]$.

A straightforward generalization of the above example (c) leads to the notion of a powersum symmetric polynomial, defined for any $k \in \mathbb{N}_0$ by

$$t_1^k + t_2^k + \cdots + t_n^k =: p_k.$$  

Applying Viète’s theorem 1.11 with $b_i = t_i$ for all $i \in \mathfrak{u}$, we obtain $(t - t_1) \cdots (t - t_n) = \sum_{k=0}^{n} (-1)^k s_{n-k} t^k$, hence $0_K = \sum_{k=0}^{n} (-1)^k s_{n-k} t_j^k$ for all $j \in \mathfrak{u}$. It follows that $\sum_{k=0}^{n} (-1)^k s_{n-k} p_k = 0_K$.

We will finish this chapter by a fourth application of the Extension Principle 2.1: Clearly, any unital subring of a field is an integral domain. Our aim is to show that all integral domains arise as subrings of fields. Given an integral domain $R$, the main step consists in the construction of a field which in a natural way emanates from the structure of $R$. We start with the observation that $R \times R$ is a commutative monoid with respect to both of the operations $\cdot$, $\star$ defined as follows:

$$\forall a, b, c, d \in R \quad (a, c) \cdot (b, d) := (ab, cd), \quad (a, c) \star (b, d) := (ad + bc, cd).$$

The neutral elements are $(1_R, 1_R)$ (with respect to $\cdot$), $(0_R, 1_R)$ (with respect to $\star$). As $R$ is an integral domain, the subset

$$B := R \times \hat{R}$$

of $R \times R$ is closed with respect to both operations. As it contains both neutral elements, it is a twofold unital submonoid, with respect to $\cdot$ and with respect to $\star$.

---

44These supplementary definitions are consistent as $s_0$ is the “sum of all empty products” over $\{t_1, \ldots, t_n\}$, i.e., $s_0 = \sum 1_K = 1_K$ while $s_k$, for $k > n$, is the empty sum in $K[t_1, \ldots, t_n]$, namely the sum of all products of $k$ mutually distinct factors over $\{t_1, \ldots, t_n\}$: there are none.

45This is another example for the phenomenon that a structure is isomorphic to a proper substructure, as in the familiar case of the non-trivial subgroups $k\mathbb{Z}$ ($k \in \mathbb{N}$) of the group $(\mathbb{Z}, +)$.

46Equivalently, $n s_n = \sum_{k \in \mathbb{N}_0} (-1)^{k-1} s_{n-k} p_k$. This is the $n$-th of the famous Newton-Girard formulae, proved in the 17th century:

$$\forall m \in \mathbb{N} \quad m s_m = \sum_{k \in \mathbb{N}_0} (-1)^{k-1} s_{m-k} p_k \quad (\text{in } K[t_1, \ldots, t_n])$$

47
Note that the mapping \( \iota : R \to B, \ a \mapsto (a, 1_R) \), is a monomorphism of the integral domain \((R, +, \cdot)\) into the double magma \((B, \star, \cdot)\).

The main idea is to define a partition \( \mathcal{M} \) on the set \( B \) and two operations on \( \mathcal{M} \) which make it into a field that allows an embedding of \( R \). Our procedure consists of six steps:

(I) The set \( \mathcal{M} \)

Let \( \sim \) be the relation on \( B \) defined by

\[(a, b) \sim (a', b') \text{ if and only if } ab' = a'b\]

for arbitrary \( a, a' \in R, b, b' \in \hat{R} \). Clearly, \( \sim \) is reflexive and symmetric. Let \( a, a', a'' \in R, b, b', b'' \in \hat{R} \) such that \((a, b) \sim (a', b'), (a', b') \sim (a'', b'')\). Then \( ab' = a'b, a'b'' = a''b' \), hence \( ab'' = a'b' = a''b' \). It follows that \( b'(ab'' - a''b) = 0_R \)

which implies that \( ab'' - a''b = 0_R \) as \( b' \neq 0_R \) and \( R \) is an integral domain. Thus \((a, a'') \sim (b, b'')\). This proves that \( \sim \) is transitive, hence an equivalence relation on \( B \).

For all \( a \in R, b \in \hat{R} \) we write \([a, b]\) for the equivalence class of \((a, b)\) with respect to \( \sim \) and \( \mathcal{M} \) for the set of all equivalence classes of \( \sim \). It is well known that \( \mathcal{M} \) is a partition of \( B \). The mapping

\[\kappa : B \to \mathcal{M}, \ (a, b) \mapsto [a, b],\]

is a surjection.

(II) The multiplication \( \Box \) on \( \mathcal{M} \)

We know from 1.2 that the multiplication \( \cdot \) in \( B \) induces a natural operation on \( \mathfrak{P}(B) \). For the product of two subsets \( U, V \) of \( B \) we write here simply \( UV = \{u \cdot v | u \in U, v \in V\} \). The problem is that, for \( U, V \in \mathcal{M} \), in general the set \( UV \) will not be an equivalence class with respect to \( \sim \) so that the ordinary subset product does not define an operation on \( \mathcal{M} \). The key observation, however, is that there exists a unique equivalence class \( W \) which contains \( UV \): As \( \mathcal{M} \) is a partition and \( UV \neq \emptyset \), there exists trivially at most one element of \( \mathcal{M} \) such that \( UV \subseteq W \).

We have to prove existence: Let \( U, V \in \mathcal{M}, a, c, a', c' \in R, b, d, b', d' \in \hat{R} \) such that \((a, b), (a', b') \in U, (c, d), (c', d') \in V\). We claim that\(^{47}\)

\[(a, b) \cdot (c, d) \sim (a', b') \cdot (c', d').\]

As \((a, b), (a', b')\) are elements of the same equivalence class we know that \( ab' = a'b \).

Similarly, \( cd' = c'd \). It follows that \( ab'cd' = a'bc'd \), i.e., \((ac, bd) \sim (a'c', b'd')\). Hence any two elements of \( UV \) are equivalent, i.e., \( UV \subseteq W \) for some \( W \in \mathcal{M} \).

The above key observation now allows us to define the desired operation on \( \mathcal{M} \):

\(^{47}\)A relation \( \sim \) on a set \( X \) is called compatible with an operation \( \cdot \) on \( X \) if \( x \sim y \), \( x' \sim y' \) implies \( x \cdot x' \sim y \cdot y' \), for arbitrary \( x, x', y, y' \in X \). Making use of this notion, the claim means that the equivalence relation \( \sim \) on \( B \) is compatible with \( \cdot \).
For all $U, V \in \mathcal{M}$ define $U \square V$ to be the unique element of $\mathcal{M}$ containing $UV$.

If $(a, b) \in U$, $(c, d) \in V$, then $(ac, bd) \in UV \subseteq U \square V$, hence $U \square V$ is the equivalence class of $(ac, bd)$. Hence we have the rule

$$\forall a, c \in R \forall b, d \in \hat{R} \quad [a, b] \square [c, d] = [ac, bd].$$

In other words, $\kappa$ is an epimorphism of $(B, \cdot)$ onto $(\mathcal{M}, \square)$. By 1.3, $\square$ is associative, commutative, and $[1_R, 1_R]$ is neutral with respect to $\square$. As usual, we put $1_\mathcal{M} := [1_R, 1_R]$. We call $\square$ the multiplication on $\mathcal{M}$.

### (III) The addition $\boxplus$ on $\mathcal{M}$

Imitating the idea in (II), we show that for every $U, V \in \mathcal{M}$ there exists a unique equivalence class which contains $U \star V$. Uniqueness is trivial as in (II). As for existence, we claim, under the same hypotheses as in (II), that

$$(a, b) \ast (c, d) \sim (a', b') \ast (c', d').$$

From $ab' = a'b$, $cd' = c'd$ we conclude that $(ad + bc)b'd' = (a'd' + b'c')bd$, hence $(a, b) \ast (c, d) = (ad + bc, bd) \sim (a'd' + b'c', b'd') = (a', b') \ast (c', d')$. Hence for every $U, V \in \mathcal{M}$ there exists a unique $S \in \mathcal{M}$ such that $U \star V \subseteq S$.

Again, this puts us in a position to define the desired operation on $\mathcal{M}$: For all $U, V \in \mathcal{M}$ define $U \boxplus V$ to be the unique element of $\mathcal{M}$ containing $U \ast V$. Similarly to the line of reasoning in II, we obtain

$$\forall a, c \in R \forall b, d \in \hat{R} \quad [a, b] \boxplus [c, d] = [ad + bc, bd].$$

Thus $\kappa$ is an epimorphism of $(B, +)$ onto $(\mathcal{M}, \boxplus)$. By 1.3, $\boxplus$ is associative, commutative, and $[0_R, 1_R]$ is neutral with respect to $\boxplus$. As usual, we put $0_\mathcal{M} := [0_R, 1_R]$. We call $\boxplus$ the addition on $\mathcal{M}$.

### (IV) The group properties of $(\mathcal{M}, \boxplus)$ and $(\mathcal{M}, \square)$

We first describe the neutral elements of $(\mathcal{M}, \boxplus)$ and of $(\mathcal{M}, \square)$:

$$0_\mathcal{M} = \{(0_R, b) | b \in \hat{R}\}, \quad 1_\mathcal{M} = \{(b, b) | b \in \hat{R}\},$$

as $(a, b) \in 0_\mathcal{M} \iff (a, b) \sim (0_R, 1_R) \iff a = 0_R, (a, b) \in 1_\mathcal{M} \iff (a, b) \sim (1_R, 1_R) \iff a = b$, for all $a \in R, b \in \hat{R}$.

It follows that $[a, b] \boxplus [-a, b] = [0_R, b^2] = 0_\mathcal{M}$ for all $a \in R, b \in \hat{R}$. Hence the commutative monoid $(\mathcal{M}, \boxplus)$ is a group. If $[a, b] \neq 0_\mathcal{M}$ we know that $a \neq 0_R$, hence $(b, a) \in B$ and $[a, b] \square [b, a] = [ab, ab] = 1_\mathcal{M}$. Hence the commutative monoid $(\mathcal{M}, \square)$ is a group. From our proofs we obtain the following formula for the inverses: $-[a, b] = [-a, b], [a, b]^{-1} = [b, a]$ if $a \neq 0_R$. 

43
2.10.2. Let $R$ quotient field of $\psi$ of $L$. The non-trivial point and aim of our construction was the existence of a isomorphism of $Q$ of $R$ in $\psi$. Obviously, $R$ is a subring of a field $\hat{\psi}$ and an extension $\hat{\varphi}$ of $\varphi$ to an isomorphism of $Q$ onto $\mathcal{M}$. We observe that our field $Q$ has the following property:

2.9.6. $\forall x \in Q \exists a \in R \exists b \in \hat{\mathcal{M}} \ x = ab^{-1}$

Proof. As $Q \cong \mathcal{M}$ via $\hat{\varphi}$ it suffices to prove that for all $U \in \mathcal{M}$ there exist $a \in R$, $b \in \hat{\mathcal{M}}$ such that $U = (a \varphi) \otimes (b \varphi)^{-1}$. Let $U \in \mathcal{M}$. Then there exist $a \in R$, $b \in \hat{\mathcal{M}}$ such that

$$U = [a, b] = [a, 1_R] \otimes [1_R, b] = [a, 1_R] \otimes [b, 1_R]^{-1} = (a \varphi) \otimes (b \varphi)^{-1},$$

making use of the last equation in (IV).

2.10 Definition. Let $R$ be an integral domain. A field $Q$ is called a quotient field of $R$ if $R$ is a subring of $Q$ with the property 2.9.6.

2.10.1. Suppose $R$ is a subring of a field $L$. Set $Q := \{ab^{-1} | a \in R, b \in \hat{\mathcal{M}}\}$. Then $Q$ is a subfield of $L$.

Proof. Clearly, $R \subseteq Q$ as $a = a1_{R}^{-1} \in Q$ for all $a \in R$. If $a, c \in R, b, d \in \hat{\mathcal{M}}$, we have $ab^{-1} - cd^{-1} = (ad - bc)(bd)^{-1} \in Q$ and, if $b \neq 0_{R}$, $ab^{-1}(cd^{-1})^{-1} = (ad)(bc)^{-1} \in Q$.

Obviously, $Q$ in 2.10.1 is a quotient field of $R$, the uniquely determined quotient field of $R$ in $\mathcal{M}$. Thus a quotient field of $R$ is immediately found if a field is given containing $R$ as a subring. The non-trivial point and aim of our construction was the existence of a quotient field of $R$ without having an enveloping field at our disposal from the beginning.

2.10.2. Let $\psi$ be an isomorphism of $R$ onto an integral domain $\hat{\mathcal{M}}$. Suppose that $(Q, +, \cdot)$ is a quotient field of $R$, $(\hat{Q}, +, \cdot)$ is a quotient field of $\mathcal{M}$. Then there is a unique extension $\hat{\psi}$ of $\psi$ to an isomorphism of $(Q, +, \cdot)$ onto $(\hat{Q}, +, \cdot)$.
Proof. Let \(a, a' \in R, b, b' \in \hat{R}\). Then
\[
a \cdot b^{-1} = a' \cdot b'^{-1} \iff ab' = a'b \iff (a\psi)(b'\psi) = (a'\psi)(b\psi) \iff (a\psi) \cdot (b\psi)^{-1} = (a'\psi) \cdot (b'\psi)^{-1}.
\]
Therefore the mapping \(\hat{\psi} : Q \to \hat{Q}, a \cdot b^{-1} \mapsto (a\psi) \cdot (b\psi)^{-1}\), is well-defined, bijective and an extension of \(\psi\). The homomorphism properties of \(\hat{\psi}\) are immediate. \(\square\)

2.11 Theorem. Let \(R\) be an integral domain.

(1) There exists a quotient field of \(R\).

(2) Any two quotient fields of \(R\) are isomorphic.

(3) Let \(Q\) be a quotient field of \(R\). Then every automorphism of \(R\) extends uniquely to an automorphism of \(Q\).

Proof. We obtain (1) by the above 6-step construction, (2) by choosing \(R = \hat{R}\), \(\psi = \text{id}_R\), (3) by choosing \(R = \hat{R}, Q = \hat{Q}, \psi \in \text{Aut} R\) in 2.10.2. \(\square\)

2.11(2) is the justification of the frequently seen use of the definite article in the formulation “the quotient field of \(R\)” which does not express an absolute uniqueness, but a uniqueness up to isomorphism. As a consequence of 2.11(3), we note

2.11.1. Let \(Q\) be a quotient field of an integral domain \(R\). Then \(\text{Aut} R\) isomorphic to a subgroup of \(\text{Aut} Q\).

As a proof, it suffices to observe that the extension of an automorphism of \(R\) to an automorphism of \(Q\) given by 2.11(3) defines a monomorphism of \(\text{Aut} R\) into \(\text{Aut} Q\). \(\square\)

If \(Q\) is a quotient field of an integral domain \(R\) and \(a \in R, b \in \hat{R}\), the product \(ab^{-1}\) is alternatively also denoted by \(\frac{a}{b}\), called fraction notation. This notation is motivated by the classical case of \(R = \mathbb{Z}\) in which a quotient field is called field of rational numbers and commonly denoted by \(\mathbb{Q}\). Traditionally, the elements of \(Q\) are written as fractions. The above steps (I) – (VI) may be viewed as a general version of a construction of \(Q\) from \(\mathbb{Z}\). The element \(\frac{a}{b}\) of \(Q\) is the element corresponding to the equivalence class \([a, b]\) defined at the end of step (I). If \((a', b') \in R \times \hat{R}\) any representative of the same class, the fraction \(\frac{a'}{b'}\) therefore denotes the same element of \(Q\). Reducing (expanding resp.) the fraction \(\frac{a}{b}\) means passing to \(\frac{a'}{b'}\) where \(\vert a'\vert < \vert a\vert\) (\(\vert a'\vert > \vert a\vert\) resp.), \((a, b) \sim (a', b')\).

From 1.1.2 it follows inductively that, for any integral domain \(R\), the polynomial ring in \(n\) variables \(R[t_1, \ldots, t_n]\) is an integral domain. If \(Q\) is a quotient field of \(R\), we write \(Q(t_1, \ldots, t_n)\) for the quotient field of \(R[t_1, \ldots, t_n]\). Its elements are commonly written as fractions \(\frac{f}{g}\) where \(f, g \in R[t_1, \ldots, t_n], g \neq 0_R\). In particular, if \(K\) is a field, \(K(t_1, \ldots, t_n)\) stands for the quotient field of \(K[t_1, \ldots, t_n]\).\(^{49}\) We know that every \(\pi \in S_n\) determines

\(^{48}\)The image of an element \(x \in Q\) is given only in dependence of the elements \(a \in R, b \in \hat{R}\) chosen to represent \(x\) in the form \(a \cdot b^{-1}\). Although this representation is not unique, we have shown that the element in \(Q\) assigned to \(x\) is unique: It is the same for any choice of \(a, b\) with \(x = a \cdot b^{-1}\).

\(^{49}\)More generally, if \(K\) is a subfield of a field \(L\) and \(b_1, \ldots, b_n \in L\), then \(K(b_1, \ldots, b_n)\) is the quotient field of \(K[b_1, \ldots, b_n]\) in \(L\) (cf. 2.10.1).
an automorphism \( \bar{\pi} \) of \( K[t_1, \ldots, t_n] \) (see p. 40). By 2.11(3), \( \bar{\pi} \) extends uniquely to an automorphism of \( K(t_1, \ldots, t_n) \). Assigning this automorphism to \( \pi \in S_n \), we obtain a monomorphism of \( S_n \) into \( \text{Aut} K(t_1, \ldots, t_n) \). Every automorphism obtained in this way induces the identity automorphism on the subring \( K_{\text{sym}}[t_1, \ldots, t_n] \), hence also on its quotient field \( K_{\text{sym}}(t_1, \ldots, t_n) \) in \( K(t_1, \ldots, t_n) \). From the definition of \( K_{\text{sym}}[t_1, \ldots, t_n] \) we obtain that \( K_{\text{sym}}(t_1, \ldots, t_n) \) is the set of all fixed points with respect to the set of automorphisms induced by the symmetric group \( S_n \). We conclude:

2.11.2. Let \( n \in \mathbb{N} \), \( K \) a field. Then there exists a monomorphism of \( S_n \) into the subgroup \( \{ \alpha \mid \alpha \in \text{Aut} K(t_1, \ldots, t_n), \forall q \in K_{\text{sym}}(t_1, \ldots, t_n) \quad q\alpha = q \} \) of \( \text{Aut} K(t_1, \ldots, t_n) \). \( \square \)

Later (see p. 73) we will see that this monomorphism is in fact an isomorphism, in other words, that every automorphism of \( K(t_1, \ldots, t_n) \) which fixes each element of \( K_{\text{sym}}(t_1, \ldots, t_n) \) must necessarily permute the set of variables \( \{ t_1, \ldots, t_n \} \).
3 Cosets and homomorphisms

Let \( A, B \) sets, \( \varphi \in B^A \). For all \( b \in A \varphi \) we set
\[
b \varphi^- := \{ a | a \in A, a \varphi = b \},
\]
called the complete pre-image of \( b \). Let \( \mathcal{M} \) be the set of all pre-images \( b \varphi^- \) where \( b \in A \varphi \).

The condition of having the same image with respect to \( \varphi \) defines an equivalence relation on \( A \), and \( \mathcal{M} \) is the set of its equivalence classes. It follows that

3.0.1. \( \mathcal{M} \) is a partition of \( A \), and \( \varphi^- \) is a bijection of \( A \varphi \) onto \( \mathcal{M} \),
because \( \varphi^- \) is a surjection of \( A \varphi \) onto \( \mathcal{M} \) by the definition of \( \mathcal{M} \), and \( b \varphi^- = b' \varphi^- \) (where \( b, b' \in A \varphi \)) implies, for an arbitrary \( a \in b \varphi^- \), that \( b = a \varphi = b' \), hence \( b = b' \). \( \square \)

3.1 Definition. Let \( (G, \circ) \) be a group, \( H \leq G \), and set \( \rho : G \to \mathfrak{P}(G), g \mapsto H \circ g \) \((:= \{ h \circ g | h \in H \})\). Let \( g \in G \). The set \( H \circ g \) is called the right coset of \( g \) with respect to \( H \). We have, for any \( g' \in G \),
\[
g' \in (H \circ g)\rho^- \iff g' \rho = H \circ g \iff H \circ g' = H \circ g \iff g' \circ g^{-1} \in H \iff g' \in H \circ g,
\]
hence \( (H \circ g)\rho^- = H \circ g \). We write \( G/H \) for the set of all right cosets of \( H \) in \( G \). The mapping \( H \to H \circ g, h \mapsto h \circ g \), is a bijection. Applying 3.0.1 with \( A = G, B = \mathfrak{P}(G), \varphi = \rho, \mathcal{M} = G/H \), we obtain:

3.1.1. \( G/H \) is a partition of \( G \) which consists of subsets of \( G \) all of which are equipotent to \( H \).

If the number of right cosets of \( H \) in \( G \) is finite, this number is called the index of \( H \) in \( G \) and denoted by \( |G : H| \). The following important statement is an immediate consequence of 3.1.1:

3.1.2 (Lagrange’s theorem). Let \( G \) be finite. Then \( |G| = |G : H||H| \). \( \square \)

\( |G| \) is called the order of \( G \). By 3.1.2, we have for every subgroup \( H \) of a finite group \( G \)
\[
|H| \ \bigg| \ \ |G|, \quad |G : H| = \frac{|G|}{|H|}.
\]

Instead of \( \rho \), we could have considered \( \lambda : G \to \mathfrak{P}(G), g \mapsto g \circ H \). The set \( g \circ H \) is called the left coset of \( g \) with respect to \( H \). The set of all left cosets of \( H \) in \( G \) is a partition
of $G$, denoted by $G/H$. We observe that the bijection $\iota : G \to G, g \mapsto g^{-1}$, induces a bijection of $G/H$ onto $G/H$ as

\[(Hg)\iota = \{(hg)^{-1}|h \in H\} = \{g^{-1}h^{-1}|h \in H\} = g^{-1}H \quad \text{for all } g \in G.\]

Hence $|G/H| = |G/H|$. But in general, $g \circ H \neq H \circ g$. For example, let $G = S_3$, $H = \{\text{id}, (1 2 3), (1 2)\}, g = \left(\begin{array}{ccc} 1 & 2 & 3 \\ 1 & 2 & 3 \end{array}\right)$. Then

\[H \circ g = \left\{ \left(\begin{array}{ccc} 1 & 2 & 3 \\ 1 & 2 & 3 \end{array}\right), \left(\begin{array}{ccc} 1 & 2 & 3 \\ 1 & 2 & 3 \end{array}\right) \right\} \neq \left\{ \left(\begin{array}{ccc} 1 & 2 & 3 \\ 1 & 2 & 3 \end{array}\right), \left(\begin{array}{ccc} 1 & 2 & 3 \\ 1 & 2 & 3 \end{array}\right) \right\} = g \circ H.

A subgroup $H$ of $(G, \circ)$ is called normal if $g \circ H = H \circ g$ for all $g \in G$ which is denoted by $H \trianglelefteq G$. For a normal subgroup, the set of right cosets coincides with the set of left cosets so that there is no need to use distinct notation. The set of all cosets of $H$ in $G$ is denoted by $G/H$ if $H \trianglelefteq G$. Clearly, the subgroups $\{1_G\}$ and $G$ are normal. The group $G$ is called simple if $G \neq \{1_G\}$ and $\{1_G\}$, $G$ are the only normal subgroups of $G$. The following remark is trivial:

3.1.3. If $(G, \circ)$ is abelian, then every subgroup of $G$ is normal. \hfill \-box

In $S_3$, the subgroup of order 3 is a first not completely trivial example of a normal subgroup. Still there is a simple general reason for this fact because a subgroup it is of index 2 in $S_3$:

3.1.4. Every subgroup of $G$ of index 2 in $G$ is normal in $G$.

Proof. Suppose $H < G, |G : H| = 2$. Certainly, $g \circ H = H = H \circ g$ for all $g \in H$. Now let $g \notin H$. Then $H \circ g \neq H = g \circ H$, hence $H \circ g = G \setminus H = g \circ H$ by 3.1.1 as $|G/H| = 2 = |G/H|$. \hfill \-box

3.1.5. Suppose $H \trianglelefteq G$, and let $\circ$ be the operation on $\mathfrak{P}(G)$ defined in 1.2. Then $\rho$ is a homomorphism of $(G, \circ)$ into $(\mathfrak{P}(G), \circ)$, ker $\rho = H$, $G\rho = G/H$.

Proof. Let $g, g' \in G$. We have, making use of the normality of $H$ in the third step,

\[(g \circ g')\rho = H \circ g \circ g' = H \circ H \circ \{g\} \circ \{g'\} = H \circ \{g\} \circ H \circ \{g'\} = (H \circ g) \circ (H \circ g') = g \rho \circ g' \rho,\]

$g \in \ker \rho \iff g \rho = 1_G \rho \iff H \circ g = H \iff g \in H$. The last assertion is trivial. \hfill \-box

Let $H \trianglelefteq G$. By 1.3(4) and 3.1.5, $G/H$ is a subgroup of $(\mathfrak{P}(G), \circ)$. The group $(G/H, \circ)$ is called the factor group of $G$ over $H$. Its elements are the cosets $H \circ g$ ($g \in G$). The mapping $\rho$ from 3.1.5 is called the canonical epimorphism of $G$ onto $G/H$. We do not exaggerate when we say that the notion of a factor group is a fundamental element of mathematics. We are now in a position to prove one of the most influential algebraic instruments:
3.2 Theorem (Homomorphism theorem for groups). Let \((G, \circ)\) be a group, \(\varphi\) a homomorphism of \((G, \circ)\) into a magma \((B, \cdot)\), \(H := \ker \varphi\).

(1) \(H \trianglelefteq G, \forall g \in G \quad (g\varphi)\varphi^{-} = H \circ g,\)

(2) \(\varphi^{-}\) is an isomorphism of \((G\varphi, \cdot)\) onto \((G/H, \circ)\).

Proof. By 1.3(4), \(G\varphi\) is a subgroup of \((B, \cdot)\), \(\ker \varphi\) a subgroup of \((G, \circ)\). Let \(g \in G\), \(b := g\varphi\). We show

\[
(*) \quad H \circ g \subseteq b\varphi^{-} \subseteq g \circ H : 
\]

For all \(h \in H\) we have \((h \circ g)\varphi = (h\varphi) \cdot (g\varphi) = (1_G\varphi) \cdot (g\varphi) = (1_G \circ g)\varphi = g\varphi = b\). Hence \(H \circ g \subseteq b\varphi^{-}\). Let \(x \in b\varphi^{-}\). Then \(x\varphi = b = g\varphi\), hence \((g^{-1} \circ x)\varphi = (g\varphi)^{-1} \cdot (x\varphi) = b^{-1} \cdot b = 1_G\varphi\) by 1.3(3). It follows that \(g^{-1} \circ x \in H\), \(x \in g \circ H\). Thus \(b\varphi^{-} \subseteq g \circ H\). This proves \((*)\).

For all \(g \in G\) it follows that \(H \circ g \subseteq g \circ H\), thus also \(H \circ g^{-1} \subseteq g^{-1} \circ H\). We conclude \(g \circ H = g \circ H \circ g^{-1} \circ g \subseteq g \circ g^{-1} \circ H \circ g = H \circ g\). Now \(H \circ g = b\varphi^{-} = g \circ H\), i.e., (1).

This implies, by 3.0.1, that \(\varphi^{-}\) is a bijection of \(G\varphi\) onto \(G/H\). Let \(b, b' \in G\varphi\) and \(g, g' \in G\) such that \(g\varphi = b\), \(g'\varphi = b'\). Then \((g \circ g')\varphi = (g\varphi) \cdot (g'\varphi) = b \cdot b'\), hence \((b \cdot b')\varphi^{-} = H \circ g \circ g' = (H \circ g)\varphi^{-} = b\varphi^{-} \cdot b'\varphi^{-}\), making use of (1) and 3.1.5.

Combining 3.1.4 and 3.2(1), we obtain the following characterization of normal subgroups:

**3.2.1. The normal subgroups of a group \(G\) are exactly the kernels of the homomorphisms of \(G\).**

The aim of 3.2 is the statement that every homomorphic image of a group is isomorphic to the factor group over the kernel (which automatically is a normal subgroup). Thus, to apply the homomorphism theorem, one determines the image \(Y\) of a group \(G\) under a homomorphism and its kernel \(H\). Then 3.2(2) implies that

\[G/H \cong Y.\]

Whenever an isomorphism between a factor group \(G/H\) of \(G\) and a group \(Y\) is to be proved, we have the method of 3.2 at our disposal: It suffices to find a homomorphism of \(G\) onto \(Y\) the kernel of which is \(H\). *Done that, there is no further step necessary*: The desired isomorphism follows directly from 3.2(2). This strategy of proving isomorphisms is the main point of the homomorphism theorem and makes it into an indispensable instrument with countless applications.

For example, in the proof of 1.3.2 and in 1.3.3 we have considered various homomorphisms. Accordingly, we obtain the following group isomorphisms:

(a) \((\mathbb{C}/K(0,1), \cdot) \cong (\mathbb{R}_{>0}, \cdot)\). (Consider \(\mathbb{C} \to \mathbb{R}, \ z \mapsto |z|\).)
(b) \((\mathbb{Q}/\mathbb{Z}, +) \cong (R(1), \cdot)\). (Consider \(\mathbb{Q} \to \mathbb{C}, \ x \mapsto \exp(x \cdot 2\pi i)\).)

49
(c) \((\hat{C}/R_n(1), \cdot) \cong (\hat{C}, \cdot)\) for all \(n \in \mathbb{N}\). (Consider \(\hat{C} \rightarrow \hat{C}, \ z \mapsto z^n\).)

In particular, \((\hat{C}, \cdot)\) is an example of a group which is isomorphic to a proper factor group.

(d) \((\mathbb{Z}/n\mathbb{Z}, +) \cong (R_n(1), \cdot)\) for all \(n \in \mathbb{N}\). (Consider \(\mathbb{Z} \rightarrow \hat{C}, k \mapsto \exp\left(\frac{k2\pi i}{n}\right)\).)

We add a few more examples:

Examples.

(1) Let \(K\) be a field, \(n \in \mathbb{N}\). The determinant defines a multiplicative epimorphism of the group \(GL(n, K)\) of invertible \(n \times n\) matrices (called the general linear group) over \(K\) onto the multiplicative group \(\hat{K}\) of \(K\). Its kernel is the set \(SL(n, K)\) of all \(n \times n\) matrices over \(K\) of determinant \(1\) \((\text{called the special linear group})\). We obtain:

\[SL(n, K) \trianglelefteq GL(n, K)\quad \text{and} \quad GL(n, K)/SL(n, K) \cong \hat{K}.\]

(2) Let \(n \in \mathbb{N}_{>1}\). Mapping every permutation \(\pi \in S_n\) to its sign is an epimorphism of the symmetric group \(S_n\) onto the multiplicative group \(\{1, -1\}\). Its kernel is the set \(A_n\) of even permutations of \(\mathbb{n}\) \((\text{called the alternating group on} \ n)\). We obtain:

\[A_n \trianglelefteq S_n\quad \text{and} \quad S_n/A_n \cong \{1, -1\}.\]

(3) Let \(X\) be a (multiplicatively written) group, \(x \in X\), \(\varphi_x : \mathbb{Z} \rightarrow X\), \(k \mapsto x^k\). Then \(\varphi_x\) is a homomorphism of the group \((\mathbb{Z}, +)\) into the group \(X\). We have

\[\ker \varphi_x = \{k|k \in \mathbb{Z}, \ x^k = 1_X\} = n\mathbb{Z}\]

for a unique \(n \in \mathbb{N}_0\), by 1.2.1. The image \(\mathbb{Z}\varphi_x\) is the smallest subgroup of \(X\) containing \(x\), also called the subgroup generated by \(x\) and denoted by \(\langle x \rangle\). By 3.2,

\[\mathbb{Z}/n\mathbb{Z} \cong \langle x \rangle.\]

If \(n \neq 0\), \(n\) is called the order of \(x\) and denoted by \(o(x)\). We then have

\[\{1_X, x, x^2, \ldots, x^{o(x)-1}\} = \mathbb{Z}\varphi_x \leq X.\]

By 3.1.2, we obtain

\[3.2.2. \text{Let} \ X \text{ be a finite group. Then} \ o(x) = \min\{k|k \in \mathbb{N}, \ x^k = 1_X\} \mid |X| \text{ for all} \ x \in X.\]

If \(n = 0\), \(x\) is called of infinite order.

For all \(T \subseteq X\) we put \(\langle T \rangle := \bigcap_{H \subseteq X} H\). As an intersection of a non-empty set of subgroups of \(X\), \(\langle T \rangle\) is a subgroup of \(X\) and obviously the smallest subgroup containing \(T\), called the subgroup generated by \(T\). If \(T = \{x_1, \ldots, x_k\}\), it is also denoted by \(\langle x_1, \ldots, x_k \rangle\).
3.3 Definition. A group \( X \) is called cyclic if there exists an element \( x \in X \) such that \( \langle x \rangle = X \). Every \( x \) with this property is called a generating element or generator of \( X \). We have seen in the above Example (3) that every cyclic group is a homomorphic image of \((\mathbb{Z}, +)\). More precisely, we have obtained:

3.3.1. Let \((X, \cdot)\) be a cyclic group. Then

\[
(X, \cdot) \cong \begin{cases} 
(\mathbb{Z}, +) & \text{if } X \text{ is infinite,} \\
(\mathbb{Z}/n\mathbb{Z}, +) & \text{if } |X| = n \in \mathbb{N}.
\end{cases}
\]

In particular, \((X, \cdot)\) is commutative.

For example, the group \((\mathbb{R}, \cdot)\) (see 1.1.8, 1.3.2) is cyclic and \(P\) its set of generators, by 1.4.1. The group \((\mathbb{Z}, +)\) is cyclic, generated by 1 and by \(-1\), and these are the only generating elements of \((\mathbb{Z}, +)\) (cf. 1.2.1). \((\mathbb{Q}, +), \{(1, 1), (1, -1), (-1, 1), (-1, -1)\}, \cdot\) are examples of non-cyclic commutative groups. The isomorphism \((\mathbb{R}, \cdot) \cong (\mathbb{Z}/n\mathbb{Z}, +)\) implies, by 3.3.1:

3.3.2. The number of generating elements of a finite cyclic group of order \( n \) is \( \varphi(n) \).

Now let \( G \) be an arbitrary finite group and \( \mathcal{C} \) the set of its cyclic subgroups. As \( x \in \langle x \rangle \) for all \( x \in G \), we have \( \bigcup \mathcal{C} = G \). We define a relation \( \approx \) on \( G \) by putting, for all \( x, y \in G \),

\[ x \approx y \text{ if and only if } \langle x \rangle = \langle y \rangle. \]

Then \( \approx \) is an equivalence relation on \( G \). The equivalence class of an element \( x \in G \) consists of the set of generating elements of \( \langle x \rangle \), hence contains exactly \( \varphi(|\langle x \rangle|) \) elements. As \( G \) is the disjoint union of all equivalence classes with respect to \( \approx \), we obtain:

3.3.3. Let \( G \) be a finite group, \( \mathcal{C} \) the set of its cyclic subgroups. Then \( |G| = \sum_{X \in \mathcal{C}} \varphi(|X|) \).

3.4 Theorem (Main theorem on finite cyclic groups.). Let \( n \in \mathbb{N} \), \( G \) a group of order \( n \), \( \mathcal{T}_n := \{d \mid d \in \mathbb{N}, d|n\} \). The following are equivalent:

(i) \( G \) is cyclic

(ii) For every \( d \in \mathcal{T}_n \) there exists exactly one subgroup of \( G \) of order \( d \).

(iii) For every \( d \in \mathcal{T}_n \) there exists at most one subgroup of \( G \) of order \( d \).

(iv) For every \( d \in \mathcal{T}_n \) there exists at most one cyclic subgroup of \( G \) of order \( d \).

Proof. (i)⇒(ii) If \( G \) is cyclic, we have \( G \cong R_n(1) \) by 3.3.1. Therefore it suffices to prove that \((R_n(1), \cdot)\) has the property (ii). If \( d \in \mathcal{T}_n \), we have \( R_d(1) \leq R_n(1) \) by the first part of 1.3.2, \( |R_d(1)| = d \). Furthermore, if \( U \leq R_n(1) \), \( |U| = d \), then \( u^d = 1 \) for all \( u \in U \), by 3.2.2. Hence \( U \leq R_d(1) \). It follows that \( U = R_d(1) \).

50where \( \varphi \) is the Eulerian function, see 1.4.
The implications (ii$\Rightarrow$(iii$\Rightarrow$(iv) are trivial.

(iv)$\Rightarrow$(i) Let $C$ as in 3.3.3 and $D := \{d | d \in \mathbb{N}, \exists X \in C \mid |X| = d\}$. Clearly $D \subseteq T_n$, by 3.1.2. It follows that

$$n = \sum_{X \in C} \varphi(|X|) = \sum_{d \in D} \varphi(d) \leq \sum_{d \in T_n} \varphi(d) = n,$$

by means of 3.3.3 (first equality), (iv) (second equality), (1.4.4) (last equality). We conclude that $D = T_n$ which, in particular, implies that $n \in D$. Hence $G$ is cyclic. \[\square\]

3.5 **Corollary.** Let $K$ be a field, $G$ a finite subgroup of $(\hat{K}, \cdot)$. Then $G$ is cyclic. In particular, if $K$ is finite, then $(\hat{K}, \cdot)$ is cyclic.

**Proof.** Let $n := |G|$. For every $d \in \mathbb{N}$, in particular for every $d \in T_n$, the polynomial $t^d - 1$ has at most $d$ zeros in $K$, by 1.6(3). By 3.2.2, the elements of a subgroup of order $d$ of $(\hat{K}, \cdot)$ are zeros of that polynomial. Hence there is at most one such subgroup. By 3.4, this implies the claim. \[\square\]

3.6 **Corollary.** Let $G$ be a cyclic group, $H \leq G$. Then $H$ and $G/H$ are cyclic.

**Proof.** Let $x \in G$ such that $G = \langle x \rangle$. Then $\langle Hx \rangle = G/H$, hence $G/H$ is cyclic. If $G$ is infinite we may w.l.o.g. assume that $G = \mathbb{Z}$. By 1.2.1, $H = n\mathbb{Z}$ for some $n \in \mathbb{N}_0$, hence $H = \langle n \rangle$ so that $H$ is cyclic. If $G$ is finite, then $T_{|H|} \subseteq T_{|G|}$, by 3.1.2. Hence 3.4(iii) is inherited from $G$ to $H$. By 3.4, $H$ is cyclic. \[\square\]

Lagrange’s theorem 3.1.2 is a main reason for a strong link between finite group theory and (multiplicative) arithmetic. Let $G$ be a finite group, $n := |G|$, $U(G)$ the set of all subgroups of $G$. The image of the mapping

$$\mu : U(G) \to \mathbb{N}, \ H \mapsto |H|,$$

is a subset of $T_n$, and $U \leq V \in U(G)$ implies that $|U| \mid |V|$, both by 3.1.2. The subgroup relation ($\leq$) defines a partial order on the set $U(G)$, the relation of being a divisor ($\mid$) on the set $T_n$, and $\mu$ constitutes a link between both partially ordered sets (also called “posets”). We now consider this situation in general:

3.7 **Definition.** Let $(M, \leq)$, $(M', \preceq')$ be partially ordered sets$^{51}$. A homomorphism of $(M, \leq)$ into $(M', \preceq')$ is a mapping $\chi$ of $M$ into $M'$ with the property that for all $m, n \in M$

$$m \preceq n \Rightarrow m\chi \preceq' n\chi.$$

If $\chi$ is bijective and $m \preceq n \iff m\chi \preceq' n\chi$ for all $m, n \in M$, $\chi$ is called an isomorphism (of the partially ordered set $(M, \leq)$ onto the partially ordered set $(M', \preceq')$)$^{52}$. If $\chi$ is

$^{51}$A partial order is a reflexive, transitive and antisymmetric relation. If $\preceq$ is a partial order on $M$, then so is the relation $\succeq$, defined by: $m \succeq n$ if and only if $n \preceq m$ ($m, n \in M$). We write $m \prec n$ if $m \preceq n$ and $m \neq n$.

$^{52}$Note that a bijective homomorphism need not be an isomorphism of partially ordered sets: The identity on $T_6$ is a bijective homomorphism of $(T_6, \leq)$ into $(T_6, \leq)$, and $2 \nmid 3, 2 \leq 3$. Hence id$_6$ is not an isomorphism.

52
bijective and \( m \leq n \iff m \chi \preceq n \chi \) for all \( m, n \in M \), \( \chi \) is called an anti-isomorphism. Let \( m, n \in M \) such that \( m \leq n \). Then the set

\[
[m, n] \leq := \{ x \mid x \in M, \ m \leq x \leq n \}
\]

is called the interval between \( m \) and \( n \). If the reference to \( \leq \) is clear from the context, we simply write \([m, n]\). Clearly, \([m, n], \leq \) is then again a partially ordered set.

We shall mainly be interested in lattices, defined as posets \((M, \leq)\) such that for all \( m, m' \in M \) there exists

(i) an element \( i \in M \) such that \( i \leq m, m' \) and \( i \geq k \) for all \( k \leq m, m' \),

(ii) an element \( s \in M \) such that \( m, m' \leq s \) and \( s \leq l \) for all \( l \geq m, m' \).

As a consequence of the antisymmetry, the element \( i \) in (i) (\( s \) in (ii) resp.) is uniquely determined. It is called the infimum (supremum resp.) of \( m, n \), and denoted by \( \inf(m, n) \) (\( \sup(m, n) \) resp.).

For example, for every group \( G \), \((\mathcal{U}(G), \leq)\) is a lattice, called the subgroup lattice of \( G \). For all \( U, V \leq G \) we have \( \inf(U, V) = U \cap V \), \( \sup(U, V) = (U \cup V) \). Similarly, the normal subgroups of \( G \) also form a lattice with respect to \( \leq \). While \( \sup(U, V) \) may be difficult to calculate for arbitrary subgroups, it turns out to be just the product set (see 1.2) if \( U \) or \( V \) is normal. More generally, we note

3.7.1. Let \( G \) be a group and \( U, V \leq G \) such that \( UV = VU \). Then \( UV \in \mathcal{U}(G) \), hence \( \sup(U, V) = UV \).

Proof. If \( UV = VU \) it follows that \((UV)(UV) = U(VU)V = U(UV)V = (UU)(VV) = UV \), i.e., \( UV \) is a submonoid. Let \( u \in U, v \in V \). Then \((uv)^{-1} = v^{-1}u^{-1} \in VU = UV \). Hence \( UV \in \mathcal{U}(G) \), \( UV \subseteq \langle U \cup V \rangle = \sup(U, V) \leq UV \). The claim follows.

Viceversa it is easily seen that, for subgroups \( U, V \) of \( G \), the product set \( UV \) is a subgroup of \( G \) only if \( UV = VU \). A glance at the definition of a normal subgroup (see 3.1) suffices to see that normality of \( U \) (or \( V \)) implies that \( UV = VU \). If both \( U \) and \( V \) are normal in \( G \), so is \( UV \): For all \( g \in G \) we then have \( gUV = UgV = UVg \).

If \( S, T \leq U \) are subgroups of \( G \) and \( U \leq S \), the hypotheses of 2.2.1 are satisfied so that “Dedekind’s law” holds. But the equations in 2.2.1 deal with equality of subsets, not of subgroups of \( G \) as long as it is not known if the involved subgroup products are subgroups. Therefore, “Dedekind’s law” is a statement about the subgroup lattice \( \mathcal{U}(G) \) only under appropriate additional hypotheses.

For every \( n \in \mathbb{N} \), \((T_n, |)\) is a lattice and \( \inf(m, m') = \gcd(m, m') \), \( \sup(m, m') = \lcm(m, m') \) for all \( m, m' \in T_n \). If \( G \) is finite and \( n = |G| \), the mapping \( \mu \) introduced before 3.7 is a lattice homomorphism of \((\mathcal{U}(G), \leq)\) into \((T_n, |)\). The equivalence of the conditions 3.4(i),(ii),(iii) may be reformulated as follows:

\[ G \text{ is cyclic } \iff \mu \text{ is a lattice isomorphism of } \mathcal{U}(G), \leq \text{ onto } (T_n, |) \iff \mu \text{ is injective.} \]

53
The mapping \( \iota : T_n \to T_n, \ d \mapsto \frac{n}{d} \), is an antiautomorphism of the divisor lattice \((T_n, |)\), or, denoting by \( \mathcal{J} \) the relation “is divisible by”, an isomorphism of \((T_n, |)\) onto \((T_n, \mathcal{J})\).

A finite poset \((M, \leq)\) determines an oriented graph with vertex set \(M\) and edges \((m, m')\) whenever \( m < m' \) and \([m, m'] = \{m, m'\} \). This graph is imaged by a system of dots and line segments as follows: Each dot represents an element of \(M\), and dots representing \( m, m' \) are joined by an ascending line segment from \( m \) to \( m' \) if \((m, m')\) is an edge. For example, the subgroup lattice of the group \((R_{12}(1), \cdot)\), its image \((T_{12}, |)\) under \( \mu \), and the image \((T_{12}, \mathcal{J})\) of \((T_{12}, |)\) under \( \iota \) are depicted as follows:

\[
\begin{align*}
(U(R_{12}(1)), \leq) & \quad (T_{12}, |) & \quad (T_{12}, \mathcal{J}) \\
\end{align*}
\]

Passing from a lattice \((M, \leq)\) to \((M, \geq)\) means, in this representation, to “turn the graph upside down”. Divisor lattices have the property that the resulting image has the same outer appearance as before; so the above behaviour of \( T_{12} \) under \( \iota \) is typical. But for lattices in general, such an effect must not be expected.

For an arbitrary finite group \( G \), the image of \( U(G) \) under \( \mu \) consists of the divisors of \(|G|\) which occur as subgroup orders. If \( G \) is not cyclic, we know from 3.4 that \( \mu \) is not injective. The example of the group \( A_4 \) shows that, moreover, \( \mu \) needs not be surjective either: The reader is strongly recommended to check that in \( A_4 \) there are exactly three subgroups of order 2, four subgroups of order 3, and unique subgroups of orders 1, 4, 12 (which are exactly the normal subgroups of \( A_4 \)), but there is no subgroup of order 6.

The cornerstone of finite group theory, to be proved in 8.4(2), is the result that, for every finite group \( G \), \( U(G) \mu \) contains all divisors of \(|G|\) which are prime powers:

**Theorem** (1st Theorem by Sylow (1872)) Let \( G \) be a finite group, \( d \) a divisor of \(|G|\) which is a power of a prime. Then there exists a subgroup of \( G \) of order \( d \).

We will not digress further but take the opportunity and define the fundamental notion which arises from this result: Let \( p \) be a prime, \( G \) a finite group and \( m \in \mathbb{N}_0 \) maximal with the property \( p^m \mid |G| \). A subgroup of order \( p^m \) of \( G \) is then called a **Sylow p-subgroup** of \( G \).

We extend the definition of a complete pre-image from the beginning of this chapter, passing from elements to subsets of the target set of a mapping. Let \( A, B \) sets, \( \varphi \in B^A \).
For all $T \subseteq B$ we set

$$T\varphi^- := \{ a | a \in A, a\varphi \in T \}.$$ 

Then $\varphi^-$ is a mapping of $\mathcal{P}(B)$ to $\mathcal{P}(A)$. For all $X \subseteq A$ we have $X \varphi \varphi^- \supseteq X$.

3.7.2. Let $\varphi$ be a surjective mapping of a set $A$ onto a set $B$. Then $T\varphi^- \varphi = T$ and $S \subseteq T \iff S\varphi^- \subseteq T\varphi^-$. Proof. For all $b \in B$ we have $(b\varphi^-)\varphi = \{ b \}$, hence the first claim. We apply this in the last of the following chain of implications:

$$S \subseteq T \Rightarrow S\varphi^- \subseteq T\varphi^- \Rightarrow S\varphi^- \varphi \subseteq T\varphi^- \varphi \Rightarrow S \subseteq T.$$ 

The claimed equivalence follows.

3.8 Theorem (Correspondence theorem for groups). Let $\varphi$ be an epimorphism of a group $G$ onto a group $\tilde{G}$. Then $\gamma := \varphi^-|_{\mathcal{U}(\tilde{G})}$ is an isomorphism of $(\mathcal{U}(\tilde{G}), \leq)$ onto $([\ker \varphi, G], \leq)$.

(The ellipse sketches the subgroup lattice of $\tilde{G}$ (the subgroup interval $[\ker \varphi, G]$ resp.). The same assertion holds regarding the lattices of normal subgroups of $\tilde{G}$ and of $G$ (containing $\ker \varphi$).

