Algebra 1

Oliver Lorscheid

Lecture notes, IMPA,
March–June 2020

Contents

Preface

These lecture notes were written during the Corona crisis in 2020 when the lectues at IMPA were suspended, and they have served as the main reference for Algebra 1 in this term, which was taught online.

Even though there are several excellent books on all themes of the lecture, the overall selection of topics is specific to the needs at IMPA and lacks a good standard reference. In so far there was a need for a source that is better adapted to the course, and this text can be seen as an attempt to fill this gap.

All the exercises from the weekly homework of the course are included at the end of the corresponding chapters. An asterisque indicates exercises that are significantly more challenging than others.

Acknowledgements: I would like to thank Rafael Ferreira, Amadeus Cabral Maldonado, Rafael Xavier, Xia Xiao and Zhifei Yan for their valuable feedback on previous versions of the text.

Chapter 1

Rings

1.1 Commutative groups and monoids

Definition 1.1.1. A **commutative group** (or **abelian group**) is a set G together with a *binary operation*, which is a map

$$\begin{array}{rccc} \mu: & G\times G & \longrightarrow & G, \\ & (a,b) & \longmapsto & ab = a\cdot b \end{array}$$

such that the following axioms hold:

(1) $(ab)c = a(bc)$ for all $a,b,c \in G$, *(associativity)*

(2) $ab = ba$ for all $a,b \in G$, *(commutativity)*

(3) there is an $e \in G$ such that $ae = a$ for all $a \in G$, *(neutral element)*

(4) for every $a \in G$, there is a $b \in G$ such that $ab = e$ *(inverses)*

where $(ab)c = \mu(\mu(a,b),c)$ and $a(bc) = \mu(a,\mu(b,c))$. A **commutative monoid** is a set G together with a map $\mu : G\times G \to G$ that satisfies (1)–(3).

Remark. A group is a set G together with a map $\mu : G\times G \to G$ that satisfies axioms (1), (3) and (4). We restrict ourselves to commutative groups in this section since we do not need the more complex and intriguing concept of a group for most parts of this course. For its importance for subsequent courses, we review the fundamental results from group theory at a later point of this course; cf. Chapter 5.

Lemma 1.1.2. *Let G be a set and $\mu : G\times G \to G$ a binary operation that satisfies axioms (1) and (2). Then there is at most one $e \in G$ such that $ae = a$ for all $a \in G$. If such a neutral element e exists, then there is at most one b for every a such that $ab = e$.*

Proof. Assume that (3) is satisfied for e and e'. Then $e = ee' = e'e = e'$, as claimed. Assume that $ab = ab' = e$ for a neutral element $e \in G$. Then $b = be = bab' = b'ab = b'e = b'$, as claimed. □

Notation. Typically, we suppress the multiplication μ of a commutative group (or monoid) and simply refer to a commutative group (or monoid) by the underlying set G. In situation where we want to specify the operation, we refer to a commutative group (or monoid) by (G,μ) or $(G,\cdot)$. Sometimes it is more natural to denote the operation of the group by "$+$". All these different notations are illustrated in Example 1.1.3.

Note that the associativity (1) allows us to write terms like abc without ambiguity. Using commutativity multiple times allows us to reorder the terms in any expression arbitrarily, e.g. $abc = cba$.

We typically denote the neutral element e by 1 and the inverse of a by a^{-1} if the binary operation μ is multiplication "$\cdot$". If μ is addition "$+$", then we denote the neutral element by 0 and the inverse of a by $-a$. By the following fact, the neutral element and inverses are unique.

Example 1.1.3. In the following, we give a series of examples of commutative groups and monoids.

(1) The integers $\mathbb{Z}$ together with addition $+$ forms a commutative group. The neutral element is 0 and the (additive) inverse of $a \in \mathbb{Z}$ is $-a$. The integers $\mathbb{Z}$ together with multiplication $\cdot$ form a commutative monoid whose neutral element is 1. The only elements with (multiplicative) inverses are 1 and -1.

(2) The natural numbers $\mathbb{N}$ together with addition form a commutative monoid with neutral element 0, which is not a commutative group since $-a$ is not in $\mathbb{N}$ unless $a = 0$. Also $(\mathbb{N},\cdot)$ is a commutative monoid.

(3) The rational numbers $\mathbb{Q}$ together with addition $+$ form a commutative group as well as the positive rational numbers $\mathbb{Q}_{>0}$ together with multiplication.

(4) The complex numbers $\{z \in \mathbb{C} \,|\, |z| = 1\}$ of absolute value 1 together with multiplication $\cdot$ form a commutative group, but $(\mathbb{C},\cdot)$ is only a commutative monoid.

We summarize this in Table 1.1.

	commutative group?	commutative monoid?
$(\mathbb{Z},+)$	✓	✓
$(\mathbb{Z},\cdot)$	no inverses	✓
$(\mathbb{N},+)$	no inverses	✓
$(\mathbb{N},\cdot)$	no inverses	✓
$(\mathbb{Q},+)$	✓	✓
$(\mathbb{Q}_{>0},\cdot)$	✓	✓
$(\{z \in \mathbb{C} \,\|\, \|z| = 1\},\cdot)$	✓	✓
$(\mathbb{C},\cdot)$	no inverses	✓

Table 1.1: Examples of commutative groups and monoids

Definition 1.1.4. Let G be a commutative group. A **subgroup of** G is a nonempty subset H of G such that $ab^{-1} \in H$ for all $a,b \in H$.

Lemma 1.1.5. *Let G be a commutative group and H a subset. Then H is a subgroup if and only if the multiplication μ of G restricts to a map $\mu_H : H \times H \to H$ such that (H, μ_H) is a commutative group.*

Proof. The proof is left as Exercise 1.1. □

We continue with the definition of the quotient of a commutative group G by a subgroup H. Since most applications of this construction in this lecture concerns additive groups, i.e. the group operation is addition $\mu(a,b) = a+b$, we formulate this result for additive groups.

Definition 1.1.6. Let $(G,+)$ be a commutative group, $a \in G$ and H a subgroup. The **coset of H in G with respect to a** is the subset

$$[a] \;=\; a+H \;=\; \{a+h \in G \,|\, h \in H\}$$

of G. The **quotient of G by H** is the set $G/H = \{[a] \,|\, a \in G\}$ of cosets of H.

Proposition 1.1.7 (Universal property of quotient groups)**.** *Let $(G,+)$ be a commutative group and H a subgroup.*

(1) Let $a,b \in G$. The following are equivalent:

- *(a)* $[a] = [b]$;
- *(b)* $a \in [b]$;
- *(c)* $[a] \cap [b] \neq \varnothing$;
- *(d)* $a - b \in H$.

(2) The map

$$\begin{array}{rccc} [+]: & G/H \times G/H & \longrightarrow & G/H \\ & ([a],[b]) & \longmapsto & [a+b] \end{array}$$

is well defined, and $(G/H, [+])$ is a commutative group.

Proof. We begin with proving the equivalence of the affirmations in (1). Assume (1a). Then $a = a+e \in [a] = [b]$, thus (1b). Since $a = a+e \in [a]$, (1b) implies (1c). Assume (1c). Then there is a $c \in [a] \cap [b]$, i.e. $c = a+h = b+h'$ for some $h, h' \in H$. Thus $a - b = h' - h \in H$, which implies (1d). Assume (1d), i.e. $h = a - b \in H$, and consider $a + h' \in [a]$ for some $h' \in H$. Since $a = b+h$ and $h+h' \in H$, this implies that $a+h' = b+h+h' \in [b]$, which shows that $[a] \subset [b]$. By the symmetry of the argument, we conclude that $[a] = [b]$, which shows (1a). This completes the proof of (1).

We continue with the proof of (2). To establish that $[+]$ is well-defined, consider cosets $[a] = [a']$ and $[b] = [b']$ of H in G, i.e. $a' = a+h$ and $b' = b+h'$ for some $h, h' \in H$. Then $a' + b' - (a+b) = h + h' \in H$, which means by (1) that $[a'+b'] = [a+b]$. This shows that the definition of $[+]$ does not depend on the choice of representative, i.e. $[+]$ is well-defined.

We verify that $(G/H,[+])$ satisfies the axioms of a commutative group. Associativity follows from

$$\Big([a][+][b]\Big)[+][c] \;=\; [(a+b)+c] \;=\; [a+(b+c)] \;=\; [a][+]\Big([b][+][c]\Big),$$

commutativity follows from

$$[a][+][b] \;=\; [a+b] \;=\; [b+a] \;=\; [b][+][a],$$

the neutral element is $[e]$ since

$$[a][+][e] \;=\; [a+e] \;=\; [a]$$

and the inverse of $[a]$ is $[-a]$ since

$$[a][+][-a] \;=\; [a-a] \;=\; [e]$$

where a, b and c are arbitrary elements of G. □

Notation. In the following, we denote the addition of the quotient G/H simply by "+", i.e. we write $[a]+[b]=[a+b]$. Sometimes we write $\bar{a}$ for $[a]$.

Example 1.1.8. We give some examples of subgroups and quotient groups.

(1) Let G be a commutative group. Then $\{0\}$ and G are subgroups, which are called the **trivial subgroup** and the **improper subgroup**, respectively. The quotient $G/\{0\}$ can be identified with G itself since $[a]=[b]$ if and only if $a=b$. The quotient G/G is equal to $\{[e]\}$ since $[a]=[e]$ for all $a\in G$.

(2) For every $n\geqslant 0$, the subset $\mathbb{Z}=\{na|a\in\mathbb{Z}\}$ of $\mathbb{Z}$ is a subgroup of $(\mathbb{Z},+)$. We have an equality $[a]=[b]$ of cosets of $n\mathbb{Z}$ if and only if $a-b\in n\mathbb{Z}$, i.e. $a-b$ is divisible by n. For $n=0$, we find the trivial subgroup $0\mathbb{Z}=\{0\}$ and the quotient $\mathbb{Z}/0\mathbb{Z}=\mathbb{Z}$. For $n=1$, we find the improper subgroup $1\mathbb{Z}=\mathbb{Z}$ and the quotient $\mathbb{Z}/1\mathbb{Z}=\{[0]\}$. For $n\geqslant 1$, the quotient $\mathbb{Z}/n\mathbb{Z}=\{[a]|a\in\mathbb{Z}\}$ consists of the cosets $[0],\ldots,[n-1]$, and the addition of $\mathbb{Z}/n\mathbb{Z}$ is *addition modulo n*, i.e.

$$[a]+[b] \;=\; \begin{cases} [a+b] & \text{if } a+b<n, \\ [a+b-n] & \text{if } a+b\geqslant n. \end{cases}$$

1.2 Rings, integral domains and fields

Definition 1.2.1. A **(commutative) ring (with one)** is a set A together with two binary operations

$$\alpha:\; \begin{array}{ccc} A\times A & \longrightarrow & A \\ (a,b) & \longmapsto & a+b \end{array} \qquad \text{and} \qquad \mu:\; \begin{array}{ccc} A\times A & \longrightarrow & A \\ (a,b) & \longmapsto & ab=a\cdot b \end{array}$$

such that

(1) $(A,+)$ is a commutative group,

(2) $(A,\cdot)$ is a commutative monoid,

(3) $a(b+c) = ab+ac$ for all $a,b,c \in A$ *(distributivity)*

where $a(b+c) = \mu\big(a,\alpha(b,c)\big)$ and $ab+ac = \alpha\big(\mu(a,b),\mu(a,c)\big)$ (i.e. multiplication is evaluated before addition).

Notation. For the purpose of these lectures, we assume that all rings are commutative and with one, unless stated otherwise. The binary operation α is called the *addition* of A and μ its *multiplication*. Often we suppress the binary operations from the notation of a ring A, but in instances where we want to explicitly refer to the addition $+$ and the multiplication $\cdot$ by $(A,+,\cdot)$; cf. Example 1.2.3 for some instances.

We denote the neutral element for the addition by 0 and the additive inverse of an element a by $-a$. We write $a-b$ for $a+(-b)$. We denote the neutral element for multiplication by 1 and the multiplicative inverse of an element a by a^{-1} if it exists. Note that by Lemma 1.1.2, all these elements are unique and thus these notations are well-defined.

Lemma 1.2.2. *The following hold true for every ring A.*

(1) We have $-0 = 0$, $0\cdot a = 0$ and $-a = (-1)\cdot a$ for all $a \in A$, as well as $(-1)^2 = 1$.

(2) If $a-b = 0$, then $a = b$.

(3) If $0 = 1$, then $A = \{0\}$.

(4) If $a,b \in A$ have multiplicative inverses a^{-1} and b^{-1}, then $a^{-1}b^{-1}$ is a multiplicative inverse of ab.

Proof. We begin with (1). Since $0+0 = 0$, we have $0 = -0$ by the uniqueness of (additive) inverses. For $a \in A$, we have $0\cdot a + 0\cdot a = (0+0)a = 0\cdot a$ and thus, after adding $-0\cdot a$ to both sides, $0\cdot a = 0$, as claimed. In consequence, $0 = 0\cdot a = (1-1)a = 1\cdot a + (-1)\cdot a$ and thus $-a = (-1)\cdot a$ by the uniqueness of (additive) inverses. Finally, we have $0 = 0\cdot(-1) = (1-1)\cdot(-1) = 1\cdot(-1)+(-1)\cdot(-1) = (-1)+(-1)^2$ and thus $(-1)^2 = 1$ by the uniqueness of the (additive) inverse of -1, which establishes all claims of (1).

We turn to (2)–(4). If $a-b = 0$, then $a = -(-b)$, and thus $a = (-1)\cdot(-1)\cdot b = b$ by (1), which establishes (2). If $0 = 1$, then $0 = 0\cdot a = q\cdot a = a$ for all $a \in A$ and thus $A = \{0\}$, which establishes (3). Finally, (4) follows from $(ab)(a^{-1}b^{-1}) = aa^{-1}bb^{-1} = 1\cdot 1 = 1$. □

Example 1.2.3. In the following, we discuss a series of examples and non-examples.

(0) The **trivial ring** $A = \{0\}$ with one element $0 = 1$ and the tautological addition $0+0 = 0$ and multiplication $0\cdot 0 = 0$ forms a ring.

(1) The integers $\mathbb{Z}$ together with the usual addition and multiplication form a ring. The same is true for the rational numbers $(\mathbb{Q},+,\cdot)$, the real numbers $(\mathbb{R},+,\cdot)$ and the complex numbers $(\mathbb{C},+,\cdot)$. The natural numbers $(\mathbb{N},+,\cdot)$ do not form a ring for the lack of additive inverses.

(2) Polynomials over rings from rings. For example the set $\mathbb{R}[T]$ of polynomials $a_nT^n + \cdots + a_1T + a_0$ with real coefficients $a_0, \ldots, a_n \in \mathbb{R}$, together with the usual addition and multiplication of polynomials, forms a ring.

(3) The set $\mathbb{R} \times \mathbb{R}$ of pairs (a,b) of real numbers $a,b \in \mathbb{R}$, together with componentwise addition and multiplication, forms a ring.