We insert a few comments before we prove this result:

From 1.3(4) we know that $U\varphi \leq \tilde{G}$ if $U \leq G$. The correspondence assertion of 3.8 means that each subgroup of $\tilde{G}$ arises in this way, and that the respective subgroup $U$ of $G$ is uniquely determined under the condition that $\ker \varphi \leq U$. By 3.1.5, for any $N \leq G$ the mapping $\rho : G \rightarrow G/N, g \mapsto Ng$, is an epimorphism. By 3.8, the subgroups of $G/N$ are exactly the quotients $X/N$ where $N \leq X \leq G$. Applying 3.8 to the canonical homomorphism $\rho$, we thus obtain

$$([N, G], \leq) \cong (\mathcal{U}(G/N), \leq).$$

For example, we may determine the subgroups of a factor group $(\mathbb{Z}/n\mathbb{Z}, +)$ of $(\mathbb{Z}, +)$: Every subgroup $X$ of $(\mathbb{Z}, +)$ has the form $k\mathbb{Z}$ for a unique $k \in \mathbb{N}_0$, by 1.2.1. We have $n\mathbb{Z} \leq k\mathbb{Z}$ if and only if $k|n$. Hence the subgroups of $(\mathbb{Z}/n\mathbb{Z}, +)$ are the sets $k\mathbb{Z}/n\mathbb{Z}$ where $k|n$. $(\mathbb{Z}/n\mathbb{Z}, +)$ has no nontrivial proper subgroup if and only if $n = 1$ or $n$ is a prime. Given an arbitrary group $G$, a subgroup $H$ of $G$ is called maximal (denoted by $H < G$) if $H < G$ and there is no subgroup $X$ such that $H < X < G$. Thus

53We re-obtain our former definition if we consider a singleton $T = \{ b \}$ with $b \in A\varphi$. The drawback that $\{ b \}$ should be distinguished from $b$ does less harm than introducing a further function name, different from $\varphi^-$, for this reason.
3.8.1. \( n\mathbb{Z} < \mathbb{Z} \) if and only if \( n \) is a prime.

\[
\text{Proof of 3.8.} \, \text{By 1.3(4), } \text{ker} \, \varphi = 1_{\tilde{\mathbb{Z}}} \varphi^{-} \subseteq Y\gamma \text{ for all } Y \leq \tilde{\mathbb{G}}. \, \text{Furthermore, for any } g, h \in Y\gamma \text{ we have } (gh^{-1})\varphi = (g\varphi)(h\varphi)^{-1} \in Y, \text{ hence } gh^{-1} \in Y\gamma. \, \text{Thus } \gamma \text{ is a mapping of } U(\tilde{\mathbb{G}}) \text{ into } [\text{ker } \varphi, \tilde{\mathbb{G}}], \text{ and injective by the first assertion in 3.7.2. We prove that } \gamma \text{ is surjective: Let } X \in [\text{ker } \varphi, \tilde{\mathbb{G}}], Y := X\varphi. \, \text{Then } Y\gamma = X\varphi^{-} \trianglerighteq X. \, \text{Let } g \in Y\gamma. \, \text{Then } g\varphi \in Y = X\varphi. \, \text{Hence there exists an element } x \in X \text{ such that } g\varphi = x\varphi. \, \text{It follows that } gx^{-1} \in \text{ker } \varphi \leq X, \text{ thus } g \in Xx = X. \, \text{This shows that } Y\gamma = X. \, \text{Thus } \gamma \text{ is a bijection. Now it suffices to apply the second assertion in 3.7.2 to complete the proof.}
\]

If \( X \leq G \) and \( Y := X\varphi \), then \((gX)\varphi = (g\varphi)Y, (Xg)\varphi = Y(g\varphi). \) Hence \( \text{ker } \varphi \leq X \trianglelefteq G \) implies that \( X = Y\gamma, Y \trianglelefteq \tilde{\mathbb{G}}. \) Conversely, if \( Y \trianglelefteq \tilde{\mathbb{G}} \), there exists a homomorphism \( \psi \) of \( \tilde{\mathbb{G}} \) into some other group \( W \) such that \( Y = \text{ker } \psi \), by 3.2.1. Then \( \varphi\psi \) is a homomorphism of \( G \) into \( W \), and \( g \in \text{ker } \varphi\psi \) if and only if \( g\varphi \in \text{ker } \varphi = Y. \) Hence \( \text{ker } \varphi\psi = Y\gamma \), thus \( \text{ker } \varphi \leq Y\gamma \trianglelefteq G \) by 3.2.1. This proves the additional assertion about the lattices of normal subgroups. \( \square \)

3.9 Proposition. Let \( G \) be a group, \( N \trianglelefteq G. \)

1. Let \( U \leq G. \) Then \( U \cap N \trianglelefteq U, \, NU \leq G \) and \( U/N \cap N \cong NU/N. \)

2. Let \( N \subseteq M \leq G, M/N \trianglelefteq G/N. \) Then \( M \trianglelefteq G \) and \( (G/N)/(M/N) \cong G/M. \)

\[
\begin{array}{c}
\text{G} \\
N & U \\
\hline
NU \\
\hline
U \cap N \\
\hline
N \quad \text{G}
\end{array}
\]

\[
\begin{array}{c}
\text{M} \\
\hline
\text{N}
\end{array}
\]

\[
\begin{array}{c}
\varphi : U \to G/N, \, u \mapsto Nu (= Nnu \text{ for all } n \in N).
\end{array}
\]

\[54\] It is useful to sketch lattices not in every detail but to depict just the part which is of interest in the context. For example, in 3.9(1), the subgroups \( N, U, U \cap N, NU \) are of interest. We do not know anything about the subgroup intervals \([U \cap N, U], [U \cap N, N], [N, NU], [U, NU]\), but we know that \( U \cap N = \inf(U, N) \) and \( NU = \sup(U, N) \) (by 3.7.1). Therefore we outline the situation as shown beside the proof although we have no information about the mentioned subgroup intervals. (In a complete representation of the subgroup lattice we would, for example, join \( U \cap N \) and \( U \) by an ascending line segment if and only if \( U \cap N \leq U \)! In such an illustration, a normal subgroup and another subgroup will always give rise to a parallelogram if it is not known whether one of the two contains the other. By 3.8, the line segment leading from \( U \cap N \) to \( U \) stands for the whole subgroup lattice of \( U/U \cap N \). The analogous comment holds for \( NU/N \). (So to speak, we have replaced the ellipses in the sketch after the formulation of 3.8 by line segments.) Thus the parallelism of the two sides of the parallelogram which correspond to these two factor groups “signalizes” the isomorphism asserted in (1). Furthermore, it is useful to distinguish (interesting) normal subgroups (of the “mother group” \( G \)) from non-normal subgroups by different kinds of dots: Our dots for \( N \) (and \( M, G \) in (2)) are slightly thicker.

56
Then $\varphi$ is a homomorphism of $U$ onto $NU/N$, $U \cap N = \{u|u \in U, Nu = N\} = \ker \varphi \triangleleft U$, $U/U \cap N = U/\ker \varphi \cong U\varphi = NU/N$ by 3.2.

(2) Let $\varphi : G \to (G/N)/(M/N)$, $g \mapsto (M/N)Ng$. Then $\varphi$ is the composition of two canonical epimorphisms:

$$G \to G/N \to (G/N)/(M/N), \ g \mapsto Ng \mapsto (M/N)Ng,$$

hence is an epimorphism. We have

$$g \in \ker \varphi \iff Ng \in M/N \iff \exists x \in M \quad Ng = Nx \iff g \in NM \iff g \in M.$$

By 3.2, it follows that $G/M = G/\ker \varphi \cong G\varphi = (G/N)/(M/N)$. 

We now turn to structures with more than one operation:

3.10 Definition. Let $K$ be a commutative unitary ring, $(A, +, \cdot)$ a $K$-algebra. An ideal (more precisely, “$K$-ideal”) of $A$ is a $K$-subspace $J$ of $A$ such that $J \cdot A \subseteq J$, $A \cdot J \subseteq J$, for which we will use the notation $J \triangleleft A$.\footnote{Thus an ideal $J$ is both a left ideal and a right ideal of $A$.} As the group $(A, +)$ is abelian, there is no danger of confusion with the notation used for normal subgroups of groups: Every subgroup of $(A, +)$ is normal (3.1.3), so the use of the symbol $\triangleleft$, regarding an algebra, should reasonably express something else: it is the ideal property. We even omit the letter $K$ if the reference to $K$ is clear, thus removing even the only distinguishing detail between the normal subgroup notation and the ideal notation. Clearly, there is a good reason for this double usage of the same symbol. In short terms, it will soon become clear that, in the theory of algebras, ideals are the analogue of normal subgroups in group theory. A first (and rather vague) observation in the direction of this analogue is that ideals are special subalgebras, distinguished by a property which depends on the containing algebra $A$: Like the notion of a normal subgroup, that of an ideal is a relative one, meaningful only with respect to $A$. The defining property of an ideal is obviously much stronger than the ordinary (internal) multiplicative closure property. An ideal behaves like a “big zero” in an algebra: Multiplying any element by an element of an ideal, the result is “swallowed” by the ideal, like $0 \cdot x = 0 = x \cdot 0$ for any algebra element $x$. Clearly, $\{0\}$ and $A$ are ideals of $A$, the so-called trivial ideals. A unitary algebra which has only the trivial ideals is called simple. Every field is an example of a simple $Z$-algebra. More in general, we have:

3.10.1. If $(\hat{A}, \cdot)$ is a group, then $A$ is unitary and associative, and $\{0_A\}$, $A$ are the only one-sided ideals of $A$. In particular, $A$ is simple.

Proof. As $A = \hat{A} \cup \{0_A\}$, the existence of a neutral element and the associative law in $(\hat{A}, \cdot)$ imply that $A$ is unitary and associative. Let $R$ be a right ideal of $A$, $0_A \neq x \in R$. Then $1_A \in xA \subseteq R$, hence $R \supseteq xA = (xA)A = A$, $R = A$. The proof for left ideals is analogous.
A is called an (associative) division algebra if \((\mathbb{A}, \cdot)\) is a group. The starting point of the theory of algebras (around 1840) was the discovery of a 4-dimensional non-commutative associative division algebra over \(\mathbb{R}\), the so-called quaternion algebra. The algebra of all \(n \times n\) matrices over a field \(K\) is an example of a simple algebra which, for \(n > 1\), is not a division algebra. In this course, however, we will mainly be interested in commutative algebras; even more specifically, in fields.

In many respects, the familiar example of the integral domain \((A = K =) \mathbb{Z}\) is typical as an algebra. It is of course an important example which should be completely understood. Clearly, an ideal (of any algebra) must be an additive subgroup. The additive subgroups of \(\mathbb{Z}\) have been determined in 1.2.1. As a trivial consequence we note:

**3.10.2.** The ideals of the \(\mathbb{Z}\)-algebra \((\mathbb{Z}, +, \cdot)\) are exactly the subgroups of \((\mathbb{Z}, +)\). \(\square\)

We write \(J \triangleleft A\) if \(J\) is a maximal ideal of \(A\), i.e., if \(J\) is a proper ideal of \(A\) and there is no ideal \(I\) of \(\mathbb{A}\) such that \(J \subseteq I \subset A\). For a simple algebra \(A\) – in particular for a field \(-\), 3.10.1 implies that \(\{0_A\} \triangleleft A\). By 3.10.2 and the discussion prior to the proof of 3.8, \(n\mathbb{Z} \triangleleft \mathbb{Z}\) if and only if \(n\) is a prime.

Let \(A\) be an arbitrary \(K\)-algebra and \(J \subseteq A\). Then \((J+a) \cdot (J+b) \subseteq J \cdot J + J \cdot b + a \cdot J + ab \subseteq J + ab\) for all \(a, b \in A\), by the ideal property of \(J\). It follows that

\[
\forall U, V \in A/J \exists W \in A/J \quad U \cdot V \subseteq W.
\]

Recall that a very similar observation enabled us in steps (II), (III) of the construction of a quotient field (preceding 2.10) to define the crucial operations \((\sqcup, \sqcap)\). Here again we have a partition (the set \(A/J\) of all additive cosets of \(J\) in \(A\)) which is compatible with an operation \((\cdot)\), in the sense of (\(\ast\)). Clearly, \(W\) in (\(\ast\)) is uniquely determined by \(U, V\). Therefore we define \(UV\) to be the additive coset of \(J\) in \(A\) in which the set \(U \cdot V\) is contained.\(^{56}\) As in step (II) on p. 42, we cannot expect \(U \cdot V\) to be a complete coset.

---

\(^{56}\)The general notion of a quotient is based on that of a partition \(\mathcal{M}\) of the underlying set \(X\) of a magma \((X, \circ)\) which is compatible (with \(\circ\)) in the following sense:

\[
\forall U, V \in \mathcal{M} \exists W \in \mathcal{M} \quad U \circ V \subseteq W
\]

As then \(W\) is uniquely determined by the couple \((U, V)\) we may define \(UV\) to be that \(W \in \mathcal{M}\) which contains the product set \(U \circ V\). This defines an operation on \(\mathcal{M}\), thus giving rise to a magma with underlying set \(\mathcal{M}\). The mapping \(x \mapsto [x]\) (where \([x]\) denotes the unique element of \(\mathcal{M}\) containing \(x\)) is obviously an epimorphism of \(X\) onto our magma \(\mathcal{M}\). A magma which is given in the described way by a compatible partition of \(X\) is called a quotient of \(X\). Without difficulty we obtain the following

**General homomorphism theorem.** Let \(\varphi\) be a homomorphism of a magma \((X, \circ)\) into a magma \((Y, \cdot)\). \(\mathcal{M} := \{b \cdot \varphi^{-1}[b] \mid b \in X\varphi\}\). Then \(X\varphi\) is a submagma of \((Y, \cdot)\), \(\mathcal{M}\) is a quotient of \(X\), and \(\varphi^{-}\) is an isomorphism of \(X\varphi\) onto \(\mathcal{M}\).

Note, however, that only for a group \((X, \circ)\) we have the satisfactory description of the partition element \([x]\) as the coset of \(x\) with respect to the kernel of \(\varphi\) (see 3.2(1)), a regularity which we may not even dream of in the case of an arbitrary magma.
But passing to the uniquely determined coset that contains $U \cdot V$ (and calling it $UV$), we obtain a (multiplicative) operation on $A/J$. In order to calculate the product $UV$, it suffices to determine just one element of it (and form its coset with respect to $J$). For any $a \in U, b \in V$, certainly $ab \in U \cdot V \subseteq UV$, hence we have
\[
\forall a, b \in A \quad (J + a)(J + b) = J + ab.
\]
In other words, the canonical $K$-space epimorphism $\kappa : A \to A/J, a \mapsto J + a$, is also a multiplicative homomorphism, hence a $K$-algebra epimorphism. By the properties proved in 1.3, the $K$-double magma $A/J$, endowed with the usual coset addition and the multiplication just defined, turns out to be a $K$-algebra, as an epimorphic image of $A$. It is called the factor algebra of $A$ by $J$. Clearly, $\ker \kappa = J$. Now we are prepared to prove the analogue of 3.2 for algebras:

3.11 Theorem (Homomorphism theorem for algebras). Let $K$ be a commutative unitary ring, $A$ a $K$-algebra, $\varphi$ a $K$-algebra homomorphism of $A$ into a $K$-double magma $B$, $J := \ker \varphi$.

1. $J \triangleleft A$,

2. $\varphi^{-}$ is a $K$-algebra isomorphism of $A\varphi$ onto $A/J$.

Proof. We shall write $+$ for the first operation (in both $A$ and $B$) and express the second operation simply by juxtaposition. In particular, $\varphi$ is a $K$-linear mapping of the $K$-space $(A, +)$ into the $K$-magma $(B, +)$. The isomorphism $\varphi^{-}$ of $(A\varphi, +)$ onto $(A/J, +)$ (see 3.2(2)) is therefore $K$-linear, hence an isomorphism of $K$-magmas. It remains to show

ad (1) $\forall x \in J \forall a \in A \quad xa, ax \in J$ : Let $x \in J, a \in A$. Then $(xa)\varphi = (x\varphi)(a\varphi) = (0_Aa)\varphi = 0_A\varphi$, similarly $(ax)\varphi = 0_A\varphi$ and $(cx)\varphi = 0_A\varphi$ for all $c \in K$.

ad (2) $\varphi^{-}$ is a multiplicative homomorphism: Let $a, a' \in A$. Making use of 3.2(1), we obtain $((a\varphi)(a'\varphi))\varphi^{-} = (aa')\varphi\varphi^{-} = J + aa' = (J + a)(J + a') = (a\varphi)(a'\varphi)\varphi^{-}$. Hence $\varphi^{-}$ is an isomorphism of the $K$-double magma $A\varphi$ onto the $K$-algebra $A/J$. This implies, in particular, that the $K$-double magma $A\varphi$ is a $K$-algebra$^{57}$. □

We observed at the end of 3.10 that every ideal $J$ of a $K$-algebra $A$ is the kernel of the canonical $K$-algebra homomorphism of $A$ onto $A/J$. Combining this with 3.11(1), we obtain the following analogue of 3.2.1:

3.11.1. The ideals of a $K$-algebra $A$ are exactly the kernels of the $K$-algebra homomorphisms of $A$. □

The theorem includes the homomorphism theorem for $K$-spaces (by considering them as $K$-algebras in which the product of any two elements is defined to be zero) and for rings (being (associative) $\mathbb{Z}$-algebras). Now let $R$ be any commutative unitary ring.

$^{57}$Thus, an – albeit trivial – separate verification of this fact is not necessary.
The mapping \( \varphi : \mathbb{Z} \to R, k \mapsto k1_R \), is an additive homomorphism. In particular, \((-k) \varphi = -(k \varphi)\) for all \( k \in \mathbb{Z} \). It is also a multiplicative homomorphism:

\[
\forall k, l \in \mathbb{Z} \quad (kl) \varphi = (kl)1_R = (k1_R)(l1_R) = (k \varphi)(l \varphi),
\]

the crucial mid equation being a consequence of the distributive law in \( R \). As in Example (3) on p. 50, define \( n \in \mathbb{N}_0 \) by the equation \( \ker \varphi = n \mathbb{Z} \). By 3.11(2), we have the isomorphism \( \mathbb{Z}/n\mathbb{Z} \cong (1_R) \) not only regarding the additive groups (as noted in Example (3) on p. 50) but also multiplicatively, hence with respect to the ring structures of \( \mathbb{Z} \) and \( R \).

**3.12 Definition.** Let \( R \) be a commutative ring and \( \varphi \) as above. The number \( n \in \mathbb{N}_0 \) such that \( \ker \varphi = n \mathbb{Z} \) is called the **characteristic** of \( R \) and denoted by \( \text{char} R \). We have \( \mathbb{Z}/n\mathbb{Z} \cong \mathbb{Z} \varphi \) and distinguish the following two cases:

1st case: \( n = 0 \). Then \( \mathbb{Z} \varphi \cong \mathbb{Z} \).

2nd case: \( n > 0 \). Then \( \mathbb{Z} \varphi = \{0_R, 1_R, 1_R + 1_R, \ldots, 1_R + \cdots + 1_R\} \).

In the second case, \( R \) is called **of positive characteristic**. In both cases, we obtain an explicit description of the subring generated by \( 1_R \) which we call the **prime ring** of \( R \).

**3.12.1.** Let \( R \) be an integral domain. Then \( \text{char} R \) equals 0 or is a prime.

**Proof.** Let \( n := \text{char} R \) and assume \( n \neq 0 \). Then \( n > 1 \) as \( R \) is unitary. Assume that there are \( k, l \in n-1 \) such that \( kl = n \). Then \( k1_R, l1_R \neq 0_R \) and \( 0_R = n1_R = (k1_R)(l1_R) \), a contradiction as \( R \) has no zero divisors \( \neq 0 \). Hence \( n \) is a prime. \( \square \)

Given a \( K \)-algebra \( A \), we write \( \mathcal{T}(A) \) for the set of all subalgebras of \( A \). As in the case of subgroups of a group, these form a lattice with respect to set-theoretic inclusion. The same holds for the set of ideals of \( A \). Writing \( \preceq \) for the relation of being a \( K \)-subalgebra, we have \( \inf(U, V) = U \cap V \), \( \sup(U, V) = \langle U + V \rangle := \bigcap \{W|U + V \subseteq W \preceq A\} \) for all \( U, V \preceq A \). Clearly, \( \sup(U, V) = U + V \) if and only if \( U + V \) is a subalgebra of \( A \). In particular, this is the case if \( U \preceq A \) or \( V \preceq A \) as follows from the following trivial remark:

**3.12.2.** Let \( U, V \in \mathcal{T}(A) \) such that \( UV, VU \subseteq U + V \). Then \( U + V \in \mathcal{T}(A) \). \( \square \)

If both \( U \) and \( V \) are ideals of \( A \), so is \( U + V \): For all \( a \in A \) we then have \( a(U + V) = aU + aV \subseteq U + V \), likewise \( (U + V)a \subseteq U + V \). Applying 3.11 like 3.2 in the proofs of 3.8, 3.9, we readily obtain the following important analogues for algebras:

---

58 This is a special case of Example (3) on p.50 with additive notation, \( x := 1_R \).

59 Applying the rule for additive inverses mentioned beforehand, we may assume w.l.o.g. that \( k \in \mathbb{N}_0 \) and proceed by a routine induction on \( k \) in this case.
3.13 Proposition. Let $K$ be a commutative unitary ring, $A$ a $K$-algebra, $J \trianglelefteq A$.

(1) (Correspondence theorem for algebras). Let $\varphi$ be an epimorphism of $A$ onto a $K$-algebra $\tilde{A}$. Then $\varphi^{-1}|_{T(\tilde{A})}$ is an isomorphism of $(T(\tilde{A}), \preceq)$ onto $(\ker \varphi, A, \preceq)$. The same assertion holds regarding the lattices of the ideals of $\tilde{A}$ and those of $A$ containing ker $\varphi$.

(2) Let $U \preceq K A$. Then $U \cap J \preceq U$, $J + U \preceq K A$, and $U/U \cap J \sim (J + U)/J$.

(3) Let $J \subseteq I \preceq K G$, $I/J \trianglelefteq A/J$. Then $I \trianglelefteq A$ and $(A/J)/(I/J) \cong A/I$.

The results 3.9, 3.13(2),(3) are commonly called the isomorphism theorems for groups, algebras resp. In the sequel, we will choose $K = \mathbb{Z}$ and consider a commutative ring $R$ in place of $A$. By 3.10.1, every field is simple. We now observe the converse:

3.13.1. Let $R$ be a commutative simple ring. The $R$ is a field.

Proof. Let $x \in \hat{R}$. Then $0_R \neq x \in xR \preceq R$, hence $xR = R$. In particular, $xy = 1_R$ for some $y \in R$. Clearly, $y \neq 0_R$. It follows that $(\hat{R}, \cdot)$ is an abelian group. \qed

3.13.2. Let $R$ be a commutative unitary ring, $J \trianglelefteq max R$. Then $R/J$ is a field.

Proof. By 3.13(1) (in the version for ideals), it follows that the commutative unitary factor ring $R/J$ is simple. The claim follows from 3.13.1. \qed

Conversely, if $R/J$ is a field for some ideal $J$ of $R$, then $J$ must be maximal, by 3.13(1). Making use of 3.8.1 and 3.10.2, we conclude

3.13.3. Let $n \in \mathbb{N}$. The factor ring $\mathbb{Z}/n\mathbb{Z}$ is a field if and only if $n$ is a prime. \qed

This remark and the description in the 2nd case of 3.12 imply

3.13.4. Let $K$ be a field such that $p := \text{char } K > 0$. Then the prime ring of $K$ is a subfield of $K$ and isomorphic to $\mathbb{Z}/p\mathbb{Z}$. \qed

We could have considered an integral domain $R$ instead of a field here but would not have gained more generality: The version for $R$ follows from 3.13.4 by choosing $K$ as a quotient field of $R$.

If $K$ is a field of characteristic 0, the prime ring of $K$ is isomorphic to $\mathbb{Z}$ (see the 1st case in 3.12). Hence its quotient field in $K$ (see 2.10.1) is isomorphic to $\mathbb{Q}$. It is obviously the smallest subfield of $K$ containing $1_K$. In any field $K$, the subfield generated by $1_K$ is called the prime field of $K$. We summarize:

3.13.5. Let $K$ be a field, $K_0$ its prime field. If char $K = 0$, then $K_0 \cong \mathbb{Q}$. If char $K = p > 0$, then $K_0 \cong \mathbb{Z}/p\mathbb{Z}$, in particular, $|K_0| = p$. \qed
The theory of fields therefore has a natural subdivision into two chapters of rather
different peculiarities: (1) Fields of characteristic 0, (2) fields of prime characteristic.
An ordered field is necessarily of characteristic 0 because a sum of 1’s can never attain
the value 0 in an ordered field. Clearly, a finite field must be of prime characteristic.
The theory of finite fields may be viewed as a natural subchapter in field theory, and
we will derive its fundamentals as applications of the results of the next chapter. There
are quite a number of theorems on fields which hold for every characteristic but hitherto
require different proofs for case (1) and case (2). In the theory of representations over
a field $K$ it makes a fundamental difference if $\text{char } K = 0$ (the so-called classical case)
or $\text{char } K$ is a prime (the so-called modular case), the latter generally imposing much
more intricate obstacles than the rather well understood first case.

3.14 Definition. Let $R$ be a commutative unitary ring. A subset $J$ of $R$ is called a
principal ideal of $R$ if there exists an element $a \in R$ such that $J = aR$. Any such
element $a$ is called an (ideal) generator of $J$.

3.14.1. $\forall a, b \in R \ (aR \subseteq bR \iff \exists x \in R \ bx = a)$,
as $aR \subseteq bR$ if and only if $a \in bR$ because $R$ is associative and unitary. □
Recall 2.2 where we introduced the notation $b | a$ for the condition on the right in 3.14.1.

3.14.2. Let $K$ be a field, $f, g \in K[t] \setminus K$ such that $g \mid f$ and $g$ of minimal degree with
this property. Then $g$ is irreducible in $K[t]$.

Proof. Let $r, s \in K[t]$ such that $rs = g$ and assume that $r \not\in K$. Then $r \mid f$ and
deg $r \leq$ deg $g$. It follows that deg $r =$ deg $g$, hence $s \in \bar{K}$ by 1.1.2. □

$R$ is called a principal ideal domain if the following holds:

(i) $R$ is an integral domain,

(ii) Every ideal of $R$ is a principal ideal.

By 3.10.2 and 1.2.1, $\mathbb{Z}$ is a principal ideal domain. The polynomial ring $\mathbb{Z}[t]$ is not a
principal ideal domain: The ideal generated by 2 and $t$ is not a principal ideal. More
generally, it is an easy exercise to prove that $R[t]$ is a principal ideal domain only if $R$
is a field. Conversely, this necessary condition is sufficient: \footnote{\text{Clearly, any principal ideal of $R$ is an ideal of $R$, thus deserves its name.}}

\footnote{\text{Polynomial rings in one variable over a field belong to the larger class of integral domains called Euclidean domains. Essentially, these are integral domains for which there exists a function into $\mathbb{N}_0$ with properties that are similar to those of the degree function for $K[t]$, in the sense that they allow the line of reasoning as in the proof of 3.14.3. It is then not really surprising that, by the same proof in general form, every Euclidean domain turns out to be a principal ideal domain. The general notion of a Euclidean domain can play a role in directions which are not pursued in this course.}}
3.14.3. Let $K$ be field, $S := K[t]$. Then $S$ is a principal ideal domain. If $\{0_K\} \neq J \leq S$, $g \in J$ of minimal degree, then $J = gS$.

*Proof.* Condition (i) holds by 1.1.2. (ii) Let $J \leq S$. If $J = \{0_S\}$, then $J = 0_S S$, hence principal. Now let $J \neq \{0_S\}$, $g \in J$ of minimal degree. Clearly, $gS \subseteq J$ as $g \in J \leq S$. We claim that $gS \supseteq J$. We have $gS = cgS$ for all $c \in K$. Therefore we may choose $c$ with the property that $cg$ is normed, then consider $cg$ instead of $g$. In other words, we may assume w.l.o.g, that $g$ is normed. Let $f \in S$. By 1.6(1), there exist $h, r \in J$ such that $f = gh + r$, $r = 0_K$ or $\deg r < \deg g$. Then $r = f - gh \in J$ so that, by the minimality of the degree of $g$, we conclude $r = 0_K$. It follows that $f = gh \in gS$. \[\square\]

3.15 Proposition. Let $K$ be a field, $0_K \neq f \in K[t]$, $J := fK[t]$. Then

(1) $\dim_K K[t]/J = \deg f$,

(2) $\dim_K K[t]/J$ is a field if and only if $f$ is irreducible.

*Proof.* (1) Let $n := \deg f$. By 1.6(1), for every $f \in K[t]$ there exists a polynomial $r \in K[t]$ of degree $< n$ such that $J + f = J + r$. Hence $\{J + 1_K, J + t, \ldots, J + t^{n-1}\}$ is a set of generators of the $K$-space $K[t]/J$. As $J$ does not contain any (nonzero) polynomial of degree $< n$, it is a $K$-linearly independent set of $n$ elements. The claim follows.

(2) We apply 3.13.2 and its converse to conclude that $K[t]/J$ is a field if and only if $J \triangleleft \max K[t]$. From 3.14.1 and 3.14.3 it follows that the maximality of the ideal $J$ is equivalent to the property that the ideal generator $f$ is irreducible. \[\square\]

Let $R$ be a principal ideal domain, $M := \hat{R}$. Then $(M, \cdot)$ is a commutative monoid. Let $q \in M$ be indecomposable. From 3.14.1 it follows that $qR \triangleleft \max R$. Suppose that $R=M \uparrow a, b$ where $a, b \in M$. Then the ideals $aR + qR, bR + qR$ properly contain $qR$, hence equal $R$. Thus both of them contain $1_R$. It follows that $1_R \in (aR + qR)(bR + qR) \subseteq abR + qR$. Hence $ab \notin qR$, i.e., $q \uparrow ab$. Thus we have shown, by contraposition, that the indecomposable elements $q$ of the multiplicative monoid $M$ of nonzero elements of any principal ideal domain have the property $(\ast)$ on p.29. From this it follows easily that the same holds with respect to the submonoid $\tilde{N}$ of $(\hat{Z}, \cdot)$ and the submonoid $N$ of $(\hat{V}, \cdot)$ in the examples (1), (2) on p.28.
4 Field extensions

4.1 Definition. A field extension is a pair \((K, M)\) of fields \(K, M\) such that \(K\) is a subfield of \(M\). Then \(M\) is called an extension field of \(K\). A basic idea of everything that follows is to view \(M\) as a \(K\)-vector space (cf. Example (2) on p. 33). A field extension \((K, M)\) is called finite if \(\dim_K M\) is finite. In particular, this is the case if \(M\) is finite. The interpretation of \(M\) as a vector space over any subfield \(K\) has the following striking consequence:

4.1.1. If \(M\) is finite, there exists an integer \(n\) such that \(|M| = |K|^n\). In particular, \(|M|\) is a power of the prime char \(M\).

Proof. As \(M\) is a finite dimensional vector space over \(K\) we have the \(K\)-space isomorphism \(M \cong K^n\) for some \(n \in \mathbb{N}\). Hence \(|M| = |K|^n\). The last assertion is the special case where \(K\) is the prime field of \(M\) (see 3.13.5). \(\Box\)

Let \(b\) be an element of an extension field \(M\) of the field \(K\), \(F_b\) the replacement homomorphism \(K[t] \to M, f \mapsto f(b)\). The element \(b\) is called algebraic over \(K\) if \(\ker F_b \neq \{0_K\}\), i.e., if there exists a nonzero polynomial \(f \in K[t]\) such that \(f(b) = 0_K\). For example, \(\sqrt{5}\) is algebraic over \(\mathbb{Q}\) as it is a zero of \(t^2 - 5\). Every root of unity is algebraic over \(\mathbb{Q}\), being a zero of \(t^n - 1\) for some \(n \in \mathbb{N}\). Trivially, every element \(b \in K\) is algebraic over \(K\) as it is a zero of \(t - b\). If \(\ker F_b = \{0_K\}\), \(b\) is called transcendental over \(K\). Equivalently, this means that the subring \(K[b]\) of \(M\) generated by \(K\) and \(b\) is a polynomial ring over \(K\) in the variable \(b\) (see 1.1 or, more generally, 2.9). The numbers \(e\) and \(\pi\) are famous examples of real numbers which are transcendent over \(\mathbb{Q}\). We set

\[ A_K(M) := \{b \in M, b \text{ is algebraic over } K\}. \]

The extension \((K, M)\) is called algebraic if \(A_K(M) = M\). For example, \(\mathbb{Q}[\sqrt{2}]\) is a field and \((\mathbb{Q}, \mathbb{Q}[\sqrt{2}])\) is algebraic as will be recognized as a very small special case of the general remark 4.1.5. \((\mathbb{Q}, \mathbb{R})\) is not algebraic: We could refer to the transcendency of \(e\) or \(\pi\) or, simpler, to the argument mentioned in footnote 63.

4.1.2. If \((K, M)\) is finite, then \((K, M)\) is algebraic.

---

62Resuming the notion introduced on p. 39, this means that the 1-tuple \((b)\) is algebraically dependent.

63The first proofs for the transcendency of \(e\), \(\pi\) resp., were discovered by Ch. Hermite (1873, for \(e\)), C. L. F. v. Lindemann (1882, for \(\pi\)) resp. The set of all transcendental real numbers is non-denumerable because \(\mathbb{R}\) is non-denumerable while the set of all algebraic real numbers is denumerable, as is not very difficult to prove. By contrast and hitherto without exception, a proof of transcendency for a specific real number requires non-trivial analytical methods. Therefore questions of this kind are usually treated within the area of Analytic number theory.
Proof. Let \( n := \dim_K M, b \in M \). The \((n+1)\)-tuple \((1_K, b, \ldots, b^n)\) is \( K \)-linearly dependent, i.e., there exist \( c_0, \ldots, c_n \in K \), not all of them zero, such that \( c_0 + c_1 b + \cdots + c_n b^n = 0_K \). Thus \( b \) is a zero of the nonzero polynomial \( \sum_{j=0}^n c_j t^j \in K[t] \).

For every \( b \in A_K(M) \) there exists a unique polynomial with zero \( b \) which is of minimal degree and normed. This polynomial is called the minimal polynomial of \( b \) over \( K \) and is denoted by \( \min_{b,K} \). From 3.14.3 we obtain

4.1.3. Let \( b \in A_K(M) \). Then \( \ker F_b = \min_{b,K} K[t] \).

4.1.4. Let \( b \in A_K(M) \), \( f \in K[t] \). Then \( f = \min_{b,K} \) if and only if \( f \) is irreducible and normed and \( f(b) = 0_K \).

Proof. Suppose \( f = \min_{b,K} \) and let \( g, h \in K[t] \) such that \( f = gh \). Then \( f \) is normed and \( g(b)h(b) = 0_K \). As \( K \) has no zero divisors \( \neq 0_K \), it follows that \( g(b) = 0_K \) or \( h(b) = 0_K \). Hence, by 4.1.3, \( f \mid g \) or \( f \mid h \). As \( f = gh \), this implies \( h \in K \) or \( g \in K \).

Thus \( f \) is irreducible. Conversely, let \( f \) be irreducible, normed, and \( f(b) = 0_K \). Then \( f \in \ker F_b \), hence \( \min_{b,K} \mid f \) by 4.1.3. As \( f \) is irreducible in \( K[t] \) and \( \min_{b,K} \notin K \), it follows that \( f = c \min_{b,K} \) for some \( c \in K \). But \( f \) and \( \min_{b,K} \) are normed, thus \( c = 1_K \), \( f = \min_{b,K} \).

Our next remark will have most important applications in the sequel:

4.1.5. Let \( b \in A_K(M) \), \( J := \min_{b,K} K[t] \). Then there exists an isomorphism \( \psi \) of \( K[b] \) onto \( K[t]/J \) with the properties \( b \psi = J + t, c \psi = J + c \) for all \( c \in K \). In particular, \( K[b] \) is a field, and \( \dim_K K[b] = \deg \min_{b,K} \).

Proof. We have \( K[b] = K[t]F_b \) and, by 4.1.3, \( J = \ker F_b \). Thus the isomorphism assertion now follows from 3.11(2). The final assertions follow both from 3.15 as \( \min_{b,K} \) generates \( J \) by 4.1.3 and is irreducible by 4.1.4.

For example, \( \mathbb{Q} \sqrt{2} \) is a field and of dimension 2 over \( \mathbb{Q} \) as \( \min_{\mathbb{Q} \sqrt{2}, \mathbb{Q}} = t^2 - 2 \). If \( b \in M \) is transcendental over \( K \), \( K[b] \) is a polynomial ring over \( K \), not a field, hence properly contained in its quotient field \( K(b) \) in \( M \). If \( b \in A_K(M) \), then \( A[b] = A(b) \) by 4.1.5. Hence, for an arbitrary \( b \in M \), we have

4.1.6. \( b \in M \) is algebraic over \( K \) if and only if \( K[b] = K(b) \).

4.1.7. \( A_K(M) \) is a subfield of \( M \).

Proof. Clearly, \( 0_K, 1_K \in A_K(M) \). Let \( b, b' \in A_K(M) \). We have \( K[b, b'] = (K[b])[b'] \),

\*\*The reader should be warned that an approach to the assertion by considering minimal polynomials in \( K[t] \) of elements \( a, b \in M \) is no good idea. It is generally hard to derive information about \( \min_{a+b,K} \)

or \( \min_{a+K}, \min_{b,K} \). We mention the following result:

**Theorem.** (Isaacs 1970) Let char \( K = 0 \) and \( \dim_K K(a+b) = \dim_K K[a] \dim_K K[b] \). Then \( K[a, b] = K[a+b] \).

The hypothesis is satisfied if \( \dim_K K[a] \) and \( \dim_K K[b] \) are relatively prime so that in this case we have \( \deg \min_{a+b,K} = \deg \min_{a,K} \deg \min_{b,K} \), by 4.1.5. The proof in [Is] should give a sufficient impression about the intricacies of the topic.

66
and the trivial inclusion \( A_K(M) \subseteq A_{K[b]}(M) \) implies that the extension \((K[b], K[b, b'])\) is finite. By the special case of 2.6, \( \dim_K K[b, b'] = \dim_K K[b] \dim_K K[b, b'] \) is finite, hence \((K, K[b, b'])\) is algebraic by 4.1.2. Thus \( b' - b \) and (if \( b' \neq 0_K \)) \( bb'^{-1} \) are elements of an algebraic extension field of \( K \) and hence algebraic over \( K \). This shows that \( A_K(M) \) is additively, \( A_K(M) \setminus \{0_K\} \) multiplicatively a group.

Generalizing inductively the argument in this proof, we observe that for a finite number of algebraic elements \( a_0, a_1, \ldots, a_n \) of an extension field of \( K \), the extension field \( K[a_0, \ldots, a_n] = K[a_0][a_1] \cdots [a_n] \) is of dimension \( \prod_{j \in \mathbb{N}_0} \dim_K K[a_j] \) over \( K \). In particular, its dimension is finite, hence it is an algebraic extension of \( K \) by 4.1.2. Now the following is a simple consequence:

4.1.8. Let \((K, L), (L, M)\) be algebraic field extensions. Then \((K, M)\) is algebraic.

**Proof.** Let \( b \in M \). As \((L, M)\) is algebraic, there exists an \( m \in \mathbb{N}_0 \) and elements \( a_0, \ldots, a_m \in L \), \( a_m \neq 0_K \), such that \( \sum_{j=0}^m a_j b^j = 0_K \). It follows that \( b \) is algebraic over the field \( K[a_0, \ldots, a_m] \). By the above argument, \( K[a_0, \ldots, a_m, b] \) is an algebraic extension of \( K \). In particular, \( b \) is algebraic over \( K \). □

We are going to pursue now questions about the zeros of polynomials as presented in our first chapter. To this end we introduce some useful notation: For any \( f \in K[t] \), the set of all zeros of \( f \) in \( M \) will be denoted by \( Z_M(f) \). Given an isomorphism \( \varphi \) of \( K \) onto a field \( \bar{K} \), we write \( \bar{\varphi} \) for the extension of \( \varphi \) to the isomorphism of \( K[t] \) onto the polynomial ring \( \bar{K}[\bar{t}] \) over \( \bar{K} \) with the properties \( t \bar{\varphi} = \bar{t} \). Explicitly,

\[
\bar{\varphi} : K[t] \to \bar{K}[\bar{t}], \quad \sum_j c_j t^j \mapsto \sum_j (c_j \varphi) \bar{t}^j
\]

4.2 Lemma. Let \( K \) be a field, \( g \in K[t] \) irreducible.

(1) There exists an extension field \( L \) of \( K \) such that \( Z_L(g) \neq \emptyset \).

(2) Let \( L \) be as in (1), \( b \in Z_L(g) \), \( \varphi \) a monomorphism of \( K \) into a field \( X \), \( \bar{K} := K \varphi \), \( \bar{g} := g \bar{\varphi} \). Then for each \( b \in Z_X(\bar{g}) \) there exists a unique monomorphism \( \varphi_b \) of \( K[b] \) into \( X \) such that \( \varphi_b|_K = \varphi \), \( b \varphi_b = \bar{b} \).

Before we embark on the proof we insert a simple remark, under the same hypotheses as in (2):

4.2.1. Let \( \psi \) be any extension of \( \varphi \) to a monomorphism of \( K[b] \) into \( X \). Then \( b \psi \in Z_X(\bar{g}) \), because if \( g = \sum_{j=0}^n c_j t^j \ (c_j \in K) \) then

\[
0_X = 0_K \psi = g(b) \psi = \left( \sum_{j=0}^n c_j b^j \right) \psi = \sum_{j=0}^n (c_j \varphi(b^j)) = \bar{g}(b \psi)
\]

Therefore we have the following consequence of (2): 67
4.2.2. The assignment \( \tilde{b} \mapsto \varphi_\tilde{b} \) is a bijection of \( Z_X(\tilde{g}) \) onto the set of extensions of \( \varphi \) to a monomorphism of \( K[b] \) into \( X \). Hence their number equals \( |Z_X(\tilde{g})| \). \( \square \)

Proof of 4.2. We assume w.l.o.g. that \( g \) is normed and put \( J := gK[t] \). By 3.15(2), \( K[t]/J \) is a field.

(1) (Kronecker’s construction) Let \( \sigma : K \to K[t]/J, c \mapsto J + c \). This is a field monomorphism, being the restriction to \( K \) of the canonical epimorphism of \( K[t] \) onto \( K[t]/J \).

\[
\begin{array}{c}
L \\
\sigma \\
K[t]/J \\
\sigma \\
K \\
\end{array}
\]

Clearly, \( J \cap K = \{0_K\} \). By 2.1, there exists an extension field \( L \) of \( K \) and an extension \( \hat{\sigma} \) of \( \sigma \) to an isomorphism of \( L \) onto \( K[t]/J \). Let \( b := (J + t)\hat{\sigma}^{-1} \), \( c_0, \ldots, c_n \in K \) such that \( g = \sum_{j=0}^n c_jb^j \). Then

\[
(g(b))\hat{\sigma} = (\sum_{j=0}^n c_jb^j)\hat{\sigma} = \sum_{j=0}^n (J + c_j)(J + t)^j = J + \sum_{j=0}^n c_jt^j = J + g = J,
\]

hence \( g(b) = 0_K \).

(2) Let \( \tilde{b} \in Z_X(\tilde{g}) \). Since \( K[b] \) is generated by \( K \) and \( b \) as a ring, there is at most one field monomorphism of \( K[b] \) into \( X \) which maps \( b \) to \( \tilde{b} \) and \( c \) to \( c\varphi \) for all \( c \in K \). It remains to prove that there is at least one such monomorphism.

As \( g \) is irreducible in \( K[t] \), \( \tilde{g} \) is irreducible in \( \tilde{K}[\tilde{t}] \). We have \( g = \min_{g,K}, \tilde{g} = \min_{\tilde{b},\tilde{K}} \), by 4.1.4. Let \( \tilde{J} := J\tilde{\varphi} = \tilde{g}\tilde{K}[\tilde{t}] \). Then \( \tilde{\varphi} \) induces an isomorphism \( \tilde{\varphi} \) of \( K[t]/J \) onto \( \tilde{K}[\tilde{t}]/\tilde{J} \) such that \( (J + t)\tilde{\varphi} = \tilde{J} + t, (J + c)\tilde{\varphi} = \tilde{J} + c\varphi \) for all \( c \in K \).

Let \( \psi \) be the isomorphism of \( K[b] \) onto \( K[t]/J \), in the same way \( \tilde{\psi} \) the isomorphism of \( \tilde{K}[\tilde{b}] \) onto \( \tilde{K}[\tilde{t}]/\tilde{J} \) which is given by 4.1.5. Put \( \varphi_\tilde{b} := \psi\tilde{\varphi}\tilde{\psi}^{-1} \):

\[
K[b] \to K[t]/J \to \tilde{K}[\tilde{b}] \to \tilde{K}[\tilde{b}]
\]

Then \( \varphi_\tilde{b} \) is a composition of isomorphisms, hence an isomorphism of \( K[b] \) onto \( \tilde{K}[\tilde{b}] \) and maps \( b \) to \( \tilde{b} \), \( c \) to \( c\varphi \) for all \( c \in K \). \( \square \)

4.3 Theorem (Splitting field theorem). Let \( K \) be a field, \( f \in K[t], f \neq 0_K \).

(1) There exists a splitting field of \( f \) over \( K \).

(2) Let \( M \) be a splitting field of \( f \) over \( K \), \( \varphi \) a monomorphism of \( K \) into a field \( X \) which is a splitting field of \( f\varphi \) over \( K\varphi \). Then there exists an extension \( \psi \) of \( \varphi \) to an isomorphism of \( M \) onto \( X \). In particular, for any two splitting fields \( M, X \) of \( f \) over \( K \) there exists an isomorphism \( \psi \) of \( M \) onto \( X \) such that \( c\psi = c \) for all \( c \in K \).

Proof by induction on deg \( f \).\(^{65}\) If deg \( f = 0 \), then \( f \in \tilde{K} \) and both (1), (2) are trivial. For the inductive step, let deg \( f > 0 \) and assume (1), (2) to be true for nonzero polynomials

\(^{65}\)This is the common short way of announcing a proof by induction on \( n \) for the assertion: “For all \( n \in \mathbb{N}_0 \) the following holds: If \( K \) is a field, \( f \in K[t], f \neq 0_K, \) deg \( f = n \), then the assertions (1), (2) follow.” Note that the assertion is not about one given field \( K \) but the universal quantifier regarding \( K \) is part of the claimed proposition. The proof starts with a field \( K \), but the inductive hypothesis will be applied to a different field!
of smaller degree over arbitrary fields. By 3.14.2, there exists an irreducible polynomial divisor \( g \) of \( f \), and we may assume \( g \) to be normed.

(1) By 4.2(1), there exists an extension field \( L \) of \( K \) such that \( Z_L(g) \neq \emptyset \). Let \( b \in Z_L(g) \).

Then \( f(b) = 0 \) over \( K \) as \( g \mid f \). By 1.6(2), there exists a polynomial \( h \in (K[b])[t] \) such that \( f = (t - b)h \). By 1.1.2, \( \deg h = \deg f - 1 \).

By our inductive hypothesis (for (1)), there exists a splitting field \( M \) of \( h \) over \( K[b] \). Since \( b \in Z_M(f) \), \( M \) contains a splitting field of \( f \) over \( K \). But the latter contains \( K[b] \) and all zeros of \( h \) in \( M \), hence equals \( M \).

(2) Put \( \tilde{f} := f \tilde{\varphi} \), \( \tilde{K} := K \varphi \), \( \tilde{g} := g \tilde{\varphi} \). Then \( Z_M(g) \subseteq Z_M(f) \), \( Z_X(\tilde{g}) \subseteq Z_X(\tilde{f}) \). Let \( \tilde{b} \in Z_M(g) \), \( \tilde{b} \in Z_X(\tilde{g}) \). Let \( \varphi_{\tilde{b}} \) as given by 4.2(2). By 1.6(2), there exist \( h \in (K[b])[t] \), \( \tilde{h} \in (\tilde{K}[^{\tilde{b}}])[\tilde{t}] \) such that \( f = (t - b)h \), \( \tilde{f} = (\tilde{t} - \tilde{b})\tilde{h} \). Then \( M \) is a splitting field of \( \tilde{h} \) over \( K[b] \), \( X \) is a splitting field of \( \tilde{h} \) over \( \tilde{K}[^{\tilde{b}}] \). Inductively there exists an isomorphism \( \psi \) of \( M \) onto \( X \) such that \( \psi|_{K[b]} = \varphi_{\tilde{b}} \). A fortiori, \( \psi|_{K} = \varphi \).

The final assertion is the special case where \( \varphi = \text{id}_K \). \( \square \)

Let \( X \) be a field. If \( (K, L) \) is a field extension, \( \varphi \) a monomorphism of \( K \) into \( X \), we write \( \varphi^L \) for the set of all monomorphisms of \( L \) into \( X \) which are extensions of \( \varphi \).