(4) The set of $n \times n$-matrices $\mathrm{Mat}_n(A)$ with coefficients in a ring A does not form a (*commutative*) ring for $n > 1$ since the multiplication of matrices is not commutative.

(5) The set $\mathcal{C}_c(\mathbb{R},\mathbb{R})$ of compactly supported (continuous / differentiable / smooth) functions $f : \mathbb{R} \to \mathbb{R}$, together with valuewise addition and multiplication, do not form a ring (*with one*) for the lack of a multiplicatively neutral element. To wit, the constant function $c_1 : \mathbb{R} \to \mathbb{R}$ that sends every $a \in \mathbb{R}$ to $c_1(a) = 1$ would be neutral for multiplication, but it does not have compact support.

Definition 1.2.4. Let A be a ring. A **subring of** A is a subset B of A such that $1 \in B$ and $a - b, ab \in B$ for all $a,b \in B$. The **unit group of** A is the set $A^\times$ of all elements $a \in A$ with a multiplicative inverse a^{-1}. A ring A is

- **without zero divisors** if $ab \neq 0$ for all $a,b \in A - \{0\}$;
- an **integral domain** if $0 \neq 1$ and if the multiplication map

$$\begin{array}{rccc} m_a : & A & \longrightarrow & A \\ & b & \longmapsto & ab \end{array}$$

  is injective for every $a \in A - \{0\}$;
- a **field** if $A^\times = A - \{0\}$, i.e. $0 \neq 1$ and every nonzero element has a multiplicative inverse.

Remark. Let $B \subset A$ be a subring of A. Then both addition α and multiplication μ of A restrict to maps $\alpha_B : B \times B \to B$ and $\mu_B : B \times B \to B$ that endow B with the structure of a ring, which justifies the term 'subring'. We leave a verification of this claim as Exercise 1.5.

By Lemma 1.2.2, the multiplication μ of a ring A restricts to a map $\mu^\times : A^\times \times A^\times \to A^\times$, which turns $A^\times$ into a commutative group. This justifies the term 'unit group'.

Lemma 1.2.5. *Every field is an integral domain, and every integral domain is without zero divisors. Conversely, a ring without zero divisors is an integral domain if* $0 \neq 1$.

Proof. Let A be a field and $a \in A - \{0\}$. Consider $b,b' \in A$ such that $m_a(b) = m_a(b')$. Since a has a multiplicative inverse a^{-1}, we have $b = a^{-1}ab = a^{-1}m_a(b) = a^{-1}m_a(b') = a^{-1}ab' = b'$, which shows that m_a is injective. Since $0 \neq 1$, this shows that A is an integral domain.

Let A be an integral domain and $a,b \in A$ with $ab = 0$, but $a \neq 0$. Then $m_a(b) = ab = 0 = a \cdot 0 = m_a(0)$ and thus $b = 0$ by the injectivity of m_a. Thus A is without zero divisors.

Let A be without zero divisors, $0 \neq 1$ and $a \in A - \{0\}$. If $m_a(b) = m_a(b')$, then $ab = ab'$ and thus $a(b-b') = ab - ab' = 0$. Since A is without zero divisors, we must have $b - b' = 0$ and thus $b = b'$. This shows that m_a is injective and that A is an integral domain. □

Example 1.2.6. We summarize which rings from Example 1.2.3 are without zero divisors, integral domains and fields in Table 1.2. The validity or failure of the defining properties of these particular types of rings can also be determined for the examples that are not rings, and we provide such an answer in brackets.

	ring?	without zero divisors?	integral domain?	field?
$\{0\}$	✓	✓	no	no
$\mathbb{Z}$	✓	✓	✓	no
$\mathbb{Q}$	✓	✓	✓	✓
$\mathbb{R}$	✓	✓	✓	✓
$\mathbb{C}$	✓	✓	✓	✓
$\mathbb{R}[T]$	✓	✓	✓	no
$\mathbb{R} \times \mathbb{R}$	✓	no	no	no
$\mathbb{N}$	no additive inverses	(✓)	(✓)	(no)
$\mathrm{Mat}_n(A)$	not commutative	(no)	(no)	(no)
$\mathcal{C}_c(\mathbb{R},\mathbb{R})$	without one	(no)	(no)	(no)

Table 1.2: Examples of rings

1.3 Ideals and quotients

Definition 1.3.1. Let A and B be rings. A **ring homomorphism from A to B** is a map $f : A \to B$ such that $f(1) = 1$, $f(a+b) = f(a) + f(b)$ and $f(ab) = f(a)f(b)$ for all $a, b \in A$. An **isomorphism of rings** is a bijective ring homomorphism.

Lemma 1.3.2. *Let A, B and C be rings. Then the following hold true.*

(1) The identity map $\mathrm{id}_A : A \to A$ is a ring homomorphism.

(2) The composition $g \circ f : A \to C$ of two ring homomorphisms $f : A \to B$ and $g : B \to C$ is a ring homomorphism.

(3) Given a ring homomorphism $f : A \to B$ and $a \in A$, we have $f(0) = 0$, $f(-a) = -f(a)$ and $f(a^{-1}) = f(a)^{-1}$, provided a has a multiplicative inverse a^{-1}.

(4) A ring homomorphism $f : A \to B$ is injective if and only if $f^{-1}(0) = \{0\}$.

(5) The image $\mathrm{im}\, f = \{f(a) | a \in A\}$ *of a ring homomorphism $f : A \to B$ is a subring of B.*

(6) A ring homomorphism $f : A \to B$ is an isomorphism if and only if there is a ring homomorphism $g : B \to A$ such that $g \circ f = \mathrm{id}_A$ and $f \circ g = \mathrm{id}_B$.

Proof. The proof is left as Exercise 1.8. □

Definition 1.3.3. Let A be a ring. An **ideal of** A is a subset I of A that contains 0, $b+c$ and ab for all $a \in A$ and $b, c \in I$.

Let S be a subset of A. The **ideal generated by** S is the subset

$$\langle S \rangle \;=\; \Big\{ \sum_{i=1}^{n} a_i s_i \,\Big|\, n \geqslant 1, a_i \in A, s_i \in S \Big\}$$

of A, with the convention that $\langle \varnothing \rangle = \{0\}$. If $S = \{a_i\}$ we also write $\langle a_i \rangle = \langle S \rangle$.

Remark. In other words, an ideal of A is an additive subgroup I of $(A,+)$ such that $AI = I$ where $AI = \{ab | a \in A, b \in I\}$. Note that, in particular, $a - b = a + (-1) \cdot b \in I$ for all $a, b \in I$.

Lemma 1.3.4. *Let A be a ring and S a subset of A. Then $\langle S \rangle$ is an ideal, and $S = \langle S \rangle$ if and only if S is an ideal.*

Proof. We begin with the proof that $\langle S \rangle$ is an ideal, which is clear in the case $\langle \varnothing \rangle = \{0\}$. Thus we assume that $S \neq \varnothing$ in the following. Clearly, $0 = 0 \cdot s$ is in $\langle S \rangle$. Consider $\sum_{i=1}^{n} a_i s_i, \sum_{i=1}^{m} b_i t_i \in \langle S \rangle$ and $c \in A$. If we define $a_i s_i = b_{i-n} s_{i-n}$ for $i = n+1, \ldots, n+m$, then both

$$\Big(\sum_{i=1}^{n} a_i s_i \Big) + \Big(\sum_{i=1}^{m} b_i t_i \Big) \;=\; \sum_{i=1}^{n+m} a_i s_i \qquad \text{and} \qquad c \cdot \Big(\sum_{i=1}^{n} a_i s_i \Big) \;=\; \sum_{i=1}^{n} (c a_i) s_i$$

are in $\langle S \rangle$, which shows that $\langle S \rangle$ is an ideal. This also implies that if $S = \langle S \rangle$, then S is an ideal.

Assume that S is an ideal. Clearly $S \subset \langle S \rangle$. To show equality, we consider $\sum_{i=1}^{n} a_i s_i \in \langle S \rangle$ with $a_i \in A$ and $s_i \in S$. Since S is an ideal, we know that $a_i s_i \in S$ for all $i = 1, \ldots, n$. Since S is an additive subgroup of A, it also contains the sum $\sum_{i=1}^{n} a_i s_i$. Thus $\langle S \rangle \subset S$, which concludes the proof. □

Example 1.3.5. We give some examples of ideals.

(1) Let A be a ring. Then subsets $\{0\}$ and A are ideals of A, which are called the **trivial ideal** and the **improper ideal** of A, respectively. For every $a \in A$, the ideal generated by $S = \{a\}$ equals

$$\langle a \rangle \;=\; \{ab \,|\, b \in A\}$$

and is called the **principal ideal generated by** a.

(2) In particular, for every $n \in \mathbb{Z}$, the subset $n\mathbb{Z} = \langle n \rangle$ is an ideal of $\mathbb{Z}$.

(3) Another example of a principal ideal is the subset

$$\langle (1,0) \rangle \;=\; \big\{ (a,0) \in \mathbb{R} \times \mathbb{R} \,\big|\, a \in \mathbb{R} \big\}$$

of $\mathbb{R} \times \mathbb{R}$.

Definition 1.3.6. Let $f : A \to B$ be a ring homomorphism. The **kernel of** f is the subset $\ker f = \{a \in A | f(a) = 0\}$ of A.

Lemma 1.3.7. *The kernel of a ring homomorphism $f : A \to B$ is an ideal of A.*

Proof. Since $f(0) = 0$, we have $0 \in \ker f$. Consider $a \in A$ and $b, c \in \ker f$, i.e. $f(b) = f(c) = 0$. Then $f(b+c) = f(b) + f(c) = 0$ and $f(ab) = f(a)f(b) = 0$, which shows that $b+c$ and ab are in $\ker f$. Thus $\ker f$ is an ideal of A. □

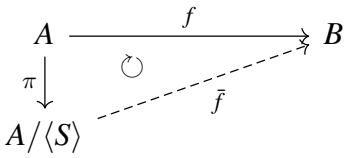
Proposition 1.3.8 (Universal property for quotient rings)**.** *Let A be a ring and I an ideal of A. For $a \in A$, define $[a] = \{a + b \in A | b \in I\}$ and $A/I = \{[a] | a \in A\}$. Then the following hold true.*

(1) The maps

$$[+]: \begin{array}{ccc} A/I \times A/I & \longrightarrow & A/I \\ ([a],[b]) & \longmapsto & [a+b] \end{array} \quad \text{and} \quad [\cdot]: \begin{array}{ccc} A/I \times A/I & \longrightarrow & A/I \\ ([a],[b]) & \longmapsto & [ab] \end{array}$$

are well-defined, and $(A/I, [+], [\cdot])$ is a ring. The association $a \mapsto [a]$ defines a ring homomorphism $\pi : A \to A/I$ with $\ker \pi = I$.

(2) Let S be a subset of A. Then $\pi : A \to A/\langle S \rangle$ satisfies the following universal property*: for every ring homomorphism $f : A \to B$ with $f(S) \subset \{0\}$, there is a unique ring homomorphism $\bar{f} : A/\langle S \rangle \to B$ such that $f = \bar{f} \circ \pi$, i.e. the diagram*



commutes.

Proof. Since an ideal is, in particular, an additive subgroup of $(A, +)$, Proposition 1.1.7 shows that $[+]$ is well-defined. We continue with $[\cdot]$. Let $[a] = [a']$ and $[b] = [b']$, i.e. $a' = a + c$ and $b' = b + d$ for $c, d \in I$. Then $ad, bc, cd \in I$ and thus $a'b' - ab = ad + bc + cd \in I$, which shows that $[ab] = [a'b']$, by Proposition 1.1.7, and that $[\cdot]$ is well-defined.

We continue with the verification that $(A/I, [+], [\cdot])$ is a ring. By Proposition 1.1.7, we know that $(A/I, [+])$ is a commutative group. That $(A/I, [\cdot])$ is a commutative monoid with neutral element $[1]$ follows from

$$\Big([a][\cdot][b]\Big)[\cdot][c] \;=\; [(ab)c] \;=\; [a(bc)] \;=\; [a][\cdot]\Big([b][\cdot][c]\Big),$$

$$[a][\cdot][b] \;=\; [ab] \;=\; [ba] \;=\; [b][\cdot][a], \qquad \text{and} \qquad [a][\cdot][e] \;=\; [ae] \;=\; [a]$$

where a, b and c are arbitrary elements of A. The distributivity of $[\cdot]$ over $[+]$ follows from

$$[a][\cdot]\Big([b][+][c]\Big) \;=\; [a(b+c)] \;=\; [ab+ac] \;=\; [a][\cdot][b][+][a][\cdot][c],$$

which concludes the proof that $(a/I, [+], [\cdot])$ is a ring.

The claim that $\pi : A \to A/I$ is a ring homomorphism follows from $\pi(1) = [1]$ and

$$\begin{aligned} \pi(a+b) &= [a+b] = [a][+][b] = \pi(a)[+]\pi(b), \\ \pi(ab) &= [ab] = [a][\cdot][b] = \pi(a)[\cdot]\pi(b) \end{aligned}$$

for all $a, b \in A$. By Proposition 1.1.7, (1), we have $[0] = I$ and $\pi(a) = [0]$ if and only if $a \in I$. Thus $\pi^{-1}([0]) = I$ as claimed, which concludes the proof of (1)

We continue with the proof of (2). Given a ring homomorphism $f : A \to B$ with $f(S) \subset \{0\}$, we claim that the association $[a] \mapsto f(a)$ does not depend on the choice of representative $a \in [a]$ and defines a ring homomorphism $A/\langle S \rangle \to B$. Once we have proven this, it is clear that from the definition of $\bar{f}$ that $f = \bar{f} \circ \pi$. Note that $f = \bar{f} \circ \pi$ implies the uniqueness of $\bar{f}$ since it requires that $\bar{f}([a]) = \bar{f}(\pi(a)) = f(a)$.

In order to show that $\bar{f}$ is well-defined, consider $a, b \in A$ such that $[a] = [b]$. By Proposition 1.1.7, we have $a - b \in \langle S \rangle$, and thus $a - b = \sum a_i s_i$ for some $a_i \in A$ and $s_i \in S$. It follows that

$$f(a) = f(b + \sum a_i s_i) = f(b) + \sum f(a_i) \underbrace{f(s_i)}_{=0} = f(b),$$

which shows that the value $\bar{f}([a]) = \bar{f}([b])$ does not depend on the choice of representative for $[a] = [b]$. The map $\bar{f}$ is a ring homomorphism since $\bar{f}([1]) = f(1) = 1$ and

$$\begin{aligned} \bar{f}([a][+][b]) &= f(a+b) = f(a) + f(b) = \bar{f}([a]) + \bar{f}([b]), \\ \bar{f}([ab]) &= f(ab) = f(a) \cdot f(b) = \bar{f}([a]) \cdot \bar{f}([b]) \end{aligned}$$

for all $a, b \in A$. This concludes the proof of the proposition. □

Notation. In the following, we denote the addition and the multiplication of a quotient ring A/I simply by $+$ and $\cdot$, respectively. We sometimes write $\bar{a}$ for $[a]$.