4.3.1. Let \( (K, L) \), \( (L, M) \) be field extensions, \( \varphi \) a monomorphism of \( K \) into \( X \). Then

\[
\varphi^M = \bigcup_{\psi \in \varphi^L} \psi^M,
\]

because the non-empty sets in this union are exactly the equivalence classes for the equivalence \( \sim \) on \( \varphi^M \) defined by: \( \omega \sim \omega' \) if and only if \( \omega|_L = \omega'|_L \) \( (\omega, \omega' \in \varphi^M) \).

4.4 Proposition. Let \( X \) be a field, \( (K, M) \) a finite field extension, \( \varphi \) a monomorphism of \( K \) into \( X \), \( \tilde{K} := K \varphi \). Let \( b \in M \), \( g := \min_{b, \tilde{K}} \), \( \tilde{g} := g \tilde{\varphi} \). Then

\[
|\varphi^M| = \sum_{\tilde{b} \in Z_X(\tilde{g})} |\varphi_b^M| \leq |Z_X(\tilde{g})| \dim_{K[b]} M \leq \deg g \dim_{K[b]} M = \dim_K M.
\]

Proof. By means of 4.3.1 and 4.2.2, we have \( |\varphi^M| = \sum_{\psi \in \varphi^{K[b]}} |\psi^M| = \sum_{\tilde{b} \in Z_X(\tilde{g})} |\varphi_{\tilde{b}}^M| \). This proves the first equality. The last equality follows from 4.1.5 and the special case of 2.6 which imply \( \deg g \dim_{K[b]} M = \dim_K K[b] \dim_{K[b]} M = \dim_K M \). We have \( \deg g = \deg \tilde{g} \), hence \( |Z_X(\tilde{g})| \leq \deg g \) by 1.6(3). It remains to prove the first inequality in our claim. This will be a consequence of the following intermediate result:

\[(*) \quad |\varphi^M| \leq \dim_K M \]

which we prove by induction on \( \dim_K M \). If \( \dim_K M = 1 \), then \( K = M \), hence \( \varphi^M = \{\varphi\} \) and the claim is trivial. For the inductive step, let \( \dim_K M > 1 \) and assume that \((*)\) holds

\[66\text{This proof deals with absolute nothingness but is nevertheless correct if } \varphi^M = \emptyset.\]
for all field extensions of smaller dimension. Let \( b \in M \setminus K \). Then \( \dim_{K[b]} M < \dim_K M \).

Making use of the parts we have already proved, we obtain from our inductive hypothesis

\[
|\varphi^M| = \sum_{\tilde{b} \in \tilde{Z}_X(\tilde{g})} |\varphi^M_{\tilde{b}}| \leq \sum_{\tilde{b} \in \tilde{Z}_X(\tilde{g})} \dim_{K[b]} M = |\tilde{Z}_X(\tilde{g})| \dim_{K[b]} M \leq \dim_K M.
\]

This proves (*) which we will now apply to the monomorphisms \( \varphi_{\tilde{b}} \) (in place of \( \varphi \)) of \( K[b] \) (in place of \( K \)) into \( X \). We obtain

\[
\sum_{\tilde{b} \in \tilde{Z}_X(\tilde{g})} |\varphi^M_{\tilde{b}}| = |\tilde{Z}_X(\tilde{g})| \dim_{K[b]} M \leq \dim_K M.
\]

\[\blacksquare\]

4.4.1. Let \( (K, M) \) be a finite field extension, \( X = M, \varphi \in \text{Aut } K \). Then \( \varphi^M \subseteq \text{Aut } M \), because for all \( \psi \in \varphi^M \) we have \( \dim_K M = \dim_K \varphi M \psi = \dim_K M \psi \) and \( M \cong M \psi \leq M \) which implies, by the finiteness of \( (K, M) \), that \( M = M \psi \). \[\blacksquare\]

In the following we consider the special case where \( X = M, \varphi = \text{id}_K \):

4.5 Definition. Let \((K, M)\) be a field extension. Then

\[
\text{Aut}_{K} M := \{ \alpha | \alpha \in \text{Aut } M, \forall c \in K \quad c\alpha = c \}
\]

is a subgroup\(^{67}\) of \( \text{Aut } M \) and is called the Galois group of \((K, M)\).

4.5.1. \( \forall \beta \in \text{Aut } M \quad \text{Aut}_{K\beta} M = \beta^{-1}(\text{Aut}_{K} M)\beta \).

Proof. Let \( \alpha, \beta \in \text{Aut } M \). Then

\[
\alpha \in \text{Aut}_{K\beta} M \iff \forall c \in K \quad (c\beta)\alpha = c\beta \iff \forall c \in K \quad c\beta\alpha^{-1} = c \\
\iff \beta\alpha^{-1} \in \text{Aut}_{K} M \iff \alpha \in \beta^{-1}(\text{Aut}_{K} M)\beta.
\]

\(\blacksquare\)

In an arbitrary group \( G \), put \( u^g := g^{-1}ug \) for all \( u, g \in G \), called the conjugate of \( u \) by \( g \).\(^{68}\) Set \( U^g := g^{-1}Ug \) for all \( U \subseteq G, g \in G \). A subgroup \( U \) is normal if and only if \( \forall c \in K \).

\(^{67}\)It contains \( \text{id}_M \), is closed with respect to composition, and \( c\alpha = c \) if and only if \( c = c\alpha^{-1} \), for all \( c \in K \).

\(^{68}\)For every \( g \in G \), the mapping \( G \to G, u \mapsto u^g \), is an automorphism of \( G \). The mapping

\[
G \to \text{Aut } G, \quad g \mapsto \begin{bmatrix} G & \to & G \\ u & \mapsto & u^g \end{bmatrix}
\]

is a homomorphism. Its kernel \( \{ z | z \in G, \forall g \in G \quad gz = zg \} \) is called the centre of \( G \) and denoted by \( Z(G) \). Its image is a normal subgroup of \( \text{Aut } G \), called the group of inner automorphisms of \( G \) and denoted by \( \text{In } G \). Thus, by 3.2,

\[
G/Z(G) \cong \text{In } G \leq \text{Aut } G.
\]

These assertions are recommended as routine exercises. They belong to the basics in group theory.
$U^g = U$ for all $g \in G$ as a glance at the definition (3.1) shows. We may express 4.5.1 in the following form:

**4.5.1'** \[ \forall \beta \in \text{Aut } M \quad \text{Aut}_{L\beta}M = (\text{Aut}_L M)^\beta. \]

**4.5.2.** Let $M$ be a splitting field of a nonzero polynomial $f \in K[t]$. Then $\text{Aut}_K M$ is isomorphic to a subgroup of $S_{Z_M(f)}$. In particular, $|\text{Aut}_K M| = k!$ where $k := |Z_M(f)|$.

**Proof.** For all $\alpha \in \text{Aut}_K M$ the restriction $\alpha|_{Z_M(f)}$ is a permutation of $Z_M(f)$ as follows from 4.2.1 in the special case $X = M$, $\phi = \text{id}_K$. The restriction mapping $\rho : \text{Aut}_K M \to S_{Z_M(f)}$, $\alpha \mapsto \alpha|_{Z_M(f)}$, is a homomorphism. Let $\alpha, \beta \in \text{Aut}_K M$ such that $\alpha \rho = \beta \rho$. Then $\alpha = \beta$ because $M$ is generated as a field by $K \cup Z_M(f)$. Hence $\rho$ is a monomorphism. \[ \square \]

As an immediate consequence of 4.4 (for $X = M$, $\phi = \text{id}_K$ (hence $\tilde{g} = g$)) we note:

**4.5.3.** If $(K, M)$ is finite, $b \in M$, then $|\text{Aut}_K M| \leq |Z_M(\min_b, K)| \leq \dim_K M$. \[ \square \]

The question as to when these inequalities become, in fact, equalities will now lead to ground-laying notions of Galois theory:

**4.5.4.** Let $(K, M)$ be finite and $|\text{Aut}_K M| = \dim_K M$. Then $|Z_M(\min_b, K)| = \deg \min_b, K$ for all $b \in M$, hence

(i) For every $b \in M$, $\min_b, K$ splits in $M[t]$ into linear factors.

(ii) There is no $b \in M$ such that $\min_b, K$ has a multiple root in $M$.

The first claim follows by specializing 4.4 as before. This means, by 1.6(3), that $M$ contains a splitting field of $\min_b, K$, hence (i), and that the number of zeros of $\min_b, K$ in $M$ attains the maximal possible value, hence (ii). \[ \square \]

$(K, M)$ is called **normal** if it is algebraic and satisfies condition (i). It is called **separable** if it is algebraic and satisfies condition (ii). A finite, normal, separable field extension is called **Galois** or a **Galois extension**. \[^{69}\]

**4.5.5.** Let $(K, L)$, $(L, M)$ be algebraic field extensions.

(1) If $(K, M)$ is normal, so is $(L, M)$.

(2) If $(K, M)$ is separable, so is $(L, M)$.

(3) If $(K, M)$ is Galois, so is $(L, M)$.

**Proof.** The first two claims follow from the fact that, by 4.1.3 (with $L$ in place of $K$), $\min_{b, L} | \min_{b, K}$ for all $b \in M$. Clearly, (3) is implied by (1) and (2). \[ \square \]

\[^{69}\]**Note that there are authors who do not require a Galois extension to be finite.**
A nonzero polynomial \( f \in K[t] \) is called **separable** over \( K \) if no irreducible factor of \( f \) has a multiple zero in any extension field of \( K \). Clearly, this is equivalent to the condition that no irreducible factor of \( f \) has a multiple zero in a splitting field of \( f \) over \( K \). The condition that \( f \) itself has this property is certainly stronger but sometimes easy to control. The following remark presents an important example:

4.5.6. **Let \( p := \text{char } K > 0 \) and \( k \in \mathbb{N} \) such that \( p|k \). Then \( t^k - t \) is separable over \( K \). If \( K \) is finite, \( k := |K| \), then \( K = Z_K(t^k - t) \). In particular, \( K \) is a splitting field of \( t^k - t \) over its prime field.**

**Proof.** We have \((t^k - t)' = -1 \) as \( p|k \), hence \( t^k - t \) is separable by 1.8: Its formal derivative is of degree 0, thus has no zero. The hypotheses on \( K \) imply that \( x^{k-1} = 1_K \) for all \( x \in K \), by 3.2.2. Therefore \( x^k = x \) for all \( x \in K \), hence \( K = Z_K(t^k - t) \), and \( t^k - t \) is a polynomial over the prime field of \( K \). 

By 4.1.1, \( k = |K| \) implies that \( k \) is a power of \( p \). Of course, it is a typically finite phenomenon that a field consists of the zeros of some polynomial. Returning to the general case, we observe the following criterion of separability:

4.5.7. **A nonzero polynomial \( f \in K[t] \) is separable if and only if \( g' \neq 0_K \) for every irreducible polynomial divisor \( g \) of \( f \) in \( K[t] \).**

**Proof.** Applying 4.3(1), we may consider a splitting field \( M \) of \( f \) over \( K \). Clearly, \( M \) contains a splitting field over \( K \) for every polynomial divisor of \( f \) in \( K[t] \). If \( g \) is such a divisor, irreducible and \( g' = 0_K \), then every zero of \( g \) in \( M \) is a multiple zero by 1.8. Hence \( f \) is not separable. Conversely, if \( f \) is not separable, there exists an irreducible polynomial divisor \( g \) of \( f \) in \( K[t] \), w.l.o.g. normed, which has a multiple zero \( b \in M \). By 4.1.4, \( g = \min_{b,K} \). But 1.8 implies that \( g'(b) = 0_K \) which forces \( g' = 0_K \). 

If \( \text{char } K = 0 \), then \( g' \neq 0_K \) for every \( g \in K[t] \setminus K \). Hence 4.5.7 implies

4.5.8. **If \( \text{char } K = 0 \), every nonzero polynomial over \( K \) is separable.**

4.6 **Proposition.** **Suppose the hypotheses of 4.3(2) and let \( f \) be separable, \( n := \dim_K M \). Then there exist exactly \( n \) extensions of \( \varphi \) to an isomorphism of \( M \) onto \( X \).**

**Proof** by induction on \( n \). The case \( n = 1 \) is trivial. For the inductive step, let \( n > 1 \) and assume the claim to be true for smaller values of \( n \). Let \( b \in Z_M(f) \setminus K \), \( g := \min_{b,K} \), \( \tilde{g} := g\tilde{\varphi} \), \( \tilde{b} \in Z_X(\tilde{g}) \) and \( \varphi_b \) according to 4.2(2).

\(^{70}\)For example, \( f = t^4 + 2t^2 + 1 = (t^2 + 1)^2 \in \mathbb{Q}[t] \) is separable but has multiple zeros in \( \mathbb{C} \). Note, however, that there are authors who define a polynomial to be separable if it has no multiple zeros in any extension field, thus excluding cases like our example \( f \). Clearly, for irreducible polynomials the two definitions coincide. An algebraic element \( b \in M \) is called **separable** over \( K \) if \( \min_{b,K} \) is separable. From 4.1.4 it is clear that this definition is independent of the choice of the definition of separability for arbitrary polynomials.
Then \( M \) is a splitting field of \( f \) over \( K[b] \), \( X \) is a splitting field of \( \tilde{f} \) over \( \tilde{K}[\tilde{b}] \), and \( \dim_{K[b]} M < n \). Clearly, \( \deg g = \deg \tilde{g} \), and \( f \) (\( \tilde{f} \) resp.) is separable over \( K[b] \) (\( \tilde{K}[\tilde{b}] \) resp.). By our inductive hypothesis, there exist exactly \( \dim_{K[b]} M \) extensions of \( \varphi_b \) to an isomorphism of \( M \) onto \( X \). Hence, by 4.3.1 and 4.2.2,

\[
|\varphi^M| = |Z_X(\tilde{g})| \dim_{K[b]} M = \deg g \dim_{K[b]} M = \dim_K K[b] \dim_{K[b]} M = \dim_K M.
\]

where the last two equations make use of 4.1.5 and the special case of 2.6. □

4.7 Theorem. Let \((K, M)\) be a field extension. The following are equivalent:

(i) \((K, M)\) is Galois,

(ii) There exists a nonzero separable polynomial over \( K \) with splitting field \( M \),

(iii) \( |\text{Aut}_K M| = \dim_K M < \infty \).

Proof. (ii)⇒(iii) follows from the special case of \( X = M, \varphi = \text{id}_X \) in 4.6, and (iii)⇒(i) is the contents of 4.5.4. Suppose (i). Let \( B \) be a \( K \)-basis of \( M \), \( f := \prod_{b \in B} \min_b K \). By 4.1.4 and our hypothesis, \( f \) is separable over \( K \). Furthermore, \( f \) splits over \( M \) into linear factors as \( B \subseteq M \) and \((K, M)\) is normal. As \( Z_M(f) \supseteq B, K \cup Z_M(f) \) cannot be contained in any proper subfield of \( M \). Hence \( M \) is a splitting field of \( f \) over \( K \). The proof is complete. □

Example. Let \( K \) be a field, \( n \in \mathbb{N} \). Then \((K_{\text{sym}}(t_1, \ldots, t_n), K(t_1, \ldots, t_n))\) is a Galois extension, and its Galois group \( G \) is isomorphic to \( S_n \): Viète’s theorem 1.11 shows that \( K(t_1, \ldots, t_n) \) is a splitting field of the separable polynomial \( \sum_{k=0}^n (-1)^k s_k t_1^{n-k} \in K_{\text{sym}}(t_1, \ldots, t_n)[t] \). By 4.7, this proves the first assertion. By 2.11.2, \( S_n \) is isomorphic to a subgroup of \( G \). Vice versa, by 4.5.2, \( G \) is isomorphic to a subgroup of \( S_n \). Hence \( G \cong S_n \).

The definition of the Galois group of a field extension is a special case of the following general notion:

4.8 Definition. Let \((M, \circ)\) be a magma, \( G \) a unital submonoid of \( \text{End}(M, \circ) \). For all \( y \in M \), the set

\[
C_G(y) := \{ \alpha \in G, \ y \circ \alpha = y \}
\]

is called the **centralizer** of \( y \) in \( G \). For all \( \alpha \in G \),

\[
\text{Fix}_M(\alpha) := \{ y \in M, \ y \circ \alpha = y \}
\]

is called the **set of fixed elements** of \( \alpha \). We observe

4.8.1. Let \( y \in M, \ \alpha \in G \). Then \( C_G(y) \) is a unital submonoid of \( G \), \( \text{Fix}_M(\alpha) \) is a submagma of \( (M, \circ) \). □

\(^{71}\)Recall 1.3.1 and its comments.
For every $U \subseteq M$ let $C_G(U) := \bigcap_{y \in U} C_G(y)$, for every $H \subseteq G$ similarly $\text{Fix}_M(H) := \bigcap_{\alpha \in H} \text{Fix}_M(\alpha)$. It follows from 4.8.1 (and 2.2.2) that $C_G(U)$ is a unital submonoid of $G$, $\text{Fix}_M(H)$ a submagma of $(M, \circ)$. The following remark is also immediate:

4.8.2. Let $U, \tilde{U} \subseteq M$, $H, \tilde{H} \subseteq G$. Then

$$U \subseteq \tilde{U} \Rightarrow C_G(\tilde{U}) \subseteq C_G(U), \quad H \subseteq \tilde{H} \Rightarrow \text{Fix}_M(\tilde{H}) \subseteq \text{Fix}_M(H).$$

\hfill \square

Let $\mathcal{T}(G)$ be the set of all unital submonoids of $G$, $\mathcal{V}(M)$ the set of all submagmas of $(M, \circ)$. Then $(\mathcal{T}(G), \subseteq)$, $(\mathcal{V}(M), \subseteq)$ are lattices, and the foregoing remarks may be reformulated as follows:

4.8.3. The mappings

$$\zeta : \mathcal{V}(M) \to \mathcal{T}(G), \ U \mapsto C_G(U), \quad \varphi : \mathcal{T}(G) \to \mathcal{V}(M), \ H \mapsto \text{Fix}_M(H)$$

are lattice anti-homomorphisms. \hfill \square

4.8.4. (1) $\forall U \in \mathcal{V}(M) \ U \zeta \varphi \supseteq U$, $U \zeta \varphi \zeta = U \zeta$,

(2) $\forall H \in \mathcal{T}(G) \ H \varphi \zeta \supseteq H$, $H \varphi \zeta \varphi = H \varphi$,

(3) $M \zeta = \{\text{id}_M\}$, $\{\text{id}_M\} \varphi = M$.

Proof. (1) Let $U \in \mathcal{V}(M)$. Every element of $U$ is fixed by every element of $U \zeta$, hence $U \subseteq U \zeta \varphi$, $U \zeta \subseteq U \zeta \varphi \zeta$. Let $\alpha \in U \zeta \varphi \zeta$. Then $\alpha$ centralizes $U \zeta \varphi$, a fortiori $U$. Hence $\alpha \in C_G(U) = U \zeta$.

(2) Let $H \in \mathcal{T}(G)$. Every element of $H \varphi$ is centralized by every element of $H$, hence $H \subseteq H \varphi \zeta$, $H \varphi \subseteq H \varphi \zeta \varphi$. Let $y \in H \varphi \zeta \varphi$. Then $y$ is a fixed element of $H \varphi \zeta$, a fortiori of $H$. Hence $y \in \text{Fix}_M(H) = H \varphi$.

(3) is trivial. \hfill \square

Now let $M$ be a field, $K$ a subfield of $M$ and $G := \text{Aut}_K M$. For every subgroup $H$ of $\text{Aut} M$, $\text{Fix}_M(H)$ is a subfield of $M$: By 4.8.1 it is additively and multiplicatively closed. By the homomorphism rule for inverses (1.3(3)), it is also closed with respect to taking inverses. Clearly, it contains $0_M$ and $1_M$. The prime field of $M$ must hence be contained in $\text{Fix}_M(H)$, in other words:

4.8.5. If $K$ is the prime field of a field $M$, then $c \alpha = c$ for all $c \in K$, $\alpha \in \text{Aut} M$. \hfill \square

If $H \leq G$, clearly $K \leq \text{Fix}_M(H)$, i.e., $\text{Fix}_M(H)$ is an element of the interval $[K, M]$ of the lattice of all subfields of $M$. As in 3.7, we write $\mathcal{U}(G)$ for the subgroup lattice of $G$. For all $L \in [K, M]$, $C_G(L)$ is a subgroup of $G$; it is just a different notation for the
Galois group $\text{Aut}_LM$ of the extension $(L, M)$. Restricting our above $\zeta$ to $[K, M]$, $\varphi$ to $U(G)$, we obtain the following lattice anti-homomorphisms:

- $\Gamma : [K, M] \to U(G), \ L \mapsto C_G(L)(= \text{Aut}_LM),$
- $\Phi : U(G) \to [K, M], \ H \mapsto \text{Fix}_M(H).$

**4.9 Proposition.** Let $(K, M)$ be a Galois extension. Then $(\text{Aut}_K M)\Phi = K$.

In other words: The fixed field of the Galois group of a Galois extension equals the ground field of the extension: It is not larger!

**Proof.** Clearly, $K\Gamma\Phi \supseteq K$ (4.8.4(1)). To prove the reverse inclusion, put $G := K\Gamma$, $L := G\Phi$ and observe that $K \leq L$, $\text{Aut}_LM = L\Gamma = K\Gamma\Phi = K\Gamma$, by 4.8.4(1). It follows that $\dim_K L \dim_L M = \dim_K M = |G| = |\text{Aut}_LM| \leq \dim_L M$, where we applied first the special case of 2.6, then 4.7, and finally 4.5.3. Hence $\dim_K L = 1$, i.e., $K = L$. □

**4.10 Corollary.** Let $(K, M)$ be a Galois extension. Then $\Gamma\Phi = \text{id}_{[K,M]}$. In particular, $\Gamma$ is injective, $\Phi$ is surjective, $[K, M]$ is finite.

**Proof.** Let $L \in [K, M]$. By 4.5.5(3), $(L, M)$ is Galois. Hence, by 4.9, $L\Gamma\Phi = L$. Thus $\Gamma\Phi = \text{id}_{[K,M]}$. From 4.5.3 we know that, in particular, the number of subgroups of $\text{Aut}_KM$ is finite. As $\Phi$ is surjective it follows that $[K, M]$ is finite. □

On our way to the main theorem of Galois theory, 4.10 may be viewed as the goal of the first part. The second part will consist in the proof of the analogous result for the reverse composition $\Phi\Gamma$. Before we embark on this aim, let us consider an illustrating example:

**Example.** Let $M$ be the splitting field of $t^3 - 2$ over $\mathbb{Q}$ in $\mathbb{C}$, $w$ a primitive 3rd root of unity. Then $M = \mathbb{Q}[\sqrt[3]{2}, w]$, by 1.5. As $M \not\subseteq \mathbb{R}$, we have $\dim_{\mathbb{Q}} \mathbb{Q}[\sqrt[3]{2}] \dim_{\mathbb{Q}[\sqrt[3]{2}]} M = 3 \cdot 2 = 6$. From 4.1.5 it follows that $\min_{\mathbb{Q}[\sqrt[3]{2}]} = t^3 - 2$, $\min_{\mathbb{Q},w} = t^2 + t + 1$ (see Example (4) on p. 10). As $M$ is generated as a field by $\mathbb{Q}$, $\sqrt[3]{2}$ and $w$, 4.2.1 implies that at most the six assignments listed in the first two lines of the following table can have an extension to an automorphism of $M$:

<table>
<thead>
<tr>
<th>$\sqrt[3]{2}$</th>
<th>$\sqrt[3]{2}w$</th>
<th>$\sqrt[3]{2}w^{-1}$</th>
<th>$\sqrt[3]{2}$</th>
<th>$\sqrt[3]{2}w$</th>
<th>$\sqrt[3]{2}w^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$w$</td>
<td>$w$</td>
<td>$w$</td>
<td>$w^{-1}$</td>
<td>$w^{-1}$</td>
<td>$w^{-1}$</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>$\alpha_2$</td>
<td>$\alpha_3$</td>
<td>$\alpha_4$</td>
<td>$\alpha_5$</td>
<td>$\alpha_6$</td>
</tr>
</tbody>
</table>

By 4.5.8 and 4.7, $M$ is a Galois extension of $\mathbb{Q}$. By 4.8.5 and again by 4.7, $|\text{Aut}_M| = |\text{Aut}_{\mathbb{Q}}M| = 6$. More precisely, 4.5.2 implies that $\text{Aut}_M \cong S_3$. Hence all six assignments given by the table extend to an automorphism of $M$. We write $\alpha_i$ $(i \in \mathbb{Z}_6)$ for the automorphism which extends the $i$-th assignment in our table (3rd line).

There exist at least the following subfields of $M$:

- $M$, $\mathbb{Q}[\sqrt[3]{2}]$, $\mathbb{Q}[\sqrt[3]{2}w]$, $\mathbb{Q}[\sqrt[3]{2}w^{-1}]$, $\mathbb{Q}[w]$, $\mathbb{Q}$. 75
In the order given, these have the dimensions 6, 3, 3, 2, 1 over \( \mathbb{Q} \). The three subfields of dimension 3 are mutually distinct (as the union of any two of them generates \( M \) as a field). By 4.10, \( \Gamma \) is injective, hence \( |[\mathbb{Q}, M]| \leq |U(S_3)| = 6 \). It follows that the above list contains every subfield of \( M \). We conclude that \( \Gamma, \Phi \) are lattice anti-isomorphisms:

\[
\text{The small numbers are the relative extension degrees (subgroup indices, resp.). The subfields which are normal extensions of \( \mathbb{Q} \) (the normal subgroups, resp.) have been marked by bigger nodes than the others.)}
\]

Our aim is to show that the exact correspondence observed in this example holds for every Galois extension. As a preparation, we resume and slightly generalize the terminology introduced in 4.8: Let \((M, \circ)\) be a magma, \( H \subseteq \text{End}(M, \circ) \), \( k \in \mathbb{N} \). We put

\[
\text{Fix}_{M^k}(H) := \{(y_1, \ldots, y_k) \mid y_i \in M, \forall \alpha \in H y_i\alpha = y_i\}.
\]

Clearly, \( \text{Fix}_{M^k}(H) = \left(\text{Fix}_M(H)\right)^k \). A subset \( T \) of \( M^k \) is called \( H \)-invariant if for all \( y_1, \ldots, y_k \in M, \alpha \in H \),

\[
(y_1, \ldots, y_k) \in T \Rightarrow (y_1\alpha, \ldots, y_k\alpha) \in T,
\]

which is denoted in short by \( T\alpha \subseteq T \). A trivial example of an \( H \)-invariant subset is \( \text{Fix}_{M^k}(H) \).

**4.11 Proposition.** Let \( M \) be a field, \( H \leq \text{Aut} M \), \( k \in \mathbb{N} \).

1. Let \( V \) be an \( H \)-invariant subspace of \( M^k \), \( V \neq \{(0_M, \ldots, 0_M)\} \). Then

\[
V \cap \text{Fix}_{M^k}(H) \neq \{(0_M, \ldots, 0_M)\}.
\]

2. Let \( H \) be finite, \( \alpha_1, \ldots, \alpha_m \) the elements of \( H \) (mutually distinct). Let \( b_1, \ldots, b_k \in M \). Then the kernel of the \( M \)-linear mapping defined by the matrix multiplication

\[
\mu : M^k \to M^m, \quad (x_1, \ldots, x_k) \mapsto (x_1, \ldots, x_k) \cdot (b_{i\alpha_j})_{j \in \mathbb{Z} \atop i \in \mathbb{Z}}
\]

is \( H \)-invariant.

Consequently, \( \ker \mu \cap \text{Fix}_{M^k}(H) \neq \{(0_M, \ldots, 0_M)\} \) if \( k > m \).
Proof. (1) Let \( x_1, \ldots, x_k \in M \), not all \( = 0_M \), such that \( (x_1, \ldots, x_k) \in V \), chosen with the property that the number \( n \) of subscripts \( j \in K \) with \( x_j \neq 0_M \) is minimal. W.l.o.g. assume \( x_1, \ldots, x_n \neq 0_M \), \( x_{n+1} = \cdots = x_k = 0_M \). As \( V \) is an \( M \)-subspace of \( M^k \) we obtain
\[
(x_n^{-1}x_1, \ldots, x_n^{-1}x_{n-1}, 1_M, 0_M, \ldots, 0_M) = x_n^{-1}(x_1, \ldots, x_k) \in V.
\]
Hence there exist \( y_1, \ldots, y_n \in M \) such that \( (y_1, \ldots, y_{n-1}, 1_M, 0_M, \ldots, 0_M) \in V \). For all \( \alpha \in H \) we have \( 1_M \alpha = 1_M, 0_M \alpha = 0_M, V \alpha \subseteq V \). Hence
\[
(y_1\alpha - y_1, \ldots, y_{n-1}\alpha - y_{n-1}, 0_M, \ldots, 0_M) = (y_1, \ldots, y_{n-1}, 1_M, 0_M, \ldots, 0_M) - (y_1, \ldots, y_{n-1}, 1_M, 0_M, \ldots, 0_M) \in V.
\]
By the choice of \( n \) this implies that \( y_j\alpha - y_j = 0_M \), hence \( (y_1, \ldots, y_{n-1}, 1_M, 0_M, \ldots, 0_M) \in V \cap \text{Fix}_{M^k}(H) \).

(2) Let \( x_1, \ldots, x_k \in M \) such that \( (x_1, \ldots, x_k) \in \ker \mu, \alpha \in H, \alpha_j' := \alpha_j \alpha \) for all \( j \in m \). Let \( j \in m \). Then \( \sum_{i \in K} x_i(b_i \alpha_j) = 0_M, \) hence \( \sum_{i \in K} (x_i \alpha)(b_i \alpha_j') = 0_M \alpha = 0_M. \) It follows that \( (x_1\alpha, \ldots, x_k\alpha) \in \ker \mu. \)

If \( k > m \), then \( \ker \mu \neq \{(0_M, \ldots, 0_M)\} \) by the dimension formula for linear mappings. Applying (1) to \( V := \ker \mu \) we obtain the claim. \( \square \)

4.12 Proposition (Artin). Let \( M \) be a field, \( H \) a finite subgroup of \( \text{Aut} M \), \( L := \text{Fix}_M H. \) Then

(1) \( \text{dim}_LM = |H|, \)

(2) \( H = \text{Aut}_LM, \)

(3) \( (L, M) \) is Galois.

Proof. (1) Our main step will be to show that \( \text{dim}_LM \leq |H|: \) Let \( \alpha, \mu \) be as in 4.11(2), \( k \in \mathbb{N} \) and \( b_1, \ldots, b_k \in M \) such that \( (b_1, \ldots, b_k) \) is \( L \)-linearly independent. We want to apply the last part of 4.11 to derive the desired inequality \( k \leq |H|. \) All we have to do is to prove that \( \ker \mu \cap L^k = \{(0_M, \ldots, 0_M)\} \), because \( L^k = \text{Fix}_{M^k}(H). \)

Let \( (x_1, \ldots, x_k) \in \ker \mu \cap L^k. \) Assuming w.l.o.g. that \( \alpha_1 = \text{id}_M, \) we consider the first component of \( (x_1, \ldots, x_k) \mu \) and obtain
\[
0_M = (x_1, \ldots, x_k) \begin{pmatrix} b_1 \\ \vdots \\ b_k \end{pmatrix} = x_1b_1 + \cdots x_kb_k,
\]
hence \( x_1 = \cdots = x_k = 0_M \) as \( (b_1, \ldots, b_k) \) is \( L \)-linearly independent.

Thus \( \text{dim}_LM \leq |H| \leq |\text{Aut}_LM|. \) By 4.5.3, this proves (1) and, moreover, (2) as \( H \leq \text{Aut}_LM. \) Furthermore, it now suffices to apply 4.7 to obtain (3). \( \square \)

4.13 Corollary. Let \( (K, M) \) be a Galois extension. Then \( \Phi \Gamma = \text{id}_{\text{id}(G)} \) \( (\)where \( G = \text{Aut}_K M). \) In particular, \( \Gamma \) is injective, \( \Phi \) is surjective.
Proof. By 4.5.3, $G$ is finite. Let $H \leq G$, $L := \text{Fix}_M(H)$. Then $H \Phi \Gamma = (\text{Fix}_M(H)) \Gamma = L \Gamma = \text{Aut}_LM = H$, by 4.12(2).

As a consequence, we obtain the following further characterization of Galois extensions:

**4.14 Corollary.** Let $(K, M)$ be a field extension. The following are equivalent:

(i) $(K, M)$ is Galois,

(ii) There exists a finite subgroup $G$ of $\text{Aut} M$ such that $\text{Fix}_M G = K$.

Proof. If (i) holds, let $G := \text{Aut}_KM$. Then $\text{Fix}_M G = G \Phi = K \Gamma \Phi = K$ by 4.9. If (ii) holds, we apply 4.12(3) to obtain (i).

**4.15 Proposition.** Let $(K, M)$ be a Galois extension, $L \in [K, M]$, $G := \text{Aut}_KM$. Then $(K, L)$ is normal if and only if $L$ is $G$-invariant (i.e., $L \alpha \subseteq L$ for all $\alpha \in G$).

Proof. For all $b \in M$, $\alpha \in G$ we have $\min_{b,K} = \min_{b_0,K}$, by 4.2.1 (with $\varphi := \text{id}_K$, $X := M$, $\psi := \alpha|_{K[b]}$). Hence, if $(K, L)$ is normal, $b \in L$, then $b_0 \in L$. Conversely, let $L$ be $G$-invariant, $b \in L$. As $(K, M)$ is normal, $[K, M]$ contains a splitting field $Y$ of $\min_{b,K}$ over $K$. Let $\tilde{b} \in Z_Y(\min_{b,K})$. By 4.2(2), there exists an isomorphism $\psi$ of $K[b]$ onto $K[\tilde{b}]$ which maps $b$ onto $\tilde{b}$ and $c$ onto $c$ for all $c \in K$. By 4.7 and 4.3(2), $\psi$ extends to an automorphism $\alpha$ of $M$. Clearly $\alpha \in G$, hence $b = b_0 = \alpha \in L \alpha \subseteq L$.

**4.16 Theorem (Main theorem of Galois theory).** Let $(K, M)$ be a Galois extension and $G := \text{Aut}_KM$. Let

$$
\Gamma : [K, M] \to \mathcal{U}(G), \quad L \mapsto \text{Aut}_LM,
\Phi : \mathcal{U}(G) \to [K, M], \quad H \mapsto \text{Fix}_M(H).
$$

(1) $\Gamma, \Phi$ are mutually inverse lattice anti-isomorphisms. If $L,L^* \in [K, M]$ such that $L \subseteq L^*$, then $\dim L^* = |L \Gamma : L^* \Gamma|$.

(2) Let $L \in [K, M]$. Then $(K, L)$ is normal if and only if $L \Gamma \subseteq G$.

(3) Let $L \in [K, M]$, $(K, L)$ normal. Then $\text{Aut}_KL \cong G/L \Gamma$.

---

---
Proof. (1) By 4.10 and 4.13, $\Gamma \Phi = \text{id}_{[K,M]}$, $\Phi \Gamma = \text{id}_{\text{Aut}(G)}$. By 4.8.2, this proves the first assertion. Let $L, L^* \in [K,M]$ such that $L \subseteq L^*$. We apply the special case of 2.6, 4.7 and 3.1.2 to conclude
\[
\dim_L L^* = \frac{\dim_L M}{\dim_L M} = \left| L^\Gamma \right| = \left| L : L^\Gamma \right|.
\]
(2) By 4.15 (including footnote 72) and (1), we obtain
\[
(K, L) \text{ normal } \iff \forall \alpha \in G \ L\alpha = L \iff \forall \alpha \in G \ (L\alpha)\Gamma = L\Gamma
\]
\[
\iff \exists \alpha \in G \ (\text{Aut}_L M)^{\alpha} = \text{Aut}_L M \iff \text{Aut}_L M \leq G,
\]
where we make use of 4.5.1 for the 3rd equivalence.
(3) Let $L \in [K,M]$, $(K, L)$ normal. By 4.15, $\alpha|_L \in \text{Aut}_K L$ for all $\alpha \in G$. Let
\[
\rho : G \to \text{Aut}_K L, \ \alpha \mapsto \alpha|_L.
\]
Then $\rho$ is a homomorphism, $\ker \rho = \{ \alpha|_L \in G, \ \alpha|_L = \text{id}_L \} = \text{Aut}_L M = L\Gamma$. Every $\beta \in \text{Aut}_K L$ is, in particular, a $K$-monomorphism of $L$ into $M$, hence extends to an automorphism $\alpha \in \text{Aut}_K M$, by 4.7 and 4.3(2). As $\beta = \alpha|_L \in G\rho$, $\rho$ is surjective. By 3.2(2), $G/L\Gamma = G/\ker \rho \cong G\rho = \text{Aut}_K L$.

The perfect correspondence between the lattice of intermediate fields of a Galois extension $(K, M)$ and the subgroup lattice of its Galois group $G := \text{Aut}_K M$ gives rise to applications in both directions: Some knowledge from group theory may help to give information about fields in the interval $[K, M]$, some knowledge from field theory may help to give information about subgroups of $G$. For example, if $k$ is a prime power dividing $\dim_K M$, there must exist a field $L \in [K,M]$ such that $\dim_L M = k$. The reason is that $G$ has a subgroup $H$ of order $k$ by Sylow’s 1st theorem (see p. 54, 8.4(2)) so that it suffices to put $L := \Phi H\Phi$, by 4.16(1). In the reverse direction, if we know that $M$ is the splitting field of a separable irreducible polynomial of degree $n$ over $K$, then $G$ must have a subgroup $H$ such that $|G : H| = n$. As a proof, we just have to observe that, by 4.1.5, $\dim_K K[b] = n$ for every $b \in \mathcal{Z}_M(f)$. Therefore it suffices to put $H := (K[b])\Gamma$, by 4.16(1).

We will see in chapter 5 that 4.16 has crucial consequences for the problem of solubility of polynomial equations by radicals. We finish this chapter with applications of 4.16 in a different direction, deriving important structural properties of finite fields.

4.17 Definition. Let $p$ be a prime, $M$ a commutative unitary ring such that $\text{char} M = p$. We set
\[
\varphi_p : M \to M, \ x \mapsto x^p.
\]

4.17.1. $\varphi_p$ is a ring endomorphism.
Proof. Let $x, y \in M$. Then $(xy)\varphi_p = (xy)^p = x^py^p = (x\varphi_p)(y\varphi_p)$, and

$$(x + y)\varphi_p = (x + y)^p = \sum_{j=0}^{p-1} \binom{p}{j} x^{p-j}y^j = x^p + y^p = x\varphi_p + y\varphi_p$$

as $p|\binom{p}{j}$ for all $j \in p - 1$.\hfill\square

The mapping $\varphi_p$ is called the Frobenius endomorphism of $M$.

4.17.2. If $M$ is a field $\mathbb{74}$, $\varphi_p$ is a monomorphism. For a finite field $M$, $\varphi_p$ is an automorphism.

Proof. With respect to addition, let $x \in \ker \varphi_p$. Then $0_M = x\varphi_p = x^p$, hence $x = 0_M$. It follows that $\varphi_p$ is injective. If $M$ is finite, this implies that $\varphi_p$ is surjective.\hfill\square

If $M$ is a finite field, $\varphi_p$ is called the Frobenius automorphism of $M$.

4.18 Theorem (1st main theorem on finite fields). Let $p$ be a prime, $n \in \mathbb{N}$.

(1) There exists a field of order $p^n$.

(2) Any two fields of order $p^n$ are isomorphic.

Proof. Recall footnote 34: (2) Let $M$ be a field of order $p^n$. By 4.5.6, $M$ is a splitting field of the polynomial $t^n - t$ over the prime field of $M$. By 3.13.5, the prime fields of any two fields of characteristic $p$ are isomorphic. Applying 4.3(2) to the polynomial $f = t^n - t$ we obtain the claim.

(1) Making use of 4.3(1), we consider a splitting field $X$ of $f := t^n - t$ over a field $K$ of order $p$. By 4.5.6, $f$ is separable over $K$, hence $|Z_X(f)| = p^n$. We show that $Z_X(f)$ is a subfield of $X$: Clearly, $0_X, 1_X \in Z_X(f)$. Let $a, b \in Z_X(f)$. By 4.17.1, $(a - b)^n - (a - b) = (a - b)\varphi_f^n - (a - b) = a\varphi_f^n - b\varphi_f^n - a + b = 0_M$, i.e., $a - b \in Z_X(f)$. If $b \neq 0_M$, $(ab^{-1})^n = a^n(b^n)^{-1} = ab^{-1}$, i.e., $ab^{-1} \in Z_X(f)$. By the definition of a splitting field, it follows that $X = Z_X(f)$. Hence $X$ is a field of order $p^n$.\hfill\square

The proof shows that a finite field $M$ (the order of which is, by 4.1.1, necessarily a power of a prime) is a splitting field of a separable polynomial over its prime field $K$, hence a Galois extension field of $K$, by 4.7. We give a second line of reasoning for this fact, considering the Frobenius automorphism $\varphi_p$ of $M$.

4.18.1. Let $p$ be a prime, $n \in \mathbb{N}$. With respect to a field $M$ of order $p^n$, $o(\varphi_p) = n$.

Proof. By 3.5 there exists an element $x \in M$ such that, multiplicatively, $M = \langle x \rangle$. Let $k \in \mathbb{N}$; then

$$\varphi_p^k = \text{id}_M \iff x \varphi_p^k = x \iff x^{b^k} = x \iff x^{b^k-1} = 1_M \iff o(x)p^k - 1,$$

(cf. Example (3) on p. 50).) Hence $o(\varphi_p) = \min \{k \mid p^n - 1 | p^k - 1\} = n$.\hfill\square

\textsuperscript{74} As the proof shows, it is sufficient that there is no nonzero element $x \in M$ such that $x^p = 0_M$. (In an arbitrary ring $R$, an element $x$ is called nilpotent if there exists a positive integer $n$ such that $x^n = 0_R$.)

80
We have $\text{Aut} M = \text{Aut}_K M$ and, by 3.2.2 and 4.5.3,

$$\dim_K M = n = o(\varphi_p) \mid |\text{Aut} M| \leq \dim_K M,$$

hence $|\varphi_p| = o(\varphi_p) = |\text{Aut} M|$. It follows that $\langle \varphi_p \rangle = \text{Aut} M$. As $\dim_K M = n = |\text{Aut}_K M|$, 4.7 implies that $(K, M)$ is Galois. We derived this fact before, but now we have even determined the Galois group of $(K, M)$: $\text{Aut} M = \langle \varphi_p \rangle$ is cyclic. Now 3.4 implies (see 3.7) that

$$\mu : U(\text{Aut} M) \to \mathbb{N}, \ H \mapsto |H|,$$

is an isomorphism of $(U(\text{Aut} M), \leq)$ onto $(T_n, |)$.

**4.19 Theorem** (2nd main theorem on finite fields). Let $M$ be a finite field, $K$ its prime field, $p := |K|$. Then there exists a positive integer $n$ such that $|M| = p^n$, $\text{Aut} M = \langle \varphi_p \rangle$, and the mapping

$$\delta : [K, M] \to \mathbb{N}, \ L \mapsto \dim_K L,$$

is a lattice isomorphism of $([K, M], \leq)$ onto $(T_n, |)$.

**Proof.** The first two assertions have already been obtained by our foregoing remarks. It remains to prove the claim about $\delta$. Let $\iota : T_n \to T_n$, $k \mapsto \frac{n}{k}$ (see p. 54). Then $\Gamma \mu \iota$ is a composition of two anti-isomorphisms and an isomorphism of lattices, hence a lattice isomorphism of $([K, M], \leq)$ onto $(T_n, |)$.

For all $L \in [K, M]$ we have, by 4.16(1), $L\Gamma \mu \iota = \frac{n}{|\text{Aut}_L M|} = \frac{\dim_K M}{\dim_L M} = \dim_K L$, thus $\Gamma \mu \iota = \delta$. 

\[81\]
Appendix

On irreducibility of polynomials over \( \mathbb{Q} \)

4.20 Definition. A polynomial \( f = \sum_{j=0}^{n} c_j t^j \in \mathbb{Z}[t] \) is called primitive if \( \gcd(c_0, \ldots, c_n) = 1 \). Clearly, this condition holds if one of the coefficients \( c_i \) equals 1 or \(-1\). In particular, normed polynomials in \( \mathbb{Z}[t] \) are primitive. A further trivial remark is the following:

4.20.1. Let \( f \in \mathbb{Z}[t] \setminus \mathbb{Z} \) be irreducible in \( \mathbb{Z}[t] \). Then \( f \) is primitive,
as a common divisor \( d \neq 1, -1 \) of \( c_0, \ldots, c_n \) allows the decomposition \( f = d \cdot \sum_{j=0}^{n} \frac{c_j}{d} t^j \) into non-units of \( \mathbb{Z}[t] \).

The set of primitive polynomials is a unital submonoid of \((\mathbb{Z}[t], \cdot)\). This is the content of the following, known as Gauss’s lemma:

4.20.2. Let \( g, h \in \mathbb{Z}[t] \) be primitive. Then \( gh \) is primitive.

Proof. Let \( k, l \in \mathbb{N}_0 \), \( a_i, b_j \in \mathbb{Z} \) for all \( i, j \in \mathbb{N}_0 \) such that \( g = \sum_{i=0}^{k} a_i t^i \), \( h = \sum_{j=0}^{l} b_j t^j \), \( a_k, b_l \neq 0 \) where we put \( a_i := 0 =: b_j \) for \( i > k, j > l \). Then \( gh = \sum_{m=0}^{k+l} (\sum_{i=0}^{m} a_i b_{m-i}) t^m \). Assume there is a prime \( p \) such that \( p|\sum_{i=0}^{m} a_i b_{m-i} \) for all \( m \). As \( g, h \) are primitive, there exists a minimal \( r \leq k \) and a minimal \( s \leq l \) such that \( p \nmid a_r \), \( p \nmid b_s \). We have

\[
p| \sum_{i=0}^{r+s} a_i b_{r+s-i} = a_0 b_{r+s} + \cdots + a_{r-1} b_{s+1} + a_r b_s + a_{r+1} b_{s-1} + \cdots + a_{r+s} b_0.
\]

Each of the first \( r \) and the last \( s \) summands in this sum is divisible by \( p \) as \( p \) divides \( a_0, \ldots, a_{r-1}, b_0, \ldots, b_{s-1} \). Hence \( p|a_r b_s \) which implies, as \( p \) is a prime, that \( p|a_r \) or \( p|b_s \) (Recall that \( p \) has the property stated for \( q \) in (*) on p. 29.) This contradiction shows that \( gh \) is primitive. \( \square \)

4.20.3. Let \( g \in \mathbb{Q}[t] \), \( g \neq 0 \), \( m := \min\{n|n \in \mathbb{N}, \ ng \in \mathbb{Z}[t]\} \). Then \( g = \frac{\hat{g}}{m} \) for some \( z \in \hat{\mathbb{Z}} \) and a primitive polynomial \( \hat{g} \in \mathbb{Z}[t] \). If \( g \) is normed, \( z = 1 \) may be chosen.

Proof. Set \( g^* := mg \). Let \( z \) be the greatest common divisor of the coefficients of \( g^* \) and \( \hat{g} \in \mathbb{Z}[t] \) such that \( g^* = z\hat{g} \). Then \( \hat{g} \) is primitive and \( g = \frac{\hat{g}}{m} \). Now suppose that \( g \) is normed and let \( d \in \mathbb{N} \) be a common divisor of all coefficients of \( g^* \). Then \( d|m \) as \( g \) is normed, and \( \frac{d}{m} g = \frac{1}{d} g^* \in \mathbb{Z}[t] \). Hence \( d = 1 \) by the minimality of \( m \). \( \square \)

4.20.4. Let \( h \in \mathbb{Z}[t] \) be primitive and \( g \in \mathbb{Q}[t] \). If \( gh \in \mathbb{Z}[t] \), then \( g \in \mathbb{Z}[t] \).

Proof. According to 4.20.3, let \( z \in \hat{\mathbb{Z}}, m \in \mathbb{N}, \hat{g} \in \mathbb{Z}[t] \) be primitive such that \( \gcd(z, m) = 1 \), \( g = \frac{\hat{g}}{m} \). As \( gh \in \mathbb{Z}[t] \), a prime divisor of \( m \) would have to divide each coefficient of the polynomial \( ugh \), hence of \( \hat{g}h \) by (*) on p. 29. But by 4.20.2, \( \hat{g}h \) is primitive. It follows that \( m \) has no prime divisor, i.e., \( v = 1 \) as \( v \in \mathbb{N} \). Thus \( g = z\hat{g} \in \mathbb{Z}[t] \). \( \square \)
Let \( g, h \in \mathbb{Q}[t] \), \( h \) normed. If \( gh \in \mathbb{Z}[t] \), then \( g \in \mathbb{Z}[t] \).