Example 1.3.9. (0) If $I = \{0\}$ is the trivial ideal of a ring A, then the quotient map $\pi : A \to A/\{0\}$ is an isomorphism of rings.

(1) Let $I = n\mathbb{Z} = \langle n \rangle$ be the principal ideal of $\mathbb{Z}$ generated by a positive integer $n \in \mathbb{Z}$. Then $\mathbb{Z}/n\mathbb{Z} = \{[0], \ldots, [n-1]\}$, as explained in Example 1.1.8. Similar to addition, multiplication is calculated *modulo n*, i.e.

$$[a] \cdot [b] = [ab - kn]$$

where $k \in \mathbb{N}$ is such that $0 \leqslant ab - kn < n$.

Proposition 1.3.10. *Let $n \geqslant 0$. Then $\mathbb{Z}/n\mathbb{Z}$ is an integral domain if and only if $n = 0$ or $n = p$ is prime. If $n = p$ is prime, then $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ is a field.*

Proof. If $n = 0$, then $\mathbb{Z}/0\mathbb{Z} = \mathbb{Z}$ is an integral domain. If $n = 1$, then $\mathbb{Z}/1\mathbb{Z} = \{0\}$ is not. If $n > 1$ is not prime, i.e. $n = k \cdot l$ for some $k, l > 1$, then $\bar{k} \neq \bar{0} \neq \bar{l}$ and $\bar{k} \cdot \bar{l} = \bar{0}$ in $\mathbb{Z}/n\mathbb{Z}$, which shows that $\mathbb{Z}/n\mathbb{Z}$ is with zero divisors and thus not an integral domain.

If $n = p$ is prime, then $\bar{0} \neq \bar{1}$ in $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$. Consider $\bar{k}, \bar{l} \in \mathbb{F}_p$ with $\bar{k} \cdot \bar{l} = \bar{0}$, but $\bar{k} \neq \bar{0}$. Then in $\mathbb{Z}$, the prime number p does divide $k \cdot l$, but not k. By the unique prime factorization of integers, p divides l, i.e. $\bar{l} = \bar{0}$ in $\mathbb{F}_p$. Thus $\mathbb{F}_p$ is an integral domain. Since $\mathbb{F}_p$ is finite, it is a field by Exercise 1.6. □

Definition 1.3.11. Let A be a ring. The **polynomial ring over** A is the set

$$A[T] \;=\; \Big\{ \sum_{i\in\mathbb{N}} a_i T^i \,\Big|\, a_i \in A,\ a_i = 0 \text{ for all but finitely many } i \in \mathbb{N} \Big\}$$

together with the addition

$$\big(\sum_{i\in\mathbb{N}} a_i T^i\big) + \big(\sum_{i\in\mathbb{N}} b_i T^i\big) \;=\; \sum_{i\in\mathbb{N}} (a_i + b_i) T^i$$

and the multiplication

$$\big(\sum_{i\in\mathbb{N}} a_i T^i\big) \cdot \big(\sum_{i\in\mathbb{N}} b_i T^i\big) \;=\; \sum_{i\in\mathbb{N}} \Big(\sum_{k+l=i} a_k \cdot b_l\Big) T^i.$$

A **polynomial over** A is an element $f = \sum_{i\in\mathbb{N}} a_i T^i$ of $A[T]$. Its **degree** is

$$\deg f \;=\; \max\{i \in \mathbb{N} \,|\, a_i \neq 0\}$$

if $a_i \neq 0$ for some i, and $\deg f = 0$ if $a_i = 0$ for all i. If $n = \deg f$, then a_n is called the **leading coefficient of** f and a_0 is called the **constant coefficient of** f. If $a_n = 1$, then f is called **monic**.

Notation. We sometimes omit the index i from the sum and write $\sum a_i T^i$. If $a_i = 0$ for $i > n$, then we also write $\sum a_i T^i = a_n T^n + \cdots + a_1 T + a_0$. In particular, we consider elements $a \in A$ as *constant polynomials* $\sum a_i T^i = a_0 = a$ where $a_i = 0$ for all $i > 0$. In particular, the *trivial polynomial* is 0, i.e. $a_i = 0$ for all $i \in \mathbb{N}$, and the *identity polynomial* is 1, i.e. $a_0 = 1$ and $a_i = 0$ for $i > 0$. We leave it as Exercise 1.10 to verify that $A[T]$ is indeed a ring with zero 0 and one 1.

Proposition 1.3.12. *Let K be a field and $f = \sum a_i T^i \in K[T]$ be a nonzero polynomial of degree $n \geqslant 1$. Then the map*

$$\begin{aligned} \Big\{ \sum_{i=0}^{n-1} b_i T^i \,\Big|\, b_i \in K \Big\} \quad &\longrightarrow \quad K[T]/\langle f \rangle \\ g = \sum_{i=0}^{n-1} b_i T^i \quad &\longmapsto \quad g + \langle f \rangle \end{aligned}$$

is a bijection.

Proof. We begin with the injectivity of the map. If g and g' are polynomials of degree $< n$ whose classes $[g] = [g']$ in $K[T]/\langle f \rangle$ are equal, then $g - g' = f \cdot h$ for some polynomial $h = \sum c_i T^i$ over K. Assume that $h \neq 0$, i.e. $c_m \neq 0$ for $m = \deg h$. Then

$$f \cdot h \;=\; \Big(\sum_{i=0}^{n} a_i T^i\Big) \cdot \Big(\sum_{i=0}^{m} c_i T^i\Big) \;=\; a_n c_m T^{n+m} + (\text{lower terms}),$$

which is a polynomial of degree $n+m \geqslant n$. But this cannot be since $\deg(g-g') \leqslant \max\{\deg g, \deg g'\} < n$. We conclude that h must be the trivial polynomial and thus $g = g'$, which establishes the injectivity of the map of the proposition.

We turn to the surjectivity of the map. Consider a class $[g] = g + \langle f \rangle$ in $K[T]/\langle f \rangle$ that is represented by a polynomial $g = \sum b_i T^i \in K[T]$ of minimal degree, i.e. $\deg g' \geqslant \deg g$ for all $g' \in [g]$. We need to show that $m = \deg g < n$. If this was not the case, i.e. $m \geqslant n$, then we could define

$$g' \;=\; g - \frac{b_m}{a_n} T^{m-n} \cdot f \;=\; \underbrace{\Big(b_m - \frac{b_m}{a_n} \cdot a_n\Big)}_{=0} T^m + (\text{terms of degree} < m),$$

which is a polynomial of degree smaller than $m = \deg g$ in the class $[g]$, a contradiction to the minimality of $\deg g$. This shows that $\deg g < n$ and concludes the proof of surjectivity. □

Definition 1.3.13. An ideal I of a ring A is

- **prime** if $S = A - I$ is a *multiplicative subset*, i.e. $1 \in S$ and $ab \in S$ for all $a, b \in S$;
- **proper** if $I \neq A$;
- **maximal** if it is proper and $I \subset J$ implies $I = J$ for all proper ideals J of A.

Lemma 1.3.14. *Let A be a ring and I and ideal of A.*

(1) The ideal I is prime if and only if A/I is an integral domain.

(2) The ideal I is maximal if and only if A/I is a field.

Proof. We begin with (1). By definition I is prime if and only if $1 \in S$ and for all $a, b \in S$ also $ab \in S$. This is equivalent with the condition that $\bar{0} \neq \bar{1}$ in A/I and $\bar{a}\bar{b} \neq \bar{0}$ for all $\bar{a}, \bar{b} \in A/I - \{\bar{0}\}$, which are precisely the defining properties of an integral domain. Thus (1).

We turn to (2). Assume that I is maximal. Since $I \neq A$, we have $A/I \neq \{\bar{0}\}$ and thus $\bar{0} \neq \bar{1}$ by Lemma 1.2.2. Consider $\bar{a} \in A/I - \{\bar{0}\}$. Then $a \notin I$ and $\langle I \cup \{a\} \rangle = A$ by the maximality of I. Thus we have $1 = c + ba$ for some $c \in I$ and $b \in A$, by Lemma 1.3.4 and noting that I is closed under addition and multiplication by elements of A. Thus we have $\bar{c} = \bar{0}$ and $\bar{1} = \bar{c} + \bar{b}\bar{a} = \bar{b}\bar{a}$ in A/I, which shows that $\bar{a}$ has a multiplicative inverse, which is $\bar{b}$. This shows that A/I is a field.

Conversely assume that A/I is a field. Since $\bar{0} \neq \bar{1}$, we conclude that I is a proper ideal. Consider an ideal J of A such that $I \subset J$. If there is an element $a \in J - I$, then

$\bar{a} \neq \bar{0}$ in A/I and thus has a multiplicative inverse $\bar{b}$, i.e. $\bar{a}\bar{b} = \bar{1}$ or $1 - ab \in I$. Since $a \in J$, also $ab \in J$. Therefore $1 = (1-ab) + ab \in J$ and $J = A$. This shows that I is maximal and concludes the proof of the lemma. □

Corollary 1.3.15. *Every maximal ideal is a prime ideal.*

Proof. Let A be a ring and I a maximal ideal of A. Then A/I is a field by Lemma 1.3.14 and thus an integral domain by Lemma 1.2.5. Again by Lemma 1.3.14, we conclude that I is a prime ideal. □

1.4 The isomorphism theorems for rings

Theorem 1.4.1 (First isomorphism theorem)**.** *A ring homomorphism $f : A \to B$ induces an isomorphism of rings*

$$\begin{array}{rccc} \bar{f}: & A/\ker f & \longrightarrow & \operatorname{im} f. \\ & \bar{a} & \longmapsto & f(a) \end{array}$$

Proof. By the definition of $\ker f$ as $f^{-1}(0)$, we have $f(\ker f) = \{0\}$. Thus by the universal property of the quotient $A/\ker f$ (Proposition 1.3.8), f factors into the quotient map $\pi : A \to A/I$ composed with a (uniquely defined) ring homomorphism $\hat{f} : A/\ker f \to B$. Since $f = \hat{f} \circ \pi$, we have necessarily that $\hat{f}(\bar{a}) = f(a)$. By the definition of the subring $\operatorname{im} f$ of B, $\hat{f}$ restricts to a surjective ring homomorphism $\bar{f} : A/\ker f \to \operatorname{im} f$, which is the map described in the theorem since $\bar{f}(\bar{a}) = \hat{f}(\bar{a}) = f(a)$. The ring homomorphism $\bar{f}$ is injective since $\ker \bar{f} = \pi(\ker f) = \{\bar{0}\}$. This shows that $\bar{f}$ is an isomorphism, as claimed. □

Theorem 1.4.2 (Second isomorphism theorem)**.** *Let A be a ring, B a subring and I an ideal of A. Then the following holds true.*

(1) $B + I = \{b + c | b \in B, c \in I\}$ *is a subring of A.*

(2) $B \cap I$ *is an ideal of B.*

(3)

$$\begin{array}{rccc} \Phi: & B/(B\cap I) & \longrightarrow & (B+I)/I \\ & a+(B\cap I) & \longmapsto & a+I \end{array}$$

is an isomorphism of rings.

Proof. We begin with (1). Since $0, 1 \in B$ and $0 \in I$, both $0 = 0+0$ and $1 = 1+0$ are in $B+I$. For all $b, b' \in B$ and $c, c' \in I$, both $(b+c) - (b'-c') = (b-b') + (c-c')$ and $(b+c)(b'+c') = bb' + (bc' + b'c + cc')$ are elements of $B+I$. Thus $B+I$ is a subring of A.

We turn to (2). Clearly $0 \in B \cap I$. Given $a \in B$ and $b, c \in B \cap I$, we have that ab and $b+c$ are in both B and I. Thus $B \cap I$ is an ideal of B.

We turn to (3). The map Φ is a ring homomorphism since $\Phi(1+B\cap I)=1+I$ and for all $a,b\in B$,

$$\begin{aligned}\Phi\big([a]+[b]\big) &= \Phi\big([a+b]\big) = [a+b] = [a]+[b] = \Phi\big([a]\big)+\Phi\big([b]\big),\\ \Phi\big([a]\cdot[b]\big) &= \Phi\big([a\cdot b]\big) = [a\cdot b] = [a]\cdot[b] = \Phi\big([a]\big)\cdot\Phi\big([b]\big).\end{aligned}$$

where $[a]$ stands for $a+B\cap I$ where it appears as an argument of Φ and for $a+I$ otherwise.

We turn to the injectivity of Φ. An element $a\in B$ is in $\ker\Phi$ if and only if $\Phi([a])=[0]=I$, i.e. $a\in I$. Since $a\in B$, this is equivalent with a being an element of $B\cap I$, i.e. $a+B\cap I=[0]$. Thus $\ker\Phi=\{[0]\}$, which shows injectivity.

To verify the surjectivity of Φ, consider an element $b+c\in B+I$ with $b\in B$ and $c\in I$. Then $[b+c]=[b+c-c]=[b]$ as classes of $B+I/I$, and thus $[b+c]=[b]=\Phi([b])$ is in the image of Φ. This completes the proof of the theorem. □

Theorem 1.4.3 (Third isomorphism theorem)**.** *Let A be a ring, I an ideal of A and $\pi:A\to A/I$ the quotient map. Then*

$$\begin{array}{rccc}\Phi: & \big\{\textit{ideals of } A \textit{ containing } I\big\} & \longrightarrow & \big\{\textit{ideals of } A/I\big\}\\ & J & \longmapsto & J/I=\pi(J)\end{array}$$

is an inclusion preserving bijection, and

$$\begin{array}{rccc}f: & A/J & \longrightarrow & (A/I)/(J/I)\\ & a+J & \longmapsto & ([a]+J)/I\end{array}$$

is a ring isomorphism for every ideal J of A containing I.

Proof. We begin with showing that $\Phi(J)=J/I$ is indeed an ideal of A/I where J is an ideal of A that contains I. Clearly $[0]\in J/I$. Consider $a\in A$ such that $[a]\in J/I$, i.e. there are $b\in J$ and $c\in I$ such that $a=b+c$. Since $I\subset J$, we conclude that $a=b+c\in J$. Thus $J=\pi^{-1}(J/I)$, which implies for $a,b,c\in A$ with $[b],[c]\in J$ that both $[a]\cdot[b]=[ab]$ and $[b]+[c]=[b+c]$ are in J/I. This shows that J/I is an ideal and thus that Φ is well-defined. Moreover, the fact that $J=\pi^{-1}(J/I)$ implies at once that Φ is injective.

By Exercise 1.13, the inverse image $\pi^{-1}(\bar{J})$ of an ideal $\bar{J}$ of A/I is an ideal of A. Note that $\pi^{-1}(\bar{J})$ contains I. Since π is surjective, we have $\bar{J}=\pi\big(\pi^{-1}(\bar{J})\big)$, which shows that Φ is surjective. It is clear that Φ is inclusion preserving, which concludes the proof of the first claim of the theorem.