**Proof.** Choosing \( m \) as in 4.20.3 (with \( h \) in place of \( g \)), we obtain that \( mh \) is primitive and \( g(mh) = mgh \in \mathbb{Z}[t] \). By 4.20.4, \( g \in \mathbb{Z}[t] \). Consequently, a product of two normed polynomials over \( \mathbb{Q} \) has integer coefficients if and only if both factors have integer coefficients.

4.20.6. Let \( b \in \mathbb{C} \) be a zero of a normed polynomial \( f \in \mathbb{Z}[t] \). Then \( \min_{b, \mathbb{Q}} \in \mathbb{Z}[t] \).

**Proof.** By 4.1.3, \( f = \min_{b, \mathbb{Q}} h \) for some \( h \in \mathbb{Q}[t] \). As \( f \) is normed, also \( h \) is normed. The claim follows from 4.20.5.

4.20.7. Let \( f \in \mathbb{Z}[t] \setminus \mathbb{Z} \) be irreducible in \( \mathbb{Z}[t] \). Then \( f \) is irreducible in \( \mathbb{Q}[t] \).

**Proof.** Let \( f \in \mathbb{Z}[t], g, h \in \mathbb{Q}[t] \) such that \( f = gh \), \( \deg g, \deg h > 0 \). According to 4.20.3, let \( q \in \mathbb{Q} \) such that \( h = qh \), for some primitive polynomial \( h \in \mathbb{Z}[t] \). Then \( qg \in \mathbb{Q}[t] \) and \( (qg)h = f \in \mathbb{Z}[t] \). By 4.20.4, it follows that \( qg \in \mathbb{Z}[t] \). As \( \deg qg = \deg g, \deg h = \deg h \), \( f \) is not irreducible in \( \mathbb{Z}[t] \).

4.21 Theorem (Eisenstein 1850). \(^{76}\) Let \( n \in \mathbb{N} \), \( f = \sum_{j=0}^{n} c j^j \in \mathbb{Z}[t] \). Suppose that there exists a prime \( p \) such that \( p | c_0, \ldots, c_{n-1}, p \nmid c_n, p^2 \nmid c_0 \). Then \( f \) is irreducible in \( \mathbb{Q}[t] \).

**Proof.** We may factor out the greatest common divisor \( d \) of the coefficients \( c_j \) and, writing \( f = df \), observe that the primitive polynomial \( f \) satisfies the same hypotheses as \( f \). As \( d \) is a unit in \( \mathbb{Q}[t] \), it suffices to prove the assertion for \( f \) in place of \( f \). In other words, we assume in the sequel, w.l.o.g., that \( f \) is primitive. By 4.20.7, it then suffices to show that \( f \) is irreducible in \( \mathbb{Z}[t] \):

Let \( g, h \in \mathbb{Z}[t] \) such that \( f = gh \). Let \( k \in \mathbb{N}_0, a_0, \ldots, a_k, b_0, \ldots, b_{n-k} \in \mathbb{Z} \) such that \( g = \sum_{i=0}^{k} a_i t^i, h = \sum_{j=0}^{n-k} b_j t^j, a_k \neq 0 \neq b_{n-k} \). We have \( p | c_0 = a_0 b_0 \). As \( p \) is a prime, it follows that \( p | a_0 \) or \( p | b_0 \) (by \(^*\) on p. 29). W.l.o.g. assume that \( p | a_0 \). Then \( p \nmid b_0 \) as \( p^2 \nmid c_0 \). Since \( f \) is primitive, \( p \nmid c_n = a_k b_{n-k} \), hence \( p \nmid a_k \).

Let \( m \in \mathbb{Z} \) be minimal with the property that \( p \nmid a_m \). Then \( p | a_{m-j} \) for all \( j \in \mathbb{Z} \). The assumption \( k < n \) would imply \( m < n \) so that, by hypothesis, \( p | c_m = \sum_{j=0}^{m} a_{m-j} b_j \), hence \( p | c_m - \sum_{j \in \mathbb{Z}} a_{m-j} b_j = a_m b_0 \). But \( p \nmid a_m, b_0 \), a contradiction as \( p \) is a prime. It follows that \( k = n \), hence \( h \in \mathbb{Z} \). This implies that \( h | c_0, \ldots, c_n \) so that \( h \in \{1, -1\} \) as \( f \) is primitive, i.e., \( h \) is a unit of \( \mathbb{Z}[t] \).

By 4.21, every polynomial \( t^n - p \) where \( n \in \mathbb{N}, p \) a prime, is irreducible in \( \mathbb{Q}[t] \); more generally, every polynomial \( t^n - pk \) where \( k \in \mathbb{Z} \) and \( p \nmid k \). The polynomials of degree 5 which lead to the equations mentioned on p. 22 are of “Eisenstein type”. While it is a

---

\(^{75}\)Numbers \( b \) with this property are called algebraic integers. They are the classical object of study in Algebraic Number Theory. For a few basic properties, see p. 126.

\(^{76}\)Frequently, a result is erroneously not called after its discoverer. Eisenstein’s theorem is no exception: In substance, it was anticipated by Schönemann in 1846.
trivial fun to produce countless examples of irreducible polynomials in \( \mathbb{Q}[t] \) by choosing suitable coefficients in accordance with Eisenstein’s theorem, there are other applications of 4.21 which are less obvious. As an example, we consider the \( p \)-th cyclotomic polynomial

\[
\Phi_p = t^{p-1} + \cdots + t + 1 = \frac{t^p - 1}{t - 1}
\]

(where \( p \) is a prime). In Example (4) on p.10, we considered the case \( p = 3 \). As \( p \) is a prime, the zeros of \( \Phi_p \) in \( \mathbb{C} \) are exactly the primitive \( p \)-roots of unity, i.e., the \( p \)-th roots of unity \( \neq 1 \). At first glance surprisingly, we may apply 4.21 to prove that \( \Phi_p \) is irreducible in \( \mathbb{Q}[t] \):

To this end we note first that the substitution endomorphism \( F_{t+1} \) of \( \mathbb{Q}[t] \) is an automorphism: Clearly, \( f \) and \( fF_{t+1} \) have the same degree, for any \( f \in \mathbb{Q}[t] \setminus \{0\} \). Hence \( \ker F_{t+1} \) is trivial and \( F_{t+1} \) is injective. Furthermore, \( t = (t - 1)F_{t+1} \) so that \( \mathbb{Q}[t]F_{t+1} \) contains \( t \). It follows that \( F_{t+1} \) is surjective. A polynomial \( f \in \mathbb{Q}[t] \) is irreducible if and only if its image with respect to any automorphism of \( \mathbb{Q}[t] \) is irreducible. In particular, the claim that \( \Phi_p \) is irreducible will follow if we succeed in showing that \( \Phi_p F_{t+1} \) is irreducible. As we have

\[
\Phi_p F_{t+1} = \frac{(t + 1)^p - 1}{(t + 1) - 1} = \sum_{j=0}^{p} \binom{p}{j} t^j - 1 = \sum_{j \in \mathbb{Z}} \binom{p}{j} t^j - 1,
\]

we see that \( \Phi_p F_{t+1} \) satisfies the hypotheses of 4.21, hence is indeed irreducible.

Recall that for arbitrary \( n \in \mathbb{N} \), the \( n \)-th cyclotomic polynomial is given by

\[
\Phi_n := \prod_{w \in \mathbb{P}_n} (t - w)
\]

(cf. p.24). The field \( \mathbb{Q}_n := \mathbb{Q}[w] \) where \( w \in \mathbb{P}_n \) is called the \( n \)-th cyclotomic field. From 1.4.4 we conclude

\[
t^n - 1 = \prod_{d|n} \Phi_d \quad \text{for all } n \in \mathbb{N}
\]

as both polynomials are normed of degree \( n \) and have the same \( n \) zeros in \( \mathbb{C} \): the \( n \)-th roots of unity.

**4.21.1.** \( \Phi_n \in \mathbb{Z}[t] \), for every \( n \in \mathbb{N} \).

*Proof by induction on \( n \).* The claim is trivial for \( n = 1 \). Let \( n > 1 \) and assume that \( \Phi_d \in \mathbb{Z}[t] \) for all \( d < n \). In \( \mathbb{C}[t] \) we have the equation

\[
t^n - 1 = \Phi_nh \quad \text{where } h := \prod_{n \neq d|n} \Phi_d.
\]
Thanks to our induction hypothesis, \( h \in \mathbb{Z}[t] \) and \( \Phi_n = \frac{t^n - 1}{h} \in \mathbb{Q}(t) \cap \mathbb{C} = \mathbb{Q}[t] \). As \( h \) is obviously normed, we conclude that \( g \in \mathbb{Z}[t] \) by 4.20.4.

Finally we show that the \( n \)-th cyclotomic polynomial is irreducible in \( \mathbb{Q}[t] \), for every \( n \in \mathbb{N} \). If \( n \) is a prime, Eisenstein’s theorem 4.21 provides a very simple proof as we have seen above. For arbitrary \( n \), the proof is more complicated. It is a nice example for a line of reasoning which makes use of rings of prime characteristic to obtain an assertion for characteristic 0. We will make use of the standard notation introduced on p. 67 in the more general framework of a commutative unitary ring \( K \) and a unital ring epimorphism \( \varphi \) of \( K \) onto a unitary ring \( K' \). Obviously \( \varphi \) extends to a ring epimorphism \( \tilde{\varphi} \) of \( K'[\tilde{t}] \) onto the polynomial ring \( K[\tilde{t}] \) with the property \( t\varphi = \tilde{t} \) (cf. 2.5.1).

**4.22 Lemma.** Let \( p \) be a prime, \( g, h \in \mathbb{Z}[t] \) normed and irreducible, \( b \in \mathbb{C} \) such that \( g(b) = 0 = h(b^p) \). Put \( \mathbb{Z}_p := \mathbb{Z}/p\mathbb{Z} \) and consider the standard extension \( \mathbb{Z}[t] \to \mathbb{Z}_p[\tilde{t}], \ f \mapsto \tilde{f} \), of the canonical epimorphism of \( \mathbb{Z} \) onto \( \mathbb{Z}_p \). Suppose that \( \tilde{g}, \tilde{g}' \) have no common zeros. Then \( \tilde{g} = h \).

**Proof.** As \( h(b^p) = 0 \), \( b \) is a zero of the polynomial \( h(t^p) \). It follows that \( h(t^p) = fg \) for some \( f \in \mathbb{Z}[t] \), by 4.1.3, 4.1.4 and 4.20.5. Applying 4.17.1, we obtain

\[
\tilde{h}^p = \tilde{h}(t^p) = \tilde{f} \tilde{g}.
\]

Let \( X \) be a splitting field of \( \tilde{h} \) over \( F_p \) (4.3(1)). Then \( Z_X(\tilde{g}) \subseteq Z_X(\tilde{h}) \). Our hypothesis on \( \tilde{g} \) implies, by 1.8, that \( |Z_X(\tilde{g})| = \deg \tilde{g} \). As \( b^p \in \mathbb{Q}[b] \), we conclude, by 4.1.5 and 4.1.4, that

\[
\deg \tilde{g} = |Z_X(\tilde{g})| \leq |Z_X(\tilde{h})| \leq \deg \tilde{h} = \deg h = \dim_\mathbb{Q} \mathbb{Q}[b^p] \leq \dim_\mathbb{Q} \mathbb{Q}[b] = \deg g = \deg \tilde{g}.
\]

It follows that \( \deg \tilde{g} = \deg \tilde{h} \), \( Z_X(\tilde{g}) = Z_X(\tilde{h}) \), hence \( \tilde{g} = \tilde{h} \) as both polynomials are normed. \( \square \)

**4.23 Theorem.** \( \Phi_n \) is irreducible in \( \mathbb{Q}[t] \), for every \( n \in \mathbb{N} \).

**Proof.** Let \( n \in \mathbb{N} \), \( w \in \mathcal{P}_n \), \( g := \min_{w, \mathbb{Q}} \). For every prime \( p \) such that \( p \nmid n \) we show: \( g(w^p) = 0 \).

To this end put \( h := \min_{w, \mathbb{Q}} \). As \( p \nmid n \) we have \( b^p \in \mathcal{P}_n \), hence \( g, h \in \mathbb{Z}[t] \) and \( g, h \mid \Phi_n \) _{\mathbb{Z}[t]} \n
\[\text{Here is a much more general reason for this statement, independent of any induction and a typical application of Galois theory:}\]

**Proposition.** Let \( (K, M) \) be a Galois extension, \( G \) its Galois group. Let \( Q \) be a finite \( G \)-invariant subset of \( M \). Then \( \prod_{b \in \mathcal{Q}(t - b) \in K'[t]} \).

**Proof.** Let \( n := |Q| \) and \( b_1, \ldots, b_n \) be the elements of \( Q \). Then \( s_k(b_1, \ldots, b_n) = s_k(b_1, \ldots, b_n, \alpha) = s_k(b_1, \ldots, b_n) \) for all \( k \in \mathbb{Z} \cup \{0\}, \alpha \in G \). By 1.11, the coefficients of the polynomial \( \prod_{b \in \mathcal{Q}} (t - b) \) are contained in \( \text{Fix}_M(G) \), hence in \( K \), by 4.9. \( \square \)

The step in the proof of 4.21.1 deals with the special case where \( K = \mathbb{Q}, M \) is a splitting field of \( t^n - 1 \) over \( \mathbb{Q} \), \( Q = \mathcal{P}_n \).
by 4.20.6, 4.1.3. We have to show that \( g = h \) and make the assumption that \( g \neq h \). Then \( gh \mid \Phi_n \), hence
\[
\tilde{g}h \mid \Phi_n \mid \tilde{t}^n - 1 \mathbb{Z}_p.
\]
The polynomials \( \tilde{t}^n - 1 \mathbb{Z}_p \) and \((\tilde{t}^n - 1)_p') = n\tilde{t}^{n-1} \) have no common zero, hence \( \tilde{t}^n - 1 \mathbb{Z}_p \) has no multiple zeros in any extension field of \( \mathbb{Z}_p \), by 1.8. The same must hold for its divisor \( \tilde{g} \) so that we may apply 4.22 and conclude that \( \tilde{g} = \tilde{h} \). It follows that \( \tilde{g}^2 \mid \Phi_n \), a contradiction because \( \tilde{t}^n - 1 \mathbb{Z}_p \) has no multiple zeros in any extension field of \( \mathbb{Z}_p \). It follows that \( g = h \), thus \( g(w^p) = 0 \).

An easy induction now shows that \( g(w^k) = 0 \) for every \( k \in \mathbb{N} \) such that \( \gcd(k, n) = 1 \): The claim is trivial for \( k = 1 \). For the inductive step, let \( k > 1 \). Let \( p \) be the smallest divisor \( > 1 \) of \( k \). Then \( p \) is a prime for which we have shown above that \( g(w^p) = 0 \). Now \( w^p \in \mathcal{P}_n \), \( g = \min_{w^p \mathbb{Q}} \) by 4.1.4, and \( \frac{n}{p} \) is prime to \( n \) and \( < k \). Hence, applying the inductive hypothesis to the primitive root of unity \( w^p \), we conclude that \( 0 = g((w^p)^\frac{n}{p}) = g(w^k) \).

By 1.4.2 and its converse, the set of all \( w^k \) where \( \gcd(k, n) = 1 \) equals \( \mathcal{P}_n \). Hence \( \mathbb{Z}_c(g) = \mathcal{P}_n \) so that \( g = \Phi_n \). The claim follows.

Consequently, \( \dim_{\mathbb{Q}} \mathbb{Q}_n = \varphi(n) \) for every \( n \in \mathbb{N} \). Now it is easy to determine the Galois group \( \text{Aut} \mathbb{Q}_n \). Proceeding as in the Example on p. 75, we know that every automorphism of \( \mathbb{Q}_n \) is uniquely determined by the image of \( w \) where \( w \) is a fixed element of \( \mathcal{P}_n \). The only candidates for this image are the elements of \( \mathcal{P}_n \), and these are the powers \( w^k \) with \( \gcd(k, m) = 1 \). It suffices to consider \( k \in \mathbb{N} \) which shows that every such assignment \( w \mapsto w^k \) extends to an automorphism \( \alpha_k \) of \( \mathbb{Q}_n \) as their number equals \( \dim_{\mathbb{Q}} \mathbb{Q}_n \), hence \( \varphi(n) \). Furthermore, the simple rule \( (w^k)^l = w^{kl} \) shows that the mapping
\[
\varphi : \{k \mid k \in \mathbb{Z}, \gcd(k, n) = 1\} \to \text{Aut} \mathbb{Q}_n, \quad k \mapsto \alpha_k
\]
is a multiplicative epimorphism. For all \( k \) with \( \gcd(k, n) = 1 \), the anti-image \( \alpha_k \varphi^{-} \) equals the coset \( n\mathbb{Z} + k \). The set of these cosets is the unit group of the factor ring \( \mathbb{Z}/n\mathbb{Z} \) and isomorphic (via \( \varphi^{-} \)) to \( \text{Aut} \mathbb{Q}_n \) (see footnote 56). Thus \( \text{Aut} \mathbb{Q}_n \) is an abelian group of order \( \varphi(n) \).

Within the universe of all Galois extensions of \( \mathbb{Q} \) with abelian Galois group, the cyclotomic fields play a dominant role because of the following famous result:

**Theorem** (Kronecker-Weber 1886) *Let \( (\mathbb{Q}, M) \) be a Galois extension with abelian Galois group. Then there exists a positive integer \( n \) such that \( M \) is a subfield of \( \mathbb{Q}_n \).*

Theorems 4.21 and 4.23 deal with important classes of irreducible polynomials over \( \mathbb{Q} \) but they do not give an answer to the question how to decide whether a given polynomial \( f \) over \( \mathbb{Q} \) is irreducible or not. Due to Kronecker, there is an algorithmic answer to this question, thanks to 4.20.7 which allows us to assume that \( f \) has integer coefficients. Let
\[ f = \sum_{j=0}^{n} c_j t^j \in \mathbb{Z}[t]. \] If \( f = gh \) where \( g, h \in \mathbb{Z}[t] \), we have \( \deg g \leq \frac{n}{2} \) or \( \deg h \leq \frac{n}{2} \). Thus it suffices to decide if \( f \) has a divisor \( g \) of degree \( k \leq \frac{n}{2} \) in \( \mathbb{Z}[t] \). Let \( k \leq \frac{n}{2} \) and consider arbitrary mutually distinct integers \( x_1, \ldots, x_{k+1} \). The assumption that \( g \mid f \) clearly implies that \( g(x_i) \mid f(x_i) \) for all \( i \in \mathbb{Z}_k \). Now each of the values \( f(x_i) \) has only finitely many divisors \( d_i \in \mathbb{Z} \). Therefore there exist only finitely many \((k+1)\)-tuples \((d_1, \ldots, d_{k+1})\) such that \( d_i \mid f(x_i) \) for all \( i \in \mathbb{Z}_k \).

Given one of these \((k+1)\)-tuples, put \( q_i := \frac{d_i}{\prod_{j \neq i} (x_i - x_j)} \) for all \( i \in \mathbb{Z}_k \). Then

\[
g_{d_1, \ldots, d_{k+1}} := \sum_{i \in \mathbb{Z}_k} q_i (t - x_1) \cdots (t - x_{i-1})(t - x_{i+1}) \cdots (t - x_{k+1})
\]

has the property that substitution of \( x_i \) gives the value \( d_i \) for all \( i \in \mathbb{Z}_k \). It is the only polynomial in \( \mathbb{Q}[t] \) with this property because the difference of any two distinct such polynomials would be nonzero, of degree \( \leq k \) and would have at least the \( k+1 \) zeros \( x_1, \ldots, x_{k+1} \) which is impossible by 1.6(3). Now it suffices to check if \( g_{d_1, \ldots, d_{k+1}} \) has integer coefficients and divides \( f \). Either this happens for a suitable choice of the divisor tuple \( (d_1, \ldots, d_{k+1}) \), or \( f \) is irreducible. Clearly it helps technically to assume that \( f \) is primitive and to choose the elements \( x_i \) such that \( f(x_i) \) has only very few integer divisors.
5 Solubility by radicals

Let $K$ be a field, $n \in \mathbb{N}$, $a \in K$, $L$ an extension field of $K$. We set $\hat{L} := L \setminus \{0_L\}$,

$$R_n(a)_L := \{b | b \in L, b^n = a\}.$$

5.0.1. $R_n(1_K)_L$ is a finite multiplicative subgroup of $\hat{L}$, $R_n(a)_L = R_n(1_K)_Lb$ for every $b \in R_n(a)_L$. In particular, $|R_n(a)_L| = |R_n(1_K)_L|$ if $a \neq 0_K$, $R_n(a)_L \neq \emptyset$.

Proof. $1_K \in R_n(1_K)_L$, and for any $b, b' \in R_n(1_K)_L$ we have $(b'b^{-1})^n = b^n(b^n)^{-1} = 1_K1_K = 1_K$, hence $b'b^{-1} \in R_n(1_K)_L$. The polynomial $t^n - a \in K[t]$ has at most $n$ roots in any extension field of $K$ which completes the proof of the first claim. If $a = 0_K$ we have $R_n(a)_L = \{0_K\} = R_n(1_K)_L \cdot 0_K$. Now let $a \in \hat{K}$, $b \in R_n(a)_L$. Then $b \neq 0_K$ and for any $x \in L$,

$$x \in R_n(1_K)_Lb \iff xb^{-1} \in R_n(1_K)_L \iff (xb^{-1})^n = 1_K \iff x^n a^{-1} = 1_K \iff x^n = a \iff x \in R_n(a)_L.$$ 

Furthermore, the multiplication by $b$ is injective as $b \neq 0_K$.

By 3.5, the group $R_n(1_K)_L$ is cyclic.

5.0.2. Let $a \neq 0_K$. Then every splitting field $L$ of $t^n - a$ over $K$ contains a splitting field of $t^n - 1_K$ over $K$.

Proof. Let $M$ be a splitting field of $t^n - 1_K$ over $L$. We have to show that $M = L$. For any $b \in R_n(a)_L$ we have, by a twofold application of 5.0.1, $R_n(1_K)_Lb = R_n(a)_L = R_n(a)_M = R_n(1_K)_M b$. As $b \neq 0_K$, we conclude $R_n(1_K)_M = R_n(1_K)_L \subseteq L$, hence $M = L$.

5.1 Theorem (on splitting fields of pure polynomials). Let $K$ be a field, $a \in \hat{K}$, $K_n$ a splitting field of $t^n - 1_K$ over $K$, $b$ a root of $t^n - a$ in an extension field of $K_n$, $L := K_n[b]$, $m := |R_n(1_K)_L|$. Then

1. $R_n(1_K)_L = \langle w \rangle$, $K_n = K[w]$ for an appropriate $w \in R_n(1_K)_L$,

2. $L$ is a splitting field of $t^n - a$ over $K$,

3. $\text{Aut}_K K_n$ is isomorphic to a subgroup of $U(\mathbb{Z}/m\mathbb{Z}, \cdot)$,$^{78}$

4. $\text{Aut}_{K_n} L$ is isomorphic to a subgroup of $(R_n(1_K)_L, \cdot)$.

In particular, $\text{Aut}_K K_n$ and $\text{Aut}_{K_n} L$ are abelian (the latter even cyclic).

$^{78}$The unit group (=group of all invertible elements) of any monoid $(M, \cdot)$ is denoted by $U(M)$. If $(M, \cdot) = (\mathbb{Z}/m\mathbb{Z}, \cdot)$ we have $a + m\mathbb{Z} \in U(\mathbb{Z}/m\mathbb{Z}) \Leftrightarrow \gcd(a, m) = 1$. Hence $|U(\mathbb{Z}/m\mathbb{Z})| = \varphi(m)$.
Proof. (1), (2) follow from 5.0.1 and 5.0.2.

(3) Let \( w \) as in (1). If \( \alpha \in \text{Aut}_K K_n \), then \( \langle w \alpha \rangle = R_n(1_K)_L \alpha = R_n(1_K)_L \). Hence there exists \( j \in \mathbb{N} \), unique modulo \( m \), such that \( \gcd(j, m) = 1 \), \( w\alpha = w^j \). We put \( \overline{\alpha} := m\mathbb{Z} + j \).

If also \( \beta \in \text{Aut}_K K_n \), \( k \in \mathbb{N} \) such that \( w\beta = w^k \), it follows that \( w\alpha\beta = (w^j\beta = (w\beta)^j = w^{jk} \), hence

\[
\overline{\alpha\beta} = m\mathbb{Z} + jk = (m\mathbb{Z} + j)(m\mathbb{Z} + k) = \overline{\alpha}\overline{\beta}.
\]

This shows that \( f : \text{Aut}_K K_n \to U(\mathbb{Z}/m\mathbb{Z}) \), \( \alpha \mapsto \overline{\alpha} \), is a multiplicative homomorphism. If \( \alpha \in \ker f \) we have \( \overline{\alpha} = m\mathbb{Z} + 1 \), hence \( w\alpha = w \), \( \alpha = \text{id}_{K[w]} \). Therefore \( f \) is injective.

(4) If \( \alpha \in \text{Aut}_K L \) there exists a unique \( w_\alpha \in R_n(1_K)_L(\subseteq K_n) \) such that \( b\alpha = w_\alpha b \) as \( b\alpha \in R_n(a)_L = R_n(1_K)_L b \) by 5.0.1. If also \( \beta \in \text{Aut}_K L \),

\[
w_{\alpha\beta}b = b\alpha\beta = (w_\alpha b\beta = (w_\alpha\beta)(b\beta) = w_\alpha w_\beta b.
\]

Hence \( g : \text{Aut}_K L \to R_n(1_K)_L \), \( \alpha \mapsto w_\alpha \), is a homomorphism. If \( \alpha \in \ker g \) we have \( b\alpha = b \) and therefore \( \alpha = \text{id}_{K_n[b]} = \text{id}_L \). Thus \( f \) is injective.

5.2 Corollary. Let \( L \) be a splitting field of a pure polynomial \( \neq t^n \) of degree \( n \) over a field \( K \). Set

\[
N := \{ \alpha | \alpha \in \text{Aut}_K L, \alpha|_{R_n(1_K)_L} = \text{id}_{R_n(1_K)_L} \}
\]

Then \( N \subseteq \text{Aut}_K L \), \( N \) is cyclic and \( \langle \text{Aut}_K L \rangle/N \) abelian.

In particular, if \( R_n(1_K)_L \subseteq K \), then \( K = K_n \) and \( \text{Aut}_K L \) is cyclic.

Proof. \( N \) is the kernel of the restriction homomorphism \( \text{Aut}_K L \to \text{Aut}_K K_n \), \( \alpha \mapsto \alpha|_{K_n} \).

By the homomorphism theorem for groups and 5.1(3), \( (\text{Aut}_K L)/N \) is abelian. By definition of \( N \), \( K_n \) is centralized by \( N \), i.e., \( N \subseteq \text{Aut}_K L \). Hence \( N \) is cyclic by 5.1(4) and 3.6. If \( K = K_n \) it follows that \( \text{Aut}_K L = N \). \( \square \)

We take a closer look at the number \( m \) in 5.1: Certainly we have \( m \leq n \). When does equality hold? Using 1.8, we obtain

\[
m = n \iff t^n - 1_K \text{ has exactly } n \text{ zeros in } K_n
\]

\[
\iff t^n - 1_K \text{ and } (t^n - 1_K)' \text{ have no common zero in } K_n
\]

\[
\iff \forall x \in R_n(1_K)_{K_n} \ \ nx^{n-1} \neq 0_K \iff n \cdot 1_K \neq 0_K \iff \text{char } K \nmid n
\]

In particular, \( |R_n(1_K)_{K_n}| = n \) if \( \text{char } K = 0 \). By 4.23, the \( n \)-th cyclotomic polynomial \( \Phi_n \) is irreducible. Hence for any primitive \( n \)-th root of unity \( w \),

\[
|\text{Aut } \mathbb{Q}_n|_{4.23} = \dim_{\mathbb{Q}} \mathbb{Q}_n = \dim_{\mathbb{Q}} \mathbb{Q}[w] = \deg \Phi_n = \varphi(n) = |U(\mathbb{Z}/n\mathbb{Z}, \cdot)|
\]

Thus we have the following

Special case of 5.1(3) \( \quad \text{Aut } \mathbb{Q}_n \cong U(\mathbb{Z}/n\mathbb{Z}, \cdot) \)

We shall need the following

Special case of 2.7 Let \( L \) be a field, \( \alpha_1, \ldots, \alpha_n \) mutually distinct automorphisms of \( L \), \( c_1, \ldots, c_n \in L \). Then there exists an element \( b \in L \) such that \( c_1(b\alpha_1) + \cdots + c_n(b\alpha_n) \neq 0_L \).
5.3 Lemma (on Lagrange resolvents). Let \( n \in \mathbb{N} \) and \((K, L)\) be a Galois field extension of degree \( n \) with a cyclic Galois group, \( \alpha \) a generator of \( \text{Aut}_K L \). Suppose that \(|R_n(1_K)_K| = n\). For all \( w \in R_n(1_K)_K \), \( b \in L \) set
\[
 r(w, b) \coloneqq b + w b \alpha + w^2 b \alpha^2 + \cdots + w^{n-1} b \alpha^{n-1},
\]
(called the Lagrange resolvent for \( w, b \), with respect to \( \alpha \)). Then we have

1. \( r(w, b) \alpha = w^{-1} r(w, b) \),
2. \( r(w, b)^n \in K \),
3. \( \exists b \in L \) \( r(w, b) \neq 0_K \),
4. If \( w \) is primitive, then \( r(w, b) = 0_K \) or \( L = K[r(w, b)] \).

Proof. Let \( w \in R_n(1_K)_K \), \( b \in L \).

1. \( r(w, b) \alpha = b \alpha + w b \alpha^2 + \cdots + w^{n-1} b \alpha^{n-1} = w^{-1} r(w, b) \) as \( w \in K = \text{Fix}_L \alpha \).

2. \( r(w, b)^n \alpha = w^{-n} r(w, b)^n = r(w, b)^n \), hence \( r(w, b)^n \in \text{Fix}_L \alpha = K \).

3. follows from the special case of 2.7 as \( \alpha^0, \ldots, \alpha^{n-1} \) are mutually distinct.

4. Let \( w \) be primitive and suppose \( r(w, b) \neq 0_K \). Then
\[
 \forall i \in n-1 \quad r(w, b) \alpha^i = w^{-i} r(w, b) \neq r(w, b).
\]

Hence \( C_{\langle \alpha \rangle}(K[r(w, b)]) = \{ \text{id}_L \} \) so that \( K[r(w, b)] = L \) by 4.9. \( \square \)

5.4 Theorem. Let \( n \in \mathbb{N} \) and \((K, L)\) be a field extension of degree \( n \). Suppose that \(|R_n(1_K)_K| = n\). Then the following are equivalent:

1. \((K, L)\) is Galois and \( \text{Aut}_K L \) is cyclic.
2. \( L \) is a splitting field of a pure irreducible polynomial over \( K \).

Proof. (i)\( \Rightarrow \) (ii) Let \( w \in R_n(1_K)_K \) be primitive and, by 5.3(3), choose \( b \in L \) such that \( r(w, b) \neq 0_K \). Put \( a \coloneqq r(w, b)^n \). By 5.3(2), \( a \in K \), and by 5.3(4), \( L = K[r(w, b)] \).

By 5.3(1), \( L \) is a splitting field of \( t^n - a \). This polynomial is irreducible as otherwise \( \dim_K K[r(w, b)] < n \).

(ii)\( \Rightarrow \) (i) Let \( a \in K \) such that \( t^k - a \) is irreducible in \( K[t] \) for some \( k \in \mathbb{N} \), with splitting field \( L \). Let \( r \in L \) such that \( r^k = a \). Then \( k = \dim_K K[r] \dim_K L = n \), hence \( R_k(1_K)_K \subseteq R_n(1_K)_K \), implying \(|R_k(1_K)_K| = k \). Thus \( K[r] = L \) by 5.0.1, \( k = n \).

As \( (t^n - a)^i = n t^{n-1} \neq 0_K \), \( t^n - a \) is separable so that the extension \((K, L)\) is Galois. By 5.2, \( \text{Aut}_K L \) is cyclic. \( \square \)

If the degree of the extension is a prime, we conclude the following variation of 5.4:
5.5 Corollary. Let \( p \) be a prime and \((K, L)\) a field extension. Suppose that \(|R_p(1_K)_K| = p\). Then the following are equivalent:

(i) \((K, L)\) is Galois and \(|\text{Aut}_KL| = p\).

(ii) \(L\) is a splitting field of a pure irreducible polynomial of degree \( p \) over \(K\).

Proof. Observing that both (i) and (ii) imply \(\dim_K L = p\), we just have to apply 5.4. \(\square\)

5.6 Definition. A field extension \((K, L)\) is called a radical extension if there exists in \(L\) a radical \(r\) over \(K\) such that \(L = K[r]\). The smallest \(n \in \mathbb{N}\) such that \(r^n \in K\) is called the order of \(r\) modulo \(K\). An element \(b\) of an extension field of \(K\) is called expressible by radicals if there exists \(k \in \mathbb{N}_0\) and fields \(L_0, \ldots, L_k\) such that \(^{\text{79}}\) \(K[b] \leq L_k\) and

\begin{itemize}
  \item[(i)] \(K = L_0 \leq L_1 \leq \cdots \leq L_k\),
  \item[(ii)] \((L_{j-1}, L_j)\) is a radical extension for all \(j \in \mathbb{N}\).
\end{itemize}

In the following, assuming (i) and (ii), we consider a radical \(r_j \in L_j\) over \(L_{j-1}\) such that \(L_j = L_{j-1}[r_j]\) and let \(n\) be a positive integer which is divisible by the orders of the radicals \(r_j\) over \(L_{j-1}\). Then the extension \((L_{j-1}, L_j)\) is separable for all \(j \in \mathbb{N}\) so that 5.1(2) implies:

5.6.1. Suppose (i), (ii), \(|R_n(1_K)_K| = n\). Then \((L_{j-1}, L_j)\) is Galois for all \(j \in \mathbb{N}\). \(\square\)

5.6.2. Suppose an element \(b\) of an extension field of \(K\) is expressible by radicals. Let \(k\) be as in (i), \(n\) as before, and assume \(\text{char } K \nmid n\). Then there are fields \(\tilde{L}_0, \ldots, \tilde{L}_{k+1}\) such that \(K[b] \leq \tilde{L}_{k+1}\) and

\begin{itemize}
  \item[(i)] \(K = \tilde{L}_0 \leq \tilde{L}_1 \leq \cdots \leq \tilde{L}_{k+1}\),
  \item[(ii)] \((\tilde{L}_{j-1}, \tilde{L}_j)\) is Galois for all \(j \in \mathbb{N}\). \text{ Aut}_L \tilde{L}_1 abelian, \text{ Aut}_{\tilde{L}_i} \tilde{L}_{i+1} cyclic for all } i \in \mathbb{N}.
\end{itemize}

Proof. Let \(M\) be a splitting field of \(t^n - 1_K\) over \(L_k\), \(w \in M\) a generator of \(R_n(1_K)_M\). Set \(\tilde{L}_0 := K\), \(\tilde{L}_j := L_{j-1}[w]\) for all \(j \in k + 1\). Then \((K, \tilde{L}_1)\) is Galois with an abelian Galois group by 5.1(3), \((\tilde{L}_i, \tilde{L}_{i+1})\) is Galois with a cyclic Galois group by 5.2, for all \(i \in \mathbb{N}\). Clearly, \(K[b] \leq L_k \leq \tilde{L}_{k+1}\). \(\square\)

\(^{\text{79}}\) The radicals \( \not\equiv 0_K \) in \(L\) over \(K\) are exactly the elements of those cosets in \(\tilde{L}/K\) which are of finite order in this (multiplicative) factor group.

The order of a radical \(r \in \tilde{L}\) modulo \(K\) is the group-theoretic order of the coset \(Kr \in \tilde{L}/K\).

The set of all elements of finite order in an abelian group is a subgroup, called its torsion subgroup. Thus the set \(\hat{T}\) of all radicals \( \not\equiv 0_K \) in \(L\) over \(K\) is the complete anti-image (in \(L\)) of the torsion subgroup of \(\tilde{L}/K\).

\(^{\text{80}}\) Here “\(\leq\)” means “subfield of”.

92
This remark shows that the chain of fields in the definition of expressibility by radicals may be replaced\textsuperscript{81} by a chain of fields with much stronger properties if the hypothesis on the characteristic in 5.6.2 is satisfied. This is certainly the case if char $K = 0$ to which we will confine ourselves in the sequel.

A non-zero polynomial $f$ over $K$ is called \textbf{soluble by radicals over $K$} if each of its zeros (in some splitting field over $K$) is expressible by radicals. Note that the fields $L_i$ in (i) \textit{need not be subfields of the splitting field}! The definition does not require that there exist such a chain within a splitting field of $f$. For example, if $f$ is an irreducible cubic polynomial over $\mathbb{Q}$ with three real zeros, its splitting field in $\mathbb{C}$ is contained in $\mathbb{R}$. While Cardano’s formula shows that $f$ is soluble by radicals over $\mathbb{Q}$, it may be proved that it is not possible to express its zeros by \textit{real} radicals. A fortiori there is no chain of fields as required for the solubility by radicals \textit{within a splitting field of $f$}.\textsuperscript{82}

Over a subfield $K$ of $\mathbb{C}$, every polynomial of degree 2 is soluble by radicals – which is the main interpretation of the solution formula taught in mathematics lessons in class rooms. Also, every polynomial of degree 3 or 4 over $K$ is soluble by radicals which is the main message of Cardano’s formulas. We shall see, however, that there are polynomials of degree 5 over $\mathbb{Q}$ which are not soluble by radicals. With this revolutionary discovery in the first decades of the 19th century, a hunt for a solution formula (for higher degree equations) came to an end which had puzzled mathematicians for hundreds of years. It was a result “in the negative”, showing that a certain mathematical desire was unreachable, a proof of an impossibility of a certain task. The main error of the preceding times had probably been to not take seriously into consideration that \textit{there might not exist a solution} to the problem which was pursued. The certainly unexpected “scandal” was that this non-existence of a solution could even be \textit{proved}.

A culture of algebraic thinking and reasoning had its start here which formed the character of the whole discipline. Even the word “algebra” (of arabic origin) comes and is inseparable from the problem of solving equations. The methods and notions (like field, group, automorphism) which had the strength to settle the (negative) answer likewise opened an area of research which seems inexhaustible even today. A wealth of developments has certainly widely enriched and extended the discipline called “Algebra” these days. Still a true understanding of its character, its specific sense of beauty, its central structural elements, is unthinkable without a thorough study of the theory which gave birth to this cornerstone of modern mathematics.

\textbf{5.7 Lemma. Let $K$ be a field, char $K = 0$. A non-zero polynomial $f \in K[t]$ is soluble by radicals over $K$ if and only if there exist $m \in \mathbb{N}_0$ and fields $K_0, \ldots, K_m$ such that}

\begin{itemize}
  \item[(i)] $K = K_0 \leq K_1 \leq \cdots \leq K_m,$
  \item[(ii)] $(K_{j-1}, K_j)$ is Galois and $\text{Aut}_{K_{j-1}}K_j$ abelian for all $j \in \{1, \ldots, m\}$.
  \item[(iii)] $K_m$ contains a splitting field of $f$ over $K$.
\end{itemize}

\textsuperscript{81}Note that the extensions $(\tilde{L}_{j-1}, \tilde{L}_j)$ in (ii) are radical extensions, for $j \geq 2$ by 5.4.

\textsuperscript{82}The case of an irreducible cubic polynomial over $\mathbb{Q}$ with three real zeros is called \textit{casus irreducibilis}. 

93
Proof. “⇒” We proceed by induction on $\deg f$, the case $\deg f \leq 1$ being trivial. For the inductive step let $\deg f > 1$, $b$ a zero of $f$ in an extension field of $K$. By 5.6.2, there exists an $h \in \mathbb{N}_0$ and fields $K_0, \ldots, K_h$ such that $K = K_0 \leq \cdots \leq K_h$, $(K_{j-1}, K_j)$ is Galois, $\text{Aut}_{K_{j-1}}K_j$ is abelian for all $j \in [h]$, $K[b]$ is a subfield of $K_h$. As $f$ is soluble by radicals over $K$, $g := \frac{f}{t-b}$ is soluble by radicals over $K_h$. We have $\deg g < \deg f$, hence inductively there exist fields $K_{h+1}, \ldots, K_m$ such that $K_h \leq K_{h+1} \leq \cdots \leq K_m$, $(K_{j-1}, K_j)$ is Galois, $\text{Aut}_{K_{j-1}}K_j$ is abelian for all $j \in [m] \setminus [h]$ and $K_m$ contains a splitting field of $g$ over $K_h$. As $K[b] \leq K_h$, $K_m$ contains a splitting field of $f$ over $K$. With reference to the chain $K = K_0 \leq \cdots \leq K_h \leq K_{h+1} \leq \cdots \leq K_m$ the claim follows.

“⇐” From (i), (ii) we conclude that every intermediate field between $K_{j-1}$ and $K_j$ is Galois over $K_{j-1}$ because $\text{Aut}_{K_{j-1}}K_j$ is abelian. We choose a chain of maximal length of subgroups $H_i$ of $\text{Aut}_{K_{j-1}}K_j$: $\{\text{id}_{K_j}\} = H_0 < H_1 < \cdots < H_s = \text{Aut}_{K_{j-1}}K_j$. Then $H_i/H_{i-1}$ is a group of prime order, for all $i \in [s]$. By the fundamental theorem of Galois theory, there is a corresponding refinement of the intermediate fields. Therefore, we may assume w.l.o.g. the following stronger version of (ii):

(ii’) $(K_{j-1}, K_j)$ is Galois and $|\text{Aut}_{K_{j-1}}K_j|$ is a prime for all $j \in [m]$.

Let $n$ be a common multiple of the orders $|\text{Aut}_{K_{j-1}}K_j|$ ($j \in [h]$), $M$ a splitting field of $t^n - 1_K$ over $K_m$, $w \in M$ a primitive $n$-th root of unity, $L_0 := K$, $L_1 := K[w]$, $L_{j+1} := K_j[w]$ for all $j \in [m]$. By 5.5, $(L_{j-1}, L_j)$ is a radical extension (of degree 1 or of prime degree), for all $j \in [m+1]$. The field $L_{m+1}$ contains $K_m$, hence a splitting field of $f$ over $K$. Thus $f$ is soluble by radicals over $K$. □

Easy examples show that iterated Galois extensions need not result in a normal extension over the ground field; e.g., consider $(\mathbb{Q}, \mathbb{Q}[\sqrt{2}])$, $(\mathbb{Q}[\sqrt{2}], \mathbb{Q}[\sqrt[4]{2}])$. Clearly, the extension $(\mathbb{Q}, \mathbb{Q}[\sqrt{2}])$ is not normal. Therefore we cannot expect the field $K_m$ in 5.7(i) to be Galois over $K$, but it is certainly contained in some Galois extension field $M$ of $K$. 

![Diagram](image-url)
Let $Z$ be the splitting field of $f$ in $K_m$ over $K$, $D_j := Z \cap K_j$ for all $j \in m$. Writing $\Gamma$ for the Galois lattice anti-isomorphism of the field interval $[\bar{K}, M]$ onto $\mathcal{U}(\text{Aut}_K M)$, we have

$$D_j \Gamma = Z \Gamma \cdot K_j \Gamma \leq Z \Gamma \cdot K_{j-1} \Gamma = D_{j-1} \Gamma$$

as $K_j \Gamma \leq K_{j-1} \Gamma$, and by 3.9,

$$D_{j-1} \Gamma/D_j \Gamma \cong K_{j-1} \Gamma/(D_j \Gamma \cap K_{j-1} \Gamma) \cong (K_{j-1} \Gamma/K_j \Gamma)/(D_j \Gamma \cap K_{j-1} \Gamma/K_j \Gamma)$$

so that $D_{j-1} \Gamma/D_j \Gamma$, being an epimorphic image of an abelian group, is abelian.

**5.8 Theorem** (Galois 1829). Let $K$ be a field, $\text{char} \ K = 0$, $f \in K[t]$, $f \not\equiv 0_K$, $Z$ a splitting field of $f$ over $K$, $G := \text{Aut}_K Z$. The following are equivalent:

(i) $f$ is soluble by radicals over $K$,

(ii) There exist $m \in \mathbb{N}_0$ and subgroups $H_0, \ldots, H_m$ of $G$ such that

$$\begin{align*}
\{1_G\} &= H_m \leq H_{m-1} \leq \cdots \leq H_1 \leq H_0 = G \\
H_{j-1}/H_j &\text{ is abelian for all } j \in m
\end{align*}$$

Proof. (i)$\Rightarrow$(ii) Let $m, D_j, \Gamma$ be as introduced above. It suffices to prove the claim for the group $KT/Z\Gamma$ as it is isomorphic to $G$, by 4.16(3). We just have to put $H_j := D_j \Gamma/Z\Gamma$ for all $j \in m$. $H_0 := KT/Z\Gamma$. Then $H_{j-1}/H_j \cong D_{j-1} \Gamma/D_j \Gamma$ is abelian for all $j \in m$.

(ii)$\Rightarrow$(i) By considering a maximal refinement of the chain given in (*), we may w.l.o.g. assume that $|H_{j-1}/H_j|$ is a prime for all $j \in m$ (cf. the proof of the equivalence of (ii) and (ii') in 5.7). Let $\Phi : \mathcal{U}(G) \to [K, Z]$, $H \mapsto \text{Fix}_Z H$. Set $K_j := H_j \Phi$ for all $j \in m$, $n := |G|$. Let $L^*$ be a splitting-field of $t^{n-1} - 1_K$ over $Z$, $w \in L^*$ such that $\langle \omega \rangle = R_n(1_K)_{L^*}$, $L_0 := K$, $L_j := K_{j-1}[w]$ for all $j \in m+1$.

Then $L_1$ is a splitting field of $t^n - 1_K$ over $K$, $\text{Aut}_K L_1$ abelian by 5.1(3). By 5.5, $L_{j+1}$ is either a splitting field of a pure irreducible polynomial of prime degree (dividing $n$) over $L_j$, or $L_j = L_{j+1}$, for all $j \in m$. Hence $\text{Aut}_{L_{j-1}} L_j$ is abelian for all $j \in m+1$. Moreover, $Z \leq L_{m+1}$.

Proof follows from 5.7. \(\blacksquare\)

---

83Without this hypothesis it may be proved that (i)$\Rightarrow$(ii) holds if $\text{char} \ K$ does not divide the number $n$ as defined in 5.6.1, 5.6.2; (ii)$\Rightarrow$(i) holds if $\text{char} \ K \nmid |\text{Aut}_K Z|$. 

95
5.9 Definition. A group $G$ is called soluble if 5.8(ii) holds. Thus 5.8 may be reformulated in short as follows:

5.8' A non-zero polynomial over a field $K$ of characteristic 0 is soluble by radicals over $K$ if and only if the Galois group of its splitting field over $K$ is soluble. □

This result is the historic reason for the group-theoretical term “soluble”.

5.9.1. If $G$ is finite, the condition in $(\ast)$ that $H_{j-1}/H_j$ is abelian may be replaced by the condition that $|H_{j-1}/H_j|$ is a prime. □

5.9.2 Examples. (1) $S_3$ is soluble because $A_3 \unlhd S_3$, $|A_3| = 3$, $|S_3/A_3| = 2$.

(2) $S_4$ is soluble because $V := \left\{ \text{id}_4, (12)(34), (13)(24), (14)(23) \right\} \unlhd S_4$, $V \unlhd A_4 \unlhd S_4$, $V$ is abelian, $|A_4/V| = 3$, $|S_4/A_4| = 2$. (In fact, $S_4/V \cong S_3$.)

(3) For $n \geq 5$, $A_n$ is not soluble:

Proof. Let $X$ be the set of all 3-cycles on $\mathbb{Z}_n$. It is easily seen that $\langle X \rangle = A_n$.\(^{84}\) We show

$$N \unlhd A_n, \ A_n/N \text{ abelian} \Rightarrow N = A_n.$$

Let $i, j, k$ be mutually distinct elements of $\mathbb{Z}_n$. As $n \geq 5$, there exist $l, m \in \mathbb{Z}_n$ such that $i, j, k, l, m$ are mutually distinct. We have

$$(ikl)N \cdot (ijm)N = (ijm)N \cdot (ikl)N,$$

hence $(ikl)^{-1}(ijm)^{-1}(ikl)(ijm) \in N$ so that $N = A_n$. A fortiori, $A_n$ is not soluble.\(^{85}\)

From the following property of solubility we may now conclude that $S_n$ is not soluble for $n \geq 5$:

5.9.3. $G$ soluble, $U \leq G \Rightarrow U$ soluble.

Proof. Let $H_0, \ldots, H_m \leq G$ such that $(\ast)$ in 5.8(ii) is satisfied. Set $U_j := U \cap H_j$ for all $j$. Then

$$\{1_U\} = U_m \unlhd U_{m-1} \unlhd \cdots \unlhd U_0 = U$$

and $U_{j-1}/U_j \cong U_{j-1}H_j/H_j \leq H_{j-1}/H_j$ is abelian. □

5.9.4. $G$ soluble, $N \unlhd G \Rightarrow G/N$ soluble.