The association $a\mapsto[a]+J/I$ defines a surjective ring homomorphism $g:A\to(A/I)/(J/I)$ with kernel $I+J=J$. By Theorem 1.4.1, it induces an isomorphism $\bar{g}:A/J\to(A/I)/(J/I)$, which maps $a+J$ to $g(a)=([a]+J)/I$. Thus $f=\bar{g}$ is an isomorphism, as claimed. □

1.5 The Chinese remainder theorem

Definition 1.5.1. Let $\{A_i\}_{i\in I}$ be a family of rings. The **product of** $\{A_i\}_{i\in I}$ is the Cartesian product

$$\prod_{i\in I}A_i = \big\{(a_i)_{i\in I}\,\big|\,a_i\in A_i\big\},$$

together with componentwise addition $(a_i)+(b_i)=(a_i+b_i)$ and componentwise multiplication $(a_i)\cdot(b_i)=(a_ib_i)$.

Proposition 1.5.2 (Universal property for product rings)**.** *Let $\{A_i\}_{i\in I}$ be a family of rings. Then the following holds.*

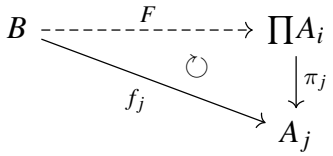
(1) The product $\prod_{i\in I}A_i$ is a ring whose zero is the element (a_i) with $a_i=0$ for all $i\in I$ and whose one is the element (a_i) with $a_i=1$ for all $i\in I$.

(2) The canonical projections

$$\begin{array}{rccc}\pi_j: & \prod_{i\in I}A_i & \longrightarrow & A_j\\ & (a_i)_{i\in I} & \longmapsto & a_j\end{array}$$

are ring homomorphisms.

(3) The product $\prod A_i$ together with the canonical projections $\{\pi_j\}_{j\in I}$ satisfies the following universal property: for every ring B and for every family of ring homomorphisms $\{f_j: B\to A_j\}_{j\in I}$, there exists a unique ring homomorphism $F: B\to\prod A_i$ such that $f_j=\pi_j\circ F$ for all $j\in I$, i.e. the diagram



commutes for every $j\in I$.

Proof. The verification of (1) and (2) is straight-forward, and we leave this as an exercise to the reader.

We turn to (3). Given ring homomorphisms $f_j: B\to A_j$, we define $F: B\to\prod A_i$ by $F(a)=(f_i(a))_{i\in I}$. Assuming that F is a ring homomorphism, we see that $f_j(a)=\pi_j\big((f_i(a))_{i\in I}\big)=\pi_j\circ F(a)$, as desired. The requirement $f_j=\pi_j\circ F$ also shows that if $(b_i)_{i\in I}=F(a)$, then $b_j=\pi_j\big(F(a)\big)=f_j(a)$, which shows that F is unique.

We continue with showing that $F: B\to\prod A_i$ is a ring homomorphism. By (1), $F(1)=(f_i(1))_{i\in I}=(1)_{i\in I}$ is the one of $\prod A_i$. Given $a,b\in B$, we have

$$\begin{aligned}F(a+b) &= \big(f_i(a+b)\big)_{i\in I} = \big(f_i(a)\big)_{i\in I}+\big(f_i(b)\big)_{i\in I} = F(a)+F(b),\\ F(ab) &= \big(f_i(ab)\big)_{i\in I} = \big(f_i(a)\big)_{i\in I}\cdot\big(f_i(b)\big)_{i\in I} = F(a)\cdot F(b),\end{aligned}$$

which shows that F is a ring homomorphism and concludes the proof. □

Definition 1.5.3. Let A be a ring and $I_1,\dots,I_n$ ideals of A. The **product of** $I_1,\dots,I_n$ is the ideal

$$\prod_{i=1}^{n}I_i = I_1\cdots I_n = \big\langle\{a_1\cdots a_n\,|\,a_i\in I_i\}\big\rangle$$

of A. The **sum of** $I_1,\dots,I_n$ is the ideal

$$\sum_{i=1}^{n} I_i \;=\; I_1+\cdots+I_n \;=\; \{a_1+\cdots+a_n \,|\, a_i\in I_i\}$$

of A. The ideals $I_1,\dots,I_n$ are **pairwise coprime** if I_i and I_j are *coprime*, i.e. $I_i+I_j=A$, for all $i\neq j$.

Definition 1.5.4. It is left as Exercise 1.13 to verify that $\sum I_i$ is indeed an ideal of A.

Theorem 1.5.5 (Chinese remainder theorem)**.** *Let A be a ring, $I_1,\dots,I_n$ pairwise coprime ideals of A and $I=\bigcap_{i=1}^n I_i$. Then the association*

$$\begin{array}{rccc} f: & A/I & \longrightarrow & \prod_{i=1}^n (A/I_i) \\ & a+I & \longmapsto & (a+I_1,\dots,a+I_n) \end{array}$$

is a well-defined isomorphism of rings.

Proof. We begin with the proof that f is a well-defined and injective ring homomorphism. By Proposition 1.5.2, the quotient maps $f_i: A\to A/I_i$, mapping a to $a+I_i$, induce a ring homomorphism $F: A\to \prod A/I_i$ such that $f_j=\pi_j\circ F$ for all $i=1,\dots,n$ where $\pi_j:\prod A/I_i\to A/I_j$ is the j-th canonical projection. Thus the kernel of F is

$$\begin{aligned} \ker F &= \{a\in A\,|\,F(a)=0\} \\ &= \{a\in A\,|\,f_j(a)=\pi_j\circ F(a)=0 \text{ for all } j=1,\dots,n\} \\ &= \bigcap_{i=1}^{n} \ker f_j \;=\; \bigcap_{i=1}^{n} I_i \;=\; I. \end{aligned}$$

By the universal property of the quotient (Proposition 1.3.12), F factors into the quotient map $\pi: A\to A/I$ followed by a uniquely determined ring homomorphism $f: A/I\to \prod A/I_i$. This is indeed the map of the theorem since we have $f(a+I)=F(a)=(a+I_1,\dots,a+I_n)$. Since $\ker f=\pi(I)=\{\bar 0\}$, the ring homomorphism $f: A/I\to\prod A/I_i$ is injective.

We turn to the surjectivity of f. Since the ideals $I_1,\dots,I_n$ are pairwise coprime, we find for every $j\neq i$ elements $a_{i,j}\in I_i$ and $a_{j,i}\in I_j$ such that $a_{i,j}+a_{j,i}=1$. We define $a_i=\prod_{j\neq i} a_{j,i}$ where j varies through $\{1,\dots,n\}-\{i\}$. By definition, $a_i\in\prod_{j\neq i} I_j$, and by Exercise 1.13, we know that $\prod_{j\neq i} I_j\subset I_{j'}$ for all $j'\neq i$. Thus $a_i+I_j=0+I_j$ for $j\neq i$. On the other hand, we have $\prod_{j\neq i}(a_{i,j}+a_{j,i})=1$ and thus

$$a_i \;=\; 1-\Big(\underbrace{\prod_{j\neq i}(a_{i,j}+a_{j,i})-\prod_{j\neq i} a_{j,i}}_{\in I_i}\Big),$$

which shows that $a_i\in 1+I_i$. Thus the inverse image of an element $(b_i+I_i)\in\prod A/I_i$ under f is $\sum_{j=1}^n a_jb_j+I\in A/I$, as can be computed directly:

$$f\Big(\sum_{j=1}^{n} a_jb_j+I\Big) \;=\; \Big(\sum_{j=1}^{n} a_jb_j+I_i\Big)_{i=1,\dots,n} \;=\; (b_i+I_i)_{i=1,\dots,n},$$

using that $a_j + I_i = \delta_{i,j} + I_i$ where the Kronecker symbol $\delta_{i,j}$ is 1 for $i = j$ and 0 for $i \neq j$. This concludes the proof of surjectivity. □

Corollary 1.5.6. *Let $e_1, \ldots, e_n \geqslant 1$ be pairwise coprime integers, i.e. $\langle e_i, e_i \rangle = \mathbb{Z}$ for all $i \neq j$. Then there is an $a \in \mathbb{Z}$ such that $a \equiv a_i \pmod{e_i}$, i.e. $a - a_i \in \langle e_i \rangle$, for all $i = 1, \ldots, n$.*

Proof. By the Chinese remainder theorem (Theorem 1.5.5), the ring homomorphism

$$\begin{array}{rccc} f: & \mathbb{Z} & \longrightarrow & \prod_{i=1}^{n} (\mathbb{Z}/e_i\mathbb{Z}) \\ & a & \longmapsto & a + e_i\mathbb{Z} \end{array}$$

is surjective. Thus there exists an $a \in \mathbb{Z}$ such that $a + e_i\mathbb{Z} = a_i + e_i\mathbb{Z}$, i.e. $a \equiv a_i \pmod{e_i}$, for all $i = 1, \ldots, n$. □

1.6 Euclidean domains and principal ideal domains

Definition 1.6.1. Let A be a ring and $a, b \in A$. A **divisor** of a is an element $d \in A$ such that $a = cd$ for some $c \in A$. We write $d|a$ if d is a divisor of a. A **common divisor of a and b** is an element $d \in A$ such that $d|a$ and $d|b$. A **greatest common divisor of a and b** is a common divisor d of a and b such that every other common divisor d' of a and b is a divisor of d. We define $\gcd(a, b)$ as the ideal of A that generated by all greatest common divisors of a and b.

Remark. Note that every element $d \in A$ is a divisor of 0. In contrast, a **zero divisor** is an element $d \in A$ such that $cd = 0$ for some $c \in A$ with $c \neq 0$. Thus 0 is always a zero divisor, which we call the *trivial zero divisor*. But it might be that there are no other zero divisors in A, which is the case if and only if A is a ring is without zero divisors, according to Definition 1.2.4.

Remark. In general, the greatest common divisor of two elements a and b does not have to exist. For example, the elements 6 and $2 + 2\sqrt{-5}$ of $\mathbb{Z}[\sqrt{-5}]$ do not have a greatest common divisor. For more details, see Exercise 1.26.

Lemma 1.6.2. *Let A be a ring and $a, b, c, d \in A$. Then the following hold true.*

(1) If $d|a$ and $d|b$, then $d|a + b$ and $d|ca$. If $d|a$ and $a|b$, then $d|b$.

(2) The following are equivalent:

   (a) $d|a$;
   (b) $a \in \langle d \rangle$;
   (c) $\langle a \rangle \subset \langle d \rangle$.

(3) The ideal $\gcd(a, b)$ is principal and equal to $\langle d \rangle$ if d is a greatest common divisor of a and b.

Proof. We begin with (1). If $d|a$ and $d|b$, then $a = da'$ and $b = db'$ for some $a', b' \in A$. Thus $a+b = d(a'+b')$ and $ca = dca'$, i.e. $d|a+b$ and $d|ca$, as claimed.

We continue with (2). Assume (2a), i.e. $d|a$. Thus $a = da'$ for some $a' \in A$, which shows that $a \in \langle d \rangle$, as claimed in (2b). Assume (2b), i.e. $a \in \langle d \rangle$. Then $a = cd$ for some $c \in A$ and thus $ba = bcd \in \langle d \rangle$ for all $b \in A$. This shows that $\langle a \rangle = \{ba | b \in A\} \subset \langle d \rangle$, as claimed in (2c). Assume (2c), i.e. $\langle a \rangle \subset \langle d \rangle$. Then $a \in \langle d \rangle$, i.e. $a = cd$ for some $c \in A$. Thus $d|a$, as claimed in (2a). This concludes the proof of (2).

We continue with (3). If a and b have no greatest common divisor, then, by definition, $\gcd(a,b) = \langle \varnothing \rangle = \langle 0 \rangle$ is principal. If the set S of greatest common divisors of a and b is not empty and $d, d' \in S$, then $d|d'$ by the definition of a greatest common divisor of a and b. Thus $d' = dc$ for some $c \in A$. This shows that $S \subset \langle d \rangle \subset \langle S \rangle = \gcd(a,b)$, which implies that $\gcd(a,b) = \langle d \rangle$. This verifies all claims of (3) and concludes the proof of the lemma. □

Definition 1.6.3. A **Euclidean domain** is an integral domain A for which there exists a *Euclidean function*, which is a function $N : A \to \mathbb{N}$ such that for all $a, b \in A$ with $b \neq 0$, there are $q, r \in A$ such that $a = bq + r$ and $N(r) < N(b)$ or $r = 0$.

Example 1.6.4. We list some examples of Euclidean domains.

(1) Every field K is a Euclidean domain with respect to any function $N : K \to \mathbb{N}$ since for every $a, b \in K$ with $b \neq 0$, we have $a = qb + r$ for $q = ab^{-1}$ and $r = 0$.

(2) The ring of integers $\mathbb{Z}$ is an Euclidean domain with respect to the usual absolute value $N(a) = |a|$. Indeed, given $a, b \in \mathbb{Z}$ with $b \neq 0$, we define q as the closest integer to the rational number a/b and $r = a - qb$. Then we have $|a/b - q| < 1$ and therefore the desired inequality

$$|r| \;=\; |a - qb| \;=\; |a/b - q| \cdot |b| \;<\; 1 \cdot |b| \;=\; |b|.$$

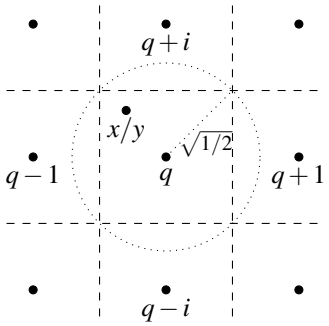
(3) The ring $\mathbb{Z}[i]$ of Gaussian integers is a Euclidean domain with respect to the function

$$\begin{array}{rccc} N: & \mathbb{Z}[i] & \longrightarrow & \mathbb{N}, \\ & a+ib & \longmapsto & a^2+b^2 \end{array}$$

as can be seen as follows. First note that if we consider $a + ib$ as an element of the complex plane $\mathbb{C} = \mathbb{R} \times \mathbb{R}$, then $N(a+ib) = a^2 + b^2$ is equal to the square $|a+ib|^2$ of the distance of $a + ib$ to 0.

Given $x = a + ib$ and $y = c + id \neq 0$, we can consider x/y as a point in the complex

plane. This point is of distance of at most $\sqrt{1/2}$ from a point $q = n + im \in \mathbb{Z}[i]$:



If we define $r = x - qy$, then

$$N(r) \;=\; |x-qy|^2 \;=\; |x/y-q|^2 \cdot |y|^2 \;\leqslant\; \tfrac{1}{2} \cdot |y|^2 \;<\; |y|^2 \;=\; N(y)$$

verifies that N is a Euclidean function.

(4) Let K be a field. Then the polynomial ring $K[T]$ is a Euclidean domain with respect to the degree map $\deg : K[T] \to \mathbb{N}$. Indeed, given two polynomials $f, g \in K[T]$ with $\deg g \geqslant 1$, we know by Proposition 1.3.12 that there is a polynomial $r \in K[T]$ of degree smaller than g such that $r + \langle g \rangle = f + \langle g \rangle$ as classes of $K[T]/\langle g \rangle$. In other words, $f - r \in \langle g \rangle$ or $f - r = qg$ for some polynomial $q \in K[T]$. Thus we have $f = qg + r$ with $\deg r < \deg g$, as desired. If $\deg g = 0$, but $g \neq 0$, then $g = a_0$ is invertible and thus $f = qg + 0$ for $q = a_0^{-1} f$.

Theorem 1.6.5 (Euclidean algorithm)**.** *Let A be a Euclidean domain with Euclidean function $N : A \to \mathbb{N}$ and $a, b \in A$ with $b \neq 0$. Then there is a sequence of elements $r_0, \ldots, r_n \in A$ and $q_2, \ldots, q_{n+1} \in A$, starting with $r_0 = a$ and $r_1 = b$, such that*

$$\begin{aligned}
r_0 &= q_2 r_1 + r_2 && \text{and} && N(r_2) < N(r_1),\\
r_1 &= q_3 r_2 + r_3 && \text{and} && N(r_3) < N(r_2),\\
&\vdots && && \vdots\\
r_{n-2} &= q_n r_{n-1} + r_n && \text{and} && N(r_n) < N(r_{n-1}),\\
r_{n-1} &= q_{n+1} r_n. &&
\end{aligned}$$

Given such a sequence, all the elements $r_1, \ldots, r_n$ are nonzero and $d = r_n$ is a greatest common divisor of a and b. Moreover $d = ca + eb$ for some $c, e \in A$, which means that $\gcd(a,b) = \langle d \rangle = \langle a, b \rangle$.

Proof. The defining property of an Euclidean domain lets us find the required elements recursively: for every pair of elements r_{i-2} and $r_{i-1} \neq 0$ of A, there are q_i and r_i in A such that $r_{i-2} = q_i r_{i-1} + r_i$ and $N(r_i) < N(r_{i-1})$ or $r_i = 0$. If $r_i = 0$, then we stop the recursion and find $n = i - 1$, which happens after at most $N(r_1) + 1$ steps. This establishes the first assertion of the theorem.

We show the latter assertions by induction on n. We begin with the claim that $r_1, \ldots, r_n$ are nonzero. If $n = 1$, then this follows from the assumption that $r_1 = b \neq 0$.

If $n = 2$, then $q_3 r_2 = r_1 \neq 0$ implies that $r_2 \neq 0$. For $n > 2$, let us assume that $r_2 = 0$. Then $r_3 = r_1 - q_3 r_2 = r_1$, which contradicts the assumption that $N(r_3) < N(r_2) < N(r_1)$. Thus we conclude that $r_2 \neq 0$. Therefore we can apply the inductive hypothesis to the sequences with $r'_i = r_{i+1}$ and $q'_i = q_{i+1}$, for which $r'_{n-1} = 0$, to conclude that $r_{i+1} = r'_i \neq 0$ for all $i = 1, \ldots, n-1$, which finishes the induction.

We continue with showing that d is a common divisor of a and b. If $n = 1$, then $d = b$ and $a = q_2 b$, thus d divides both a and b. If $n > 1$, then can apply the inductive hypothesis to the sequences with $r'_i = r_{i+1}$ and $q'_i = q_{i+1}$ to conclude that d divides both $a' = r'_0 = r_1 = b$ and $b' = r'_1 = r_2$. Therefore d also divides $q_2 r_1 + r_2 = r_0 = a$, which finishes the induction.

We continue with showing that every common divisor d' of a and b divides d. If $n = 1$, then $d = b$, and we have $d' | b = d$ by assumption. If $n > 1$, then d' divides $b = r_1$ and $q_2 a + b = r_2$. Thus we can apply the inductive hypothesis to the sequences with $r'_i = r_{i+1}$ and $q'_i = q_{i+1}$ to conclude that d' divides d, which finishes the induction.

We continue with showing that $d = ca + eb$ for some $c, e \in A$. If $n = 1$, then $d = r_2 = r_0 - q_2 r_1 = 1 \cdot a + (-q_2) b$, as desired. If $n > 1$, then we can apply the inductive hypothesis to the sequences with $r'_i = r_{i+1}$ and $q'_i = q_{i+1}$ to conclude that

$$d \;=\; c' r'_0 + e' r'_1 \;=\; c' r_1 + e'(r_0 - q_2 r_1) \;=\; ca + eb$$

for $c = e'$ and $e = c' - q_2 e'$, which finishes the induction.

To conclude the proof, we observe that $\gcd(a,b) = \langle d \rangle$ by Lemma 1.6.2 (3), that $a, b \in \langle d \rangle$ since $d|a$ and $d|b$, and thus $\langle a, b \rangle \subset \langle d \rangle$, and that $\langle d \rangle \subset \langle a, b \rangle$ since $d = ca + eb$. □

Definition 1.6.6. A **principal ideal domain** (often just **PID**) is an integral domain A for which every ideal is a principal ideal.

Proposition 1.6.7. *Every Euclidean domain is a principal ideal domain.*

Proof. Let A be a Euclidean domain with Euclidean function $N : A \to \mathbb{N}$ and I an ideal of A. Since the trivial ideal $\{0\} = \langle 0 \rangle$ is always principal, we can assume that I is not trivial. Let $b \in I$ be a nonzero element with minimal value $N(b)$, i.e. $N(b) \leqslant N(b')$ for all $b' \in I - \{0\}$. Given $a \in I$, there are $q, r \in A$ such that $a = qb + r$ and $N(r) < N(b)$ or $r = 0$. Since $r = a - qb \in I$, the minimality of $N(b)$ implies that $r = 0$. Thus $a = qb$ is an element of $\langle b \rangle$, which shows that $I = \langle b \rangle$ is principal. □

Remark. Two examples of principal ideal domains that are not Euclidean domains are

$$\mathbb{R}[x,y]/\langle x^2 + y^2 + 1 \rangle \qquad \text{and} \qquad \mathbb{Z}[T]/\langle T^2 - T + 5 \rangle.$$

It requires in both cases some effort to prove this, which we will not do here. In general, there seem to be no easy examples of such rings.

Lemma 1.6.8. *Let A be a principal ideal domain and $a, b, d \in A$. Then d is a greatest common divisor of a and b if and only if $\langle a, b \rangle = \langle d \rangle$. In particular, every pair of elements of A has a greatest common divisor.*

Proof. The proof is left as Exercise 1.18. □

1.7 Unique factorization domains

One of the basic theorems in number theory is the unique factorization of a positive integer into prime factors. Let us consider this statement in some more detail. A **prime number** is a positive integer p different from 1 whose only positive integer divisors are 1 and p. In the following, we use the convention that the empty product is defined as 1.

Theorem 1.7.1 (Fundamental Theorem of Arithmetic)**.** *Every positive integer n has a factorization $n = p_1 \cdots p_n$ into uniquely determined prime numbers $p_1, \ldots, p_n$, up to a permutation of indices.*

This theorem was first stated and proven by Euclid around 300 BC in his influential work "The elements". The most difficult part of the proof is the uniqueness claim, which follows from what is called today Euclid's Lemma.

Theorem 1.7.2 (Euclid's Lemma)**.** *If a prime number p divides the product ab of two positive integers a and b, then p must divide at least one of a and b.*

The original proof of Euclid is elementary, but technically quite involved. This elementary method of proof was simplified later using on Bézout's Lemma. As a motivation for this section, the reader is encouraged to reflect about these results and to think about a proof.

In this section, we generalize the factorization of a positive integer into prime numbers from the (positive) integers to arbitrary rings. There are two types of generalizations of prime numbers, one that captures the defining property of a prime number, the other capturing the property exhibited by Euclid's Lemma.

The core of this section is a characterization of the validity unique factorization in an integral domain in several equivalent ways. From this, we deduce a proof of the Fundamental Theorem of Arithmetic and Euclid's Lemma.

Definition 1.7.3. Let A be a ring. An element $a \in A$ is

- **irreducible** if $a \neq 0$, if $a \notin A^\times$ and if $a = bc$ for some $b, c \in A$ implies that $b \in A^\times$ or $c \in A^\times$;
- **prime** if $a \neq 0$, if $a \notin A^\times$ and if $a|bc$ for some $b, c \in A$ implies that $a|b$ or $a|c$.

Example 1.7.4. We examine some examples of irreducible and prime elements.

(1) A prime number is the same thing as an positive integer that is irreducible in the sense of Definition 1.7.3. Indeed, a prime number p is apparently not zero and not a unit. Given an equality $p = cd$ with $d > 0$, we have $d|p$ and thus either $d = 1 \in \mathbb{Z}^\times$ or $d = p$, which implies that $c = 1 \in \mathbb{Z}^\times$. If $d < 0$, then we can use the same argument for $-d$ and $-c$ in place of d and c, respectively. Conversely, assume that $p \in \mathbb{Z}$ is a positive and irreducible integer. Then clearly $p > 1$. Given $d|p$ for some $d > 0$, i.e. $p = cd$ for some $c \in \mathbb{Z}$, we conclude that $d \in \mathbb{Z}^\times$ or $c \in \mathbb{Z}^\times$. Since $\mathbb{Z}^\times = \{\pm 1\}$, we have $d = 1$ in the former case and $d = p$ in the latter case.

By Euclid's Lemma, a prime number is also a prime element in the sense of Definition 1.7.3. A proof of this is more difficult, and we postpone it to the end of this section; cf. Corollary 1.7.16.

(2) Let A be ring and $a \in A$. Then $f = T - a$ is irreducible in $A[T]$. Indeed, assume that $f = gh$ for two polynomials $g, h \in A[T]$. Then $\deg g + \deg h = \deg f = 1$, which means that one of g and h is a linear polynomial and the other is a constant. By the symmetry of g and h, we assume that $g = bT - c$ and $h = d$. Then $f = gh = bdT - cd$, and thus $bd = 1$, which shows that $h = d$ is a unit. Thus f is irreducible, as claimed.

If $A[T]$ is a unique factorization domain, then $f = T - a$ is also prime, and this is the case if A itself is a unique factorization domain, which we will only prove in Theorem 1.11.9. At the end of this section, we will prove this fact in the case that $A = K$ is a field, cf. Proposition 1.6.7 and Corollary 1.7.15).

(3) The element $(1,0) \in \mathbb{R} \times \mathbb{R}$ is not irreducible since $(1,0) = (1,0) \cdot (1,0)$, but $(1,0)$ is not a unit. It is prime since if $(1,0)|(a,b) \cdot (c,d)$, then either $b = 0$ or $d = 0$ and thus $(1,0)|(a,b)$ or $(1,0)|(c,d)$.

(4) Let $\mathbb{Z}[\sqrt{-5}]$ be the subring of $\mathbb{C}$ that consists of all complex numbers of the form $a + b\sqrt{-5}$ with $a, b \in \mathbb{Z}$. The element $2 \in \mathbb{Z}[-\sqrt{-5}]$ is irreducible, but not prime. The proof is left as Exercise 1.26.

Lemma 1.7.5. *Let A be a ring. Then the following hold true.*

(1) An element $a \in A$ is prime if and only if $\langle a \rangle$ is a nontrivial prime ideal.

(2) If a prime element $a \in A$ divides a product $b_1 \cdots b_n$, then a divides b_i for some $i \in \{1, \ldots, n\}$.

(3) If A is an integral domain, then every prime element of A is irreducible.

Proof. We begin with the proof of (1). Assume that $a \in A$ is prime, i.e. $a|bc$ implies $a|b$ or $a|c$. By Lemma 1.6.2, this is means, equivalently, that $bc \in \langle a \rangle$ implies $b \in \langle a \rangle$ or $c \in \langle a \rangle$. Thus $S = A - \langle a \rangle$ is closed under multiplication. Since $a \notin A^\times$, $\langle a \rangle$ is a proper ideal and thus $1 \in S$, which shows that $\langle a \rangle$ is a prime ideal. Since $a \neq 0$, this prime ideal is not trivial.

Assume conversely that $\langle a \rangle$ is a nontrivial prime ideal. Then $a \neq 0$ since $\langle a \rangle$ is nontrivial and $a \notin A^\times$ since $\langle a \rangle$ is a proper ideal. Since $S = A - \langle a \rangle$ is closed under multiplication, we conclude that $a|bc$ implies $a|b$ or $a|c$, which shows that a is prime. This concludes the proof of (1).

We continue with (2), which can be proven by induction on the number of factors n. Note that necessarily $n \geqslant 1$ since a prime element a cannot divide 1. The claim is trivial for $n = 1$. For $n > 1$, we have that $a|b_1 \ldots b_n$ implies that $a|b_1 \ldots b_{n-1}$ or $a|b_n$. In the former case, $a|b_i$ for some $i \in \{1, \ldots, n-1\}$ by the inductive hypothesis. Thus the claim.

We continue with (3) and assume that A is an integral domain. Let a be a prime element. Then $a \neq 0$ and $a \notin A^\times$. Consider $a = bc$ for some $b, c \in A$. Then $a|bc$ and thus $a|b$ or $a|c$, by the definition of a prime element. By symmetry of the argument in b

and c, we can assume that $a|c$, i.e. $c = da$ for some $d \in A$. Then we have $a = bc = bda$, i.e. $m_a(1) = 1 \cdot a = bd \cdot a = m_a(bd)$. Since A is an integral domain, the multiplication by a is injective and thus $bd = 1$, which shows that $b \in A^\times$. This shows that a is irreducible, as claimed, and concludes the proof of the lemma. □

Definition 1.7.6. Let A be a ring. Two elements $a, b \in A$ are **associated** if there is an $u \in A^\times$ such that $a = ub$. We write $a \sim b$ if a and b are associated.

Lemma 1.7.7. *Let A be a ring. The relation $\sim$ is an equivalence relation on A. If $a \sim b$, then a is irreducible if and only if b is irreducible, and a is prime if and only if b is prime.*

Proof. We begin with the verification that $\sim$ is an equivalence relation. Since $a = 1 \cdot a$, we have $a \sim a$, which shows that $\sim$ is reflexive. If $a \sim b$, i.e. $a = ub$ for some $u \in A^\times$, then $b = u^{-1}a$ and thus $b \sim a$, which shows that $\sim$ is symmetric. If $a \sim b$ and $b \sim c$, i.e. $a = ub$ and $b = vc$ for some $u, v \in A^\times$, then $uv \in A^\times$ and $a = ub = (uv)c$. Thus $a \sim c$, which shows that $\sim$ is transitive. This shows that $\sim$ is an equivalence relation.

Let $a \sim b$, i.e. $a = ub$ for some $u \in A^\times$. Assume that a is irreducible. Since $a \neq 0$, we have $b \neq 0$, and since $a \notin A^\times$, we have $b \notin A^\times$. Consider an equality $b = cd$ with $c, d \in A$. Then $a = ub = (uc)d$. Since a is irreducible, either $(uc) \in A^\times$ or $d \in A^\times$. Note that if $uc \in A^\times$, then also $c = u^{-1}uc \in A^\times$. This concludes the proof that b is irreducible. The inverse implication follows by the symmetry of the argument in a and b.