\(^{84}\) $A_n$, defined as the kernel of the sign homomorphism, is generated by all permutations which are products of two transpositions. But $(ij)(kl) = (ijk)(kil)$ and $(ij)(ik) = (ijk)$ where $i, j, k, l$ are mutually distinct elements of $\mathbb{Z}_n$.

\(^{85}\) $A_n$ is in fact simple for $n \geq 5$. For a proof in the case $n = 5$, see 8.1.4.
Let \( n \) be a positive integer. Let \( H_0, \ldots, H_m \leq G \) as before. Set \( V_j := NH_j/N \) for all \( j \). Then \( V_j \leq G/N \),

\[
\{1_{G/N}\} = V_m \leq V_{m-1} \leq \cdots \leq V_0 = G/N,
\]

and \( V_{j-1}/V_j \cong H_{j-1}/(NH_j \cap H_{j-1}) \) is isomorphic to a factor group of \( H_{j-1}/H_j \), hence abelian. \( \square \)

5.9.5. \( N \trianglelefteq G \) and \( N, G/N \) soluble \( \Rightarrow \) \( G \) soluble.

Proof. Let \( H_0, \ldots, H_m \leq N \) such that \( \{1_G\} = H_m \trianglelefteq \cdots \trianglelefteq H_1 \trianglelefteq H_0 = N, H_{j-1}/H_j \) abelian for all \( j \). Similarly \( \mathcal{V}_0, \ldots, \mathcal{V}_k \leq G/N \) such that \( \{1_{G/N}\} = \mathcal{V}_k \trianglelefteq \mathcal{V}_{k-1} \trianglelefteq \cdots \trianglelefteq \mathcal{V}_0 = G/N \), \( \mathcal{V}_{j-1}/\mathcal{V}_j \) abelian for all \( j \). Let \( U_0, \ldots, U_k \leq G \) such that \( N = U_1 \), \( U_i/N = \mathcal{V}_i \) for all \( i \geq 0 \). Then \( U_i \trianglelefteq U_{i-1} \) and \( U_{i-1}/U_i \cong \mathcal{V}_{i-1}/\mathcal{V}_i \) is abelian, for all \( i \). Now it suffices to consider the chain \( \{1_G\} = H_m \trianglelefteq \cdots \trianglelefteq H_0 = N = U_k \trianglelefteq \cdots \trianglelefteq U_0 = G \) and to observe that all successive quotients from this chain are abelian. \( \square \)

A general solution formula (by radicals) for polynomials of a given degree \( n \) would be a formula for the solutions in terms of (radical expressions) of its coefficients, following the model of the solution formula for quadratic equations (Cardano’s formulas in the case of equations of degree 3 or 4, resp.). More precisely, the coefficients are variables in this context. The question of a solution formula for an arbitrary (normed) polynomial of degree \( n \) over a field \( K \) thus leads to considering a polynomial over the polynomial ring in \( n \) variables over \( K \). Let \( s_j \) be the \( j \)-th elementary symmetric polynomial in \( K[t_1, \ldots, t_n] \), i.e., \( s_0 := 1_K \),

\[
\forall j \in \mathfrak{m} \quad s_j := \sum_{1 \leq i_1 < \cdots < i_j \leq n} t_{i_1} \cdots t_{i_j}.
\]

Recall Viète’s theorem: Let \( c_0, \ldots, c_n, b_1, \ldots, b_n \in K \), \( \sum_{j=0}^n c_j t^{n-j} = (t - b_1) \cdots (t - b_n) \). Then \( c_0 = 1_K, c_j = (-1)^j s_j(b_1, \ldots, b_n) \) for all \( j \).

Furthermore, \( K[t_1, \ldots, t_n] \cong K[s_1, \ldots, s_n] \) by the main theorem on symmetric polynomials.

5.10 Theorem. Let \( K \) be a field, \( n \in \mathbb{N} \). Let \( L \) be a quotient field of the polynomial ring in \( n \) variables \( K[u_1, \ldots, u_n] \), \( F_n := t^n + u_1 t^{n-1} + \cdots + u_{n-1} t + u_n \in L[t] \), \( M \) a splitting field of \( F_n \) over \( L \). Then the extension \( (L,M) \) is Galois and \( \text{Aut}_L M \cong S_n \).

Corollary (Abel 1826). If \( \text{char} K = 0 \), then \( F_n \) is soluble by radicals if and only if \( n \leq 4 \).

Proof. Let \( r_1, \ldots, r_n \in M \) such that \( F_n = (t - r_1) \cdots (t - r_n) \). By Viète’s theorem, \( u_k = (-1)^k s_k(r_1, \ldots, r_n) \) for all \( k \in \mathfrak{m} \cup \{0\} \). Hence

\[
M = L[r_1, \ldots, r_n] = K(u_1, \ldots, u_n, r_1, \ldots, r_n) = K(r_1, \ldots, r_n).
\]
The mapping \( \{u_1, \ldots, u_n\} \rightarrow K[s_1, \ldots, s_n], \ u_j \mapsto (-1)^j s_j \), extends uniquely to a \( K \)-algebra isomorphism \( \psi \) of \( K[u_1, \ldots, u_n] \) onto \( K[s_1, \ldots, s_n] \). Let \( \overline{\psi} \) the canonical extension of \( \psi \) to the polynomial algebra \( (K[u_1, \ldots, u_n])[t] \). Then \( F_n \overline{\psi} = \sum_{j=0}^n (-1)^j s_j t^n - t = (t - t_1) \cdots (t - t_n) \). By 4.3(2), \( \psi \) extends to an isomorphism of the splitting field \( M \) of \( F_n \) onto the splitting field \( K(t_1, \ldots, t_n) \) of \( F_n \overline{\psi} \). Every permutation \( \pi \) of \( \{r_1, \ldots, r_n\} \) extends to an automorphism of \( M \). Let \( \pi \) such that the elements 1, \( \ldots \), \( n \) are mutually distinct is called a \( \pi \)-group contains an element \( \pi \) of order \( n \). As \( \pi \) is indeed generated by \( \pi \) and a transposition, it follows that \( \text{Aut}_L M \cong S_n \) extends uniquely to a \( \overline{S_n} \) is soluble by radicals over \( L \) if \( \text{Aut}_L M \) is soluble \( \iff S_n \) is soluble \( \iff n \leq 4 \) making use of 5.9.2, 5.9.3 in the last step.

Let \( f \) be an irreducible polynomial of prime degree \( p \) over \( \mathbb{Q} \). If \( f \) has exactly \( p - 2 \) real zeros, the two non-real zeros in \( \mathbb{C} \) must necessarily be interchanged under complex conjugation \( \kappa \). Hence the automorphism given by restriction of \( \kappa \) to the splitting field \( L \) of \( f \) induces a transposition \( \tau \) on \( \mathbb{Z}_L(f) \). Furthermore, the hypotheses on \( f \) imply that \( p \mid \dim_{\mathbb{Q}} L = |\text{Aut} L| = |G| \) where \( G \) is the group of the permutations on the set \( \mathbb{Z}_L(f) \) induced by \( \text{Aut} L \). A subgroup of \( S_p \) the order of which is divisible by \( p \) and which contains a transposition equals \( S_p \), hence \( \text{Aut} L \cong S_p \). It follows that \( f \) is not soluble by radicals if \( p \geq 5 \). For example, the polynomial \( f = 2t^5 - 10t + 5 \) is irreducible in \( \mathbb{Q}[t] \) by Eisenstein’s theorem, and has exactly three real zeros. The Galois group of its splitting field over \( \mathbb{Q} \) is therefore isomorphic to \( S_5 \), hence not soluble by what we have observed in 5.9.2(3). Over \( \mathbb{Q} \), its zeros are therefore not expressible by radicals.

A completely different and much more shallow question is how the zeros may be approximated. If the aim is just to visualize the real polynomial function

\[
 f : \mathbb{R} \to \mathbb{R}, \quad x \mapsto 2x^5 - 10x + 5,
\]

simple applications of Newton’s method quickly reveal the decimals of the three real zeros, with an error as small as desired \((-1, 6005 \ldots, 0, 5066 \ldots, 1, 3289 \ldots)\). This (and a routine calculation of the extremal points \((-1, 13\), \((1, -3)\) and the point of inflection \((0, 5)\)) is sufficient to sketch the curve:

By the universal property of the polynomial ring, \( \pi \) first extends to an endomorphism of \( K[r_1, \ldots, r_n] \). Then it suffices to apply the general observation that any isomorphism between integral domains \( R, S \) extends to an isomorphism between their quotient fields: If \( \varphi \) is an isomorphism of \( R \) onto \( S \), the assignment of \( \varphi \) to \( \overline{\varphi} \) is well-defined and defines an isomorphism as requested.

A proof of this statement is simple with the aid of Sylow’s theorem (8.4(2)): It implies that the subgroup contains an element \( \pi \) of order \( p \). As \( p \) is a prime, the elements \( 1, \pi, \pi^2, \ldots, \pi^{p-1} \) are mutually distinct. For arbitrary \( n \in \mathbb{N} \), a permutation \( \pi \in S_n \) such that the elements \( 1, \pi, \pi^2, \ldots, \pi^{n-1} \) are mutually distinct is called a cycle of length \( n \), and a simple induction argument shows that such a cycle and the transposition interchanging the two elements 1 and \( \pi \) generate the whole group \( S_n \). This general remark may be applied to the special case of \( n = p \) to see that \( S_p \) is indeed generated by any subset containing an element of order \( p \) and a transposition. It follows that \( \text{Aut} L \cong S_p \)
Using the term „number“ for an element of $\mathbb{C}$, we have a first and fundamental distinction between *algebraic* and *transcendental* numbers. The set $\mathbb{A}$ of all algebraic numbers is denumerable and a subfield of $\mathbb{C}$ while its complement, the set of all transcendental numbers, is non-denumerable. Highly non-trivial results proved in the last decade of the 19th century are that the Eulerian number $e$ and the circle number $\pi$ are transcendental. Classical field theory as developed in the preceding chapters, however, deals with subfields of $\mathbb{A}$, frequently even of finite dimension over their prime field $\mathbb{Q}$ and then called *algebraic number fields*. The set of all numbers which are expressible over $\mathbb{Q}$ by radicals is a subfield of $\mathbb{A}$ for which we write $\mathbb{A}_{rad}$. The results of this chapter show that $\mathbb{A}_{rad} \neq \mathbb{A}$. The zeros of $2t^5 - 10t + 5$ are examples of elements of $\mathbb{A} \setminus \mathbb{A}_{rad}$. The picture on the right gives a sketch of the *edifice of numbers*, including a few typical representatives for each of its parts – except for $\mathbb{A} \setminus \mathbb{A}_{rad}$. There is no well-known standard example and no familiar notation for numbers which would be directly associated with this intriguing set (like the root notation for real elements of $\mathbb{A}_{rad}$ and the fractional notation for elements of $\mathbb{Q}$).

Note that $\mathbb{Q}$ arises as the field of fractions of the integral domain $\mathbb{Z}$. Furthermore, the additive group $(\mathbb{Z}, +)$ is the difference group of the monoid $(\mathbb{N}_0, +)$.

---

88 This is a counterpart of the general notion of *quotient field of an integral domain*, proved similarly
Classical Algebra thus helps a lot to enlighten the fascinating yet mysterious world of numbers. In particular, it gives an idea of its subtlety and shows how small is the share of all (even of all algebraic) numbers which is given by real square roots. Who would expect how dramatic is the first step out of the safe harbour of $\mathbb{Q}$ when it is revealed in a school lesson that the length of the diagonal of the unit square cannot be expressed as a quotient of natural numbers? Who would expect which monumental area opens up with the school lesson in which solving a quadratic equation is reduced to determining some real square root? A thorough study of Algebra is obviously indispensable for a didactically adequate and righteous teaching of these crucial contents – as long as we share the conviction that pupils deserve this.

---

by means of the extension principle: Every commutative monoid in which the law of cancellation holds (i.e., $a + b = a + c \Rightarrow b = c$ for arbitrary elements $a, b, c$) is contained in an abelian group the elements of which may be written as differences of the monoid given. This group is uniquely determined up to isomorphism and called the difference group of the monoid.
6 Representations

Let $K$ be a commutative unitary ring, $V$ a $K$-space. Then
\[
\text{End}_K V := \{ \alpha | \alpha \in \text{End}(V, +), \forall c \in K \forall v \in V \quad (cv)\alpha = c(v\alpha) \}
\]
is a $K$-space, even an associative unitary $K$-algebra. If $V$ has a $K$-basis of $n$ elements, we have a $K$-algebra isomorphism $\text{End}_K V \cong K^{n \times n}$, by Linear Algebra.

6.1 Definition. Let $K$ be a commutative unitary ring and $A$ a $K$-space. A $K$-space representation of $A$ is a $K$-linear mapping $\delta : A \to \text{End}_K V$ where $V$ is a $K$-space. The couple $(V, \delta)$ is then called an $A$-module. 89 For a subset $T$ of $V$ we write $T \subseteq \delta V$ if $T(a\delta) \subseteq T$ for all $a \in A$, and $T$ is then called invariant with respect to $\delta$. If $T$ is an invariant $K$-subspace of $V$, $T$ is called a submodule of $V$, and then we write $T \leq A V$. In most cases the discussion refers to a single action $\delta$ of $A$ on $V$, and then it is customary to write $T \leq V$ instead of $T \leq \delta V$. We make a number of trivial but useful remarks:

6.1.1. If $n \in \mathbb{N}$ and $V_1, \ldots, V_n$ are $A$-modules via $\delta_1, \ldots, \delta_n$ resp., then $V_1 \times \cdots \times V_n$ is an $A$-module via “componentwise action”:

\[
\delta^* : A \to \text{End}_K(V_1 \times \cdots \times V_n), \ a \mapsto \begin{bmatrix}
V_1 \times \cdots \times V_n & \to & V_1 \times \cdots \times V_n \\
(v_1, \ldots, v_n) & \mapsto & (v_1(a\delta), \ldots, v_n(a\delta))
\end{bmatrix}
\]

In particular, for any $A$-module $V$ the space $V^n$ is made into an $A$-module by componentwise action. \[\square\]

6.1.2. Let $B$ be a $K$-space and $V$ a $B$-module via $\delta$, $\alpha$ a $K$-linear mapping of $A$ into $B$. Then $V$ is an $A$-module via $\alpha \delta$. \[\square\]

For a $K$-algebra, the distributive laws ensure that we have two natural ways to view the underlying additive $K$-space as a $K$-space module:

6.1.3. Let $B$ be a $K$-algebra, $\lambda : B \to \text{End}_K(B, +), \ b \mapsto [B \to B, x \mapsto bx]$, $\rho : B \to \text{End}_K(B, +), \ b \mapsto [B \to B, x \mapsto xb]$. Then $\lambda$, $\rho$ are $K$-space representations of $B$. \[\square\]

89In less formal language: “$A$ acts on $V$ via $\delta$”, or “$V$ is an $A$-module via $\delta$”. If there is no doubt about the significance of $\delta$, it is simply suppressed so that one says: “$A$ acts on $V$”, “$V$ is an $A$-module” resp. The idea of a representation is that, by $\delta$, any element of $A$ is “made into” an endomorphism of a space, and that this “transformation” happens in a nicely controlled way, i.e., linearly. While the behaviour of a space $V$ under a single endomorphism is an important chapter of a course on Linear Algebra, Representation Theory studies the behaviour of $V$ under a space of endomorphisms of $V$, i.e., the simultaneous actions of its elements on $V$ are considered.
The representation \( \lambda \) in 6.1.3 is called the \textit{left regular}, \( \rho \) the \textit{right regular} representation of the algebra \( B \). We simply have \( x(b\lambda) = bx, \ x(bp) = xb \) for all \( x, b \in B \). Applying this to the \( K \)-algebra \( B \) of all endomorphisms of a \( K \)-space \( V \), we obtain:

\textbf{6.1.4.} Let \((V, \delta)\) be an \( A \)-module. Then \( \text{End}_K V \) is an \( A \)-module in two ways: via \( \delta \lambda \) and via \( \delta \rho \). \( \square \)

Here we have \( \varphi(\delta \lambda) = (a \delta) \varphi, \ \varphi(\delta \rho) = \varphi(\delta) \) for all \( \varphi \in \text{End}_K V, \ a \in A \), the right hand sides being compositions of endomorphisms of \( V \).

\textbf{6.1.5.} \textit{Intersections and sums of submodules of an \( A \)-module \( V \) are submodules of \( V \).}\( \square \)

Here the second assertion is a consequence of the first as, for any set \( \mathcal{X} \) of submodules of an \( A \)-module \( V \), \( \sum \mathcal{X} = \bigcap \{ U \mid U \leq V, \ \forall T \in \mathcal{X} \ T \subseteq U \} \). An element \( v \in V \) is contained in \( \sum \mathcal{X} \) if and only if there exists a finite set \( \{(U_1, u_1), \ldots, (U_n, u_n)\} \) of couples \( (U_j, u_j) \) with \( u_j \in U_j \in \mathcal{X}, u = u_1 + \cdots + u_n \). We may obviously assume that \( u_1, \ldots, u_n \) are mutually distinct (moreover, that \( u_1, \ldots, u_n \neq 0_V \)). Then, by adding formally the couples \( (U, 0_V) \) where \( U \in \mathcal{X}, U \neq U_j \) for all \( j \in \mathbb{N} \), the given set may be viewed as a "function of choice with finite support" for \( \mathcal{X} \). The sum \( \sum \mathcal{X} \) is called \textit{direct} if for every \( v \in \sum \mathcal{X} \) there exists a \textit{unique} function of choice with finite support for \( \mathcal{X} \) such that \( v \) is the sum of its nonzero images. If the sum over \( \mathcal{X} \) is direct and amounts to all of \( V \), we write \( \bigoplus \mathcal{X} = V \), and \( \mathcal{X} \) is then called a \textit{direct \( A \)-decomposition} of \( V \). More generally, these considerations apply in the case of abelian groups (instead of \( A \)-modules), and it is easily seen that the sum over \( \mathcal{X} \) is direct if and only if

\[ \forall U \in \mathcal{X} \ U \cap \sum (\mathcal{X} \setminus \{U\}) = 0. \]

By 6.1.5, for any subset \( T \) of an \( A \)-module \( V \) there is a smallest submodule of \( V \) containing \( T \), given by

\[ \bigcap_{T \subseteq U \leq V} U =: \langle T \rangle_A, \]

called the submodule \textit{generated by} \( T \). If \( \langle T \rangle_A = V \), \( T \) is called a \textit{system} (or a \textit{set} of \textit{generators}) of \( V \), and \( V \) is said to be generated as an \( A \)-module by \( T \).

A submodule \( U \) of \( V \) is called a \textit{direct \( A \)-summand} of \( V \) if there is a submodule \( W \) such that \( V = U \oplus W \). We then call \( W \) a \textit{direct co-summand} of \( U \) in \( V \) and write \( V = U \oplus_A W \).

The submodules \( 0 \) and \( V \) are always \( A \)-summands of \( V \) (with co-summand \( V \), \( 0 \) resp.). An \( A \)-module \( V \) is called \textit{directly \( A \)-indecomposable} if \( V \neq 0 \) and \( 0, V \) are the only direct \( A \)-summands of \( V \).

The basic special case of these notions is \( A = K \), \( K \) a field. For all \( c \in K \), set \( c\delta : V \to V, v \mapsto cv \). The notion of "\( A \)-module" then just means "\( K \)-vector space". Thus the submodules are exactly the \( K \)-subspaces. In this case, every submodule is a direct \( A \)-summand as any linearly independent set of vectors (hence any basis of a \( K \)-subspace) may be extended to a \( K \)-basis of \( V \). It follows that a submodule \( U \) is \( A \)-indecomposable if
and only if \( \text{dim}_K U = 1 \). If \( B \) is a \( K \)-basis of \( V \), \( \{ \langle v \rangle_K \mid v \in B \} \) is a direct \( A \)-decomposition of \( V \). Each \( \langle v \rangle_K \) is not only directly \( A \)-indecomposable but does not even contain a submodule \( \neq \{0_A\} \), \( \langle v \rangle_K \), an obviously in general much stronger property.

In general, such a simple behaviour cannot be expected any longer. An \( A \)-module \( V \) is called irreducible\(^90\) if \( V \neq 0 \) and \( 0, V \) are its only submodules.

6.1.6. Every irreducible \( A \)-module is directly \( A \)-indecomposable. □

A trivial example to see that the converse of 6.1.6 is indeed hopeless is the following: \( A = K = Z = V \). As a \( Z \)-module, \( Z \) is directly indecomposable, but not irreducible: The submodules are exactly the subgroups of \( (Z, +) \). There is no irreducible submodule in \( Z \): On the contrary, every subgroup \( \neq \{0\} \) is of the form \( kZ \ (k \in \mathbb{N}) \), hence again isomorphic to \( (Z, +) \).

An \( A \)-module \( V \) is called completely reducible\(^91\) if there exists a set \( X \) of irreducible submodules of \( V \) such that \( \sum X = V \).\(^92\)

6.1.7. Let \( V \) be an \( A \)-module, \( U \leq A V \leq \bar{A} V \). If \( U \) is a direct \( A \)-summand of \( V \), then also of \( V' \).

Proof. Let \( W \leq A V \) such that \( V = U \oplus W \). We set \( W' := W \cap V' \) and obtain \( U \cap W' \leq A U \cap W = 0, U + W' = U + (W \cap V') = (U + W) \cap V' = V' \), by Dedekind’s law. □

Let us return to the example of the two regular representations of a \( K \)-Algebra \( B \), and let \( T \leq \bar{B} \). Then

\[
\begin{align*}
T \leq \bar{B} & \iff \forall b \in \bar{B} \ bT \subseteq T, \\
T \leq \bar{B} & \iff \forall b \in B \ TB \subseteq T.
\end{align*}
\]

In the first case, \( T \) is called a left ideal of \( B \), in the second case a right ideal of \( B \). Left ideals are exactly the submodules of \( (B, \lambda) \), right ideals the submodules of \( (B, \rho) \). If \( S, T \) are \( K \)-subspaces of \( B \), we write \( ST \) for the smallest \( K \)-subspace containing all elements \( uv \) where \( u \in S, v \in T \).\(^93\) Then \( T \leq \bar{B} \) (\( T \leq \bar{B} \) resp.) if and only if \( BT \subseteq T \) (\( TB \subseteq T \) resp.). A subset \( T \) of \( B \) is called a one-sided ideal of \( B \) if \( T \leq B \) or \( T \leq B \).

6.1.8. Let \( R \) be a right ideal, \( S \) a left ideal of \( B \). Then \( RS \subseteq R \cap S \). In particular, if \( R \cap S = \{0_B\} \), then \( RS = \{0_B\} \). □

A right ideal \( T \) of \( B \) is called minimal if \( T \neq \{0_B\} \) and there is no right ideal \( S \) of \( B \) such that \( \{0_B\} < S < T \) (analogously for left ideals). This means:

\(^{90}\)alternatively, also called simple
\(^{91}\)alternatively, also called semisimple
\(^{92}\)If \( V = 0 \), \( V \) is completely reducible: Choose \( X = \emptyset \).
\(^{93}\)This is the additive closure of the set of these elements.
6.1.9. The minimal right (left resp.) ideals of $B$ are exactly the irreducible submodules of the $B$-module $(B, \rho)$ ($(B, \lambda)$ resp.).

The ideals of $B$ (already known as the kernels of the algebra homomorphisms of $B$) clearly are the subsets of $B$ which are both a right and a left ideal of $B$. For any $K$-subspace $T$ of $B$, $\sum_{k=0}^{\infty}((TB)B)\cdots B$ is the smallest right ideal, $\sum_{k=0}^{\infty}B(\cdots(B(B(T))\cdots)$ the smallest left ideal of $B$ containing $T$.\footnote{If $B$ is associative and unitary, everything is much simpler as then those complicated sums reduce to $TB$, $BT$ resp., and the smallest ideal containing $T$ is $BTB$.}

For any $X \subseteq B$ we write $X^{\leq B}$ for the smallest ideal of $B$ containing $X$ and set $x^{\leq B} := \{x\}^{\leq B}$. We write $J \leq B$ if $J$ is a minimal ideal of $B$, i.e. if $\{0_B\} < J \leq B$ and there is no ideal $I$ of $B$ such that $\{0_B\} < I < J$. Clearly, if $I, J$ are distinct minimal ideals, $IJ = \{0_B\}$ by 6.1.8.

If we consider the new multiplication $\cdot$ on $B$ given by

$$\forall x, y \in B \quad x \cdot y := yx,$$

then $B$ is an algebra with respect to this multiplication again. We write $B^-$ for it\footnote{This algebra is sometimes called the opposite algebra of $B$ and then denoted by $B^{\text{opp}}$ or $B^{\text{opp}}$.} to distinguish it from the given algebra $B$. Clearly, the left ideals of $B^-$ are the right ideals of $B$, and the right ideals of $B^-$ are the left ideals of $B$.

6.2 Lemma. Let $K$ be a field, $A$ a $K$-vector space, $V$ a finite dimensional $A$-module, $\mathcal{X}$ a set of irreducible submodules of $V$ such that $\sum_{A} \mathcal{X} = V$. Then for any $U \leq V$ there exists a subset $\mathcal{T}$ of $\mathcal{X}$ such that

(i) $\sum \mathcal{T}$ is a direct sum,

(ii) $V = U \oplus \sum \mathcal{T}$.

Proof. The set of all subsets $\mathcal{T}$ of $\mathcal{X}$ satisfying the two conditions

$$(i), \quad U \cap \sum \mathcal{T} = 0$$

is non-empty as shown by the possible choice $\mathcal{T} = \emptyset$, and all those subsets $\mathcal{T}$ of $\mathcal{X}$ have the property that $|\mathcal{T}| \leq \dim_K V$. We now consider such a subset $\mathcal{T}$ of $\mathcal{X}$ of maximal cardinality. Our aim is to show that $U + \sum \mathcal{T} = V$.

Assume $U + \sum \mathcal{T} \neq V$. Then there exists $W \in \mathcal{X}$ such that $W \nsubseteq U + \sum \mathcal{T}$. It follows that $W \neq W \cap (U + \sum \mathcal{T}) \leq W$, hence $W \cap (U + \sum \mathcal{T}) = 0$, $W$ being irreducible.

We claim that $U \cap (W + \sum \mathcal{T}) = 0$: Let $u \in U$, $w \in W$, $s \in \sum \mathcal{T}$ such that $u = w + s$. Then $w = u - s \in W \cap (U + \sum \mathcal{T})$, hence $u - s = 0_V$, i.e., $u = s \in U \cap \sum \mathcal{T} = 0$.

It follows that the sum over $\mathcal{T}' := \mathcal{T} \cup \{W\}$ is a direct sum and $U \cap \sum \mathcal{T}' = 0$, contradicting the choice of $\mathcal{T}$ as $|\mathcal{T}'| = |\mathcal{T}| + 1$. \qed
6.3 Theorem. Let \( K \) be a field, \( A \) a \( K \)-vector space, \( V \) a finite-dimensional \( A \)-module. The following are equivalent:

(i) \( V \) is completely reducible.

(ii) There exists a direct decomposition of \( V \) into irreducible submodules.

(iii) Every submodule of \( V \) is a direct \( A \)-summand of \( V \).

Corollary If \( V \) is completely reducible then every submodule of \( V \) is completely reducible.

Proof. The implication (ii)\( \Rightarrow \) (i) is trivial, while its converse follows from 6.2 (with \( U := 0 \)). The implication (i)\( \Rightarrow \) (iii) is again a consequence of 6.2. We now show (iii)\( \Rightarrow \) (i) by induction on \( \dim K V \), observing first that condition (iii) is hereditary by every submodule of \( V \), by 6.1.7. Suppose (iii) and assume inductively that every proper submodule of \( V \) is completely reducible. If \( V \) is not irreducible, \( 0 < U < V \), then there exists a submodule \( W \) such that \( V = U \oplus W \). Then \( U, W < V \), hence \( U, W \) are completely reducible. It follows that \( V \) is completely reducible. If \( V \) is irreducible, (i) holds trivially.

The observation on the heredity of (iii) implies the Corollary.

If \( K \) is a commutative unitary ring, \( V \) a \( K \)-space, \( v \in V \), \( \alpha \in \text{End}_K V \) and \( U \subseteq K V \) such that \( U \alpha \subseteq U \), we write \( \alpha \mathcal{V}/U \) for the \( K \)-endomorphism of \( V/U \) induced by \( \alpha \).

6.4 Definition. Let \( K \) be a commutative unitary ring, \( A \) a \( K \)-space, \( (V, \delta) \) an \( A \)-module, \( U \subseteq A V \). The mapping \( \delta_{V/U} : A \to \text{End}_K V/U, a \mapsto (a\delta)_{V/U} \), is obviously \( K \)-linear, hence makes \( V/U \) into an \( A \)-module, called the factor module of \( V \) by \( U \). Explicitly,

\[
\forall a \in A \quad a(\delta_{V/U}) = (a\delta)_{V/U}.
\]

Let \( (V', \delta') \) be a further \( A \)-module. A \( K \)-linear mapping \( \varphi : V \to V' \) with the property

\[
\forall a \in A \quad \varphi(a\delta') = (a\delta)\varphi
\]

is called an \( A \)-module homomorphism.

6.4.1 Examples. (1) Let \( U \subseteq A V \) and \( \varphi \) be the canonical vector space epimorphism \( V \to V/U \), \( v \mapsto U + v \). Then \( \varphi \) is an \( A \)-module homomorphism:

\[
\forall v \in V, \forall a \in A \quad (v\varphi)(a\delta_{V/U}) = (U + v)(a\delta)_{V/U} = U + v(a\delta) = (v(a\delta))\varphi.
\]

96The image of \( U + v \) under \( \alpha \) is completely contained in the coset \( U + va \), hence the mapping \( U + v \mapsto U + va \) is well-defined. Obviously it is \( K \)-linear.
(2) Let \( U, W \leq V \) such that \( V = U \oplus W \). For all \( v \in V \) write
\[
u_v \text{ for the (uniquely determined) element of } U \text{ for which there exists an element } w \in W \text{ such that } v = u_v + w. \]
The mapping
\[
\pi : V \to U, \ v \mapsto u_v,
\]
is called the projection of \( V \) onto \( U \) with respect to the direct decomposition \( \{U, W\} \).
It is an \( A \)-module homomorphism: If \( v, v' \in V \), \( v = u_v + w, \ v' = u_{v'} + w' \) where \( w, w' \in W \), we have for all \( c \in K \)
\[
v + cw' = u_v + w + cu_{v'} + w' = (u_v + cu_{v'}) + (w + cw'),
\]
hence, by uniqueness, \( u_{v+cw'} = u_v + cu_{v'} \). Therefore \( \pi \) is \( K \)-linear. Furthermore,
\[
\forall v \in V \forall a \in A, \ (v(a\delta))\pi = ((u_v + w)a\delta)\pi = (\underbrace{u_v(a\delta)}_{\in U} + \underbrace{w(a\delta)}_{\in W})\pi
\]
\[
= u_v(a\delta) = (v\pi)(a\delta).
\]

In both examples, the \( A \)-module homomorphism was surjective. As usual, a surjective (injective, bijective resp.) module homomorphism is called a module epimorphism (monomorphism, isomorphism). We write \((V, \delta) \cong (V', \delta')\) if there exists an \( A \)-module isomorphism of \((V, \delta)\) onto \((V', \delta')\). In that case, the modules \((V, \delta), (V', \delta')\) are called isomorphic. As usual, the symbols for the actions of \( A \) are omitted if no confusion will occur. It is then customary to write \( V \cong V' \). An \( A \)-module homomorphism of \( V \) into \( V \) is called an \( A \)-module endomorphism. An \( A \)-module automorphism is a bijective \( A \)-module endomorphism. The set of all \( A \)-module endomorphisms of \((V, \delta)\) is denoted by \( \text{End}(V, \delta) \) or by \( \text{End}_AV \) if there is no need for an explicit reference to \( \delta \). We write \( \text{Aut}_AV \) for the group of all module automorphisms. Clearly, \( \text{End}_AV \) is an associative \( K \)-algebra:

**6.4.2.** \( \text{End}_AV \) is a unital \( K \)-subalgebra of \( \text{End}_KV \). \( \blacksquare \)

A representation \( \delta' \) of \( A \) is called equivalent to \( \delta \) if the module for \( \delta' \) is \( A \)-isomorphic to \((V, \delta)\), otherwise inequivalent.

**6.4.3.** Let \((V, \delta), (V', \delta')\) be isomorphic \( A \)-modules. Then \( \ker \delta = \ker \delta' \), \( \text{End}_AV \cong \text{End}_AV' \).

**Proof.** Let \( \varphi \) be an \( A \)-module isomorphism of \( V \) onto \( V' \). Writing \( O, O' \) resp., for the zero endomorphism of \( V, V' \) resp., we have \( a\delta = O \iff (a\delta)\varphi = O' \iff \varphi(a\delta') = O' \iff a\delta' = O' \). Furthermore, if \( \varphi \) is an \( A \)-isomorphism of \( V \) onto \( V' \), then \( \varphi^{-1}\alpha \varphi \in \text{End}_AV' \) for all \( \alpha \in \text{End}_AV \), and by \( \alpha \mapsto \varphi^{-1}\alpha \varphi \) we obtain an isomorphism of \( K \)-algebras. \( \blacksquare \)

\( ^{97}\)More precisely, \( \text{End}_AV \) is the centralizer of the subspace \( A\delta \) in \( \text{End}_KV \), consisting of all \( K \)-endomorphisms of \( V \) which commute with each element of \( A\delta \).
It is a routine exercise to extend the basic homomorphism theorem for vector spaces (groups resp.) to the analogous theorem for modules. Therefore we state without further details this important instrument and its standard consequences:

6.5 Theorem. (Homomorphism Theorem for modules) Let $K$ be a commutative unitary ring, $A$ a $K$-space. Let $V$, $V'$ be $A$-modules and $\varphi$ an $A$-module homomorphism of $V$ into $V'$. Then

1. $V\varphi \leq V'$,
2. $\ker\varphi \leq \underbrace{V}_A$,
3. $V/\ker\varphi \cong \underbrace{V\varphi}_A$ (by the correspondence $\ker\varphi + v \leftrightarrow v\varphi$.)

Supplements

1. (“Correspondence Theorem”) Let $\mathcal{X}$ the set of all submodules of $V$ containing $\ker\varphi$, $\mathcal{Y}$ the set of all submodules of $V\varphi$. Then the mapping $\mathcal{Y} \rightarrow \mathcal{X}$, $T \mapsto T\varphi^-$, is a lattice isomorphism (in particular, a bijection).

2. Let $U, W \leq \underbrace{V}_A$. Then $(U + W)/\underbrace{U}_A \cong W/(U \cap W)$.

3. Let $S \leq T \leq \underbrace{V}_A$. Then $(V/S)/(T/S) \cong \underbrace{V/T}_A$.

Corollary Factor modules of completely reducible $A$-modules are completely reducible.

Proof. Let $\mathcal{X}$ be a set of irreducible submodules of $V$ such that $\sum \mathcal{X} = V$, $U \leq \underbrace{V}_A$. For all $W \in \mathcal{X}$ we have $W \cap U \leq \underbrace{V}_A$, hence $W \cap U = 0$ or $W \subseteq U$ as $W$ is irreducible. Let $\mathcal{X}' := \{W | W \in \mathcal{X}, W \not\subseteq U\}$. Then $V = U + \sum \mathcal{X}'$. Let $\varphi$ be the canonical epimorphism of $V$ onto $V/U$ (see 6.4.1(1)). Then $V \cong \underbrace{W}_A \varphi$ for all $W \in \mathcal{X}'$ by Supplement (2), and $V/U = V\varphi = \sum_{W \in \mathcal{X}'} W\varphi$. The claim follows.

Let $K$ be a field, $A$ a $K$-vector space. By 6.3, a finite dimensional $A$-module is completely reducible if and only if there exists a direct decomposition into irreducible submodules. Replacing the word “irreducible” by “directly $A$-indecomposable”, we clearly have, by a straightforward induction on the dimension:

6.5.1. Every finite dimensional $A$-module has a decomposition into directly $A$-indecomposable submodules.

In the direction of uniqueness, this simple existence statement is accompanied by the following classical and nontrivial result which, however, will not be needed in the sequel:

Theorem (Krull, Schmidt) Let $K$ be a field, $A$ a $K$-vector space, $V$ a finite dimensional $A$-module. Let $\mathcal{X}$, $\mathcal{Y}$ be direct $A$-decompositions of $V$ into directly $A$-indecomposable submodules. Then there exists an element $\alpha \in \text{Aut}_A V$ with the property
\{U\alpha|U \in \mathcal{X}\} = \mathcal{Y}.

Thus \(|\mathcal{X}| = |\mathcal{Y}|\). By \(\alpha\), a bijection of \(\mathcal{X}\) onto \(\mathcal{Y}\) is induced which assigns to each \(U \in \mathcal{X}\) an \(A\)-isomorphic \(U' \in \mathcal{Y}\). The main instrument in the proof of this theorem are the projections considered in 6.4.1(2). These will also play an important role in our next result:

6.6 Proposition. Let \(K\) be a commutative unitary ring, \(A\) a \(K\)-space, \(V\) an \(A\)-module.

(1) Let \(n \in \mathbb{N}\), and consider \(V^n\) as an \(A\)-module via componentwise action (6.1.1). Then the \(K\)-algebras \(\text{End}_A V^n\) and \((\text{End}_A V)^{n \times n}\) are isomorphic.

(2) (Schur’s Lemma) If \(V\) is irreducible, then every nonzero element of \(\text{End}_A V\) is invertible in \(\text{End}_A V\).

Proof. (1) For all \(j \in \mathbb{N}\) put \(V^{(j)} := \{ (0_V, \ldots, 0_V, v, 0_V, \ldots, 0_V) | v \in V \}\). Then \(V^{(j)}\) is a direct \(A\)-summand of \(V^n\), with co-summand \(\sum_{i \neq j} V^{(i)}\). Furthermore, the sum of all projections

\[ \pi_j : V^n \to V^n, \quad (v_1, \ldots, v_n) \mapsto (0_V, \ldots, 0_V, v_j, 0_V, \ldots, 0_V) \quad (6.4.1(2)) \]

is the identity on \(V^n\): \(\sum_{j \in \mathbb{N}} \pi_j = \text{id}_{V^n}\). Clearly,

\[ \varphi_j : V^{(j)} \to V, \quad (0_V, \ldots, 0_V, v, 0_V, \ldots, 0_V) \mapsto v, \]

is an \(A\)-isomorphism. For all \(\alpha \in \text{End}_A V^n\) set \(\alpha_{ij} := \varphi_j^{-1} \alpha \pi_j \varphi_j\) for all \(i, j \in \mathbb{N}\). Then \(\alpha_{ij} \in \text{End}_A V\) for all \(i, j \in \mathbb{N}\) which gives rise to the following mapping:

\[ \Phi : \text{End}_A V^n \to (\text{End}_A V)^{n \times n}, \quad \alpha \mapsto (\alpha_{ij})_{i,j \in \mathbb{N}}. \]

We show that \(\Phi\) is a bijection: The calculation

\[
(v_1, \ldots, v_n)\alpha = \sum_{i \in \mathbb{N}} (0_V, \ldots, 0_V, v_i, 0_V, \ldots, 0_V)\alpha = \sum_{i \in \mathbb{N}} (v_i \varphi_i^{-1})\alpha \text{id}_{V^n}
\]

\[
= \sum_{j \in \mathbb{N}} \sum_{i \in \mathbb{N}} v_i \varphi_i^{-1} \alpha_{ij} = \sum_{j \in \mathbb{N}} \sum_{i \in \mathbb{N}} v_i \alpha_{ij} \varphi_j^{-1} = \left( \sum_{i \in \mathbb{N}} v_i \alpha_{1i}, \ldots, \sum_{i \in \mathbb{N}} v_i \alpha_{in} \right)
\]

shows that \(\alpha\) is uniquely determined by the \(\alpha_{ij}\). Hence \(\Phi\) is injective. Now let \(\alpha(i, j) \in \text{End}_A V\) for all \(i, j \in \mathbb{N}\). Set

\[ \alpha : V^n \to V^n, \quad (v_1, \ldots, v_n) \mapsto \left( \sum_{k \in \mathbb{N}} v_k \alpha(k, 1), \ldots, \sum_{k \in \mathbb{N}} v_k \alpha(k, n) \right). \]

Then \(\alpha \in \text{End}_A V^n\), and for all \(v \in V\)

\[ v\alpha_{ij} = (0_V, \ldots, 0_v, v, 0_V, \ldots, 0_V)\alpha_{ij} \pi_j \varphi_j \]

\[ = (v \alpha(i, 1), \ldots, v \alpha(i, n)) \pi_j \varphi_j = v \alpha(i, j), \]

108
hence $\alpha_{ij} = \alpha(i, j)$, i.e., $\alpha \Phi = (\alpha(i, j))_{i,j \in B}$ so that $\Phi$ is surjective.

For all $i, j \in \mathbb{N}$, the mapping $\text{End}_A V^n \to \text{End}_A V$, $\alpha \mapsto \alpha_{ij}$, is $K$-linear. It follows that $\Phi$ is $K$-linear. Furthermore, for any $\alpha, \beta \in \text{End}_A V^n$, we have $\alpha_{ij} \beta_{jk} = (\varphi_i^{-1} \alpha \varphi_j)(\varphi_j^{-1} \beta \varphi_k) = \varphi_i^{-1} \alpha \varphi_j \varphi_k$ for all $i, j, k \in \mathbb{N}$. Hence

$$\sum_{j \in \mathbb{N}} \alpha_{ij} \beta_{jk} = \varphi_i^{-1} \alpha \sum_{j \in \mathbb{N}} \pi_j \beta \pi_k \varphi_k = \varphi_i^{-1} \alpha \beta \pi_k \varphi_k = (\alpha \beta)_{ik},$$

i.e., $((\alpha \beta)_{ij})_{i,j \in \mathbb{N}} = (\alpha_{ij})_{i,j \in \mathbb{N}} (\beta_{ij})_{i,j \in \mathbb{N}}$. $\Phi$ is an isomorphism of $K$-algebras. \hfill $\square$

(2) If $\alpha \in \text{End}_A V \setminus \{0_{\text{End}_A V}\}$, then $V \alpha \neq 0$, i.e., $\ker \alpha \neq V$, hence $\ker \alpha = 0$, furthermore $V \alpha = V$ – both by 6.5(1),(2) because $V$ is irreducible. Therefore, $\alpha$ is bijective. The inverse function $\alpha^{-1}$ is $K$-linear, and for all $a \in A$ we have $\alpha(a \delta) = (a \delta) \alpha$, hence $(a \delta) \alpha^{-1} = \alpha^{-1} (a \delta)$ (where $\delta$ is the representation). Thus $\alpha^{-1} \in \text{End}_A V$. \hfill $\square$

6.7 Definition. Let $K$ be a field. A division algebra over $K$ is an associative unitary $K$-algebra $D$ in which every element $\neq 0_D$ is multiplicatively invertible. We re-formulate Schur’s Lemma 6.6(2):

6.7.1. Let $A$ be a $K$-vector space, $V$ an irreducible finite-dimensional $A$-module. Then $\text{End}_A V$ is a division algebra over $K$. \hfill $\square$

6.7.2. An associative unitary $K$-algebra is a division algebra if and only if its nonzero elements form a multiplicative group. \hfill $\square$

6.7.3. The only one-sided ideals of a division algebra $D$ are $\{0_D\}$, $D$. In particular, $D$ is simple.

Proof. If $J$ is a nonzero right ideal of $D$, we have $J = JD = D$. \hfill $\square$

6.7.4. Let $D$ be a division algebra over $K$, $n \in \mathbb{N}$. Then $D^{n \times n}$ is simple.

This follows from 6.7.3 by means of the following assertion:

6.8 Proposition. Let $K$ be a field, $B$ a simple associative (unitary) $K$-algebra, $n \in \mathbb{N}$. Then the $K$-algebra $B^{n \times n}$ is simple.

Before we give a proof, we introduce some useful notation concerning the standard basis of a matrix space, more generally over some algebra instead of a field, the case usually studied in Linear Algebra. For any unitary $K$-algebra $B$, $n \in \mathbb{N}$, let $e(k, l)$ be the $n \times n$ matrix over $B$ having the $(k, l)$ coordinate equal to $1_B$ and all other coordinates equal to $0_B$. More generally, let $b e(k, l) = e(k, l) b$ be the $n \times n$ matrix over $B$ in which the

$\text{A simple algebra is commonly understood as unitary by definition but we have not introduced this}

convention.\hfill 99$

In rigorous language, $e(k, l)$ is the mapping $\mathbb{N} \times \mathbb{N} \to B$, $(i, j) \mapsto \begin{cases} 1_B & \text{if } (i, j) = (k, l) \\ 0_B & \text{otherwise} \end{cases}$.
$(k,l)$ coordinate equals $b$ and all other coordinates are $0_B$, for any $b \in B$. This is the product of the “constant diagonal matrix” $\text{diag}[b, \ldots, b]$ and $e(k,l)$. Then the matrices $e(k,l)$ form a $B$-basis of $B^{n \times n}$, and their multiplication is described by the following remark (where $[O]$ is the $n \times n$ zero matrix over $B$):

6.8.1. $\forall i,j,k,l \in \mathbb{N} \quad e(i,j) e(k,l) = \begin{cases} e(i,l) & \text{if } j = k \\ [O] & \text{if } j \neq k \end{cases}$. \[\Box\]

Now let $B$ be associative, $n \in \mathbb{N}$, $(b_{ij})_{i,j} \in J \subseteq B^{n \times n}$. Let $i_0, j_0 \in \mathbb{N}$. Then $B^* := b_{i_0,j_0}^d$. Then

(*)

$B^* e(k,l) \subseteq J$ for all $k, l \in \mathbb{N}$.

Proof. We have $b_{i_0,j_0} c(k,l) = e(k,i_0)(b_{ij})_{i,j} e(j_0,l) \in J$, hence

$$\forall b, b' \in B \quad b b_{i_0,j_0} b' e(k,l) = \text{diag}[b, \ldots, b] b_{i_0,j_0} e(k,l) \text{diag}[b', \ldots, b'] \in J.$$ But $B^* = B b_{i_0,j_0} B$ so that we obtain (*). \[\Box\]

Proof of 6.8. Let $B$ be simple and $J$ a non-zero ideal of $B^{n \times n}$. The ideal of $B$ generated by a non-zero coordinate of a matrix in $J$ is all of $B$ so that, by (*), $B e(k,l) \subseteq J$ for all $k, l \in \mathbb{N}$. It follows that $J = B^{n \times n}$. \[\Box\]

One of our aims for the sequel is to prove that, in the finite-dimensional case, the $K$-algebras $D^{n \times n}$ where $D$ is a division algebra over $K$ (see 6.7.4) are, up to isomorphism, the only simple associative unitary $K$-algebras. Thus the problem of a classification of all simple objects in the class of all finite-dimensional unitary associative $K$-algebras will find a most satisfactory answer. Our approach to obtain this result will, more generally, consist in an analysis of the possible irreducible modules for an associative algebra. It will turn out that the regular representation $\rho$ (see 6.1.3) is of utmost importance. An algebra originates from a $K$-space by the introduction of a (doubly) distributive multiplication. It is useful to have various equivalences at hand to say that that multiplication is associative:

6.8.2. Let $K$ be a commutative unitary ring, $A$ a $K$-algebra. The following are equivalent:

(i) $A$ is associative : $\forall x, y, z \in A \quad (xy)z = x(yz),$ 

(ii) $\rho$ is a multiplicative homomorphism $A \to \text{End}_K A$ : $\forall y, z \in A \quad (yp)(z \rho) = (yz) \rho,$ 

(iii) $\lambda$ is a multiplicative anti-homomorphism $A \to \text{End}_K A$ : $\forall x, y \in A \quad (xy) \lambda = (y \lambda)(x \lambda),$ 

(iv) $A \lambda, A \rho$ commute elementwise : $\forall x, z \in A \quad (x \lambda)(z \rho) = (z \rho)(x \lambda),$ 

(v) $\forall x \in A \quad x \lambda \in \text{End}(A, \rho)$ : (as in (iv))

110
Clearly, it is preferable to work with homomorphisms instead of anti-homomorphisms. Therefore, in view of 6.8.2(ii),(iii), we will mostly work with the right regular representation when dealing with associative algebras.\footnote{It is here where the proper choice (between the left and the right regular representation) begins to have influence on the transparency of the theory. The preference of \( \rho \) to \( \lambda \) corresponds, of course, to the general decision of writing mappings on the right-hand side of the arguments.}

### 6.9 Proposition

Let \( K \) be a commutative unitary ring, \( A \) an associative unitary \( K \)-algebra.