Assume that a is prime. As before, this implies that $b \neq 0$ and $b \notin A^\times$. Consider a relation $b|cd$, i.e. $cd = eb$ for some $e \in A$. Then $ucd = ea$, i.e. $a|(uc)d$. Since a is prime, we have $a|uc$ or $a|d$, i.e. $uc = e'a$ or $d = e''a$ for some $e', e'' \in A$. If $uc = e'a$, then $c = e'b$ and thus $b|c$. If $d = e''a$, then $d = u^{-1}e''b$ and thus $b|d$. This shows that b is prime. The inverse implication follows by the symmetry of the argument in a and b. □

Remark. Note that $a \sim b$ implies that $\langle a \rangle = \langle b \rangle$. If A is an integral domain, then the converse is also true: if $\langle a \rangle = \langle b \rangle$, then $a \sim b$. The proof is left as Exercise 1.16.

Definition 1.7.8. Let A be a ring and $a \in A$. A **factorization of a (into irreducible elements)** is an equation $a = u\prod_{i=1}^n f_i$ where $u \in A^\times$ and $f_1, \ldots, f_n$ are irreducible. A **factorization into primes** is a factorization $a = u\prod_{i=1}^n f_i$ for which $f_1, \ldots, f_n$ are prime (and irreducible). A factorization $a = u\prod_{i=1}^n f_i$ is **unique (up to associates)** if for every other factorization $a = v\prod_{i=1}^m g_i$, there is a bijection $\sigma : \{1, \ldots, n\} \to \{1, \ldots, m\}$ such that $f_i \sim g_{\sigma(i)}$ for all $i = 1, \ldots, n$.

Lemma 1.7.9. *Let A be an integral domain and $a, b \in A - \{0\}$. A factorization of a into primes is unique. The product ab has a factorization into primes if and only if both a and b have factorizations into primes.*

Proof. We begin with the uniqueness of a factorization into primes. Consider a factorization $a = u\prod_{i=1}^n f_i$ into primes f_i and a factorization $a = v\prod_{i=1}^m g_i$ (into irreducible g_i). We prove by induction on n that $f_i \sim g_{\sigma(i)}$ for some bijection $\sigma : \{1, \ldots, n\} \to \{1, \ldots, m\}$. If $n = 0$, then $a = u = v\prod g_i$ is a unit and therefore $\prod g_i = uv^{-1} \in A^\times$, which is only possible if $m = 0$ and $a = v$. This establishes the case $n = 0$.

If $n > 0$, then f_n divides $u\prod f_i = a = v\prod g_i$. Since f_n is prime, Lemma 1.7.5 implies that f_n divides g_k for some k, i.e. $g_k = u_n f_n$ for some $u_n \in A$. Since g_k is irreducible and $f_n \notin A^\times$, we conclude that $u_n \in A^\times$. This shows that $f_n \sim g_k$. Since A is an integral domain, we can cancel the term f_n from the equation $u\prod_{i=1}^n f_i = vu_n f_n \prod_{i\neq k} g_i$, which yields $u\prod_{i=1}^{n-1} f_i = (vu_n)\prod_{i\neq k} g_i$. By the inductive hypothesis there is a bijection $\sigma' : \{1,\dots,n-1\} \to \{1,\dots,m\} - \{k\}$ such that $f_i \sim g_{\sigma(i)}$ for $i \in \{1,\dots,n-1\}$. If we extend σ' to $\sigma : \{1,\dots,n\} \to \{1,\dots,m\}$ with $\sigma(n) = k$, then the induction claim follows for n.

We turn to the second claim of the lemma. It is clear that if a and b have factorizations into primes, then the product of the factorizations is a factorization of ab into primes. For the converse implication, we consider a factorization $ab = u\prod_{i=1}^n f_i$ into primes. We prove the claim by induction on n. If $n = 0$, then $ab = u \in A^\times$, i.e. a and b are units and have the tautological factorizations $a = a$ and $b = b$ into (an empty product of) primes.

If $n > 0$, then the prime element f_n divides ab, and thus $f_n | a$ or $f_n | b$. By the symmetry of the argument in a and b, we can assume that $f_n | a$, i.e. $a = a' f_n$ for some $a' \in A$. Since A is an integral domain, we can cancel the term f_n in the equation $a' f_n b = u\prod_{i=1}^n f_i$ and obtain $a'b = u\prod_{i=1}^{n-1} f_i$. By the inductive hypothesis, we have factorizations $a' = u'\prod f_i'$ and $b = u''\prod f_i''$ into primes. Thus also $a = a' f_n = u' f_n \prod f_i'$ is a factorization into primes, which concludes the proof. □

Definition 1.7.10. A ring A satisfies the **ascending chain condition for principal ideals** (for short, **ACCP**) if every sequence

$$\langle a_1\rangle \subset \langle a_2\rangle \subset \langle a_3\rangle \subset \cdots$$

of inclusions of principal ideals **becomes stationary**, i.e. that there exists an $N \geqslant 1$ such that $\langle a_i\rangle = \langle a_{i+1}\rangle$ for all $i \geqslant N$.

Remark. By Lemma 1.6.2, the property ACCP can be reformulated as follows: for every sequence of elements $a_1, a_2, \dots \in A$ such that $a_{i+1} | a_i$ for all $i \geqslant 0$, there is an $N \geqslant 1$ such that $a_i | a_{i+1}$ for all $i \geqslant N$. If A is an integral domain, then $a_{i+1} | a_i$ and $a_i | a_{i+1}$ implies $a_i \sim a_{i+1}$; cf. Exercise 1.16.

Lemma 1.7.11. *Let A be an integral domain that satisfies ACCP. Then every nonzero element of A has a factorization.*

Proof. Define $S = \big\{\langle a\rangle \big| a \in A - \{0\} \text{ does not have a factorization}\big\}$. Note that if $a \sim b$, then a has a factorization if and only if b does. Thus if S is empty, then every nonzero element of A has a factorization, which is what we intend to prove.

Let us assume, by contradiction, that S is nonempty. Then S must contain a maximal element since otherwise we could chose for every ideal $\langle a_i\rangle \in S$ a larger ideal $\langle a_{i+1}\rangle$, which defines an infinite properly growing sequence $\langle a_1\rangle \subsetneq \langle a_2\rangle \subsetneq \cdots$. But such a sequence cannot exist since A satisfies ACCP. This shows that S contains a maximal element $\langle a\rangle$, i.e. if $\langle a\rangle \subsetneq \langle b\rangle$, then b has a factorization.

Since every irreducible element f has a factorization, namely $f = f$, the element a is not irreducible. Thus there exist nonunits $b, c \in A$ such that $a = bc$. In other words,

$\langle a\rangle \subsetneq \langle b\rangle$ and $\langle a\rangle \subsetneq \langle c\rangle$. Thus $\langle b\rangle$ and $\langle c\rangle$ are not in S, i.e. b and c have respective factorizations $b = u\prod f_i$ and $c = v\prod g_i$. But then $a = bc = \prod f_i \cdot \prod g_i$ is a factorization of a, which contradicts our assumption that $\langle a\rangle \in S$. This shows that S is empty and that every nonzero element of A has a factorization, as claimed. □

Without additional assumptions, the inverse implication does not hold, i.e. there are integral domains for which every nonzero element has a factorization, but which fail to satisfy ACCP. This gap is closed, however, for the following class of rings, which ties together several concepts that we have introduced in the last sections.

Definition 1.7.12. A **unique factorization domain** (for short, **UFD**) is an integral domain for which every nonzero element has a unique factorization.

Theorem 1.7.13. *Let A be an integral domain. Then the following are equivalent.*

(1) A is a unique factorization domain.

(2) Every nonzero element of A has a factorization into primes.

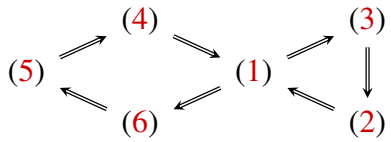
(3) Every nonzero prime ideal contains a prime element.

(4) A satisfies ACCP, and every irreducible element of A is prime.

(5) A satisfies ACCP, and every pair of elements of A has a greatest common divisor.

(6) A satisfies ACCP, and for all $a, b \in A$ and all factorizations $a = u\prod_{i=1}^{n} f_i$ and $b = v\prod_{i=1}^{m} g_i$, we have $a|b$ if and only if there is an injection $\sigma : \{1,\dots,n\} \to \{1,\dots,m\}$ such that $f_i \sim g_{\sigma(i)}$ for all $i = 1,\dots,n$.

Proof. We show the equivalence of the assertions of the theorem by first establishing the circle of implications (1)⇒(6)⇒(5)⇒(4)⇒(1) and then establishing the circle of implications (2)⇒(1)⇒(3)⇒(2):



We begin with (1)⇒(6). Assume that A is a unique factorization domain. In order to verify ACCP, we consider an increasing sequence $\langle a_1\rangle \subset \langle a_2\rangle \subset \cdots$ of principal ideals in A. We first observe that if all a_i are zero, then ACCP is satisfied. Otherwise there is a smallest i such that $a_i \neq 0$. Then $\langle a_1\rangle \subset \langle a_2\rangle \subset \cdots$ satisfies ACCP if and only if $\langle a_i\rangle \subset \langle a_{i+1}\rangle \subset \cdots$ satisfies ACCP. Thus we can assume without loss of generality that $a_1 \neq 0$.

Let $a_1 = u\prod_{i=1}^{n} f_i$ be a factorization. We prove ACCP by induction on n. If $n = 0$, then $\langle a_1\rangle = A$, and thus $\langle a_i\rangle = A$ for all $i \geqslant 1$, which establishes ACCP for $n = 0$.

If $n > 1$, then ACCP is clear if $\langle a_1\rangle = \langle a_i\rangle$ for all $i \geqslant 1$. If $\langle a_1\rangle \subsetneq \langle a_i\rangle$ for some $i \geqslant 1$, then we have $a_1 = ca_i$ for some $c \in A$ such that c is neither zero, since $a_1 \neq 0$, nor a unit, since $\langle a_1\rangle \neq \langle a_i\rangle$. Thus a factorization $c = v\prod_{i=1}^{m} g_i$ has $m > 0$ irreducible factors. By the uniqueness of factorizations, this means that every factorization of a_i

has $n-m<n$ factors. Thus the sequence $\langle a_i\rangle \subset \langle a_{i+1}\rangle \subset \cdots$ satisfies ACCP by the inductive hypothesis, which implies ACCP for $\langle a_1\rangle \subset \langle a_2\rangle \subset \cdots$, as desired.

In order to prove the characterization of divisibility in (6), we consider $a,b \in A$ with respective factorizations $a = u\prod_{i=1}^{n} f_i$ and $b = v\prod_{i=1}^{m} g_i$. If there is an injection $\sigma : \{1,\dots,n\} \to \{1,\dots,m\}$ such that $f_i \sim g_{\sigma(i)}$, i.e. $g_{\sigma(i)} = w_i f_i$ for some $w_i \in A^\times$, then

$$b \;=\; v\prod_{i=1}^{m} g_i \;=\; vu^{-1}u\prod_{i=1}^{n}(w_i f_i)\prod_{i\notin \mathrm{im}(\sigma)} g_i \;=\; \Big(vu^{-1}\prod_{i=1}^{n} w_i \prod_{i\notin \mathrm{im}(\sigma)} g_i\Big)\cdot a,$$

which shows that $a|b$.

Conversely, assume that $a|b$, i.e. $b = ca$ for some $c \in A$. Let $c = w\prod_{i=1}^{l} h_i$ be a factorization of c. Then we obtain a factorization $b = ca = (uw)\prod h_i \prod f_i$ of b. The uniqueness of the factorization in A implies that $f_i \sim g_{\sigma(i)}$ for some injection $\sigma : \{1,\dots,n\} \to \{1,\dots,m\}$. This completes the proof of (1)⇒(6).

We turn to (6)⇒(5). By assumption, A satisfies ACCP, and by Lemma 1.7.11, every element has a factorization. For establishing the existence of a greatest common divisor, we investigate the notion of divisibility, under the conditions of (6), in more detail. To begin with, we consider $n \geqslant 0$, an irreducible element $f \in A$ and an arbitrary nonzero element $b \in A$ with factorization $b = v\prod_{i=1}^{m} g_i$. By (6), $f^n|b$ if and only if there is an injection $\sigma : \{1,\dots,n\} \to \{1,\dots,m\}$ such that $f \sim g_{\sigma(i)}$ for all $i = 1,\dots,n$. This shows that the largest n such that f^n divides b is $n = \#\big\{i \in \{1,\dots,m\}\big|\, g_i \sim f\big\}$.

For an irreducible element $f \in A$ and $b \in A-\{0\}$, we define $\mathrm{ord}_f(b) = \max\big\{n \geqslant 0\,\big|\, f^n|b\big\}$ and conclude that f^n divides b if and only if $n \leqslant \mathrm{ord}_{\langle f\rangle}(b)$.

If $f \sim g$, then we have $f^n|b$ if and only if $g^n|b$. This leads to a characterization of divisibility of b by an arbitrary nonzero element a with factorization $a = u\prod_{i=1}^{n} f_i$. Namely, the injection $\sigma : \{1,\dots,n\} \to \{1,\dots,m\}$ in the characterization of $a|b$ in (6) restricts to injections

$$\sigma_f :\; \big\{i \in \{1,\dots,n\}\,\big|\, f_i \sim f\big\} \;\longrightarrow\; \big\{i \in \{1,\dots,m\}\,\big|\, g_i \sim f\big\}$$

for every irreducible element $f \in A$. This shows for nonzero a and b that $a|b$ if and only if $\mathrm{ord}_f(a) \leqslant \mathrm{ord}_f(b)$ for all irreducible $f \in A$.

We conclude that given two nonzero elements $a,b \in A$, an element $d \in A$ is a common divisor of a and b if and only if $\mathrm{ord}_f(d) \leqslant \mathrm{ord}_f(a)$ and $\mathrm{ord}_f(d) \leqslant \mathrm{ord}_f(b)$ for all irreducible $f \in A$, and d is a greatest common divisor if $\mathrm{ord}_f(d) = \min\big\{\mathrm{ord}_f(a), \mathrm{ord}_f(b)\big\}$ for all irreducible $f \in A$. Such an element d exists: we define S as the set of all irreducible $f \in A$ such that $\mu_f = \min\big\{\mathrm{ord}_f(a), \mathrm{ord}_f(b)\big\} > 0$ and let $S' \subset S$ be a complete set of representatives for all association classes $[f] = \{g \in |f \sim g\}$ of elements $f \in S$. Then $d = \prod_{f\in S'} f^{\mu_f}$ is a greatest common divisor of a and b, which concludes the proof (6)⇒(5).

We turn to (5)⇒(4). By assumption, A satisfies ACCP, and by Lemma 1.7.11, every element has a factorization. Let $a \in A$ be an irreducible element that divides a product bc of two elements $b,c \in A$, but not b. In order to prove that a is prime, we need to show that $a|c$.

Let d be a common divisor of a and b, i.e. $a = a'd$ and $b = b'd$. If $d \notin A^\times$, then $a' \in A^\times$ since a is irreducible, and thus $b = b'd = a(a')^{-1}b'$, which is not possible since $a \nmid b$. We conclude that $d \in A^\times$, and therefore it is in fact a greatest common divisor of a and b, i.e. $\gcd(a,b) = \langle d \rangle = \langle 1 \rangle$.

By contradiction, let us assume that $a \nmid c$. Then we also have $\gcd(a,c) = \langle 1 \rangle$, which has the following consequences. If d is a common divisor of a and bc, then also $d|ba$. By Exercise 1.18, we have $\gcd(ba,bc) = \langle b \rangle \gcd(a,c) = \langle b \rangle$, which implies that $d|b$. Since $\gcd(a,b) = \langle 1 \rangle$, we have $d \sim 1$, i.e. $\gcd(a,bc) = \langle 1 \rangle$. But our assumption $a|bc$ implies that $\gcd(a,bc) = \langle a \rangle \neq \langle 1 \rangle$, which is a contradiction. Thus we conclude that $a|c$, which shows that a is prime and concludes the proof of (5)⇒(4).

We turn to (4)⇒(1). By Lemma 1.7.11, ACCP implies that every element has a factorization. Since all irreducible elements are prime, a factorization $a = u\prod f_i$ is in fact a factorization into primes. By Lemma 1.7.9, it is unique, which shows that A is a unique factorization domain and completes (4)⇒(1). Similarly, the implication (2)⇒(1) follows at once from Lemma 1.7.9.

We turn to (1)⇒(3). Let I be a nonzero prime ideal of A and $a \in I - \{0\}$, with factorization $a = u\prod f_i$. Since I is a prime ideal, it must contain one of the factors f_i. We have already proven the implication (1)⇒(4), which shows that f_i is prime, which exhibits the desired prime element in I. Thus (1)⇒(3).

We turn to (3)⇒(2). Let T be the set of all nonzero elements of A that do not have a factorization into primes, together with 0. We need to show that $T = \{0\}$.

Consider $a \in T$. By Lemma 1.7.9, we have $ab \in T$ for all $b \in A$, which shows that $\langle a \rangle \subset T$. Let S be the set of of ideals I of A such that $\langle a \rangle \subset I \subset T$, which is partially ordered by inclusion. Then every chain C in S has an upper bound in S, namely $J = \bigcup_{I \in C} I$. Indeed, it is clear that $\langle a \rangle \subset J \subset T$ and it is easily proven that J is an ideal since C is totally ordered by inclusion. Thus Zorn's Lemma (Theorem A.1.2) applies, which shows that S has a maximal element $I_{\max}$.

We claim that $I_{\max}$ is a prime ideal. By contradiction assume that there are $a,b \in A - I_{\max}$ with $ab \in I_{\max}$. Then S contains neither $\langle I_{\max} \cup \{a\} \rangle$ nor $\langle I_{\max} \cup \{b\} \rangle$, which means that there are elements $c_1, c_2 \in I_{\max}$ and $d_1, d_2 \in A$ such that both $s_1 = c_1 + d_1 a$ and $s_2 = c_2 + d_2 b$ are in $A - T$, i.e. both s_1 and s_2 have factorizations into primes, and so does the product $s_1 s_2$. On the other hand,

$$s_1 s_2 = \underbrace{c_1 c_2}_{\in I_{\max}} + \underbrace{d_2 b c_1}_{\in I_{\max}} + \underbrace{d_1 a c_2}_{\in I_{\max}} + \underbrace{d_1 d_2 ab}_{\in I_{\max}}$$

is an element of $I_{\max}$, which contradicts our assumption that $I_{\max} \subset T$. Thus we conclude that $I_{\max}$ is a prime ideal.

Since $f = f$ is tautologically a factorization into primes if f is prime, $I_{\max}$ cannot contain any prime element. By (3), this is only possible if $I_{\max} = \{0\}$, which shows that $a = 0$ and $S = \{0\}$. This completes the step (3)⇒(2) and concludes the proof of the theorem. □

As a first consequence, we highlight the following auxiliary result that we have established in the proof of Theorem 1.7.13. Let $f \in A$ be irreducible. We recall the

definition

$$\operatorname{ord}_f(a) \;=\; \max\{n \geqslant 0 \,|\, f^n | a\}$$

for $a \in A$, and that $\operatorname{ord}_f(a) = \operatorname{ord}_g(a)$ if $f \sim g$. By Exercise 1.25, we have $f \sim g$ if $\langle f \rangle = \langle g \rangle$, which reasons that $\operatorname{ord}_f(a)$ depends only on $\langle f \rangle$. This yields for every principal prime ideal $\mathfrak{p} = \langle f \rangle$ of A the function $\operatorname{ord}_\mathfrak{p} : A - \{0\} \to \mathbb{N}$ that sends $a \in A - \{0\}$ to the value $\operatorname{ord}_\mathfrak{p}(a) = \operatorname{ord}_f(a)$, which is called the *order of a in $\mathfrak{p}$*. We define $\mathcal{P}(A)$ as the set of principal prime ideals of A and summarize the insights from the proof of Theorem 1.7.13 as follows.

Corollary 1.7.14. *Let A be a unique factorization domain and $a, b, d \in A$. Then the following hold:*

(1) $a|b$ if and only if $\operatorname{ord}_\mathfrak{p}(a) \leqslant \operatorname{ord}_\mathfrak{p}(b)$ for all $\mathfrak{p} \in \mathcal{P}(A)$;

(2) d is a greatest common divisor of a and b if and only if

$$\operatorname{ord}_\mathfrak{p}(d) = \min\{\operatorname{ord}_\mathfrak{p}(a), \operatorname{ord}_\mathfrak{p}(b)\}$$

for all $\mathfrak{p} \in \mathcal{P}(A)$. □

Corollary 1.7.15. *Every principal ideal domain is a unique factorization domain.*

Proof. Let A be a principal ideal domain. Then every nonzero prime ideal I of A is generated by a single element a, which is prime by Lemma 1.7.5. Thus by Theorem 1.7.13, A is a unique factorization domain. □

As we will see at a later point of this lecture, there are unique factorization domains that are not principal ideal domains, as, for instance, $\mathbb{Z}[T]$ or polynomial rings in several variables over a field. It takes some work, however, to prove that factorizations in these rings are unique, which is why we postpone these examples to a later point. As an exercise, we encourage the reader to show the easier parts of our claim: $\mathbb{Z}[T]$ is an integral domain in which every element has a factorization (into irreducible elements), but that is not a principal ideal domain.

To conclude this section, we use our results about unique factorization domains to deduce a proof for the Fundamental Theorem of Arithmetic (Theorem 1.7.1) and Euclid's Lemma (Theorem 1.7.2).

Corollary 1.7.16. *Every positive integer n can be written as a product $n = p_1 \cdots p_n$ for uniquely determined prime numbers $p_1, \ldots, p_n$, up to a permutation of indices, and every prime number is a prime element of $\mathbb{Z}$.*

Proof. As a Euclidean domain, $\mathbb{Z}$ is a principal ideal domain by Proposition 1.6.7 and thus a unique factorization domain by Corollary 1.7.15. Therefore every positive integer $a \in \mathbb{Z}$ admits a unique factorization $n = u \prod_{i=1}^n p_i$ into irreducible elements $p_i \in \mathbb{Z}$. Note that $\mathbb{Z}^\times = \{\pm 1\}$, and thus $[p] = \{\pm p\}$ for every irreducible element p. As explained in Example 1.7.4, prime numbers are the same thing as positive irreducible elements in $\mathbb{Z}$, and thus $[p]$ has a unique representative that is a prime number, namely the positive integer among p and $-p$. After multiplying all negative p_i with -1, we can assume that

all p_i are positive and thus prime numbers. Consequently also the unit $u \in \mathbb{Z}^\times = \{\pm 1\}$ is positive, i.e. $u = 1$. This shows that $n = p_1 \cdots p_n$ is a decomposition of n into prime numbers that is unique up to a permutation of the indices, which establishes our first claim, which is the Fundamental Theorem of Arithmetic.

By Theorem 1.7.13, every irreducible element of $\mathbb{Z}$ is a prime element. With this, Euclid's Lemma follows from the very definition of a prime element. □

1.8 Localizations

In this section, we generalize the construction of the rational numbers from the integers to arbitrary rings, which is called a *localization*. Before we explain this in general, let us review the construction of the rational numbers.

The rational numbers $\mathbb{Q}$ are defined as all fractions $\frac{a}{s}$ of integers $a, s \in \mathbb{Z}$ where $s \neq 0$, subject to the following rules:

$$\frac{a}{s} = \frac{ta}{ts}, \qquad \frac{a}{s} + \frac{b}{t} = \frac{ra+sb}{st} \qquad \frac{a}{s} \cdot \frac{b}{t} = \frac{ab}{st}$$

for all $a, b \in \mathbb{Z}$ and $s, t \in \mathbb{Z} - \{0\}$. The first rule identifies certain fractions, the latter two rules define addition and multiplication, respectively. We observe that the association $a \mapsto \frac{a}{1}$ defines a ring homomorphism $\iota : \mathbb{Z} \to \mathbb{Q}$, and that the image $\frac{a}{1}$ of every nonzero element $a \in \mathbb{Z}$ in $\mathbb{Q}$ has a multiplicative inverse, which is $\frac{1}{a}$.

In the following definition, we generalize this to arbitrary rings. Recall from Definition 1.3.13 that a multiplicative subset of a ring A is a subset S that contains 1 and ab for all $a, b \in S$.

Definition 1.8.1. Let A be a ring and S a multiplicative subset. The **localization of A at S** is the set

$$S^{-1}A = (S \times A)/\sim = \left\{ \tfrac{a}{s} \,\middle|\, a \in A, s \in S \right\}$$

where $\sim$ is the equivalence relation on $A \times S$ that is defines by $(s,a) \sim (s',a')$ if and only if there is an $t \in S$ such that $tsa' = ts'a$ and where $\frac{a}{s} = [(s,a)]$ denotes the equivalence class of (s,a), together with the addition

$$\begin{array}{rccc} \hat{+}: & S^{-1}A \times S^{-1}A & \longrightarrow & S^{-1}A \\ & (\frac{a}{s}, \frac{b}{t}) & \longmapsto & \frac{ta+sb}{st} \end{array}$$

and the multiplication

$$\begin{array}{rccc} \hat{\cdot}: & S^{-1}A \times S^{-1}A & \longrightarrow & S^{-1}A. \\ & (\frac{a}{s}, \frac{b}{t}) & \longmapsto & \frac{ab}{st} \end{array}$$

Lemma 1.8.2. *Let A be a ring and S a multiplicative subset. Then the localization $S^{-1}A$ is a well-defined ring, and the association $a \mapsto \frac{a}{1}$ defines a ring homomorphism $\iota_S : A \to S^{-1}A$ with $\iota_S(S) \subset (S^{-1}A)^\times$.*

Proof. We begin with the verification that $\sim$ is indeed an equivalence relation on $A \times S$. Let $a, a', a'' \in A$ and $s, s', s'' \in S$. Since $1sa = 1sa$, we have $(s,a) \sim (s,a)$, which shows that $\sim$ is reflective. Assume that $(s,a) \sim (s',a')$, i.e. $tsa' = ts'a$ for some $t \in S$. Then $ts'a = tsa'$ and $(s',a') \sim (s,a)$, which shows that $\sim$ is symmetric. Assume that $(s,a) \sim (s',a')$ and $(s',a') \sim (s'',a'')$, i.e. $tsa' = ts'a$ and $t's'a'' = t's''a'$ for some $t, t' \in S$. Then $tt's' \in S$ and $(tt's')s''a = tt's''sa' = (tt's')sa''$, which shows that $(s,a) \sim (s'',a'')$ and that $\sim$ is transitive. This shows that $\sim$ is indeed an equivalence relation.

We continue with the verification that $\widehat{+}$ and $\widehat{\cdot}$ are well-defined as maps. Consider $\frac{a}{s} = \frac{a'}{s'}$ and $\frac{b}{t} = \frac{b'}{t'}$, i.e. $usa' = us'a$ and $vtb' = vt'b$ for some $u, v \in S$. Then

$$uv(ta+sb)s't' = vtt'(us'a) + uss'(vt'b) = vtt'(usa') + uss'(vtb') = uv(t'a'+s'b')st,$$

which shows that $\frac{a}{s} \hat{+} \frac{b}{t} = \frac{a'}{s'} \hat{+} \frac{b'}{t'}$, and

$$uvs't'ab \;=\; uvsta'b',$$

which shows that $\frac{a}{s} \hat{\cdot} \frac{b}{t} = \frac{a'}{s'} \hat{\cdot} \frac{b'}{t'}$. This verifies that both $\hat{+}$ and $\hat{\cdot}$ are well-defined maps.

We turn to the verification that $S^{-1}A$ is a ring with respect to the addition $\hat{+}$ and the multiplication $\hat{\cdot}$. The additive unit is $\frac{0}{1}$ since

$$\frac{a}{s} \hat{+} \frac{0}{1} \;=\; \frac{s \cdot 0 + 1 \cdot a}{s \cdot 1} \;=\; \frac{a}{s},$$

and the multiplicative unit is $\frac{1}{1}$ since

$$\frac{a}{s} \hat{\cdot} \frac{1}{1} \;=\; \frac{a \cdot 1}{s \cdot 1} \;=\; \frac{a}{s}$$

for all $a \in A$ and $s \in S$. The additive inverse of $\frac{a}{b}$ is $\frac{-a}{b}$ since

$$\frac{a}{s} \hat{+} \frac{-a}{s} \;=\; \frac{as - as}{ss} \;=\; \frac{0}{ss} \;=\; \frac{0}{1}$$

where we use in the last step that $1 \cdot 1 \cdot 0 = 1 \cdot ss \cdot 0$. We leave the associativity and commutativity of both operations as an exercise, as well as the distributivity.