1. \( \text{End}_A A = A\lambda \cong A^- \). \footnote{In particular, \( \text{End}_K K \cong K \). Thus 6.6(1) contains the special case \( \text{End}_K K^n \cong K^{n \times n} \), usually analyzed in Linear Algebra for a field \( K \).}

2. \( R^{\leq A} = AR \) for every right ideal \( R \) of \( A \). If \( R, S \) are minimal right ideals of \( A \) and \( S \subset R^{\leq A} \), then \( R \cong S \).

#### Proof.

(1) By 6.8.2(v) we have \( x\lambda \in \text{End}_A A \) for all \( x \in A \). Hence, by 6.8.2(iii), \( \lambda \) is a homomorphism of \( A^- \) into \( \text{End}_A A \). If \( x \in \ker \lambda \), then \( x = x \cdot 1_A = 1_A(x\lambda) = 0_A \), hence \( \ker \lambda = \{0_A\} \). We claim that \( \lambda \) is onto: Let \( \alpha \in \text{End}_A A \) and put \( x := 1_A\alpha \). Then

\[
\forall a \in A : a(x\lambda) = ax = (1_A\alpha)a = (1_Aa)\alpha = a\alpha,
\]

hence \( \alpha = x\lambda \in A\lambda \).

(2) The first assertion is trivial (see also footnote 94). Now let \( R \leq A, \mathcal{X} := \{xR|x \in A, xR \neq \{0_A\}\} \). Then \( R^{\leq A} = AR = \sum \mathcal{X} \). By 6.8.2(v), \( x\lambda \) is an \( A \)-homomorphism, for all \( x \in A \), hence either \( R(x\lambda) \cong R \) or \( R(x\lambda) = \{0_A\} \) as \( R \) is an irreducible \( A \)-module. Therefore \( R' \cong R \) for all \( R' \in \mathcal{X} \) whence \( R^{\leq A} \) is completely reducible. If \( \{0_A\} \neq S \subset R^{\leq A} \), there exist \( x_1, \ldots, x_k \in A \) such that \( x_1R + \cdots + x_kR \cap S \neq \{0_A\} \). For an irreducible \( S \) this implies \( S \leq x_1R + \cdots + x_kR \). Choosing \( k \) minimal we have \( S \subset \sum_{i \neq k} x_iR =: T \), and the sum of the \( x_iR \) is a direct sum (6.2). Then \( S \cap T = \{0_A\} \), hence \( S \cong S + T/T \cong x_kR + T/T \cong A \). This implies \( R \cong S \) as \( R \) is irreducible. \qed

### 6.10 Definition

Let \( K \) be a commutative unitary ring, \( A \) a unitary associative \( K \)-algebra, \( (V, \delta) \) an \( A \)-module. If \( \delta \) is a multiplicative homomorphism, it is called an algebra representation, and \( (V, \delta) \) an algebra module of \( A \), also called an \( A \)-right module.\footnote{An \( A \)-left module is then defined to be an \( A \)-module \( (V, \delta) \) such that \( \delta \) is a multiplicative anti-homomorphism.} \( \delta \) is called unital if \( 1_A\delta = \text{id}_V \). We now fix a common convention:

\[
\forall z \in A : z\rho \in \text{End}(A, \lambda) \quad (\text{as in (iv)}) \quad \square
\]
Under the hypothesis that $A$ is a unitary associative $K$-algebra, the term “$A$-module” is used in the sense of “unital algebra module of $A$”, the term “representation of $A$” in the sense of “unital algebra representation of $A$”. A prominent example is the (right) regular representation $\rho$, making the underlying $K$-space of a unitary associative $K$-algebra $A$ into the (right) regular module $(A, \rho)$. The following simple remark exhibits a fundamental connection between this particular $A$-module and an arbitrary $A$-module:

6.10.1. Let $(V, \delta)$ be an $A$-module and $v \in V$. Then

$$\varphi_v : A \to V, \ a \mapsto v(a\delta)$$

is an $A$-module homomorphism.\(^{103}\) As a consequence, if $V$ is generated by $n$ elements as an $A$-module, $V$ is $A$-isomorphic to a factor module of $A^n$. \(\bigcap_{v \in V} \ker \varphi_v = \ker \delta.\)

**Proof.** Obviously, $\varphi_v$ is $K$-linear. Furthermore, we have for all $x, a \in A$

$$(a(x\rho)) \varphi_v = (ax) \varphi_v = v((ax)\delta) = v((a\delta)(x\delta)) = (v(a\delta))(x\delta) = (a\varphi_v)(x\delta),$$

hence $(x\rho) \varphi_v = \varphi_v(x\delta)$. Finally,

$$a \in \ker \delta \iff \forall v \in V \ v(a\delta) = 0 \iff \forall v \in V \ a \in \ker \varphi_v.$$

As a consequence, we note

6.10.2. Let $(V, \delta)$ be an irreducible $A$-module. Then

$$\ker \delta = \bigcap_{\rho \in \rho_{\text{max}}} \{M | M \leq A, \ A/M \cong _A V \}.$$  

**Proof.** For all $v \in V \setminus 0$ we have $A/\ker \varphi_v \cong _A V$. Hence, by 6.10.1,

$$\ker \delta = \bigcap_{v \in V} \ker \varphi_v \supseteq \bigcap_{\rho \in \rho_{\text{max}}} \{M | M \leq A, \ A/M \cong _A V \}$$

If $\ker \delta \not\subseteq M \leq A$ we have $M + \ker \delta = A$, hence $A \cdot \ker \delta = \ker \delta \not\subseteq M$ so that $A/M$ is not annihilated by $\ker \delta$. In particular, $A/M \not\cong _A V$. \( \square \)

6.10.3. An $A$-module $V$ is irreducible if and only if there exists a maximal right ideal $M$ of $A$ such that $V \cong _A A/M$.

**Proof.** Let $V$ is irreducible, $0_V \neq v \in V$, $\varphi_v$ as in 6.10.1. Then $V \cong _A A/\ker \varphi_v$, and $\ker \varphi_v$ is a maximal right ideal of $A$ by 6.5, Suppl.(1) which also implies the converse. \( \square \)

\(^{103}\)More generally: Let $(V, \delta)$ be an $A$-module, $v_1, \ldots, v_n \in V$. Then $\varphi : A^n \to V, \ (a_1, \ldots, a_n) \mapsto \sum_{i \in \mathbb{N}} v_i(a_i\delta)$, is an $A$-module homomorphism.
6.10.4. Let \( R \leq \rho, \min A \), \( J := R^{\leq A} \). Let \( \delta \) be an irreducible representation of \( A \) which is inequivalent to the representation given by \( R \) (via \( \rho \)). Then \( J \subseteq \ker \delta \).

Proof. By 6.10.3, we may assume w.l.o.g. that \( \delta = \rho_{A/M} \) for some maximal right ideal \( M \) of \( A \). Assume \( J \not\subseteq M \). Then, by 6.9(2), there exists an element \( x \in A \) such that \( xR \not\subseteq M \), hence \( A = M \oplus xR \), implying \( A/M \cong xR \cong R \), a contradiction. It follows that \( J \subseteq M \) and therefore \( AJ \subseteq M \), i.e., \( J \subseteq \ker \rho_{A/M} \).

6.10.5. Let \( R \leq \rho, \min A \), \( I \subseteq A \) such that \( I < R^{\leq A} \). Then \( I \subseteq \ker \delta \) for every irreducible representation \( \delta \) of \( A \). In particular, \( SI = \{0_A\} \) for every \( S \leq \rho, \min A \).

The ideal \( J(A) := \bigcap \{ \ker \delta \mid \delta \text{ irreducible representation of } A \} \) is called the Jacobson radical of \( A \).

6.10.6. If \( J(A) = \{0_A\} \), then \( R^{\leq A} \leq \rho, \min A \) for every \( R \leq \rho, \min A \).

6.10.7. If \( (A, \rho) \) is completely reducible, then \( J(A) = \{0_A\} \).

A is called semisimple if it has a direct decomposition into simple ideals. Equivalently, \( A \) is semisimple if it is isomorphic to a direct sum of simple algebras. Our next aim may be viewed as the main theorem on semisimple finite-dimensional associative algebras over a field. We prepare it by two further remarks:

6.10.8. Let \( R \leq A, S \leq A \) such that \( A = R \oplus S \). Then \( TA = TR \) for all \( T \leq \rho, K \). If \( e \in R, f \in S \) such that \( 1_A = e + f \), then \( xe = x \) for all \( x \in R \), \( fy = y \) for all \( y \in S \).

6.10.9. Let \( K \) be a field, \( A \) a finite-dimensional semisimple associative \( K \)-algebra. Then \( (A, \rho) \) is completely reducible.

Proof. There are simple algebras \( A_1, \ldots, A_h \subseteq A \) such that \( A = A_1 \oplus \cdots \oplus A_h \). Being finite-dimensional, each \( A_i \) has a minimal right ideal \( R_i \). It follows from 6.10.8 that \( R_i \leq A \). Thus \( A_i = R_i^{\leq A} = AR_i = \sum_{x \in A} xR_i \) and \( A = \sum_{i \in I} \sum_{x \in A} xR_i \). Thus \( A \) is a sum of (finitely many) minimal right ideals.

6.11 Theorem (Wedderburn 1907). Let \( K \) be a field, \( A \) a finite-dimensional unitary associative \( K \)-algebra,

\[ X \text{ the set of all minimal right ideals of } A, \]
\[ Y \text{ the set of all minimal ideals of } A. \]

Let \( (A, \rho) \) be completely reducible. Then the following holds:

(1) Every finitely generated \( A \)-module \( V \) is completely reducible. If \( V \) is irreducible, there exists an element \( R \in X \) such that \( V \cong R \).

\[ \text{cf. footnote 98} \]
(2) \( \{ R^{\geq A} | R \in \mathcal{X} \} = \mathcal{Y} = \{ J | J \leq A, J \text{ is a simple algebra} \}. \)

(3) \( \mathcal{Y} \) is finite and \( A = \bigoplus \mathcal{Y} \). Every \( J \in \mathcal{Y} \) is unitary.

(4) \( \forall R, S \in \mathcal{X} \quad R \cong_A S \Leftrightarrow R^{\geq A} = S^{\geq A} \).

(5) For every \( J \in \mathcal{Y} \) there exists a positive integer \( n_J \) such that \( J \cong_A R^{n_J} \) for any \( R \in \mathcal{X} \) with \( R \subseteq J \).

(6) Let \( A \) be simple. Then there exists a positive integer \( n \) such that \( A \cong D^{n \times n} \) where \( D = (\text{End}_A R) \) for any \( R \in \mathcal{X} \).

**Corollary** Let \( \mathcal{Y} = \{ J_1, \ldots, J_h \} \) (\( J_i \) mutually distinct) and \( R_1, \ldots, R_h \in \mathcal{X} \) such that \( R_i \subseteq J_i \) for all \( i \in [h] \). Then, up to \( A \)-isomorphism, \( \{ R_1, \ldots, R_h \} \) is a complete set of mutually non-isomorphic irreducible \( A \)-modules. Moreover,

\[
A \cong D_1^{n_1 \times n_1} \oplus \cdots \oplus D_h^{n_h \times n_h} \quad \text{where} \quad D_i := (\text{End}_A R_i)^-, \; n_i := n_{J_i} \quad (\text{cf. (5)}).
\]

Putting \( d_i := \text{dim}_K D_i \), we have

\[
\forall i \in [h] \quad \text{dim}_K R_i = d_i n_i, \quad \text{dim}_K A = d_1 n_1^2 + \cdots + d_h n_h^2.
\]

In view of the first claim in the Corollary, \( h \) is called the **number of irreducible representations** of \( A \), furthermore the numbers \( d_i n_i \) \((i \in [h])\) their **degrees**.

**Proof.** (1) Let \( n \in \mathbb{N} \) and \( B \) be generated as an \( A \)-module by \( n \) elements. Then (cf. footnote 103) \( B \) is \( A \)-isomorphic to a factor module of \( A^n \). This module is completely reducible as \( A \) is completely reducible. The first claim in (1) now follows from the Corollary of 6.5.

If \( M <_A G \), there must exist some \( R \in \mathcal{X} \) such that \( R \not\subseteq M \) as \( \sum \mathcal{X} = A \). It follows that \( R \cap M = \{ 0_A \} \) and \( R + M = A \) (as \( A/M \) is irreducible). Hence \( A/M \cong R \in \mathcal{X} \).

The claim follows from 6.10.3.

(2), (3) Clearly, \( \{ R^{\geq A} | R \in \mathcal{X} \} \supseteq \mathcal{Y} \supseteq \{ J | J \leq A, J \text{ is a simple algebra} \}. \) To obtain the first of the two reverse inclusions it suffices to combine 6.10.7 and 6.10.6. We may now show (3) as follows: \( A = \sum \mathcal{X} = \sum_{R \in \mathcal{X}} R^A = \sum \mathcal{Y}, \) and the sum over \( \mathcal{Y} \) is a direct sum: Let \( J_1, \ldots, J_h \in \mathcal{Y} \) such that \( A = J_1 + \cdots + J_h \), \( h \) minimal. Then \( J_i \cap \sum_{k \neq i} J_k = \{ 0_A \} \) as \( J_i \in \mathcal{Y} \), hence \( A = J_1 \oplus \cdots \oplus J_h \). There is no \( J \in \mathcal{Y} \) distinct from \( J_1, \ldots, J_h \) because otherwise \( J J_i \subseteq J \cap J_i = \{ 0_A \} \) for all \( i \in [h] \) whence \( J = J \cdot 1_A \subseteq J(J_1 + \cdots + J_h) = \{ 0_A \} \), a contradiction. – We now derive the remaining second inclusion in (2): Let \( k \in [h] \) and \( \{ 0_A \} \neq I \subseteq J_k \). Then \( I \subseteq A \) by (3) and two-fold application of 6.10.8, hence \( I = J_k \) showing that \( J_k \) is simple. By the second part of 6.10.8, \( J_k \) is unitary.

(4) Let \( R, S \in \mathcal{X} \). From 6.9(2) we obtain the implication "\( \Leftarrow \)". For "\( \Rightarrow \)" suppose \( R^{\geq A} \not= S^{\geq A} \). Then \( SR \subseteq S^{\geq A} R^{\geq A} \subseteq S^{\geq A} \cap R^{\geq A} = \{ 0_A \} \). The assumption \( R \cong_A S \) would
now imply \( RR = \{0_A\} \), hence \( R^2 A R^2 A = ARAR = ARR = \{0_A\} \) by 6.9(2). This is absurd as \( R^2 A \) is unitary, by (3).

(5) Let \( J \in \mathcal{Y} \). By (2), there exists an \( R \in \mathcal{X} \) such that \( J = R^2 A \), hence \( J = x_1 R \oplus \cdots \oplus x_n R \) for appropriate \( x_i \in A \). It follows that \( J \cong R^n \) as \( A \) modules. In particular, \( \dim_K J = n \dim_K R \). If \( S \in \mathcal{X} \) and \( S \subseteq J \), (4) implies \( \dim_K R = \dim_K S \) so that \( n \) is independent of the choice of \( R \).

(6) By (5), \( A \cong R^n \) for every \( R \in \mathcal{X} \). It follows that

\[
A^- \cong \text{End}(A, \rho) \cong \text{End}(R^n, \rho) \cong (\text{End}_A R)^{n \times n}
\]

whence \(^{105} A \cong (A^-)^- \cong D^{n \times n}.

Proof of the Corollary: By (3), \( A = J_1 \oplus \cdots \oplus J_h \), and every \( J_i \) is a simple algebra by (2). By (6), \( J_i \cong ((\text{End}_A R_i^-)^-)^{n_i \times n_i} \) for some \( n_i \in \mathbb{N} \), where \( R_i \) is a minimal right ideal of \( J_i \). It follows that \( R_i \) is a minimal right ideal of \( A \) and \( A \cong \bigoplus_{i \in \mathcal{A}} D_i^{n_i \times n_i} \). The assertions about the \( R_i \) follow from (4). A minimal right ideal \( R_i \) of \( D_i^{n_i \times n_i} \) is, e. g., given by \( D_i e(1, 1) + \cdots + D_i e(1, n_i) \)\(^{106} \) so that \( \dim_K R_i = d_i n_i \), \( \dim_K A = \sum_{i \in \mathcal{A}} n_i \). This is a minimal right ideal of \( A \) and \( A \cong \bigoplus_{i \in \mathcal{A}} D_i^{n_i \times n_i} \).

\[ \square \]

6.12 Definition. Let \( A \) be an associative algebra. We define the centre of \( A \) by

\[ Z(A) := \{ z | z \in A, \forall a \in A \quad za = az \}. \]

6.12.1. \( Z(A) \) is a subalgebra of \( A \). \( \square \)

6.12.2. If \( B \) is any associative algebra, then \( Z(A \oplus B) = Z(A) \oplus Z(B) \). \( \square \)

6.12.3. Let \( A \) be unitary. Then

\[ \forall n \in \mathbb{N} \quad Z(A^{n \times n}) = \{ \text{diag}[z, \ldots, z] | z \in Z(A) \} \cong Z(A). \]

Proof. For any \( a_{ij} \in A \) \((i, j \in \mathcal{N})\), we have

\[
\sum_{i,j \in \mathcal{N}} a_{ij} e(i,j) \in Z(A^{n \times n})
\]

\[
\iff \forall b \in A \quad \forall k, l \in \mathcal{N} \quad b e(k, l) \sum_{i,j \in \mathcal{N}} a_{ij} e(i,j) = \sum_{i,j \in \mathcal{N}} a_{ij} e(i,j) b e(k, l)
\]

\[
\iff \forall b \in A \quad \forall k, l \in \mathcal{N} \quad \sum_{j \in \mathcal{N}} b a_{ij} e(k, j) = \sum_{i \in \mathcal{N}} a_{ik} b e(i, l)
\]

\[
\iff \forall b \in A \quad \forall k, l \in \mathcal{N} \quad b a_{ij} = a_{ik} b, \quad \forall i, j \in \mathcal{N} \quad b a_{ij} = 0_A = a_{ij} b
\]

\[
\iff \forall k, l \in \mathcal{N} \quad a_{kk} = a_{ll} \in Z(A), \quad \forall i, j \in \mathcal{N} \quad a_{ij} = 0_A.
\]

The finally claimed algebra isomorphism is obvious. \( \square \)

\(^{105}\)For every algebra \( B \), it is an easy exercise to prove \((B^{n \times n})^- \cong (B^-)^{n \times n} \).

\(^{106}\)It is a routine exercise to show this. (The notation was introduced on p. 110.)
6.12.4. Let $A$ be unitary. Then the mapping $\iota : K \to A$, $c \mapsto c \cdot 1_A$ is a unital $K$-algebra homomorphism and $K\iota \subseteq Z(A)$.

Proof. For all $c,c',d \in K$ we have $(c + dc')\iota = (c + dc')1_A = c1_A + d'(1_A) = c\iota + d'(\iota)$, $(cd)\iota = (cd)1_A = (cd)(1_A1_A) = c1_A d1_A = c\iota d\iota$, $1_K\iota = 1_K 1 = 1_A$, and $(ca)\iota = (c1_A)a = c(1_A a) = ca = c(a 1_A) = a(\iota)$ for all $a \in A$. \hfill $\square$

6.12.5. Let $K$ be an algebraically closed field, $D$ a finite-dimensional division algebra over $K$. Then $\iota$ is an isomorphism of $K$ onto $D$. In particular, the numbers $d_i$ in the Corollary of 6.11 are all equal to 1.

Proof. Set $\bar{K} := K\iota$. Then $K \cong \bar{K} \subseteq Z(D)$. For all $y \in D$, the $K$-algebra $\bar{K}[y]$ is a domain, and $\dim_K \bar{K}[y](\leq \dim_K D)$ is finite, hence $\bar{K}[y]$ is a field and an algebraic extension of $\bar{K}$. But $\bar{K}(\cong K)$ is algebraically closed which implies $y \in \bar{K}$. It follows that $\bar{K} = D$. \hfill $\square$

6.13 Theorem. Let $K$ be an algebraically closed field, $A$ a finite-dimensional semisimple associative $K$-algebra, $h$ the number of its irreducible representations, $n_1, \ldots, n_h$ the dimensions of the irreducible $A$-modules. Then

1. $A \cong K^{n_1 \times n_1} \oplus \cdots \oplus K^{n_h \times n_h}$.

2. $h = \dim_K Z(A)$.

3. $n_1^2 + \cdots + n_h^2 = \dim_K A$.

4. $A$ commutative $\iff \forall j \in [h] n_j = 1 \iff \dim_K A = h$.

Proof. (1) By the Corollary of 6.11, $A \cong D_{i_1}^{n_1 \times n_1} \oplus \cdots \oplus D_{i_h}^{n_h \times n_h}$. For division algebras $D_i$ over $K$. As $\dim_K D_i \leq \dim_K A$, we have $D_i \cong K$ by 6.12.5.

(2) By (1), $Z(A) \cong 6.12.2 Z(K^{n_1 \times n_1}) \oplus \cdots \oplus Z(K^{n_h \times n_h}) \cong 6.12.3 \underbrace{K \oplus \cdots \oplus K}_{h}$.

(3) and the first equivalence in (4) follow from (1), the second equivalence in (4) from (3). \hfill $\square$

Thus for any $n \in \mathbb{N}$, the only commutative $n$-dimensional semisimple associative algebra over an algebraically closed field $K$ is, up to isomorphism, given by $\underbrace{K \oplus \cdots \oplus K}_{n}$.
7 Characters

Let $K$ be a commutative unitary ring, $(M, \cdot)$ a monoid. Then the space $KM$ of formal $K$-linear combinations over $M$ is made into a unitary associative $K$-algebra by distributive extension of $\cdot$ to $KMG$ as our monoid, that is, we will study the group algebra $KG$. For every finite subset $T$ of $KG$ set $\widehat{T} := \sum T$. As $G$ is a $K$-basis of $KG$, we have for all $S, T \subseteq G, y \in G$:

7.0.1. $\widehat{T} y = \widehat{Ty}, \quad \widehat{S} = \widehat{T} \Leftrightarrow S = T$. □

7.1 Definition. A subset $C$ of $G$ is called a conjugacy class of $G$ if there exists an element $g \in G$ such that $C = \{g^x | x \in G\}$. The number of conjugacy classes of $G$ is called the class number of $G$. Conjugacy is an equivalence relation on $G$ the equivalence classes of which are just the conjugacy classes. Thus we have:

7.1.1. The set of all conjugacy classes of $G$ is a partition of $G$. □

7.1.2. For any $C \subseteq G$ and $y \in G$, $y$ and $\widehat{C}$ commute in $KG$ if and only if $C$ is normalized by $y$ in $G$. In particular, $\widehat{C} \in Z(KG)$ for every conjugacy class $C$ of $G$.

Proof. For all $y \in G$ we have $C^y = C \Leftrightarrow Cy = yC \Leftrightarrow \widehat{C} y = y \widehat{C}$. □

7.2 Proposition. Let $K$ be a commutative unitary ring, $G$ a finite group, $K$ the set of all conjugacy classes of $G$, $\widehat{K} := \{\widehat{C} | C \in K\}$. Then $\widehat{K}$ is a $K$-basis of $Z(KG)$.

Corollary 1 The additive closure of $\widehat{K}$ is multiplicatively closed. $\langle \widehat{K} \rangle_Z$ is a subring of $Z(KG)$.

Corollary 2 Let $K$ be an algebraically closed field and $KG$ semisimple. Then $h(G) = h$.\(^{109}\)

Proof. $G$ is a $K$-basis of $KG$ and $K$ is a partition of $G$ (7.1.1), hence $\widehat{K}$ is $K$-linearly independent. By 7.1.2, $\langle \widehat{K} \rangle_K \leq Z(KG)$. Now let $c_g \in K$ such that $\sum_{g \in G} c_g g \in Z(KG)$. Then for all $x \in G$

$$\sum_{g \in G} c_g g^x = \left(\sum_{g \in G} c_g g\right)^x = \sum_{g \in G} c_g g,$$

\(^{107}\)In short, this set is also denoted by $g^G$ where $g \in C$; $g^x = x^{-1} g x$.

\(^{108}\)g \sim h \Leftrightarrow \exists x \in G \quad g^x = h$.

\(^{109}\)h: number of irreducible representations of $KG$, cf. Corollary of 6.11.
hence \( c_{gx} = c_g \) for all \( g, x \in G \): Conjugate group elements have the same coefficient. It follows that
\[
\sum_{g \in G} c_g g = \sum_{C \in K} \sum_{g \in C} c_g g \in \langle \mathcal{R} \rangle_K.
\]
The corollaries are easy consequences: If \( C_1, C_2 \in K \), then there exist positive integers \( n_g (g \in G) \) such that \( \widehat{C_1 C_2} = \sum_{g \in G} n_g g \). On the other hand, \( \mathcal{R} \subseteq K \) implies that there exist elements \( b_C \in K \) \( (C \in K) \) such that \( \widehat{C_1 C_2} = \sum_{C \in K} b_C \). As \( G \) is a \( K \)-basis of \( KG \), \( \mathcal{R} \subseteq \mathcal{K} \) implies that \( \{n_g 1_K | g \in G\} = \{b_C | C \in \mathcal{K}\} \). This proves the first corollary. The second corollary follows from \( \mathcal{R} \subseteq \mathcal{K} \) by means of \( \mathcal{R} \subseteq \mathcal{K} \).

For which fields \( K \) the algebra \( KG \) is semisimple? As a start for an approach to a fundamental criterion for this question, we remind of \( \mathcal{R} \subseteq \mathcal{K} \): If \( (V, \delta) \) is a \( KG \)-module, \( \text{End}_K V \) is a two-fold \( K \)-space module for \( KG \): by \( \delta \lambda \) and by \( \delta \rho \). Multiplicatively, the latter is a homomorphism, hence a proper algebra representation, while the former is an anti-homomorphism. Applying first the anti-isomorphism of \( G \) given by inversion, however, we obtain again a multiplicative homomorphism by linear extension to \( KG \).

The composition of both representations by defining
\[
\forall \sigma \in \text{End}_K V \forall g \in G \quad \sigma^g := (g^{-1} \delta) \sigma (g \delta)
\]
therefore satisfies the rule \( (\sigma^h)^h = \sigma^{(gh)} \) for all \( \sigma \in \text{End}_K V, g, h \in G \). Hence the \( K \)-linear extension of the mapping
\[
G \to \text{End}_K (\text{End}_K V), \quad g \mapsto \left[ \begin{array}{c} \text{End}_K V \\ \sigma \end{array} \rightarrow \frac{\text{End}_K V}{\sigma^g} \right]
\]
is an algebra representation of \( KG \). Now put
\[
\check{\sigma} := \sum_{g \in G} \sigma^g \text{ for all } \sigma \in \text{End}_K V.
\]
Then for all \( x \in G \), \( \check{\sigma} x = \check{\sigma} \), i.e., \( \check{\sigma} (x \delta) = (x \delta) \check{\sigma} \), which means that \( \check{\sigma} \in \text{End}_{KG} V \). By passing from \( \sigma \) to \( \check{\sigma} \), an arbitrary endomorphism of the \( K \)-space \( V \) is transformed into a \( KG \)-endomorphism.

If \( U \leq V \) such that \( V \sigma \subseteq U \), then also \( V \check{\sigma} \subseteq U \), hence \( \check{\sigma} \in \text{Hom}_{KG} (V, U) \), for
\[
\forall v \in V \quad v \check{\sigma} = \sum_{g \in G} v (g^{-1} \delta) \sigma (g \delta) \in U.
\]
Now let \( U \leq V \) and \( W \leq V \) with the property that \( V = U \oplus W \). We apply the foregoing idea to the projection \( \pi_U \) (in place of \( \sigma \)) with respect to the direct decomposition \( \{U, W\} \) of \( V \). Then \( \check{\pi}_U \in \text{Hom}_{KG} (V, U) \), and for all \( u \in U \) we have
\[
u \check{\pi}_U = \sum_{g \in G} \frac{u (g^{-1} \delta)}{\pi_U (g \delta)} = |G| u.
\]

118
For brevity, we shall use from now on the term KG-module in the sense of unital KG-algebra module, finitely generated over K, and the term representation accordingly, unless otherwise stated.

In this meaning, every representation \( \delta \) of KG induces and, vice versa, is induced by a homomorphism of G into the automorphism group of a K-space. \( \delta \) is called faithful if \( \ker \delta|_G = \{1_G\} \), i.e., if \( \delta \) is injective on G.

**7.3 Proposition.** Let \( K \) be a commutative unitary ring, \( G \) a finite group, \( (V,\delta) \) a KG-module.

(1) If \( |G| \cdot 1_K \) is a unit in \( K \), \( U \leq V \) \( \text{KG} \)-summand of \( V \), and \( U \) is a direct \( K \)-summand of \( V \), then \( U \) is a direct \( \text{KG} \)-summand of \( V \).

(2) (Maschke’s Theorem (1899)) If \( K \) is a field such that \( \text{char} \ K \nmid |G| \), then \( V \) is completely reducible.

**Proof.** (1) Let \( W \leq V \) such that \( V = U \oplus W \) and \( \pi_U \) as before. Set \( \tilde{W} := \ker \tilde{\pi}_U \). By 6.5(1), \( \tilde{W} \leq V \) as \( \tilde{\pi}_U \in \text{End}_{\text{KG}}(V) \). For all \( u \in U \) we have \( u\tilde{\pi}_U = |G|u \), so \( \tilde{\pi}_U|_U \in \text{Aut}_{\text{KG}}U \) implying \( U \cap \tilde{W} = 0 \), \( U + \tilde{W} = V \).

(2)110 follows from (1) and 6.3, as in the case of a field \( K \) every \( K \)-subspace is a direct \( K \)-summand so that (1) applies.

It is not hard to see that the converse of 7.3(2) holds:

**7.3.1.** If \( K \) is a field and \( \text{char} \ K \nmid |G| \), then \((KG,\rho)\) is not completely reducible.

For \( \sum G \) generates a 1-dimensional ideal with zero multiplication. By 6.10.4, it is contained in \( J(KG) \) which, by 6.10.7, implies the claim. \( \square \)

Maschke’s Theorem gives rise to a clear distinction of two parts of the representation theory of finite groups \( G \) over a field \( K \): Ordinary Representation Theory, characterized by the condition that \( \text{char} \ K \nmid |G| \) so that all indecomposable KG-modules are irreducible, and Modular Representation Theory where \( \text{char} \ K \mid |G| \) so that there exists an indecomposable KG-module which is not irreducible. We will confine ourselves in the following to ordinary representations. In that case we have two very different types of \( K \)-bases of \( KG \):

Let \( G \) be a finite group, \( K \) a field such that \( \text{char} \ K \nmid |G| \).

(\( I \)) Any tuple \( \mathfrak{B} = (g_1, \ldots, g_n) \) consisting of the elements of \( G \) is a \( K \)-basis tuple of \( KG \).

The endomorphisms of \( KG \) induced by the elements of \( G \) via right multiplication allow a particularly simple description in terms of \( \mathfrak{B} \) as for \( y \in G \) and all \( i \in \mathfrak{B} \) there exists a unique \( j \in \mathfrak{B} \) such that \( g_i y = g_j \). Unless \( y = 1_G \), we have \( i \neq j \). The matrix representation of \( y \rho \) with respect to \( \mathfrak{B} \) is a permutation matrix which is the unit matrix if \( y = 1_G \).

---

110Applying Zorn’s Lemma, it is easily seen that this proof works for modules of arbitrary dimensions.
and has a zero main diagonal otherwise.

\(\text{(II)}\) The minimal ideals \(J_1, \ldots, J_h\) of \(KG\) constitute a direct decomposition of \(KG\) which may be refined by decomposing each \(J_i\) directly into \((KG\text{-isomorphic})\) minimal right ideals:

\[
KG = \bigoplus_{i \in \mathbb{H}} \bigoplus_{n_i} (R_{i,1} \oplus \cdots \oplus R_{i,n_i}).
\]

\[
KG = \bigoplus_{i \in \mathbb{H}, J_i \leq KG} J_i \cong KG \bigoplus_{J_i \leq KG} R_i \leq KG
\]

Any tuple \(\mathcal{E}\) consisting of the elements of bases of all the \(R_{i,j} (i \in \mathbb{H}, j \in \mathbb{N})\) is a \(K\)-basis tuple of \(KG\).

Let \(i \in \mathbb{H}\) and \(e_i\) be the unit element of \(J_i\), \(R_i\) a minimal right ideal of \(J_i\), \(\delta_i\) the irreducible representation of \(KG\) given by right multiplication on \(R_i\). Clearly,

\[
e_i \delta_j = \begin{cases} \text{id}_{R_i} & \text{if } i = j \\ O & \text{otherwise} \end{cases} \quad \text{for all } i, j \in \mathbb{H}.
\]

There are uniquely determined scalars \(c_g \in K\) such that \(e_i = \sum_{g \in G} c_g g\). We will express these in terms of the irreducible representation \(\delta_i\). To this end fix \(x \in G\) and represent the endomorphism \((e_i x^{-1})\rho\) of \(KG\) by matrices with respect to \(\mathfrak{B}\) and \(\mathfrak{C}\):

\[
(e_i x^{-1})\rho = \sum_{g \in G} c_g (gx^{-1})\rho \quad = \quad c_x \cdot \text{id}_{KG} + \sum_{g \in G \setminus \{x\}} c_g (gx^{-1})\rho
\]

\[
\mathfrak{B} \quad \sim \quad \begin{pmatrix} c_x & 0 & \cdots \\ O & \ddots & \cdots \\ \cdots & \ddots & c_x \end{pmatrix} + \begin{pmatrix} 0_K & \star & \cdots \\ \star & \ddots & \cdots \\ \cdots & \ddots & 0_K \end{pmatrix} = \begin{pmatrix} c_x & \star & \cdots \\ \star & \ddots & \cdots \\ \cdots & \ddots & c_x \end{pmatrix}
\]

\[
\mathfrak{C} \quad = \quad (e_i \rho)(x^{-1}\rho) \sim \begin{pmatrix} [O] \\ & \cdots & M(x^{-1}\delta_i) \\ & O & \cdots \\ & \cdots & \ddots & \cdots \\ O & \cdots & \cdots & [O] \end{pmatrix},
\]

where \(M(x^{-1}\delta_i)\) is a matrix which represents the endomorphism \(x^{-1}\delta_i\) of the irreducible module \(R_i\): \(e_i \rho\) annihilates all minimal ideals \(\neq J_i\) and induces the identity mapping on \(J_i\) (7.3.2). Therefore the composition \((e_i \rho)(x^{-1}\rho)\) is represented with respect to \(\mathfrak{C}\) by a block diagonal matrix with exactly \(n_i\) blocks \(M(x^{-1}\delta_i)\) and zero matrices \([O]\) otherwise.

120
Both representing the same endomorphism \((e, x^{-1})\rho\), our matrices must be similar and, in particular, have the same trace. It follows that \(|G|e_z = n_i \text{tr}(x^{-1} \delta_i)\), hence
\[
\forall g \in G \quad c_g = n_i (|G|1_K)^{-1} \text{tr}(g^{-1} \delta_i).
\]

**7.4 Definition.** Let \(G\) be a finite group, \(K\) a field, \(\delta\) a representation of \(KG\). The linear mapping
\[
\chi : KG \to K, \; a \mapsto \text{tr}(a \delta)
\]
is called the **character** of \(\delta\). A linear mapping of \(KG\) into \(K\) is called an (irreducible) character \([\text{of degree} \; n]\) of \(G\) over \(K\) if there exists an \(\text{(irreducible) representation of} \; KG\; \text{[of degree} \; n]\) with that character function. The character associated with the one-dimensional representation given by the trivial homomorphism of \(G\) into \(\hat{K}\), \(g \mapsto 1_K\) for all \(g \in G\), is called the **trivial character** and denoted by \(\chi_1\). As we consider only unital representations, the character of \(1_G\) is always given by the trace of the identity endomorphism of the module, i.e.,

**7.4.1. If \(\chi\) is the character of \(\delta\), then \(\chi(1_G) = (\text{deg} \; \delta)1_K\).**

By 7.3(2), \((*)\) and 7.3.2 we obtain

**7.4.2. If \text{char} \; K \nmid |G| \; \text{and, for each} \; i \in \mathbb{N}, \; \varepsilon_i \; \text{is the unit element of the minimal ideal} \; J_i \; \text{of} \; KG, \; R_i, \; \delta_i, \; n_i \; \text{as in the corollary of 6.11 (with} \; A = KG\), \(\chi_i\) the corresponding irreducible character, then

\[
n_i (|G|1_K)^{-1} \sum_{g \in G} \chi_i(g^{-1})g = e_i,
\]
\[
(*)
\]
\[
n_i (|G|1_K)^{-1} \sum_{g \in G} \chi_i(g^{-1})g \delta_j = \begin{cases} 
\text{id}_{R_i} & \text{if} \; i = j, \\
O & \text{otherwise}
\end{cases}
\]

If, additionally, \(K\) is algebraically closed, then \(\chi_i(1_G) = n_i \cdot 1_K\), \text{char} \; K \nmid n_i, \; \text{and the} \; \chi_i \; \text{are mutually distinct.}

The last assertion holds by 7.4.1 because \(J_i \cong K^{n_i \times n_i}\), if \(K\) is algebraically closed, by 6.12.5, and therefore \(n_i = \dim_K R_i\). Now by \((*)\), \(e_i \) is determined by the values of \(\chi_i\). Therefore \(\chi_i = \chi_j \) implies \(e_i = e_j\), i.e. \(i = j\). As \(e_i \neq 0_{KG}\), \text{char} \; K \nmid n_i.

**7.4.3. If \(\delta, \delta'\) are equivalent representations of \(KG\) and \(\chi, \chi'\) their characters, then \(\chi = \chi'\).**

*Proof.* Let \(V, V'\) be the modules of \(\delta, \delta'\) and \(\varphi\) a \(KG\)-isomorphism of \(V\) onto \(V'\), \(g \in G\). Then \((g \delta) \varphi = \varphi(g \delta')\), hence \(\varphi^{-1}(g \delta) \varphi = g \delta'\) showing that the matrices representing \(g \delta, g \delta'\) are similar. In particular, \(\chi(g) = \chi'(g)\).

**7.4.4. If \(\chi\) is a character of \(G\) over \(K\) and \(g, g'\) are conjugate elements of \(G\), then \(\chi(g) = \chi(g')\).**
Proof. Let $\delta$ be a representation of $KG$ with character $\chi$, $x \in G$ such that $g^\prime = g^x$. Then

$$\chi(g^\prime) = \text{tr}(g^x \delta) = \text{tr}((x\delta)^{-1}(g \delta)(x\delta)) = \text{tr}(g \delta) = \chi(g).$$

\[\square\]

A class function of $G$ in $K$ is a $K$-linear mapping of $KG$ into $K$ which is constant on every conjugacy class of $G$. The set of all class functions of $G$ into $K$ is denoted by $\text{Cl}_K(G)$.

7.4.5. $\text{Cl}_K(G)$ is a $K$-subspace of the dual space $(KG)^*$ containing all characters of $G$ over $K$, and $\dim_K \text{Cl}_K(G) = h(G)$.

Proof. Sums and scalar multiples of class functions are obviously class functions. Let $C_1, \ldots, C_{h(G)}$ be the conjugacy classes of $G$ and for each $i \in h(G)$ define $f_i$ to be the $K$-linear mapping of $KG$ into $K$ with

$$g \mapsto \begin{cases} 1_K & \text{if } g \in C_i \\ 0_K & \text{otherwise} \end{cases} \text{ for all } g \in G.$$ 

Then $\{f_1, \ldots, f_{h(G)}\}$ is a $K$-basis of $\text{Cl}_K(G)$. The last assertion holds by 7.4.4. \[\square\]

If $N \trianglelefteq G$ and $f$ is a function defined on $G/N$, there is a function $f_G$ induced by $f$ on $G$ in a natural way: Set $gf_G := (Ng)f$ for all $g \in G$, i.e., $f_G$ is the composition of the canonical epimorphism of $G$ onto $G/N$ and $f$. As conjugacy of elements in $G$ implies conjugacy of their cosets modulo $N$, we have:

7.4.6. Let $N \trianglelefteq G$, $f, f' \in \text{Cl}_K(G/N)$. Then $f_G \in \text{Cl}_K(G)$. If $f$ is an [irreducible] character of $G/N$, then $f_G$ is an [irreducible] character of $G$. Furthermore, $(f + f')_G = f_G + f'_G$.

Proof. We need to consider only the statements about characters: If $(V, \delta)$ is a $G/N$-module and $f$ its character, we set $g\delta_G := (Ng)\delta$ for all $g \in G$. Then $(V, \delta_G)$ is a $G$-module and $f_G$ its character. \[\square\]

In the sequel, we will confine ourselves to the case $K = \mathbb{C}$. Then $h(G) = h$, the number of irreducible representations of $\mathbb{C}G$, and $n_1, \ldots, n_h$ are their degrees (see Corollary of 7.2 and 7.3(2)). We derive quite a number of properties of the characters:

The characters of degree 1 are obviously the homomorphisms of $G$ into the multiplicative group $\mathbb{C}$. For a cyclic group we therefore obtain:

7.4.7. If $G$ is cyclic, $x$ a generator of $G$, $n := |G|$ and $w \in \mathbb{C}$ a primitive $n$-th root of unity, then the irreducible complex characters of $G$ are given by

$$\chi_i : x^k \mapsto w^{(i-1)k} \text{ for all } k \in \mathbb{Z},$$

where $i$ ranges over $\{1, \ldots, n\}$. \[\square\]
7.5 Proposition. Let \( G \) be a finite group, \((V, \delta)\) a \(\mathbb{C}G\)-module of dimension \(m\) with character \(\chi\). Let \( x \in G, \ n \in \mathbb{N} \) such that \(x^n = 1\), \( w \) a primitive \(n\)-th root of unity. Then we have:\(^{111}\)

1. \(x\delta\) is diagonalizable: There exists a basis tuple \(\mathfrak{B}\) of \(V\) such that there are \(i_1, \ldots, i_m \in \mathbb{N}\) with the property that \(x\delta \sim \mathfrak{B} \begin{pmatrix} w^{i_1} & \cdots & O \\ \vdots & \ddots & \vdots \\ O & \cdots & w^{i_m} \end{pmatrix}\).

2. If \(\sigma \in \text{Aut} \mathbb{Q}[w]\) and \(j \in \mathbb{N}\) such that \(w\sigma = w^j\), then \(\chi(x)\sigma = \chi(x^j)\). In particular, \(|\chi(x)\sigma| \leq m\) for all automorphisms \(\sigma\) of \(\mathbb{Q}[w]\).

3. \(|\chi(x)| = m \Leftrightarrow \exists i \in \mathbb{N} \ x\delta \sim \mathfrak{B} \begin{pmatrix} w^i & \cdots & O \\ \vdots & \ddots & \vdots \\ O & \cdots & w^i \end{pmatrix}\)\(^{112}\).\(^{113}\)

4. \(\chi(x) = m \Leftrightarrow x\delta = \text{id}_V \Leftrightarrow \chi(x) = \chi(1_G)\).

5. \(\chi(x^{-1}) = \overline{\chi(x)}\).

Proof. (1) Without loss of generality we assume \(o(x) = n\). By 7.3(2), the restriction of \(\delta\) to \(\langle x \rangle\) is completely reducible which, by 7.4.7, implies that there are 1-dimensional \(\langle x \rangle\)-submodules \(V_1, \ldots, V_m\) of \(V\) such that \(V = V_1 \oplus \cdots \oplus V_m\). On each \(V_k\), \(x\delta\) induces the multiplication by some \(n\)-th root of unity. This proves the first claim in (1).

(2) By (1), \(\chi(x)\sigma = (w^{i_1} + \cdots + w^{i_m})\sigma = w^{i_1j} + \cdots + w^{imj} = \chi(x^j)\). The second assertion follows by means of the triangle inequality.

(3) The implication "\(\Leftarrow\)" is trivial. If \(w^{i_j} \neq w^{i_m}\) for some \(j < m\) in the matrix representation of \(x\delta\) from (1), then \(\frac{w^{i_j} + w^{i_m}}{2}\) is an inner point of the unit circle. Thus \(|w^{i_j} + w^{i_m}| < 2\) and \(|w^{i_1} + \cdots + w^{i_m}| \leq \sum_{k \neq j,m} |w^{i_k}| + |w^{i_j} + w^{i_m}| < (m - 2) + 2 = m\).

(4) As \(\chi(1_G) = m\), we have the following chain of trivial equivalences:

\[
\chi(x) = \chi(1_G) \Leftrightarrow \chi(x) = m \Leftrightarrow |\chi(x)| = m \text{ and } \chi(x) \in \mathbb{R}_{>0}
\]

\[
\Leftrightarrow x\delta \sim \begin{pmatrix} 1 & \cdots & O \\ \vdots & \ddots & \vdots \\ O & \cdots & 1 \end{pmatrix} \Leftrightarrow x\delta = \text{id}_V.
\]

(5) Let \(i_1, \ldots, i_m\) as in (1). As \(w^{-1} = \overline{w}\), we obtain

\[
\chi(x^{-1}) = w^{-i_1} + \cdots + w^{-i_m} = \overline{w^{i_1}} + \cdots + \overline{w^{i_m}} = \overline{\chi(x)}.
\]

\(^{111}\)The field \(\mathbb{C}\) could be replaced here by any subfield containing \(w\).

\(^{112}\)i.e., \(x\delta\) is the multiplication by the scalar \(w^i\).

\(^{113}\)
7.6 Corollary. Let $G$ be a finite group, $\chi, \chi^{(1)}, \ldots, \chi^{(k)}$ characters of $G$ such that $\chi = \sum_{i \in H} \chi^{(i)}$. Let $x \in G$ such that $\chi(x) = \chi(1_G)$. Then $\chi^{(i)}(x) = \chi^{(i)}(1_G)$ for all $i \in H$.

Proof. $\chi(1_G) = \chi(x) = \sum_{i \in H} \chi^{(i)}(x) \leq \sum_{i} |\chi^{(i)}(x)| \leq \sum_{i} \chi^{(i)}(1_G) = \chi(1_G)$ by 7.5(4),(2) and 7.4.1. Hence $\chi^{(i)}(x) = \chi^{(i)}(1_G)$ for all $i$. 

Let $C_1, \ldots, C_h$ be the conjugacy classes of $G$, $x_1 \in C_1, \ldots, x_h \in C_h$, and $x_{h+1}, \ldots, x_{|G|}$ the remaining elements of $G$. Let $\chi_1, \ldots, \chi_h$ be the irreducible characters of $G$. The matrix $(\chi_r(x_s))_{r \in H, s \in G}$ is then called the complete (reduced resp.) character table of $G$. If we write, as usual, $\chi_1$ for the trivial character and put $n_r := \deg \chi_r$, $C_1 := \{1_G\}$, we know at least the following of its entries:

|   | $1_G$ | $x_2$ | $\cdots$ | $x_h$ | $x_{h+1}$ | $\cdots$ | $x_{|G|}$ |
|---|---|---|---|---|---|---|---|
| $\chi_1$ | 1 | 1 | $\cdots$ | 1 | 1 | $\cdots$ | 1 |
| $\chi_2$ | $n_2$ | | | | | | |
| $\vdots$ | | $\vdots$ | | | | | |
| $\chi_h$ | $n_h$ | | | | | | |

The rows and columns of the character table have a remarkable property with respect to the standard hermitian form on $\mathbb{C}^n$ (where $n \in \mathbb{N}$),

$$\langle \cdot, \cdot \rangle : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}, ((c_1, \ldots, c_n), (c'_1, \ldots, c'_n)) \mapsto c_1 \overline{c_1} + \cdots + c_n \overline{c_n}$$

7.7 Theorem (Frobenius 1896). Let $G$ be a finite group, $C_i$, $x_i$, $\chi_i$ as above, $H := \{\chi_1, \ldots, \chi_h\}$.

(1) (The first orthogonality relations)

$$\frac{1}{|G|} \sum_{g \in G} \chi_i(g) \chi_j(g) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases} \text{ for all } i, j \in H$$

(2) (The second orthogonality relations)

$$\frac{1}{|G|} \sum_{\chi \in H} \chi(x_i) \overline{\chi(x_j)} = \begin{cases} \frac{1}{n_i} & \text{if } i = j \\ 0 & \text{otherwise} \end{cases} \text{ for all } i, j \in H$$

Proof. (1) Let $i, j \in H$, $n_j := \deg \chi_j$. For the unit element of the minimal ideal $J_j$ of $\mathbb{C}G$ we have, by 7.4.2 and 7.5(5),

$$e_j = \frac{n_j}{|G|} \sum_{g \in G} \chi_j(g)g.$$ 

Hence

$$\frac{1}{|G|} \sum_{g \in G} \chi_i(g) \chi_j(g) = \chi_i \left( \frac{1}{|G|} \sum_{g \in G} \chi_j(g)g \right) = \frac{1}{n_j} \chi_i(e_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

124
as $\chi_j(e_j) = n_j$, $R_i e_j = 0_{CG}$ if $i \neq j$ (where $R_i$ is a minimal right ideal of $CG$ contained in $J_i$).

(2) By (1) and 7.4.4 we have for all $i, j \in h$

$$\frac{1}{|G|} \sum_{s \in h} |C_s| \chi_i(x_s) \overline{\chi_j(x_s)} = \frac{1}{|G|} \sum_{g \in G} \chi_i(g) \overline{\chi_j(g)} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}.$$  

This may be read as the calculation of a matrix product:

$$\left( \chi_i(x_j) \right)_{(i,j) \in h \times h} \cdot \left( \frac{|C_i|}{|G|} \chi_j(x_i) \right)_{(i,j) \in h \times h} = \begin{pmatrix} 1 & 0 \\ \vdots & \ddots \\ 0 & \cdots & 1 \end{pmatrix}.$$  

But for quadratic matrices, a right inverse is also a left inverse\footnote{If $A, B$ are quadratic matrices over a field and $AB$ is the unit matrix, then $\det A, \det B \neq 0$, hence $A, B$ are elements of the group of invertible matrices in which a one-sided inverse is always a two-sided inverse.}. Hence

$$\left( \frac{|C_i|}{|G|} \chi_j(x_i) \right)_{(i,j) \in h \times h} \cdot \left( \chi_i(x_j) \right)_{(i,j) \in h \times h} = \begin{pmatrix} 1 & 0 \\ \vdots & \ddots \\ 0 & \cdots & 1 \end{pmatrix},$$

i.e.,

$$\frac{|C_i|}{|G|} \sum_{s \in h} \chi_s(x_i) \overline{\chi_j(x_j)} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$  

for all $i, j \in h$

which proves our claim. \hfill \Box

**7.8 Corollary.** With respect to the standard hermitian form, the column vectors of the reduced character table form an orthogonal basis of $\mathbb{C}^h$. $H$ is an orthonormal basis of the subspace $Cl_C(G)$ of the vector space $\mathbb{C}^G$, the latter being endowed with the following non-degenerate hermitian form:

$$\langle \cdot | \cdot \rangle_G : (f, f') \mapsto \frac{1}{|G|} \sum_{g \in G} f(g) \overline{f'(g)}.$$  

**Proof.** Both statements follow from 7.7 because sets of mutually orthogonal non-zero vectors in a vector space with a non-degenerate hermitian form are linearly independent, and $\dim \mathbb{C} Cl_C(G) = h$ by 7.4.5. \hfill \Box

The first statement of 7.8 implies that also the row vectors of the reduced character table form a basis of $\mathbb{C}^h$. The second statement means that the row vectors of the complete character table are mutually orthogonal.
7.9 Corollary. Let \( c_1, \ldots, c_h \in \mathbb{C} \) and \( f = \sum_{i \in \mathcal{H}} c_i \chi_i \). Then \( f \) is a character of \( G \) if and only if \( c_i \in \mathbb{N}_0 \) for all \( i \in \mathcal{H} \).

Proof. For all \( i \in \mathcal{H} \), let \( V_i \) be a \( \mathbb{C}G \)-module with character \( \chi_i \). Suppose first that \( c_i \in \mathbb{N}_0 \) for all \( i \in \mathcal{H} \). Clearly, the character of the \( \mathbb{C}G \)-module \( \bigoplus_{i \in \mathcal{H}} V_i \) is \( f \). Conversely, let \( f \) be the character of a \( \mathbb{C}G \)-module \( V \). By 7.3(2), \( V \) is completely reducible. Let \( n_i \in \mathbb{N}_0 \) be the number of modules which are \( \mathbb{C}G \)-isomorphic to \( V_i \) in a direct decomposition of \( V \) into irreducible submodules. Then \( f = \sum_{i \in \mathcal{H}} n_i \chi_i \). By 7.8, \( \mathcal{H} \) is linearly independent. Hence \( c_i = n_i \in \mathbb{N}_0 \) for all \( i \in \mathcal{H} \).

By 7.5(1), character values are sums of roots of unity, hence algebraic numbers of a rather special form. We digress for a moment to insert a few basic remarks about a classical concept which will turn out to be most useful in our context:

An insertion about algebraic integers

1. E Definition A complex number \( x \) is called an algebraic integer if there exists a normed polynomial \( f \in \mathbb{Z}[t] \) such that \( f(x) = 0 \). The set of all algebraic integers is denoted by \( \mathbb{G} \).

1.1 E \( \mathbb{G} \cap \mathbb{Q} = \mathbb{Z} \).

Proof. “\( \supseteq \)” is obvious. “\( \subseteq \)”: Let \( x \in \mathbb{G} \cap \mathbb{Q} \). Then there exist \( a \in \mathbb{Z}, b \in \mathbb{N} \) such that \( \gcd(a, b) = 1 \), \( x = \frac{a}{b} \), furthermore \( n \in \mathbb{N}, c_0, \ldots, c_{n-1} \in \mathbb{Z} \) with the property that

\[
\left(\frac{a}{b}\right)^n + c_{n-1}\left(\frac{a}{b}\right)^{n-1} + \cdots + c_1 a + c_0 = 0.
\]

It follows that \( a^n = -c_{n-1}a^{n-1} - \cdots - c_1 ab^{n-2} - c_0 b^{n-1} \in \mathbb{Z} \), implying \( b = 1 \) as \( \gcd(a^n, b) = 1 \). Thus \( x \in \mathbb{Z} \).

1.2 E If \( L \) is a subfield of \( \mathbb{C} \), \( x \in \mathbb{G} \cap L \), then \( x\sigma \in \mathbb{G} \) for every monomorphism \( \sigma \) of \( L \) into \( \mathbb{C} \). In particular, \( x \in \mathbb{G} \).

This is clear because if \( f \in \mathbb{Z}[t] \) and \( f(x) = 0 \), then also \( f(x\sigma) = 0 \).

2 E Proposition For any \( x \in \mathbb{C} \), the following conditions are equivalent:

(i) \( x \in \mathbb{G} \),

(ii) There exists a finitely generated subgroup \( U \neq \{0\} \) of \( (\mathbb{C},+) \) such that \( xU \subseteq U \).

\(^{114}\) The number \( n_i \) is commonly called the multiplicity of the irreducible character \( \chi_i \) in \( f \), of \( V_i \) in \( V \) resp. By 7.8, the multiplicity is given by \( \langle f|\chi_i\rangle_G \).
Proof. (i)⇒(ii): If \( c_0, \ldots, c_{n-1} \in \mathbb{Z} \) such that \( x^n + c_{n-1}x^{n-1} + \cdots + c_1x + c_0 = 0 \), set \( U := \langle x^n-1, \ldots, x, 1 \rangle_{\mathbb{Z}} \). We have \( U \neq \{0\} \) and \( x \cdot x^n-1 = x^n \in U \) whence \( xU \subseteq U \).

(ii)⇒(i): Let \( u_1, \ldots, u_n \in U \) such that \( U = \langle u_1, \ldots, u_n \rangle_{\mathbb{Z}}, u_1 \neq 0 \). As \( xU \subseteq U \), for each \( i \in \mathbb{U} \) there exist integers \( z_1, \ldots, z_m \) such that \( xu_i = \sum_{j \in \mathbb{U}} z_{ij} u_j \). Hence the \( n \)-tuple \( (u_1, \ldots, u_n) \) is a nontrivial solution of the following system of linear equations:

\[
\begin{align*}
(x - z_{11})t_1 &\quad -z_{12}t_2 &\quad \cdots &\quad -z_{1n}t_n = 0 \\
-z_{21}t_1 &\quad +(x - z_{22})t_2 &\quad \cdots &\quad -z_{2n}t_n = 0 \\
\vdots &\quad \ddots &\quad \ddots &\quad \vdots \\
-z_{n1}t_1 &\quad -z_{n2}t_2 &\quad \cdots &\quad (x - z_{nn})t_n = 0
\end{align*}
\]

It follows that the determinant of the matrix of its coefficients is 0. Hence \( x \) is a root of the polynomial

\[
\det \begin{pmatrix}
t - z_{11} & \cdots & -z_{1n} \\
\vdots & \ddots & \vdots \\
-z_{n1} & \cdots & t - z_{nn}
\end{pmatrix}
\]

which is a normed polynomial with integer coefficients. \( \square \)

3 E Corollary \( \mathbb{G} \) is a subring of \( \mathbb{C} \).

Proof. Let \( x, y \in \mathbb{G} \). We have to show that \( x - y, xy \in \mathbb{G} \). By E2, there exist \( m, n \in \mathbb{N}, u_1, \ldots, u_n, v_1, \ldots, v_m \in \mathbb{G} \) such that

\[
x \langle u_1, \ldots, u_n \rangle_{\mathbb{Z}} \subseteq \langle u_1, \ldots, u_n \rangle_{\mathbb{Z}}, \quad y \langle v_1, \ldots, v_m \rangle_{\mathbb{Z}} \subseteq \langle v_1, \ldots, v_m \rangle_{\mathbb{Z}}.
\]

Set \( T := \langle \ldots, u_i v_j, \ldots \rangle_{\mathbb{Z}} (i \in \mathbb{W}, j \in \mathbb{W}) \). Then

\[
(x - y)u_i v_j = (xu_i)v_j - u_i(yv_j) \in T, \quad (xy)u_i v_j = (xu_i)(yv_j) \in T
\]

for all \( i, j \). As \( T \neq \{0\} \), this proves the claim, by E2. \( \square \)

It is easily seen that for every algebraic number \( a \) there exists a positive integer \( n \) such that \( na \in \mathbb{G} \). In particular, the quotient field of \( \mathbb{G} \) (in \( \mathbb{C} \)) is the field of all algebraic numbers. More generally, if \( L \) is an algebraic subfield

\[
G \cap L \subseteq \mathbb{Q} \subseteq L \subseteq \mathbb{C}
\]

of \( \mathbb{C} \), then \( L \) is the quotient field of \( \mathbb{G} \cap L \). The multiplicative structure of \( \mathbb{G} \cap L \) for finite extensions \( L \) of \( \mathbb{Q} \) is of particular interest and the object of Algebraic Number Theory. While the ground case of \( L = \mathbb{Q} \) is settled by the theorem on the unique prime decomposition, its obvious generalization to extension fields \( L \) of \( \mathbb{Q} \) does not hold true, which opens a wide area of study with many deep and satisfying results. We restrict our digression to these few remarks and conclude it with the following detail to be applied in our discussion of characters:

4 E Proposition Let \( L \) be a Galois extension field of \( \mathbb{Q} \) and \( y \in \mathbb{G} \cap L \). Then

\[
\prod_{\sigma \in \text{Aut } L} y\sigma \in \mathbb{Z}, \quad \forall r \in \mathbb{R}_{\geq 0} \quad \sum_{\sigma \in \text{Aut } L} |y\sigma|^r \geq \text{dim}_\mathbb{Q} L.
\]
Proof. The first assertion follows from E1.1 because the product is in \( G \) by E3 and E1.2, and in \( Q \) by 4.9 because it is fixed by each automorphism of \( L \). Now \( y \neq 0 \) implies that \( \prod_{\sigma \in \text{Aut}_L} |y\sigma| \geq 1 \). Putting \( n := \dim_Q L \), the inequality between the geometric and the arithmetic mean values implies \( \frac{1}{n} \sum_{\sigma \in \text{Aut}_L} |y\sigma|^r \geq \sqrt[n]{\prod_{\sigma \in \text{Aut}_L} |y\sigma|^r} \geq 1 \) for all \( r \in \mathbb{R}_{\geq 0} \). \( \square \)

We return to our discussion of characters by observing the following obvious consequence of 7.5(1) and E3:

7.9.1. Each character value of a group element is an algebraic integer. \( \square \)

7.10 Proposition. Let \( H \) be a finite cyclic group, \( E \) the set of its generators, \( \chi \) a character such that \( \chi(E) \neq \{0\} \). Then

\[
(1) \sum_{g \in E} |\chi(g)|^r \geq |E| \text{ for all } r \in \mathbb{R}_{\geq 0}.
\]

\[
(2) \text{If } x \in E, m = \deg \chi \text{ and } \frac{\chi(x)}{m} \in \mathbb{G}, \text{ then } |\chi(x)| = m.
\]

Proof. Let \( x \in E, n := |H|, w \) a primitive \( n \)-th root of unity, \( L = Q[w] \). Then \( E = \{ x^i | i \in \mathbb{Z} \text{ gcd}(i, n) = 1 \} \). For each \( i \in \mathbb{Z} \) such that \( \gcd(i, n) = 1 \), there is a unique automorphism \( \sigma_i \) of \( L \) with the property \( w \sigma_i = w^i \). Furthermore, \( \text{Aut}_L \) is the set of these \( \sigma_i \). Hence, in view of 7.5(2), \( \chi(x) \neq 0 \), \( \sum_{\sigma \in E} |\chi(g)|^r = \sum_{\gcd(i,n)=1} |\chi(x)\sigma_i|^r \geq \dim_Q L = |E| \) for all \( r \in \mathbb{R}_{\geq 0} \), making use of E4 and 7.9.1. This proves (1).

(2) If \( m = \deg \chi \) and \( y := \frac{\chi(x)}{m} \in \mathbb{G} \), E4 implies that \( \sum_{\sigma \in \text{Aut}_L} |y\sigma| \geq \dim_Q L = |\text{Aut}_L| \) by 4.7.

But \( |y\sigma| \leq 1 \) for all \( \sigma \in \text{Aut}_L \), by 7.5(2), hence \( |\chi(x)| = m \). \( \square \)

7.11 Theorem. Let \( G \) be a finite group, \( \delta \) an irreducible representation of \( CG \), \( m \) its degree and \( \chi \) its character. Let \( x \in G \) and \( k := |x^G| \).

\[
(1) \frac{k}{m} \chi(x) \in \mathbb{G}.
\]

\[
(2) \text{If } \delta \text{ is faithful, } k \neq 1 \text{ and } \gcd(k, m) = 1, \text{ then } \chi(x) = 0.
\]

\[
(3) \ |\{g|g \in G, \chi(g) = 0\}| \geq m^2 - 1.
\]

Proof. \( (1) \) Let \( V \) be a \( CG \)-module with character \( \chi \). Each \( \alpha \in \text{End}_{KG} V \), in particular, each \( \alpha \in (Z(CG))^{\delta} \), is then the multiplication by some complex number \( z_\alpha \), by 6.7.1 and 6.12.5. We obtain an algebra epimorphism of \( Z(CG) \) onto \( \mathbb{C} \) by the composition

\[
\zeta : y \mapsto y^\delta \mapsto z_y^\delta \quad (y \in Z(CG)).
\]

Let \( K \) be the set of all conjugacy classes of \( G \), \( U := \langle K \rangle Z, C := x^G \). By Corollary 1 of 7.2, \( U\zeta \) is a finitely generated subgroup \( \neq \{0\} \) of \( (\mathbb{C}, +) \), and \( (C \zeta)U \zeta \subseteq U\zeta \). By E2,
\( \hat{C} \zeta \in G \). As the endomorphism \( \hat{C} \delta \) is given by the multiplication by the scalar \( \hat{C} \zeta \), we obtain \( \chi(\hat{C}) = m \cdot \hat{C} \zeta \). Hence \( \frac{k \chi(x)}{m} = \frac{\chi(\hat{C})}{m} = \hat{C} \zeta \in G \).

(2) The hypothesis implies that there are \( a, b \in \mathbb{Z} \) such that \( am + bk = 1 \). By (1) and E3,

\[
\frac{\chi(x)}{m} = \frac{(am + bk)\chi(x)}{m} = a\chi(x) + b\frac{k\chi(x)}{m} \in G.
\]

Assume \( \chi(x) \neq 0 \). Then, by 7.10(2), \( |\chi(x)| = m \) so that, by 7.5(3), \( x\delta \) commutes with every module endomorphism. As \( \delta \) is faithful, \( x \) commutes then with each element of \( G \), hence \( k = 1 \).

(3) Writing \( E_H \) for the set of generators of a cyclic subgroup \( H \) of \( G \), we have either \( \chi(E_H) = \{0\} \) or \( 0 \notin \chi(E_H) \) in view of the argument given in the proof of 7.10(1). By 7.7(1) (case \( i = j \)), it follows that

\[
|\{g | g \in G, \chi(g) = 0\}| = \sum_{g \in G, \chi(g) \neq 0} (|\chi(g)|^2 - 1) = \sum_{H \in C'} (\sum_{g \in E_H} |\chi(g)|^2) - |E_H|
\]

where \( C' \) is the set of all cyclic subgroups \( H \) of \( G \) such that \( 0 \notin \chi(E_H) \). By 7.10(1) (with \( r = 2 \)), all differences which occur as summands in the last sum are non-negative, and one of them (for \( H = \{1_G\} \)) is \( m^2 - 1 \).

**7.12 Corollary.** Let \( G \) be a finite group and \( m \) the degree of an irreducible representation of \( CG \). Then \( m | |G| \).

**Proof.** From 7.7 we obtain \( \frac{1}{|G|} \sum_{g \in G} \chi_j(g)\overline{\chi_j(g)} = 1 \) for all \( j \in B \). Hence

\[
\frac{|G|}{m} = \sum_{j \in B} \sum_{g \in C_j} \frac{\chi_j(g)}{m} \overline{\chi_j(g)} = \sum_{j \in B} \sum_{g \in E_{C_j}} \frac{|C_j|\chi(x_j)}{m} = \sum_{g \in G} \frac{\chi(x)}{m} 
\]

by E3. By E1, \( \frac{|G|}{m} \in \mathbb{Z} \).

We mention without proof the following stronger result:

**Theorem (Ito 1951)** Under the hypotheses of 7.12, \( m ||G/A|| \) for every abelian normal subgroup \( A \) of \( G \).

**7.13 Corollary.** Let \( G \) be a finite simple non-abelian group. Then there is no conjugacy class \( C \neq \{1_g\} \) of \( G \) for which \( |C| \) is a prime power.

**Proof.** Assume that \( C \neq \{1_g\} \) and \( |C| \) is a power of a prime \( p \). The hypotheses on \( G \) imply that \( |C| \neq 1 \) and every irreducible nontrivial representation is faithful. Let \( x \in C \) and \( \chi \) be an irreducible character of \( CG \). From 7.11(2) we obtain

\[
(*) \quad \chi \neq \chi_1, \; p \nmid \deg \chi \Rightarrow \chi(x) = 0.
\]
Now 7.7(2) implies that

\[ 0 = \sum_{j \in \mathcal{H}} \chi_j(x) \bar{\chi}_j(1_G) = \sum_{j \in \mathcal{H}} n_j \chi_j(x) = 1 + \sum_{p \mid n_j} n_j \chi_j(x), \]

where 1 is the summand for \( j = 1 \). It follows that \( \frac{1}{p} = -\sum_{p \mid n_j} \frac{n_j}{p} \chi_j(x) \in \mathbb{G} \), by 7.9.1 and E3. By E1.1, we obtain the contradiction \( \frac{1}{p} \in \mathbb{Z} \).

\[ \square \]

The results 7.11 and 7.13 are essentially due to Burnside (1904, 1911). They play an important role in various deeper investigations in finite group theory. Our next chapter will provide illustrating and important examples of applications (8.3, 8.8).
8 Group actions

For every set $X$, the set $X^X$ of all self-mappings of $X$ is a monoid with respect to the composition of mappings. The unit group of the monoid $X^X$ is the group $S_X$ of all permutations of $X$.

8.1 Definition. Let $M$ be a monoid, $X$ a set. A (right) action of $M$ on $X$ is a homomorphism $\varphi$ of $M$ into $X^X$. An action $\varphi$ of $M$ on $X$ is called unital if $1_M \varphi = \text{id}_X$. We say “$M$ acts on $X$ (via $\varphi$)” if an action $\varphi$ of $M$ on $X$ is given. If $\varphi$ is injective, the action is called faithful. Set $\tilde{M} := M \varphi$. For every $x \in X$ the set $x\tilde{M} := \{y | y \in X, \exists a \in M \ y = x(a\varphi)\}$ is called the orbit of $x$ under $M$ (via $\varphi$), $|x\tilde{M}|$ its length, and

$$C_{M,\varphi}(x) := \{a | a \in M, \ x(a\varphi) = x\}, \ C_{M,\varphi}(T) := \bigcap_{x \in T} C_{M,\varphi}(x) \text{ for all } T \subseteq X$$

the centralizer of $x$, $T$ resp. in $M$.

8.1.1. Let $\varphi$ be an action of $M$ on $X$, $x \in X$. Then $C_{M,\varphi}(x)$ is a unital submonoid of $M$. If $M$ is a group, then $C_{M,\varphi}(x)$ is a subgroup of $M$ and the set of all orbits under $M$ is a partition of $X$.

Proof. The first statement is trivial. If $M$ is a group, then $\tilde{M}$ is a subgroup of $X^X$ containing $\text{id}_X = 1_M \varphi$, hence a subgroup of $S_X$. For all $x \in X$, $a \in C_{M,\varphi}(x)$, we have $x = x \text{id} = x(\text{aa}^{-1})\varphi = x(a\varphi)(a^{-1}\varphi) = x(a^{-1}\varphi)$, hence $a^{-1} \in C_{M,\varphi}(x)$. The relation $\sim$, defined by $x \sim y \iff \exists a \in M \ x(a\varphi) = y$, is reflexive and transitive for every monoid $M$, in the case of a group $M$, however, also symmetric, thus an equivalence relation. Its equivalence classes are exactly the orbits of $X$ under $M$. □

8.1.2. Every action $\varphi$ of $M$ on a set $X$ induces an action $\tilde{\varphi}$ on the power set $\mathcal{P}(X)$ by setting

$$a\tilde{\varphi} : \mathcal{P}(X) \to \mathcal{P}(X), \ T \mapsto \{x(a\varphi) | x \in T\}.$$ 

$\square$

Mostly, the set $T(a\tilde{\varphi})$ is simply written as $T(a\varphi)$. The following warning, however, is important: $C_{M,\varphi}(T) = \{a | a \in M, \ T(a\varphi) = T\}$ is not the same as $C_{M,\tilde{\varphi}}(T)$ because the equation $T(a\varphi) = T$ is considerably weaker than the condition that $x(a\varphi) = x$ for all $x \in T$. In terms of $\varphi$ (instead of $\tilde{\varphi}$), $C_{M,\tilde{\varphi}}(T)$ is called the normalizer of $T$ with respect
to \( \varphi \) and denoted by \( N_{M,\varphi}(T) \). The elements which the action \( \tilde{\varphi} \) refers to are sets of elements which the action \( \varphi \) refers to. As \( N_{M,\varphi}(T) = C_{M,\varphi}(T) \), by applying 8.1.1 to \( \tilde{\varphi} \) we obtain that \( N_{M,\varphi}(T) \) is a unital submonoid of the monoid \( M \), and a subgroup if \( M \) is a group.\(^{118}\)

In the sequel we will always consider group actions, i.e., the monoid acting on \( X \) will be a group \( G \). If there is just one action \( \varphi \) in the discussion so that no confusion may arise, the index \( \varphi \) is usually suppressed. We then write \( C_G(x) \) instead of \( C_{G,\varphi}(x) \) etc.

8.1.3 Examples. (1) Let \((K,L)\) be a field extension, \( f \in K[t] \), \( X = \{ b | b \in L, f(b) = 0_K \} \), \( G := \text{Aut}_K L \). Then \( G \) acts on \( X \) via

\[
\varphi : G \rightarrow S_X, \quad \alpha \mapsto \alpha|_X.
\]

(2) Assumptions as before, but \( X := L \). Then \( G \) acts “naturally” on \( X \) (i.e., via \( \varphi = \text{id} \)).

For every \( Z \in \mathfrak{B}(K,L) \) we have \( C_G(Z) = \text{Aut}_Z L \), and the main theorem of Galois theory implies that each subgroup of \( G \) is a centralizer of an intermediate field of the given extension \((K,L)\) if it is Galois.

(3) Let \( K \) be a commutative unitary ring and \((V,\delta)\) a \( KG\)-algebra module. Then \( \delta|_G \) is an action of \( G \) on \( V \).

Vice versa, as \( G \) is a \( K \)-basis of \( KG \), every homomorphism of \( G \) into the automorphism group of a \( K \)-space gives rise to a representation of \( KG \) by linear extension. Therefore, such a homomorphism is also called a representation of the group \( G \) over \( K \) (cf. p. 119). There are important cases where \( V \) has a \( K \)-basis \( B \) which is invariant under the action of \( G \). Defining \( g\varphi := g\delta_B \) for all \( g \in G \), we then obtain an action of \( G \) on \( B \). A prominent example for this is the regular \( KG \)-module \((KG,\rho)\) with its basis \( G \). In fact, important \( KG \)-modules arise the other way round: If a group \( G \) acts on a set \( X \), this action extends linearly to an algebra representation of \( KG \) with the module \( KX \). As \( X \) is a \( K \)-basis of \( KX \) which is permuted by the action of \( G \), representations which arise in this way are called permutation representations.

(4) Let \( G \) be a group, \( X := G \),

\[
\varphi : G \rightarrow S_G, \quad g \mapsto \kappa_g : \begin{bmatrix} G & \rightarrow & G \\ x & \mapsto & x^g \end{bmatrix}
\]

Clearly, \( x^{(gh)} = h^{-1}g^{-1}xgh = (x^g)^h \) so that \((gh)\varphi = (g\varphi)(h\varphi)\) for all \( x, g, h \in G \). As \( x\kappa_g \) is the conjugate of \( x \) by \( g \), \( G \) is said to act via conjugation. The orbits of \( G \) are the conjugacy classes of \( G \).

As an example, we consider the conjugacy classes of the alternating group \( A_5 \): These form a partition of \( A_5 \) (see 8.1.1) of subsets of 1, 12, 12, 15, 20 resp. elements.

\(^{118}\)Of course, a direct proof of these trivialities is less sophisticated – but also less tasteful.
(represented by \(i_{12} \), (12345), (12354), (12)(34), (123) resp.). It is easily seen that the only divisors of 60 which arise as sums of some of these numbers including 1 are 1 and 60. In particular, apart from \( \{ i_{12} \} \) and \( A_5 \), no union of conjugacy classes of \( A_5 \) is a subgroup of \( A_5 \). This shows that \( A_5 \) has no nontrivial normal subgroups:

8.1.4. The group \( A_5 \) is simple.\(^{119} \)

The kernel of \( \varphi \) in (4) is called the centre of \( G \), usually denoted by \( Z(G) \). We have the following 4-fold characterization of the elements of \( Z(G) \) (also called “central elements” of \( G \)) all of whose parts are straightforward:

8.1.5. Let \( G \) be a group and \( g \in G \). Then

\[ \kappa_g = \text{id}_G \iff \forall x \in G \quad x^g = x \iff \forall x \in G \quad xg = gx \iff \forall x \in G \quad g^x = g, \]

\[ Z(G) = \bigcap_{x \in G} C_G(x) = C_G(G). \]

Here we have already used the following group theoretic convention: If in the context concerning a given group \( G \) neither a set \( X \) nor a specific action is mentioned, then “\( C_G(x) \)” (analogously “\( C_G(T) \)”, “\( N_G(T) \)” for \( T \subseteq G \)) refers to the action in (4), i.e., the action of \( G \) on \( X = G \) given by conjugation. In words, 8.1.5 means that for any group element it is equivalent to say that it leaves all group elements fixed, that it is fixed by all group elements (both with respect to conjugation), or that it commutes with all group elements. The last mentioned characterization immediately implies

8.1.6. \( Z(G \times H) = Z(G) \times Z(H) \) for arbitrary groups \( G, H \).

\(^{119}\)In fact, \( A_n \) is simple for all \( n \geq 5 \). These alternating groups form an infinite family within the class of all finite simple groups. The complete classification of the latter has been a gigantic world-wide research programme since the Feit-Thompson odd order theorem (see p. 146) and its proof methods opened new ways.

(5) Let \( G \) be a group, \( H \leq G \), \( X = G/\rho H \) the set of right cosets of \( H \) in \( G \). We have an action of \( G \) on \( X \) via right multiplication:

\[ \varphi : G \rightarrow S_X, \quad g \mapsto \rho_g : \begin{bmatrix} X & \rightarrow & X \\ H_y & \mapsto & Hyg \end{bmatrix} \]

as \( (Hy)(gg') = (Hyg)g' \), i.e., \( (gg') \varphi = (g\varphi)(g'\varphi) \) for all \( y, g, g' \in G \).

8.1.7. Let \( \varphi \) be as in (5). Then \( \ker \varphi = \bigcap_{y \in G} H^y \).

called the core of \( H \) in \( G \), the largest normal subgroup of \( G \) which is contained in \( H \):

We have \( g \in \ker \varphi \iff \forall y \in G \quad Hyg = Hy \iff \forall y \in G \quad ygy^{-1} \in H \iff \forall y \in G \quad g \in H^y. \)

Obviously, \( (Hy)(G\varphi) = X \) for any \( y \in G \). A group action \( \varphi : G \rightarrow S_X \) is called transitive if \( X \) consists of just one orbit under the action of \( G \).
8.1.8. For every subgroup $H$ of $G$, the action of $G$ on $G/_{\rho}H$ via right multiplication is transitive. ∎

If a group $G$ acts on sets $X$, $Y$ via $\varphi$, $\psi$ resp., then $\varphi$, $\psi$ are called equivalent (or similar) if there exists a bijection $\beta$ of $X$ onto $Y$ such that $(x(g\varphi))\beta = (x\beta)(g\psi)$ for all $x \in X$, $g \in G$. If this is the case, then in particular $(x(G\varphi))\beta = (x\beta)(G\psi)$, i.e., the $\beta$-image of the orbit of $x$ with respect to $\varphi$ is the orbit of $x\beta$ with respect to $\psi$. Furthermore, $C_{G,\varphi}(x) = C_{G,\psi}(x\beta)$ for all $x \in X$.

8.2 Proposition. Let $G$ be a group, $\varphi$ an action of $G$ on a set $X$, $\tilde{G} := G\varphi$, $x \in X$. Then

(1) $G$ acts transitively on the orbit $x\tilde{G}$.

(2) Let $H := C_{G,\varphi}(x)$. Then the action of $G$ on $x\tilde{G}$ is equivalent to the action of $G$ on $G/_{\rho}H$ via right multiplication. In particular, $|x\tilde{G}| = |G : H|$.

Special case: If $\varphi$ is transitive, then $\varphi$ is equivalent to the action of $G$ on $G/_{\rho}H$ via right multiplication.

(3) Let $R$ be a set of representatives for the set of all orbits in $X$ under $G$. Then $X = \bigcup_{x \in R} x\tilde{G}$. If $X$ is finite, we have the class equation:

$$|X| = \sum_{x \in R} |x\tilde{G}| = \sum_{x \in R} |G : C_{G,\varphi}(x)|.$$

Corollary 1 If $G$ is finite and $B$ is an orbit under an action of $G$, then $|B||G|$. For the action considered in Ex. (4) (conjugation), we obtain

$$|x^G| = \frac{|G|}{|C_G(x)|} \text{ for all } x \in G.$$

Corollary 2 If $G$ is finite, $U < G$, $|G : U|$ is the smallest divisor $\neq 1$ of $|G|$, then $U \trianglelefteq G$. (In particular, every subgroup of index 2 is normal.)

Corollary 3 Let $G$ be finite and $d$ a divisor of $|G|$ such that $d||x^G|$ for all $x \in G \setminus Z(G)$. Then $d||Z(G)||$. 

Proof. (1) is clear by the definition of an orbit. (2) Let

$$\beta : x\tilde{G} \to \Psi(G), \ y \mapsto \{g|g \in G, x(g\varphi) = y\}.$$

We show that $x(g\varphi)\beta = Hg$ for all $g \in G$: Set $y := x(g\varphi)$. Then for all $g' \in G$ we have $g' \in Hg \iff x((g'g^{-1})\varphi) = x \iff x(g'\varphi) = y \iff g' \in y\beta$, thus $y\beta = Hg$. Hence $\beta$ is a mapping of $x\tilde{G}$ onto $G/_{\rho}H$. If $y, y' \in x\tilde{G}$ such that $y\beta = y'\beta$, then $y = x(g\varphi) = y'$ for any $g \in y\beta$. Finally, given such an element $g \in G$, we have for all $g' \in G$

$$y(g'\varphi)\beta = x((gg')\varphi)\beta = Hgg' = (Hg)g' = (y\beta)g'.$$
(3) The first assertion is clear by 8.1.1. The first equality in the class equation is a direct consequence, the second is clear by (2).

Corollary 1 is obvious by (2). If \( U < G \), then \( U \) acts by right multiplication on the set \( G/U \cup \{U\} \). All orbits under \( U \) have a length which, by Corollary 1, is a divisor of \( |U| \), hence of \( |G| \), but smaller than \( |G : U| \). Under the hypothesis of Corollary 2, this implies \( Ugu = Ug \) for all \( g \in G \setminus U, u \in U \). Moreover, this equality is obvious if \( g \in U \). It follows that \( gug^{-1} \in U \) for all \( g \in G, u \in U \), i.e., \( U \trianglelefteq G \). By 8.1.5, \( Z(G) \) is the union of all conjugacy classes of length 1. Hence the class equation implies \( |G| = |Z(G)| + \sum B |B| \) where the sum ranges over all conjugacy classes \( B \) of length > 1. Under the hypothesis of Corollary 3, all these lengths and \( |G| \) are multiples of \( d \). The claim follows. \( \blacksquare \)

We combine Corollary 1 with 7.13 in the following remark:

8.2.1. Let \( G \) be a finite group, \( x \in H \leq G \) such that \( |x^G| \) is a prime power and \( o(x) \) is a power of a (possibly different) prime \( p \). Then \( x \in N \) for all \( N \trianglelefteq H \) such that \( |H/N| \neq p \).

Proof. Let \( N \trianglelefteq H \). The index of the centralizer of \( Nx \) in \( H/N \) is a divisor of \( |G : C_G(x)| \), hence a prime power. Applying 7.13 to the simple group \( H/N \) we obtain that \( N \trianglelefteq H \) or \( |H/N| = p \). \( \blacksquare \)

By 7.13, a finite group \( G \) with a conjugacy class \( C \neq \{1G\} \) of prime power length cannot be simple. We derive a more detailed consequence: \( G \) must necessarily contain a normal subgroup with a non-trivial \( p \)-factor group for every prime divisor \( p \) of the order of an element of \( C \). More precisely, we obtain: \(^{120}\)

8.3 Proposition. Let \( G \) be a finite group and \( H_0, \ldots, H_k \leq G \) such that \( \{1G\} = H_0 < \cdots < H_k = G \) and \( H_j/H_{j-1} \) has a direct decomposition into simple groups, for all \( j \in \mathbb{N} \). Suppose there exists an element \( x \in G \) such that \( o(x) = p^m \) for a prime \( p \), \( m \in \mathbb{N} \), and \( |x^G| \) is a prime power. Then \( m \leq k \) and there exist \( j_1, \ldots, j_m \in \mathbb{N} \) such that \( j_1 < \cdots < j_m \) and \( H_{j_i}/H_{j_i-1} \) has a direct factor of order \( p \), for all \( i \in \mathbb{N} \).

Proof. For all \( i \in \mathbb{N} \) we have \( \bigcap \{N|H_{i-1} \leq N \trianglelefteq H_i \} = H_{i-1} \). Let \( j_1 \) be the smallest index such that \( H_{j_1} \cap \langle x \rangle \neq \{1G\} \). Then there is a maximal normal subgroup \( N \) of \( H_{j_1} \) such that \( x^{p^{m-1}} \notin N \). The powers of \( x \) satisfy the hypotheses of 8.2.1. It follows that \( |H_{j_i}/N| = p \). Either \( m = 1 \) and the proof is complete, or \( m > 1 \) and \( x^{p^{m-2}} \notin H_{j_1} \) because otherwise \( x^{p^{m-1}} \in N \). But then we may proceed by induction, considering the element \( H_{j_i}x \) of order \( p^{m-1} \) in the factor group \( G/H_{j_1} \). \( \blacksquare \)

Thanks to 8.2 we will soon obtain fundamental results about finite groups. We need one more preparation:

\(^{120}\) Note that the centralizer of an element \( x \) is contained in the centralizer of all powers of \( x \).

\(^{121}\) A canonical way to obtain such a chain is to consider the intersection \( G_1 \) of all maximal normal subgroups of \( G_0 := G \), then the intersection \( G_2 \) of all maximal normal subgroups of \( G_1 \) etc. When \( \] for some \( k \in \mathbb{N} \), the process terminates, and we put \( H_j := G_{k-j} \) for \( 0 \leq j \leq k \).

135
8.3.1. Let \( A \) be a finite abelian group, \( p \) a prime divisor of \( |A| \), \( A_p := \{ g | g \in A, g^p = 1_A \} \).
Then \( \{ 1_A \} < A_p \leq \text{End } A \).

Proof. Being abelian, \( A \) is the product of its cyclic subgroups. Hence \( |A| \) is a divisor of the product of their orders\(^{122}\). As \( p ||A| \), there must exist a cyclic subgroup \( C \) of \( A \) the order of which is divisible by \( p \). But then \( C \), hence \( A \), contains an element of order \( p \). Furthermore, if \( g, h \in A_p, \alpha \in \text{End } A \), then \( (gh^{-1})^p = g^p(h^p)^{-1} = 1_A \), \((ga)^p = (g^p)\alpha = 1_A\).

\[\square\]

8.4 Theorem. Let \( G \) be a finite group.

(1) If \( |G| \) is a prime power \( \neq 1 \), then \( |Z(G)| \neq 1 \).

(2) (1. Theorem by Sylow (1872)) If \( k ||G| \) and \( k \) is a prime power, then \( G \) has a subgroup of order \( k \).

Proof. (1) Let \( p \) be the prime divisor of \( |G| \). By Corollary 1 of 8.2 and 8.1.5, \( |x^G| \) is a power \( \neq 1 \) of \( p \) for any \( x \in G \setminus Z(G) \). Hence \( p ||Z(G)| \) by Corollary 3 of 8.2.

(2) We proceed by induction on \( |G| \), the case of \( |G| = 1 \) being trivial. Let \( |G| > 1 \) and assume the claim holds for all groups \( G^* \) such that \( |G^*| < |G| \). As the case \( k = 1 \) is trivial, we suppose that \( k \) is a power \( \neq 1 \) of a prime \( p \).

Case 1: There exists an element \( x \in G \setminus Z(G) \) such that \( p \nmid |x^G| \). By 8.1.5, \( H := C_G(x) \) is a proper subgroup of \( G \). By Corollary 1 of 8.2(2), \( p \nmid |G : H| \). Hence \( k |||H| \), and inductively, there exists a subgroup of order \( k \) in \( H \), a fortiori in \( G \).

Case 2: \( p ||x^G| \) for all \( x \in G \setminus Z(G) \). Then, by Corollary 3 of 8.2, \( p ||Z(G)| \), so that \( Z(G) \) has a subgroup \( N \) of order \( p \), by 8.3.1. Now \( N \leq Z(G) \) implies \( N \leq G \), and \( |G/N| = \frac{|G|}{p} \) is divisible by the prime power \( k^* := \frac{k}{p} \). As \( |G/N| < |G| \), we know inductively that there exists a subgroup of \( G/N \) of order \( k^* \). Hence there exists a subgroup \( H \) of \( G \) such that \( N \leq H \) and \( |H/N| = k^*, \text{ i.e., } |H| = k \).

\[\square\]

8.5 Definition. Let \( G \) be a finite group, \( p \) a prime, \( n \in \mathbb{N}_0 \) maximal with \( p^n |||G| \). Then a subgroup of \( G \) of order \( p^n \) is called a Sylow \( p \)-subgroup of \( G \). We write \( Syl_p(G) \) for the set of all Sylow \( p \)-subgroups of \( G \) and put \( Syl(G) := \bigcup_{p |||G|} Syl_p(G) \). A consequence of 8.4(2) is the following remark which will be refined considerably in 8.9(3):

8.5.1. \( Syl_p(G) \neq \emptyset \) for all primes \( p \).

8.5.2. Let \( P \in Syl_p(G) \). Then \( P \alpha \in Syl_p(G) \) for all \( \alpha \in \text{Aut } G \). We have

\[\text{Syl}_p(G) = \{ P \} \iff P \trianglelefteq G \iff \forall \alpha \in \text{Aut } G \ P\alpha = P.\]

Proof. The first assertion is clear because \( P\alpha \) is a subgroup of \( G \) of order \( |P| \). We have to show that \( Syl_p(G) = \{ P \} \) if \( P \trianglelefteq G \); the other implications are then obvious. But if \( P \trianglelefteq G \) and \( Q \) is any \( p \)-subgroup of \( p \)-power order of \( G \), then \( PQ \) is a subgroup of \( p \)-power order of \( G \) containing \( P \). As \( P \) is a Sylow subgroup of \( G \), we conclude that \( PQ = P \), hence \( Q \leq P \). In particular, \( Syl_p(G) = \{ P \} \).

\[\square\]

\(^{122}\) by induction, making use of the formula \( |UV| = \frac{|U||V|}{|U \cap V|} \) for arbitrary subgroups \( U, V \).
For example, \(|\text{Syl}_2(S_3)| = 3\), \(|\text{Syl}_3(S_3)| = 1\), \((\text{Syl}_3(S_3) = \{A_3\})\), \(|\text{Syl}_2(A_4)| = 1\), \((\text{Syl}_2(A_4) = \{\text{id}_3, (12)(34), (13)(24), (14)(23)\})\), \(|\text{Syl}_3(A_4)| = 4\). We shall soon give an answer to the question which groups have a unique Sylow p-subgroup for all primes p. An easy remark in that direction is the following:

8.5.3. \(|\text{Syl}_p(G)| = 1\) for all primes p if and only if \(\text{Syl}(G)\) is a direct decomposition of G.

This is true because a trivial induction shows that a finite product of normal subgroups of pairwise coprime orders is always a direct product. □

8.6 Definition. For every group G let \(Z_0 := \{1_G\}\) and define, for \(n \in \mathbb{N}\), inductively the \(n\)-th centre \(Z_n\) to be the subgroup of G containing the normal subgroup \(Z_{n-1}\) such that \(Z_n/Z_{n-1} = Z(G/Z_{n-1})\). Then

\[
\{1_G\} = Z_0 \leq Z_1 \leq Z_2 \leq \cdots
\]

The sequence\(^{123}\) of normal subgroups \((Z_n)_{n \in \mathbb{N}_0}\) is called the upper central chain of G. \(^{124}\) G is called nilpotent if there exists an index \(k\) such that \(Z_k = G\). The smallest such \(k\) is then called the nilpotency class of G. Being central, each factor group \(Z_n/Z_{n-1}\) is a fortiori abelian. Hence, if G is nilpotent, the upper central chain is a chain of normal subgroups as required in the definition of solubility. Therefore we have

8.6.1. Every nilpotent group is soluble. □

A simple adaption of the proofs of 5.9.3 and 5.9.4 gives the following corresponding properties of nilpotent groups:

8.6.2. Every subgroup and every factor group of a nilpotent group is nilpotent. □

But 5.9.5 does not hold analogously for the property of nilpotency as may be seen by the example of the group \(S_3\) already: \(S_3/A_3\) and \(A_3\) are of prime order (hence cyclic, hence abelian, hence nilpotent) but \(S_3\) is not nilpotent: \(Z(S_3) = \{\text{id}_3\}\).

8.6.3. A group G is nilpotent if and only if there exists a tuple \((Y_0, \ldots, Y_m)\) of normal subgroups of G such that \(\{1_G\} = Y_0 \leq Y_1 \leq \cdots \leq Y_m = G\) and \(Y_j/Y_{j-1} \leq Z(G/Y_{j-1})\) for all \(j \in \mathbb{N}\).

Proof. Given a tuple \((Y_0, \ldots, Y_m)\) with the stated property, a simple induction shows that \(Y_j \leq Z_j\) for all \(j \in \mathbb{N}^+\). If, inductively, \(Y_{j-1} \leq Z_{j-1}\) and \(x \in Y_j\), then \(Y_{j-1}xg = Y_{j-1}gx\) for all \(g \in G\), i.e., \(x \in Z_j\). Therefore, \(Y_m = G\) implies \(Z_m = G\). □

\(^{123}\)If G is finite, there must exist an index \(n\) such that \(Z_n = Z_{n+1} = \cdots\). If \(n\) is minimal with this property, the \((n+1)\)-tuple \((Z_0, \ldots, Z_n)\), instead of the sequence \((Z_n)_{n \in \mathbb{N}_0}\), is in finite group theory also commonly called the upper central chain of G. Its last term \(Z_n\) is called the hypercentre of G, usually denoted by \(Z_\infty(G)\).

\(^{124}\)As the name suggests, there is also a lower central chain \((Y_n)_{n \in \mathbb{N}_0}\) of G, defined inductively by \(Y_0 := G\) and, given \(Y_{n-1}\), letting \(Y_n\) be the smallest normal subgroup of G contained in \(Y_{n-1}\) such that \(Y_{n-1}/Y_n \leq Z(G/Y_n)\). \((Z_n)_{n \in \mathbb{N}_0}\) is an ascending, \((Y_n)_{n \in \mathbb{N}_0}\) a descending chain of normal subgroups. It is not difficult to show that \(Z_n = G\) if and only if \(Y_n = \{1_G\}\).
A straightforward generalization of 8.1.6 to an arbitrary number of direct factors implies

\[ 8.6.4. \text{ A direct product of a finite number of nilpotent groups is nilpotent.} \]

A finite group the order of which is a power of a prime \( p \) is called a \( p \)-group.\(^{125}\)

\[ 8.6.5. \text{ Every finite } p \text{-group is nilpotent.} \]

**Proof.** Every factor group of a \( p \)-group \( G \) is a \( p \)-group. By 8.4(1), for every \( j \in \mathbb{N}_0 \) either \( Z_j = G \) or \( Z_j < Z_{j+1} \). \( G \) being finite, there must exist an index \( n \) such that \( Z_n = G \). \( \Box \)

Combining 8.6.5 and 8.6.4, we obtain that a direct product of a finite number of \( p \)-groups (where \( p \) may vary) is nilpotent. We shall now see that in the finite case, every nilpotent group is of this type:

**8.7 Theorem.** For a finite group \( G \), the following are equivalent:

(i) \( G \) is nilpotent.

(ii) Every Sylow subgroup of \( G \) is normal in \( G \).

(iii) \( \text{Syl}(G) \) is a direct decomposition of \( G \).

**Proof.** In view of the preceding remarks and 8.5.3 it remains to prove that (i) implies (ii). We proceed by induction on \( |G| \). The claim is trivial for \( |G| = 1 \). Let \( G \) be nilpotent of order \( > 1 \) and assume the implication holds for all groups of smaller order. Then \( Z(G) \neq \{1_G\} \), and, inductively, every Sylow subgroup of the nilpotent group \( G/Z(G) \) is normal in \( G/Z(G) \). Let \( P \in \text{Syl}(G) \). Then \( PZ(G)/Z(G) \in \text{Syl}(G/Z(G)) \), hence \( PZ(G) \leq G \). The properties of \( Z(G) \) trivially imply that \( P \leq PZ(G) \). Obviously, \( P \in \text{Syl}(PZ(G)) \). Now 8.5.2 implies that \( P\alpha = P \) for all \( \alpha \in \text{Aut}(PZ(G)) \). In particular, \( P^g = P \) for all \( g \in G \). \( \Box \)

We now apply the last result of the previous chapter to obtain a famous theorem:

**8.8 Theorem** (Burnside 1904). Let \( G \) be a finite group with a nilpotent subgroup of prime power index. Then \( G \) is soluble. In particular, if \( |G| \) has not more than two prime divisors, \( G \) is soluble.\(^{126}\)

**Proof by induction on the group order.** The claim is trivial if \( |G| = 1 \). Let \( H \leq G \), \( H \) nilpotent, and \( |G : H| \) a prime power. Assume the claim holds for all groups of order smaller than \( |G| \) satisfying the hypotheses. If \( H = \{1_G\} \), \( G \) is a \( p \)-group and the claim follows from 8.6.5 and 8.6.1. If \( H \neq \{1_G\} \), let \( x \in Z(H) \setminus \{1_G\} \). Then \( H \leq C_G(x) \). By Corollary 1 of 8.2, \( |x^G| \) is a prime power. Now 7.13 \( N \) such that \( \{1_G\} < N < G \). Then \( H \cap N \) is a nilpotent subgroup of prime power index in \( N \), and \( HN/N \) is a nilpotent subgroup of \( G/N \) of prime power index. Inductively we conclude that \( N \) and \( G/N \) are soluble. It follows that \( G \) is soluble, by 5.9.5.

If \( |G| = p^aq^b \) for distinct primes \( p, q \) and \( a, b \in \mathbb{N}_0 \), let \( P \in \text{Syl}_p(G) \) (by 8.5.1). Then \( P \) is nilpotent by 8.6.5, and \( |G : P| = q^b \). By what we have shown before, \( G \) is soluble. \( \Box \)

\(^{125}\)Even if the prime \( p \) has not been specified before, the term “\( p \)-group” always refers to a group of prime power order.

\(^{126}\)More precisely, Burnside proved this last statement, commonly called “Burnside’s \( p^aq^b \)-theorem”.

138
One detail of 8.5.2 says that, for each prime \( p \), every finite group \( G \) acts via conjugation on the set \( Syl_p(G) \) of its Sylow \( p \)-subgroups. We now prove the fundamental result that this action is always transitive: All Sylow \( p \)-subgroups (\( p \) fixed) are conjugate under \( G \). More precisely, we show:

**8.9 Theorem.** Let \( G \) be a finite group, \( p \) a prime.

(1) If \( P \in Syl_p(G) \) and \( Q \) is a \( p \)-subgroup of \( G \), then there exists an element \( g \in G \) such that \( Q \leq P^g \).

(2) (2. Theorem by Sylow (1872).) If \( P, Q \in Syl_p(G) \), there exists an element \( g \in G \) such that \( Q = P^g \).

(3) \(|Syl_p(G)| = |G : N_G(P)| \equiv 1 \mod p \), for any \( P \in Syl_p(G) \).

**Proof.** (1) As \( G \) acts on the set \( Y := G/P \) via right multiplication, this also holds for the subgroup \( Q \). If there were no \( Q \)-orbit of length 1, the length of every \( Q \)-orbit in \( Y \) would be divisible by \( p \) (Corollary 1 of 8.2), hence \( p \mid |Y| = \frac{|G|}{|P|} \), by 8.2(3). As \( P \in Syl_p(G) \), this is a contradiction. Therefore there exists an element \( g \in G \) such that \( (P)Q = P^g \). This implies \( P^gQ \leq PG \), hence \( Q \subseteq P^g \). This proves (1), and the claim in (2) is an immediate consequence as \(|Q| \equiv |P| \mod p \). \( \square \)

(3) By (1), \( G \) acts transitively by conjugation on the set \( X := Syl_p(G) \). The centralizer of the element \( P \in X \) with respect to this action is \( N_G(P) \) (see the comment on p. 132 concerning 8.1.2). The equation in our claim now follows from 8.2(2). With respect to the action of the subgroup \( P \) (via conjugation on \( X \)), \( \{P\} \) is an orbit of length 1. Conversely, let \( Q \in X \) such that the orbit of \( Q \) under \( P \) is of length 1. Then \( Q^g = Q \), i.e., \( Qg = gQ \), for all \( g \in P \). In particular, \( QP(= PQ) \) is a subgroup of \( G \), \( Q \subseteq PQ \), and \( |PQ| = |Q||PQ/Q| = |Q||P/(P \cap Q)| \) is a power of \( p \). But the Sylow \( p \)-subgroups \( P, Q \) are subgroups of \( PQ \), hence \( P = PQ = Q \). Thus \( \{P\} \) is the only orbit of \( P \) of length 1, and all other orbits must have a length which is a divisor \( \neq 1 \) of \(|P| \), by Corollary 1 of 8.2. Now 8.2(3) implies the claim: \(|X| \equiv 1 \mod p \). \( \square \)

By 8.9(3), the number of Sylow \( p \)-subgroups of a finite group \( G \) must satisfy both of the following two conditions which form a rather restrictive combination:

\[
(*) \quad 1 \equiv |Syl_p(G)| \mid |G| \quad \text{for every prime } p.
\]

There are many cases in which, using \((*)\), a simple inspection of the divisors of the group order reveals strong structural properties of the group. For example, let us consider the following basic case of Burnside’s \( p^aq^b \)-theorem 8.8: \(|G| = pq \) for primes \( p, q \), \( p < q \). A Sylow \( q \)-subgroup \( Q \) of \( G \) must be normal by Corollary 2 of 8.2, hence \( Syl_q(G) = \{Q\} \) by 8.5.2, and \( 1 + kp = |Syl_p(G)| |pq \) for some \( k \in \mathbb{N}_0 \), by \((*)\). It follows that \(|Syl_p(G)| = 1 \) or \( q \), and the last possibility arises only if \( p|q - 1 \). If this is not the case, we obtain that both Sylow subgroups of \( G \) are normal subgroups (of orders \( p, q \) resp.), hence, by 8.5.3, \( G \) is isomorphic to their direct product. But this means that \( G \) is cyclic. Summarizing,
Q ≤ G in every case, and G is cyclic if p ∤ q − 1. For example, all groups of orders 15, 33, 35, 51 are necessarily cyclic. Methods of this kind, frequently combined with certain counting arguments (for elements of some special order for example) and/or with character relations, can reach a considerable depth and then be of invaluable help in non-trivial structural investigations on finite groups.

8.10 Definition. Let G be a finite group. A subgroup H is called a Hall subgroup \(^{127}\) of G if gcd(|H|, |G : H|) = 1. For any set π of primes, a finite group is called a \(\pi\)-group if every prime divisor of its order belongs to π. A Hall \(\pi\)-subgroup of \(G\) is a \(\pi\)-subgroup \(H\) of \(G\) such that |\(G : H\)| has no prime divisor in π. The set of all Hall \(\pi\)-subgroups of \(G\) is denoted by \(Hall_\pi(G)\). If |π| = 1 and \(p\) is the prime such that π = \{p\}, we clearly have \(Hall_\pi(G) = Syl_p(G)\).

Thus the notion of a Hall subgroup is a natural extension of the notion of a Sylow subgroup. But a general existence theorem like 8.5.1 does not hold: The alternating group \(A_n\) has Hall \(\pi\)-subgroups for π = \{2, 3\}, but \(Hall_{\{2,3\}}(A_5) = \emptyset = Hall_{\{3,5\}}(A_5)\) as a Hall \{2, 5\}-subgroup, a Hall \{3, 5\}-subgroup resp., would be of index 3, 4 resp., in \(A_5\).

By 8.1.3(8.1), \(A_5\) would then have a nontrivial homomorphism into the soluble group \(S_3\), \(S_4\) resp., in contradiction with 8.1.4.

8.10.1. Let \(N \trianglelefteq G\) and π a set of primes, \(H \in Hall_\pi(G)\). Then \(H \cap N \in Hall_\pi(N), HN/N \in Hall_\pi(G/N)\).

Proof. |\(H \cap N\)| and |\(HN/N\)| are divisors of |\(H\)|, while |\(N : (H \cap N)\)|, |\(G : HN\)| are divisors of |\(G : H\)|. \(\Box\)

8.10.2. Let \(A\) be a minimal normal finite soluble subgroup of a group. Then \(A\) is an abelian \(p\)-group for some prime \(p\), and \(x^p = 1_A\) for all \(x \in A\).\(^{128}\)

Proof. For the given group \(G\), \(A' \trianglelefteq G\), hence \(A' = \{1_A\}\) as \(A' < A\) by the solubility of \(A\). Thus \(A\) is abelian. Furthermore, the subgroup \(A_p\) considered in 8.3.1 must coincide with \(A\) as it is, in particular, invariant under all conjugations in \(G\). \(\Box\)

8.11 Theorem (P. Hall 1937). Let \(G\) be a finite group. If \(Hall_\pi(G) \neq \emptyset\) for every set π of primes, \(G\) is soluble.

In view of 8.5.1, this result may be read as a far-reaching extension of Burnside’s \(p^aq^b\)-theorem. But the latter will be the crucial point in its proof:

\(^{127}\)after the famous English group theorist Philip Hall (1904 – 1982)

\(^{128}\)Abelian \(p\)-groups with this property are called \((p)\)-elementary abelian. They are obviously exactly the groups which decompose directly into groups of order \(p\). Remark 8.10.2 is easily generalized to the more general result that any minimal normal finite subgroup of a group has a direct decomposition into isomorphic simple groups. If \(G\) is a finite group, \(H_0, \ldots, H_k \trianglelefteq G\), \(\{1_G\} = H_0 < \cdots < H_k = G\), \(H_j/H_{j-1} \trianglelefteq G/H_{j-1}\) for all \(j \in \mathbb{N}\), then the tuple \((H_0, \ldots, H_k)\) is called a chief series, any quotient \(H_j/H_{j-1}\) a chief factor of \(G\). Every chief series of \(G\) satisfies the hypotheses on the \(H_i\) in 8.3. Consequently, if a finite group has an element of prime power order \(p^m\) the centralizer of which is of prime power index, any chief series of \(G\) involves at least \(m\ p\)-elementary abelian chief factors.
Proof by induction on $|G|$. If $|G|$ has not more than two prime divisors, the claim follows from 8.8. For the inductive step we may assume that $G$ satisfies the hypothesis and $|G|$ has at least three prime divisors. It suffices to show that $G$ is not simple: If $\{1_G\} < N < G$, $N$ and $G/N$ again satisfy the hypothesis, by 8.10.1, and are of a smaller order than $G$. Inductively, $N$, $G/N$ are soluble, hence $G$ is soluble by 5.9.5.

Let $p, q$ be distinct prime divisors of $|G|$, $U \in Hall_{\{p,q\}}(G)$. By 8.8, $U$ is soluble. Let $A$ be a minimal normal subgroup of $U$. By 8.10.2, $A$ is w.l.o.g. a $p$-group. Let $Q \in Syl_q(U)$. As $q \mid |G/U|$, $Q \in Syl_q(G)$. Let $\pi$ be the set of all prime divisors $\neq q$ of $|G|$ and $H \in Hall_\pi(G)$, $P \in Syl_p(G)$. As $p \in \pi$, $P \in Syl_p(G)$. By 8.9(1), $A \leq P^q \leq H^q$ for some $g \in G$. As $q \mid |H|$, $Q \cap H^q = \{1_G\}$. Hence $|QH^q| = |Q| \cdot |H| = |G|$, i.e., $QH^q = G$.

Now $A^g = A$ for all $x \in Q \subseteq U$ and $A^y \leq H^y$ for all $y \in H^y$ which implies that $A^z \leq H^y$ for all $z \in G$. Hence the normal subgroup of $G$ generated by $A$ is contained in $H^y \leq G$. Therefore $G$ is not simple, and the proof is complete.

Our next aim is a fundamental result on finite soluble groups the first part of which will be the converse of 8.11. For its further parts we need the following preparation:

8.12 Lemma. Let $G$ be a finite group, $N$ a normal subgroup of $G$ of prime power index. Let $\pi$ be a set of primes such that any two Hall $\pi$-subgroups of $N$ are conjugate in $N$.

Then any two Hall $\pi$-subgroups of $G$ are conjugate in $G$.

Proof. If $p \not\in \pi$, every Hall $\pi$-subgroup of $G$ is contained in $N$ so that the claim is trivial.

Let $p \in \pi$ and $H$, $J \in Hall_\pi(G)$, $D := H \cap N$. By 8.10.1, $NH = G = NJ$ and $D$, $J \cap N \in Hall_\pi(N)$, hence $D = (J \cap N)^x = J^x \cap N$ for some $x \in N$. It follows that $H, J^x \leq N_G(D)$. Let $P \in Syl_p(J^x)$, $Q \in Syl_p(H)$. Then $P, Q \in Syl_p(G)$ as $J^x, H$ are Hall subgroups of $G$, hence $P, Q \in Syl_p(N_G(D))$. By 8.9(2), $P^y = Q$ for some $y \in N_G(D)$. Now $J^{xy} = (DP)^y = DQ = H$.

8.13 Theorem (P. Hall 1928). Let $G$ be a finite soluble group, $\pi$ a set of primes.

(1) $Hall_\pi(G) \neq \emptyset$.

(2) If $H \in Hall_\pi(G)$ and $Q$ is a $\pi$-subgroup of $G$, there exists an element $g \in G$ such that $Q \leq H^g$.

(3) $G$ acts transitively by conjugation on $Hall_\pi(G)$.

Proof. We first show (1), (3) simultaneously by induction on the group order, the case $|G| = 1$ being trivial. Let $|G| > 1$ and assume that the claim holds for all soluble groups of smaller order. Let $M < G$. Inductively, $Hall_\pi(M) \neq \emptyset$ and $M$ acts transitively by conjugation on $Hall_\pi(M)$. $G/M$ is soluble and simple, hence of prime order $p$. Now the inductive step with respect to (3) follows directly from 8.12.

As for (1), if $p \not\in \pi$, $Hall_\pi(G) = Hall_\pi(M)$ and the claim is trivial. Let $p \in \pi$, $J \in Hall_\pi(M)$. For any $g \in G$, we have $J^g \in Hall_\pi(M)$, hence there exists an element $x \in M$ such that $(J^g)^x = J$. It follows that $gx \in N_G(J)$ and $g \in N_G(J)M$. We conclude
that \( N_G(J)M = G \). Let \( P \in Syl_p(N_G(J)) \). Then \( PJ \) is a \( \pi \)-subgroup \( > J \) of \( G \) and \((PJ) \cap M = J\), hence \( |G : PJ| = |M : J| \) so that \( PJ \in Hall_\pi(G) \). This completes the inductive step for (1).

Now we prove (2), again by induction on \(|G|\). For the inductive step, let \(|G| > 1\), \( H \in Hall_\pi(G) \), \( Q \) be a \( \pi \)-subgroup of \( G \), \( A \subseteq G \), \( Q^* := QA \). Then \( HA/A \in Hall_\pi(G/A) \), and \( Q^*/A \) is a \( \pi \)-subgroup of \( G/A \). As \(|G/A| < |G|\) and \( G/A \) is soluble there exists inductively an element \( x \in G \) such that \( Q^*/A \leq (HA/A)^{A\pi} \), i.e., \( Q^* \leq (HA)^\pi \). By 8.10.2, \( A \) is an elementary abelian \( p \)-group for some prime \( p \).

If \( p \in \pi \), \( HA \) is a \( \pi \)-subgroup of \( G \), hence \( HA = H \) since \( H \in Hall_\pi(G) \). It follows that \( Q \leq Q^* \leq H^\pi \).

If \( p \notin \pi \), we put \( R := Q^* \cap H^\pi \) and observe that \( Q, R \in Hall_\pi(Q^*) \). By (3), there exists an element \( y \in Q^* \) such that \( Q = R^y \leq H^{xy} \), completing the proof of (2). □

The so-called “Sylow structure” of a finite group \( G \) is most simple if \( G \) is nilpotent: By 8.7, any product of Sylow subgroups is then a normal subgroup. It is remarkable that there is a related regular behaviour of Sylow subgroups even in the far more general case of a soluble group \( G \). Clearly, if \( \pi \) is a set of primes and \( \pi' \) the set of the primes not belonging to \( \pi \), \( H \in Hall_\pi(G) \) and \( J \in Hall_{\pi'}(G) \), then \( H \cap J = \{1_G\} \) and therefore \(|HJ| = |H||J| = |G|\), hence \( HJ = G \). But in general, the product of Hall subgroups for disjoint sets of primes will not be a subgroup, not even the product of Sylow subgroups for different primes. A \textbf{Sylow system} of a finite group \( G \) is a set \( \mathcal{X} \) of Sylow subgroups \( \neq \{1_G\} \) of \( G \) such that \( PQ \) is a subgroup for all \( P, Q \in \mathcal{X} \) and \( \prod \mathcal{X} = G \). Clearly, a Sylow system of \( G \) is a set of representatives for the set \( \{Syl_p(G) | p \text{ a prime divisor of } |G|\} \).

It is easily seen that a Sylow system \( \mathcal{X} \) of \( G \) has the property that every partial product is a subgroup (and then necessarily a Hall subgroup). It follows that a group with a Sylow system has \( \pi \)-Hall subgroups for every set \( \pi \) of primes, hence is soluble by 8.11. This is a part of the following beautiful and classical result:

\textbf{Theorem} (P. Hall 1937). A finite group is soluble if and only if it has a Sylow system. Every finite soluble group acts transitively by conjugation on the set of its Sylow systems.

As Sylow subgroups are nilpotent, the first assertion of this result is a special case of the following:

\textbf{Theorem} (Wielandt, Kegel 1961). A finite group \( G \) is soluble if and only if there exists a set \( \mathcal{Y} \) of nilpotent subgroups of \( G \) such that \( PQ \) is a subgroup for all \( P, Q \in \mathcal{Y} \) and \( \prod \mathcal{Y} = G \).

A paper by Wielandt (1951) shows that the core of the difficult implication of this theorem is the case \(|\mathcal{Y}| = 2\): It must be shown that a finite group is soluble if it is a product of two nilpotent subgroups. From 8.8 we obtain immediately that this is true if one of the two factors is a \( p \)-group.

The structure of a finite nilpotent group \( G \) is completely determined by the structure of its Sylow subgroups as these form a direct decomposition of \( G \) (8.7). For a finite soluble

\footnote{More recently also called a “Sylow basis”}
group $G$, however, the structure of its Sylow subgroups does by no means determine the structure of $G$ although there exists a Sylow system. We may certainly fix an order of the groups in the Sylow system and then write each element of $G$ uniquely as a product of elements of our selected Sylow subgroups in our chosen order. But, in the non-nilpotent case, it is in no way clear how to obtain this representation for a product of group elements from the representations of its single factors. In other words, a Sylow system (given in a certain order) gives rise to a unique representation of the elements of the group as a product of elements of prime power order, but this does not describe the structure of the group. For example, there are two non-isomorphic groups of order 6, but both have Sylow systems consisting of a group of order 3 and a group of order 2. Clearly, the only nilpotent group of order 6 is the direct product of its Sylow subgroups (hence the cyclic group of order 6).

In general, let $H$, $J$ be subgroups of a group $G$ such that $H \cap J = \{1_G\}$. Then each element of the subset $HJ$ has a unique representation as a product $xy$ where $x \in H$, $y \in J$. It may well happen that $HJ$ is not a subgroup (for example, let $H$, $J$ be two different subgroups of order 2 of $S_3$). But even if this should be the case it is unclear how to write a product $(xy)(x'y')$ (where $x, x' \in H$, $y, y' \in J$) again as a product of a factor in $H$ by a factor in $J$. Certainly, if $HJ$ is a subgroup, we have $HJ = JH$, hence there exist $\tilde{x} \in H$, $\tilde{y} \in J$ such that $yx' = \tilde{x}\tilde{y}$, and then $(xy)(x'y') = \tilde{x}\tilde{y}$. But apart from the comfortable case of two normal subgroups $H$, $J$ where $y, x'$ simply commute, we have no method to calculate $\tilde{x}, \tilde{y}$. There is, however, an important “intermediate” case in which products of this kind can be dealt with satisfactorily:

8.14 Definition. Let $G$ be a group. A set of two subgroups of $G$ is called a decomposition of $G$ if their product is $G$ and their intersection trivial. Then these subgroups are called complements of each other. A decomposition is called direct if both subgroups are normal. A semidirect decomposition of $G$ is a decomposition $\{H, N\}$ of $G$ in which (at least) one of the two subgroups, say $N$, is normal in $G$.

8.14.1. Let $G$ be group, $H \leq G$, $N \trianglelefteq G$. Then

$$\forall h, \tilde{h} \in H, \forall n, \tilde{n} \in N \quad (hn)(\tilde{h}\tilde{n}) = (h\tilde{h})(n\tilde{n}),$$

and $h\tilde{h} \in H$, $n\tilde{n} \in N$. □

Therefore, if a decomposition $\{H, N\}$ of $G$ is semidirect (with $N \trianglelefteq G$), we not only have a unique product representation of the elements of $G$ with a first factor in $H$ and a second factor in $N$, but also a reduction of the operation in $G$ (hence, the structure of $G$) to the following components:

130 $xy = x'y'$ implies $x'^{-1}x = y'y^{-1}$, and this is an element of $H \cap J$, hence it is $1_G$. It follows that $x = x'$, $y = y'$.

131 If $U, V \leq G$ and $UV = G$, then also $VU = G$ as $(uv)^u = vu$ for all $u \in U$, $v \in V$. More generally, for arbitrary subgroups $U, V$, the set $UV$ is a subgroup if and only if $UV = VU$.

132 It is well known that then $G$ is isomorphic to their direct product.
• operation in $H$,
• operation in $N$,
• action of $H$ on $N$ (via conjugation).

The action of $H$ on $N$ is in this case not just a homomorphism of $H$ into the symmetric group of the set $N$ but, more specifically, into the automorphism group of the group $N$. This analysis of the interior of a given group $G$ with a semidirect decomposition allows an important converse, a group construction in form of a synthesis of two group structures and an action as automorphism group:

Let $U, V$ be groups and
\[ \varphi : U \to \text{Aut} V, \quad u \mapsto \alpha_u, \]
a homomorphism. Let $G$ be the cartesian product of $U$ and $V$. We define an operation on $G$:
\[ \forall u \in U \forall v \in V \quad (u, v) \cdot (\tilde{u}, \tilde{v}) := (u\tilde{u}, (v\alpha_{\tilde{u}})\tilde{v}) \]

Then $(1_U, 1_V)$ is neutral, $(u, v) \cdot (u^{-1}, v^{-1}\alpha_{u^{-1}}) = (1_U, 1_V)$ for $u \in U$, $v \in V$, and $\cdot$ is associative: The second component of $((u_1, v_1) \cdot (u_2, v_2)) \cdot (u_3, v_3)$ is $((v_1\alpha_{u_2}v_2)\alpha_{u_3})v_3 = (v_1\alpha_{u_2}v_2)(v_2\alpha_{u_3})v_3$ which is the second component of $(u_1, v_1) \cdot ((u_2, v_2) \cdot (u_3, v_3))$. The group $(G, \cdot)$ is called the semidirect product of $U$ and $V$ with respect to the action $\varphi$ (of $U$ on $V$ as automorphism group), and is denoted by $U \ltimes V$. 

Set $H := \{(u, 1_V) | u \in U\}$, $N := \{(1_U, v) | v \in V\}$. Then we have obvious isomorphisms $U \to H$, $u \mapsto (u, 1_V)$, $V \to N$, $v \mapsto (1_U, v)$, and the equation $(u, 1_V) \cdot (1_U, v) = (u, v)$ shows that $HN = G$. Clearly, $H \cap N = \{1_G\}$ so that $H$ is a complement of $N$ in $G$. Keeping in mind that $HN = G$, the following equation implies that $N \leq G$ whence $\{H, N\}$ is indeed a semidirect decomposition of $G$:

\[ \forall u \in U \forall v \in V \quad (1_U, v)^{(u, 1_V)} = (1_U, v\alpha_u) \]

To prove $(\ast)$ it is sufficient to observe, for $u \in U$, $v \in V$, that $(u, 1_V) \cdot (u, 1_V) = (u, 1_V) \cdot (1_U, v\alpha_u)$ which is true since both sides equal $(u, v\alpha_u)$. In the sense of $(\ast)$ we may say that conjugacy in $G = U \ltimes V$ is determined by the action $\varphi$ of $U$ on $V$.

\[ \text{133} \text{H is said to act as an automorphism group on N.} \]
\[ \text{134} \text{Alternatively, by V} \ltimes U. \text{ Note that the symbol for the semidirect product is used in a way that it contains a “normal subgroup triangle” with respect to the group V which is acted upon by U.} \]

Confusingly, some authors use it in the opposite direction.
8.14.2 Example. Let $D$ be a division algebra, $U = \hat{D}$ its multiplicative group, $V$ its additive group. By 6.8.2, we have the natural action $\rho$ of $U$ on $V$ given by right multiplication in $D$:

$$\rho : U \to \text{Aut} V \quad u \mapsto \begin{bmatrix} D \\ v \mapsto vu \end{bmatrix} =: \alpha_u.$$ 

Let $u, u' \in U$, $v, v' \in V$. Then $(v + v')u = vu + v'u$ by the right distributive law in $D$, $u\rho$ is inverted by $u^{-1}\rho$ so that $u\rho \subseteq \text{Aut} V$ for all $u \in U$. Furthermore, by the associative law in $D$, $v(uu') = (vu)u'$ whence $\rho$ is a homomorphism. Thus we have a semidirect product of the multiplicative group of $D$ with the additive group of $D$ and the following operation:

$$\forall u, \tilde{u} \in \hat{D} \forall v, \hat{v} \in D \quad (u, v) \cdot (\tilde{u}, \hat{v}) = (u\tilde{u}, v\tilde{u} + \hat{v}).$$

In particular, we may choose a finite field $K$ for $D$ to obtain a finite group $G$ by this construction.\(^{135}\) If $|K| = 3$, then $G$ is non-commutative of order 6, hence $G \cong S_3$. If $|K| = 4$, then $G$ is non-commutative of order 12 and has an elementary abelian normal subgroup of order 4, hence $G \cong \mathcal{A}_4$.\(^{137}\) If $|K| = 5$ (as in general if $|K|$ is a prime), the automorphism group of $(K, +)$ is of order $|K| - 1$, hence the (injective!) homomorphism $\rho$ is an isomorphism of $(\hat{K}, \cdot)$ onto $\text{Aut}(K, +)$.\(^{138}\) The semidirect product is then of order 20 and contains the dihedral group of order 10 as a normal subgroup.

The quaternion group or any cyclic group of prime power order are examples of groups which are not isomorphic to any semidirect product of groups of strictly smaller order. If, however, a group is a non-trivial semidirect product $U \ltimes V$, its structure is essentially reduced to that of $U$ and $V$. Therefore, theorems which assert the existence of a semidirect decomposition of a given group under not too restrictive hypotheses are of particular interest. An extremely important theorem in this direction is the following:

**Theorem** (Schur, Zassenhaus 1937).\(^{139}\) Let $G$ be a finite group and $N$ a normal Hall subgroup of $G$.

\(^{135}\)We shall only consider the additive and the multiplicative structure of $D$ while the underlying field and the vector space structure do not play a role here. Thus, rather than an algebra, we consider a pure ring structure. A division algebra, considered in this purely ring-theoretic way, is called a skew field.

\(^{136}\)By a classical theorem of Wedderburn, a finite division algebra is commutative, hence a finite field.

\(^{137}\)We may even add a second construction of a semidirect product now: $K$ is a Galois extension of its prime field (of order 2), hence has a Galois group $A$ of order 2 which acts on $G$ because it acts as an automorphism group on both $(K, +)$ and $(\hat{K}, \cdot)$ and respects their “interplay” in the operation rule of $\cdot$ in $G$. Hence we have a semidirect product of $A$ with $G$ where $A$ acts componentwise on $G$.

Then $A \times G$ is a group of order 24 and isomorphic to a (non-direct) semidirect product of $\mathcal{A}_4$ and a group $(A)$ of order 2. It follows that $A \times G \cong S_4$. More generally, for any field $K$ and $A \leq \text{Aut} K$, the analogous repeated semidirect product construction is possible.

\(^{138}\)Given a group $X$, the semidirect product of $\text{Aut} X \ltimes X$ (with respect to the natural automorphism action) is called the holomorph of $X$.

\(^{139}\)A first special case of this result was proved by Schur in 1907 while Zassenhaus gave a general proof, but still under the hypothesis that $N$ or $G/N$ be soluble. Only since the “odd order theorem” of 1963 it is clear that this hypothesis is not required.
(1) There exists a complement of \( N \) in \( G \).

(2) Any two complements of \( N \) in \( G \) are conjugate under \( N \).

For the case that \( N \) is abelian, a proof of exercise type is available (although quite a number of group theorists undauntedly keep making a fuss of it even these days). Moreover, for part (1) in its general form, there exists a well-known routine reduction to that well-understood\(^{140}\) case. For part (2) in its general form, however, the only known proof makes use of the following deep and most influential result, known as “the odd order theorem”. It was conjectured by Burnside more than 50 years before its complete proof (of 255 journal pages) was indeed obtained:

**Theorem** (Feit, Thompson 1963). Every finite group of odd order is soluble.

Thus it should be noted that the hard part of the Schur-Zassenhaus Theorem is the conjugacy statement (2) the difficulties of which are not even remotely comparable to those of the existence part (1). The odd order theorem had important predecessors where groups with abelian centralizers of nontrivial elements (Suzuki 1957), later groups with nilpotent centralizers of nontrivial elements (Feit, Hall, Thompson 1960) were studied.

Finite groups \( G \) arising from 8.14.2 always have a normal Hall subgroup as is clear from the construction, simply because the additive group of a finite field \( K \) is of order \(|K|\), the multiplicative group of order \(|K| - 1\). The complements of the normal subgroup, being Hall subgroups of the same order of a soluble (even metabelian) finite group, are all conjugate under \( G \) by 8.13(3), alternatively by part (2) of the Schur-Zassenhaus Theorem. They have a further remarkable property which we will focus upon in the sequel. We consider the general case of a semidirect product \( G \) of the multiplicative group \( U \) and the additive group \( V \) of a division algebra \( D \). Let \( H \) be the subgroup of \( G \) which is canonically isomorphic to \( U \), as on p. 144. Then

\[
\forall g \in G \setminus H \quad H \cap H^g = \{1_G\}.
\]

**Proof.** Let \( u, \tilde{u} \in U, v \in V \) and suppose \((u, 0_D)^{(\tilde{u}, v)} \in H\). Then \((u\tilde{u}, v) = (u, 0_D) \cdot (\tilde{u}, v) \in (\tilde{u}, v) \cdot H = \{(u\tilde{t}, vt) | t \in \tilde{D}\} \) which implies \( u\tilde{u} = \tilde{u}, u = 1_D \). \( \Box \)

Now let \( H \) be a proper subgroup of an arbitrary group \( G, H \neq \{1_G\} \). If condition (**) is satisfied, \( G \) is called a Frobenius group with respect to \( H \). Thus we have verified that the groups arising from the construction in 8.14.2 are examples of Frobenius groups if \(|D| \neq 2\).

**8.14.3.** Let \( G \) be a Frobenius group with respect to the subgroup \( H \). Then \( H = N_G(H)^{141}\)

---

\(^{140}\)Let \( N \) be abelian, \( R \) a transversal of \( N \) in \( G \). For each \( r \in R, g \in G \), let \( n_r(g) \in N \) such that \( rg \in Rn_r(g) \). Set \( gw := \prod_{r \in R} n_r(g) \). Then \( (gg')w = (gw)'(g'w) \) for \( g, g' \in G, nw = n^{[R]} \) for \( n \in N \). Hence \( \{g | g \in G, gw = 1_G\} \) is a complement of \( N \) in \( G \).

\(^{141}\)A subgroup with this property is commonly called self-normalizing.
because, for any \( g \in G \), \( H = H^g \) implies that \( H \cap H^g = H \neq \{1_G\} \), hence, by (**), \( g \in H \).

In our examples from 8.14.2, the subgroup \( H \) satisfying (**) was part of a semidirect decomposition of \( G \) right from the definition. But a celebrated result of Frobenius (which gave birth to the above terminology) asserts that condition (**) alone implies indeed that \( G \) is isomorphic to a semidirect product with \( H \) as its non-normal factor:

**Theorem** (Frobenius 1902). Let \( G \) be a finite Frobenius group with respect to the subgroup \( H \). Then there exists a normal complement \( F \) of \( H \) in \( G \). Furthermore, \(|H||F| = 1\). In particular, \( H \) and \( F \) are Hall subgroups of \( G \).

We will prove this famous theorem in a more general form discovered more than half a century later by Wielandt. For a subset \( M \) of a group \( G \) we set \( M^G := \{x^g | x \in M, \, g \in G\} \).

**8.15 Theorem** (Frobenius, Wielandt 1958). Let \( H \) be a subgroup of a finite group \( G \). Write \( H_{\cap} \) for the subgroup of \( H \) generated by all intersections \( H \cap H^g \) where \( g \in G \setminus H \), and set \( F := G \setminus (H \setminus H_{\cap})^G \). Then

(1) \( F \leq G \),

(2) \( H \cap F = H_{\cap} \),

(3) \( HF = G \).

If \( g \in G \setminus H \), \( h \in H \), then \( gh \notin H \) and \((H \cap H^g)^h = H \cap H^{gh} \). It follows that \( H_{\cap} \leq H \).

The theorem is trivial if \( H_{\cap} = H \). In particular, it is of interest only if \( H = N_G(H) \).

Although it is clear from the definition that \( F \) is a normal subset of \( G \), the lion’s share of the proof is to obtain part (1): Hitherto no direct proof for the multiplicative closure of \( F \) is known. Only a witty application of character theory will show that \( F \) is a subgroup. The claim will be proved by showing that \( F \) is the kernel of some group homomorphism. For this reason, \( F \) is commonly referred to as the Frobenius kernel of \( G \).

The desired group homomorphism will arise as the restriction of a representation of \( CG \) to \( G \). Rather than discussing group representations (in the sense described in 8.1.3(3)) directly, we shall argue with their characters. Recall that, by 7.5(4), the kernel of the group representation associated with a given character \( \chi \) is given by

\[
K(\chi) := \{x | x \in G, \chi(x) = \chi(1_G)\}.
\]

Our aim is to show that \( F = K(\chi) \) for a suitable character \( \chi \) of \( G \).

**8.16 Proposition.** Let \( G, H, H_{\cap}, F \) as in 8.15.

(1) Let \( h \in H \setminus H_{\cap}, \, x \in G \). Then \( h^x \in H \iff x \in H \). In particular, \( h^G \cap H = h^H \).

(2) Let \( R \) be a right transversal of \( H \) in \( G \), \( g \in (H \setminus H_{\cap})^G \). Then there exists a unique pair \((h, r)\) such that \( h \in H \setminus H_{\cap}, \, r \in R, \, g = h^r \). In particular, \( \frac{|F||H|}{|H_{\cap}|} = |G| \).

\[\text{This set is therefore commonly denoted by } \ker \chi \text{ although it is not to be read literally as the “kernel of } \chi \text{” but the kernel of the homomorphism of the group } G \text{ obtained by restriction from the representation of } CG \text{ with character } \chi.\]
Proof. (1) “⇐” is trivial. “⇒”: \( h^x \in H \) implies that \( h \in H \cap H^{x^{-1}} \). But \( h \not\in H \cap \), hence \( x \in H \). The second claim is now obvious.

(2) Existence: Let \( h_0 \in H \setminus H_\cap \), \( x \in G \) such that \( g = h_0^x \). Moreover, let \( h_1 \in H \), \( r \in R \) such that \( x = h_1 r \). Setting \( h := h_0 h_1 \) we obtain \( g = h^r \). Uniqueness: Let \( h, h' \in H \setminus H_\cap \), \( r, s \in R \) such that \( h^r = h^s \). Then \( h^{rs^{-1}} = h' \in H \), hence \( rs^{-1} \in H \) by (1). It follows that \( r = s \) and \( h = h' \). Consequently,

\[
|F| = |G| - |G : H|(|H| - |H_\cap|) = \frac{|G|}{|H|} |H_\cap|,
\]

hence the claim. \( \square \)

Let \( \psi \in Cl(H) \) be constant on \( H_\cap \). If \( g \in G \setminus F \), the definition of \( F \) and 8.16(1) show that \( g^G \cap H = hH \) for some \( h \in H \setminus H_\cap \). Hence the value \( \psi(h) \) is independent of the choice of \( h \in g^G \cap H \). It follows that

\[
\hat{\psi} : G \to \mathbb{C}, \quad g \mapsto \begin{cases} \psi(1_G) & \text{if } g \in F \\ \psi(h) & \text{if } g \not\in F, \ h \in g^G \cap H \end{cases}
\]
is well-defined and its linear extension obviously a class function of \( G \) with the property \( \psi|_H = \hat{\psi} \). This way of extending any \( H_\cap \)-constant class function of \( H \) to an \( F \)-constant class function of \( G \) will be of utmost importance for the proof of 8.15. The mapping \( \psi \mapsto \hat{\psi} \) is obviously linear.

8.17 Lemma. Let \( G, H, H_\cap, F \) as in 8.15.

(1) Let \( \mu \in Cl(G) \) such that \( \mu(x) = 0 \) for all \( x \in F \). Then

\[
\forall \sigma \in Cl(G) \quad \langle \mu|\sigma \rangle_G = \langle \mu|\sigma|_H \rangle_H.
\]

(2) Let \( \psi \) be a character of \( H \) such that \( H_\cap \leq K(\psi) \). Then \( \hat{\psi} \) is a character of \( G \) and \( K(\hat{\psi}) \cap H = K(\psi) \). If \( \psi \) is irreducible, then \( \hat{\psi} \) is irreducible.

Proof. (1) Let \( R \) be a right transversal of \( H \) in \( G \). Then

\[
\langle \mu|\sigma \rangle_G = \frac{1}{|G|} \sum_{g \in G \setminus F} \mu(g) \sigma(g) = \frac{1}{|G|} \sum_{r \in R} \sum_{h \in H \setminus H_\cap} \mu(h^r) \sigma(h^r) = \frac{|R|}{|G|} \sum_{h \in H \setminus H_\cap} \mu(h) \sigma(h) = \langle \mu|\sigma|_H \rangle_H = \frac{|R|}{|G|} \cdot \frac{|R|}{|H|}
\]

as \( \frac{|R|}{|G|} = \frac{1}{|H|} \).

(2) We assume first that \( \psi \) is irreducible. Denoting by \( \psi_1, \chi_1 \), the trivial character of \( H, G \) resp., we have \( \hat{\psi}_1 = \chi_1 \). Now let \( \psi \neq \psi_1 \). Put \( n := \psi(1_G), \mu := \hat{\psi} - n \chi_1 \). Then \( \mu(x) = 0 \) for all \( x \in F \), hence

\[
\langle \mu|\mu \rangle = \langle \mu|\mu|_H \rangle_H = \langle \psi - n \psi_1 | \psi - n \psi_1 \rangle_H \equiv 1 + n^2.
\]

148
On the other hand, by the second assertion in 7.8, there exist \( c_1, \ldots, c_h \in \mathbb{C} \) such that \( \mu = \sum_{i \in h} c_i \chi_i \) where, as usual, the irreducible characters of \( G \) are denoted by \( \chi_i \) (\( i \in h \)). It follows that
\[
\langle \mu | \chi_i \rangle_G = \langle \psi - n \psi_1 | \chi_i | H \rangle_H = \langle \psi | \chi_i | H \rangle_H - n \langle \psi_1 | \chi_i | H \rangle_H \in \mathbb{Z},
\]
because \( \chi_i | H \) is a character of \( H \). In particular, \( c_1 = \langle \psi - n \psi_1 | \psi_1 \rangle_H = -n \). We obtain
\[
\langle \mu | \mu \rangle_G = \sum_{i \in h} c_i \langle \chi_i | H \rangle = \sum_{i \in h} c_i^2 = n^2 + c_2^2 + \cdots + c_h^2.
\]
Comparing with (*) we conclude that there exists a unique \( j > 1 \) such that \( c_j \neq 0 \), and we have \( c_j^2 = 1 \). Hence \( \mu = c_j \chi_j - n \chi_1 \), and the equation \( 0 = \mu(1_G) = c_j \chi_j(1_G) - n \) shows that \( c_j > 0 \). Thus \( c_j = 1 \). Now \( \hat{\psi} - n \chi_1 = \mu = \chi_j - n \chi_1 \), i.e., \( \hat{\psi} = \chi_j \). Thus \( \hat{\psi} \) is an irreducible character of \( G \).

Now let \( \psi \) be an arbitrary character of \( H, H \cap L \leq K(\psi) \). By 7.9, there exist (not necessarily distinct) irreducible characters \( \psi_j \) of \( H \) such that \( \psi = \sum_j \psi_j \). By 7.6, \( H \cap L \leq K(\psi_j) \) for all \( j \). By what we have shown above, each \( \hat{\psi}_j \), hence also \( \hat{\psi} = \sum_j \hat{\psi}_j \) is a character of \( G \). For \( g \in H \) we have \( g \in K(\hat{\psi}) \iff \hat{\psi}(g) = \hat{\psi}(1_G) \iff \psi(g) = \psi(1_G) \iff g \in K(\psi) \).

**Proof of 8.15.** Choose a character \( \psi \) of \( H \) such that \( K(\psi) = H \cap L \). (For example, \( \psi := \omega_H \) (cf. 7.4.6) where \( \omega \) is the character of the (right) regular representation of \( H/H \cap L \).) Clearly, \( F \subseteq K(\hat{\psi}) \). By 8.17(2) and 8.16(2),
\[
|G| \geq |K(\hat{\psi})H| = \frac{|K(\hat{\psi})||H|}{|H \cap L|} \geq \frac{|F||H|}{|H \cap L|} = |G|.
\]
Hence all numbers in this chain must coincide so that \( |K(\hat{\psi})| = |F| \). It follows that \( F = K(\hat{\psi}) \leq G \).

The special case \( H \cap L = \{1_G\} \) gives the main assertion in Frobenius’s theorem. The statement on the orders of \( H \) and \( F \) in it is, by 8.2(3), a direct consequence of the following simple observation:

**8.17.1.** Let \( G \) be a Frobenius group with respect to a finite subgroup \( H, F := G \setminus (H \setminus \{1_G\})^G \). Then \( |x^H| = |H| \) for all \( x \in F \setminus \{1_G\} \).

**Proof.** Let \( x \in F \setminus \{1_G\} \) and \( h, h' \in H \) such that \( x^h = x^{h'} \). Then \( h' h^{-1} x = x h' h^{-1} \), hence \( h' h^{-1} = (h' h^{-1}) x \in H \cap H x = \{1_G\} \) as \( x \neq 1_G \). Thus \( h = h' \).

Finally the following deep result should be mentioned by which a long-standing conjecture by Burnside was settled:

**Theorem** (Thompson 1959). The Frobenius kernel of a finite Frobenius group is nilpotent.

Making use of this result, we prove:

\[143\text{By passing from } \psi_j \text{ to } \hat{\psi}_j, \text{ the decomposition of } \psi \text{ into irreducible characters of } H \text{ is transformed into the decomposition of } \hat{\psi} \text{ into irreducible characters of } G.\]
8.18 Proposition. Let $G$ be a finite group, $p$ a prime such that $|\text{Syl}_p(G)| > 1$, $k$ the largest divisor of $|G|$ with $p \mid k$. Suppose that $d \neq 1$ for all non-trivial\(^\text{144}\) divisors $d$ of $k$.

Then $G$ has a nilpotent subgroup of order $k$. In particular, $G$ is soluble.

**Proof.** Let $P, Q \in \text{Syl}_p(G)$ such that $P \neq Q$ and $S := P \cap Q$ is of maximal order. Put $G^* := N_G(S)$, $P^* := G^* \cap P$. Then $S < P^* \leq P$, and $P^*$, $G^* \cap Q \in \text{Syl}_p(G^*)$.\(^\text{145}\) Hence $|\text{Syl}_p(G^*)| > 1$. Now $1 < |G^*: N_{G^*}(P^*)| = 1$. As $|G^*: N_{G^*}(P^*)|$ is a divisor of $k$ it follows from our hypothesis that $|G^*: N_{G^*}(P^*)| = k$. Therefore $PG^* = G$ and $N_{G^*}(P^*) = P^*$. For all $g \in G^* \setminus P^*$ we have $P^* \cap P^*g = S$ because $P \cap P^* = S$ by the maximality of $|S|$. Hence $G^*/S$ is a Frobenius group. Let $S \leq F \leq G^*$ such that $F/S$ is the Frobenius kernel of $G^*/S$. The $p$-group $S$ is a normal Hall subgroup of $F$ so that $S$ has a complement $K$ in $F$, by the Schur-Zassenhaus theorem. Then $\{P, K\}$ is a decomposition of $G$, $|K| = k$. By Thompson’s theorem, $K \cong F/S$ is nilpotent. The last assertion follows from 8.8.

We conclude with a few remarks without proof: Concerning the Frobenius complement $H$ of a finite Frobenius group, Burnside showed that all Sylow $p$-subgroups of $H$ (for $p \mid |H|$) have a unique subgroup of order $p$, hence are cyclic for $p \neq 2$, cyclic or generalized quaternion groups for $p = 2$. Despite these strong restrictions on the Sylow structure of $H$, there exist examples of Frobenius groups in which the Frobenius complement is not soluble. This, however, can occur only if the Sylow 2-subgroups of $H$ are non-cyclic. If every Sylow subgroup of a finite group is cyclic, the group is soluble.

\(^\text{144}\)I.e., $1 \neq d \neq k$\(^\text{145}\)If $P^*$ were properly contained in some $p$-subgroup $R$ of $G^*$, the intersection of $P$ and a Sylow $p$-subgroup of $G$ containing $R$ would be of larger order than $|S|$, a contradiction. – In general, given $P \in \text{Syl}_p(G)$ for a finite group $G$ and a prime $p$, a Sylow $p$-subgroup $Q$ of $G$ is said to have a tame intersection with $P$ if $N_P(P \cap Q)$, $N_Q(P \cap Q) \in \text{Syl}_p(N_G(P \cap Q))$. This notion is important because the following refinement of 8.9(2) holds:

**Theorem** (Alperin 1967). Let $p$ be a prime and $P, R$ Sylow $p$-subgroups of a finite group $G$. Then there exists a tuple $(Q_1, \ldots, Q_n)$ of Sylow $p$-subgroups of $G$ each of which has a tame intersection with $P$ such that for all $i \in \mathbb{N}$ there are elements $x_i$ of $p$-power order of $N_G(P \cap Q_i)$ with the properties $(P \cap R)^{x_1 \cdots x_{i-1}} \leq P \cap Q_i$ for all $i \in \mathbb{N}$, $R^{x_1 \cdots x_n} = P$.  

150
Bibliography

[Bo]  Bosch, Siegfried, Algebra, Springer 2006


[St]  Stroth, Gernot, Algebra, de Gruyter 1998
Index

action of a $K$-space, 32
algebra
  associative, 7
  commutative, 7
  homomorphism, 7
  $K$-, 7
  simple, 57
algebraic
  -ally closed, 22
  -ally dependent, 39
  -ally independent, 39
  element, 65
  field extension, 65
  integer, 83
algebraic integer, 126
algebraic number field, 99
alternating group, 50
basis
  $K$-, 6
  of a commutative monoid, 28
  of a module, 34
  of an algebra, 37
canonical epimorphism, 48
casus irreducibilis, 93
central chain
  lower, 137
  upper, 137
centralizer, 73, 131
centre
  of a group, 133
  of an associative algebra, 115
centre of a group, 70
character, 121
character table, 124
characteristic, 60
chief factor, 140
chief series, 140
class function, 122
class number of a group, 117
compatible
  relation, 42
  with an operation, 58
complement of a subgroup, 143
completely reducible, 103
conjugacy class, 117
conjugate, 70
core of a subgroup, 133
coset
  left, 47
  right, 47
cycle, 98
cycloic group, 51
cyclotomic field, 84, 128
decomposition, 143
  direct, 143
  semidirect, 143
Dedekind’s law, 27
degree, 8
degree of an irreducible algebra representation, 114
difference group, 100
direct co-summand, 102
direct decomposition, 102
direct sum, 102
direct summand, 102
directly indecomposable, 102
division algebra, 58
minimal right ideal, 103
module
\( K \)-space-, 101
algebra, 111
automorphism, 106
depth of an irreducible submodule, 126
endomorphism, 106
epimorphism, 106
homomorphism, 105
isomorphic, 106
isomorphism, 106
left, 111
monomorphism, 106
right, 111
semisimple, 103
simple, 103
module
\( K \)-space left-, 32
\( K \)-algebra left-, 32
\( K \)-algebra right-, 32
\( K \)-space right-, 32
homomorphism, 34
unital, 32
Moivre’s formula, 11
monoid, 13
monomial, 7
monomorphism, 7
multiplicative closure in a monoid, 27
multiplicity of a zero, 18
multiplicity of an irreducible submodule, 126
nilpotency class, 137
nilpotent element, 80
nilpotent group, 137
norm with respect to a Galois extension, 128
normal field extension, 71
normalizer, 131
number of irreducible representations of an algebra, 114
opposite algebra, 104
orbit, 131
order of a group, 47
order of a group element, 50
order of a radical, 92
\( \pi \)-group, 140
p-group, 138
permutation representation, 132
polynomial, 7
polynomial
elementary symmetric, 40
function, 39
minimal, 66
normed, 8
power-sum symmetric, 41
primitive, 82
pure, 9
ring, 7
ring over a set, 37
separable, 72
symmetric, 40
power series, 8
power series algebra, 8
prime field, 61
prime ring, 60
primitive root of unity, 15
principal ideal, 62
principal ideal domain, 62
product set, 13
projection onto a direct summand, 106
quotient field, 44
quotient of a magma, 58
radical, 9
radical extension, 92
regular
left, 102
right, 102
replacement homomorphism, 39
representation
\( K \)-space-, 101
algebra, 111
group, 132
root
of a real number, 10
of unity, 10
self-normalizing subgroup, 146
semidirect product, 144
semigroup, 13
semigroup algebra, 37
semigroup ring, 37
semisimple algebra, 113
separable
    element, 72
    field extension, 71
similar group actions, 134
simple group, 48
skew field, 17, 145
soluble
    by radicals, 93
    group, 96
space, $K-$, 6
special linear group, 50
splitting field, 22
subgroup
    lattice, 53
    maximal, 55
    normal, 48
submagma, 13
submodule, 101
submonoid
    generated unital, 27
    unital, 27
subsemigroup, 13
substitution, 9
supremum, 53
Sylow subgroup, 54, 136
Sylow system, 142
symmetric group, 14
symmetric mapping, 24
tame intersection, 150
torsion subgroup, 92
transcendental, 65
transitive group action, 133
trivial character, 121
unit in a ring, 29
unital, 13
    monoid action, 131
representation, 111
unitary, 5
Viète’s theorem, 23
zero
    multiple, 18
    of a polynomial, 9, 40
zero divisor, 8