We continue with the verification that $\iota_S : A \to S^{-1}A$ is a ring homomorphism with $\iota_S(S) \subset (S^{-1}A)^\times$. We have already shown that $\iota_S(1) = \frac{1}{1}$ is the multiplicative unit of $S^{-1}A$. For $a, b \in A$, we have

$$\iota_S(a+b) \;=\; \frac{a+b}{1} \;=\; \frac{1 \cdot a + 1 \cdot b}{1 \cdot 1} \;=\; \frac{a}{1} \hat{+} \frac{b}{1} \;=\; \iota_S(a) \hat{+} \iota_S(b)$$

and

$$\iota_S(ab) \;=\; \frac{ab}{1} \;=\; \frac{a \cdot b}{1 \cdot 1} \;=\; \frac{a}{1} \hat{\cdot} \frac{b}{1} \;=\; \iota_S(a) \hat{\cdot} \iota_S(b),$$

which verifies that ι_S is a ring homomorphism. For $s \in S$, we have

$$\iota_S(s) \hat{\cdot} \frac{1}{s} \;=\; \frac{s}{1} \hat{\cdot} \frac{1}{s} \;=\; \frac{1 \cdot s}{s \cdot 1} \;=\; \frac{1}{1}$$

where we use in the last equality that $1 \cdot 1 \cdot (1 \cdot s) = 1 \cdot (s \cdot 1) \cdot 1$. Thus $\iota_S(S) \subset (S^{-1}A)^\times$, which concludes the proof of the lemma. $\square$

Remark. In the following, we simply write sums and products in $S^{-1}A$ simply as $\frac{a}{b}+\frac{b}{t}$ and $\frac{a}{s}\cdot\frac{b}{t}$, respectively.

Example 1.8.3. We describe some examples of localizations.

(1) To begin with consider $A=\mathbb{Z}$ and the multiplicative subset $S=\mathbb{Z}-\{0\}$. Then $S^{-1}A$ recovers the construction of the rational numbers $\mathbb{Q}$.

(2) The integers have also other multiplicative subsets. For $h\in\mathbb{Z}$, we can consider $S=\{1,h,h^2,\dots,\}$ and find

$$S^{-1}\mathbb{Z} \;=\; \left\{\tfrac{a}{h^i}\in\mathbb{Q}\,\middle|\,a\in\mathbb{Z},i\geqslant 0\right\}.$$

For a prime number $p\in\mathbb{Z}$, we can consider $S=\mathbb{Z}-\langle p\rangle$, which is multiplicatively closed since $\langle p\rangle$ is a prime ideal, and find

$$S^{-1}\mathbb{Z} \;=\; \left\{\tfrac{a}{s}\in\mathbb{Q}\,\middle|\,a,s\in\mathbb{Z}\text{ such that }p\nmid s\right\}.$$

(3) The Gaussian numbers $\mathbb{Q}[i]$ are the localization of the Gaussian integers $\mathbb{Z}[i]$ at $S=\mathbb{Z}[i]-\{0\}$.

Lemma 1.8.4. *Let A be a ring and S a multiplicative subset. Then the following hold true.*

(1) For all $a,b\in A$ and $s,t\in S$, we have

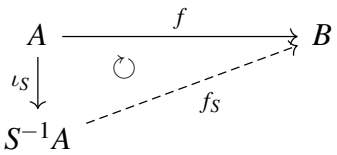
$$\frac{a}{s}\;=\;\frac{ta}{ts},\qquad \frac{a}{s}\cdot\frac{s}{a}\;=\;1,\qquad \frac{a}{s}+\frac{b}{s}\;=\;\frac{a+b}{s}.$$

(2) If $0\in S$, then $S^{-1}A=\{0\}$.

(3) If A is an integral domain and $0\notin S$, then $\iota_S:A\to S^{-1}A$ is an injection.

Proof. The proof is left as Exercise 1.32. □

Proposition 1.8.5. *Let A be a ring and S a multiplicative subset. Then the localization $S^{-1}A$ together with $\iota_S:A\to S^{-1}A$ satisfies the following universal property: for every ring B and every ring homomorphism $f:A\to B$ with $f(A)\subset B^\times$, there is a unique ring homomorphism $f_S:S^{-1}A\to B$ such that $f=f_S\circ\iota_S$, i.e. the diagram*



commutes.

Proof. We begin with the claim of uniqueness. Assume that $f_S : S^{-1}A \to B$ is a ring homomorphism such that $f = f_S \circ \iota_S$. Since $(\frac{1}{s})^{-1} = \frac{s}{1}$ in $S^{-1}A$, we have $f_S(\frac{1}{s}) = f_S(\frac{s}{1})^{-1}$ and thus

$$f_S(\tfrac{1}{s}) \;=\; f_S(\tfrac{1}{s}\cdot\tfrac{a}{1}) \;=\; f_S(\tfrac{s}{1})^{-1} f_S(\tfrac{a}{1}) \;=\; f(s)^{-1} f(a)$$

for all $\frac{a}{s} \in S^{-1}A$, which shows that f_S is uniquely determined by f.

We claim that the association $\frac{a}{s} \mapsto f(s)^{-1}f(a)$ describes a well-defined ring homomorphism $f_S : S^{-1}A \to B$. Once we have proven this, it follows that $f_S \circ \iota_S(a) = f_S(\frac{a}{1}) = f(1)^{-1}f(a) = f(a)$ for all $a \in A$, as desired.

We begin with the verification that f_S is well-defined as a map. Consider $\frac{a}{s} = \frac{a'}{s'}$, i.e. $tsa' = ts'a$ for some $t \in S$. Then $f(t)f(s)f(a') = f(t)f(s')f(a)$ in B, and after multiplying this equality with $f(t)^{-1}f(s)^{-1}f(s')^{-1}$, which is possible since $f(S) \subset B^\times$, we obtain

$$f_S(\tfrac{a}{s}) \;=\; f(s)^{-1}f(a) \;=\; f(s')^{-1}f(a') \;=\; f_S(\tfrac{a'}{s'}).$$

This shows that the definition of f_S does not depend on the choice of representative (s,a) for a class $\frac{a}{s} = [(s,a)] \in S^{-1}A = S \times A/\sim$.

We continue with the verification that f_S is a ring homomorphism. Clearly $f_S(\frac{1}{1}) = f(1)^{-1}f(1) = 1$. Given $\frac{a}{s}, \frac{b}{t} \in B$, we have

$$\begin{aligned} f_S(\tfrac{a}{s}+\tfrac{b}{t}) \;&=\; f_S(\tfrac{ta+sb}{st}) \;=\; f(s)^{-1}f(t)^{-1}\big(f(t)f(a)+f(s)f(b)\big) \\ &=\; f(s)^{-1}f(a)+f(t)^{-1}f(b) \;=\; f_S(\tfrac{a}{s})+f_S(\tfrac{b}{t}), \\ f_S(\tfrac{a}{s}\cdot\tfrac{b}{t}) \;&=\; f_S(\tfrac{ab}{st}) \;=\; f(s)^{-1}f(t)^{-1}f(a)f(b) \;=\; f_S(\tfrac{a}{s})\cdot f_S(\tfrac{b}{t}), \end{aligned}$$

which shows that f_S is a ring homomorphism and concludes the proof. □

There are certain types of localizations that are of particular interest, and generalize constructions from Example 1.8.3.

Definition 1.8.6. Let A be a ring. For $h \in A$, we define the **localization of** A **at** h as $A[h^{-1}] = S^{-1}A$ where $S = \{h^i\}_{i\in\mathbb{N}}$. For a prime ideal $\mathfrak{p}$ of A, we define **localization of** A **at** $\mathfrak{p}$ as $A_\mathfrak{p} = S^{-1}A$ where $S = A - \mathfrak{p}$. If A is an integral domain, and thus $\langle 0\rangle$ a prime ideal, we define the **field of fractions of** A as $\mathrm{Frac}A = A_{\langle 0\rangle}$.

Remark. In the context of Definition 1.8.6, we remark that h is multiplicatively invertible in $A[h^{-1}]$ with inverse $h^{-1} = \frac{1}{h}$. The field of fractions $\mathrm{Frac}A$ is indeed a field since if $\frac{a}{s}$ is nonzero in $\mathrm{Frac}A$, then a is nonzero in the integral domain A. Thus $\frac{s}{a}$ is an element of $\mathrm{Frac}A$ and a multiplicative inverse of $\frac{a}{s}$.

In the final part of this section, we study the structure of localizations at prime ideals.

Definition 1.8.7. A **local ring** is a ring A with a unique maximal ideal.

Lemma 1.8.8. *Let A be a local ring with maximal ideal $\mathfrak{m}$. Then $A^\times = A - \mathfrak{m}$.*

Proof. Since $\mathfrak{m}$ is a proper ideal, it does not contain a unit, and thus $A^\times \subset A-\mathfrak{m}$. Conversely, consider $a \in A-\mathfrak{m}$. Then $\langle a\rangle$ is not contained in $\mathfrak{m}$. Since every proper ideal is contained in the unique maximal ideal $\mathfrak{m}$ by Exercise 1.23, we conclude that $\langle a\rangle = \langle 1\rangle$, i.e. $ab = 1$ for some $b \in A$. This shows that $a \in A^\times$. Thus $A-\mathfrak{m} \subset A^\times$, which concludes the proof. □

Lemma 1.8.9. *Let A be a ring and $\mathfrak{p}$ a prime ideal of A. Then $A_\mathfrak{p}$ is a local ring with maximal ideal $\mathfrak{p}A_\mathfrak{p} = \left\{\frac{a}{s} \,\middle|\, a \in \mathfrak{p}, s \in A-\mathfrak{p}\right\}$.*

Proof. We begin with the verification that $\mathfrak{p}A_\mathfrak{p} = \left\{\frac{a}{s} \,\middle|\, a \in \mathfrak{p}, s \in A-\mathfrak{p}\right\}$ is an ideal. Clearly, it contains the zero $\frac{0}{1}$ of $A_\mathfrak{p}$. Given $\frac{a}{s}, \frac{b}{t} \in \mathfrak{p}A_\mathfrak{p}$ and $\frac{c}{r} \in A_\mathfrak{p}$, both

$$\frac{a}{s}+\frac{b}{t} = \frac{ta+sb}{st} \qquad \text{and} \qquad \frac{a}{s}\cdot\frac{c}{r} = \frac{ac}{sr}$$

are in $\mathfrak{p}A_\mathfrak{p}$ since $ta+sb, ac \in \mathfrak{p}$ and $st, sr \in A-\mathfrak{p}$. Thus $\mathfrak{p}A_\mathfrak{p}$ is an ideal.

If $\frac{a}{s} \in \mathfrak{p}A_\mathfrak{p}$, then $\frac{a}{s} = \frac{a'}{s'}$ for some $a' \in \mathfrak{p}$ and $s' \in A-\mathfrak{p}$. Thus $tsa' = ts'a$ for some $t \in A-\mathfrak{p}$. Since $a' \in \mathfrak{p}$, also $ts'a = tsa' \in \mathfrak{p}$. Since $t, s' \notin \mathfrak{p}$ and $\mathfrak{p}$ is a prime ideal, we conclude that $a \in \mathfrak{p}$. As a result, we conclude that $\mathfrak{p}A_\mathfrak{p}$ does not contain a unit and thus is a proper ideal.

On the other hand, if $\frac{a}{s} \in A_\mathfrak{p}-\mathfrak{p}A_\mathfrak{p}$, then $a \in A-\mathfrak{p}$, and thus $A_\mathfrak{p}$ contains the element $\frac{s}{a}$, which is a multiplicative inverse of $\frac{a}{s}$. This shows that $\mathfrak{p}A_\mathfrak{p} = A_\mathfrak{p}-A_\mathfrak{p}^\times$. We conclude that every ideal I of $A_\mathfrak{p}$ that is not contained in $\mathfrak{p}A_\mathfrak{p}$ contains a unit, i.e. $I = A_\mathfrak{p}$, which shows that $\mathfrak{p}A_\mathfrak{p}$ is the unique maximal ideal of $A_\mathfrak{p}$. □

1.9 Polynomial rings in several variables

In this section, we introduce polynomial rings in several variables. This requires some notation. Let I be a set. We define

$$\bigoplus_{i\in I}\mathbb{N} = \left\{(e_i)_{i\in I} \in \prod_{i\in I}\mathbb{N} \,\middle|\, e_i = 0 \text{ for all but finitely many } i \in I\right\},$$

which is a monoid with respect to the componentwise addition $(e_i)+(f_i) = (e_i+f_i)$. We often write $\underline{e} = (e_i)_{i\in I}$ for elements of $\bigoplus_{i\in I}\mathbb{N}$. An **indexed set** is a bijection $\tau : I \to S$ between sets I and S where I is called the **index set**. Equivalently, we can write S as $\{T_i\}_{i\in I}$ where $T_i = \tau(i)$. From a given indexed set $\{T_i\}_{i\in I}$ we derive another indexed set $\{\underline{T}^{\underline{e}}\}_{\underline{e}\in\bigoplus_{i\in I}\mathbb{N}}$.

Definition 1.9.1. Let A be a ring and $\{T_i\}_{i\in I}$ an indexed set. The **polynomial ring in $\{T_i\}_{i\in I}$ over A** is the set

$$A[T_i|i\in I] = \left\{\sum_{\underline{e}\in\bigoplus_{i\in I}\mathbb{N}} a_{\underline{e}}\underline{T}^{\underline{e}} \,\middle|\, a_{\underline{e}} \in A, a_{\underline{e}} = 0 \text{ for all but finitely many } \underline{e} \in \bigoplus_{i\in I}\mathbb{N}\right\},$$

together with the addition

$$\begin{array}{rccc} + : & A[T_i|i\in I]\times A[T_i|i\in I] & \longrightarrow & A[T_i|i\in I] \\ & (\sum a_{\underline{e}}\underline{T^e}, \sum b_{\underline{e}}\underline{T^e}) & \longmapsto & \sum (a_{\underline{e}}+b_{\underline{e}})\underline{T^e} \end{array}$$

and the multiplication

$$\begin{array}{rccc} \cdot : & A[T_i|i\in I]\times A[T_i|i\in I] & \longrightarrow & A[T_i|i\in I]. \\ & (\sum a_{\underline{e}}\underline{T^e}, \sum b_{\underline{e}}\underline{T^e}) & \longmapsto & \displaystyle\sum_{\underline{e}\in\bigoplus\mathbb{N}} \Big(\sum_{\underline{f}+\underline{g}=\underline{e}} a_{\underline{f}} b_{\underline{g}}\Big)\underline{T^e} \end{array}$$

Lemma 1.9.2. *Let A be a ring and $\{T_i\}_{i\in I}$ an indexed set. Then $A[T_i|i\in I]$ is a ring. For $a\in A$, let $\iota(a)=\sum a_{\underline{e}}\underline{T^e}$ be the element of $A[T_i|i\in I]$ with $a_{\underline{e}}=a$ if $\underline{e}=(e_i)_{i\in I}$ with $e_i=0$ for all $i\in I$ and $a_{\underline{e}}=0$ otherwise. This defines a ring homomorphism $\iota: A\to A[T_i|i\in I]$.*

Proof. We leave the proof as Exercise 1.33. □

Remark. We make a few observations.

(1) If the indexed set $\{T_i\}_{i\in I}$ has only one element $T_i=T$, then we recover the definition of the polynomial ring in one variable T from Definition 1.3.11. More precisely, the map

$$\begin{array}{ccc} A[T_i|i\in I] & \longrightarrow & A[T] \\ \sum a_{\underline{e}}\underline{T^e} & \longmapsto & \sum a_i T^i \end{array}$$

is an isomorphism of rings for every ring A if $\#I=1$.

(2) If $I=\{1,\dots,n\}$, then we also write $A[T_1,\dots,T_n]$ for $A[T_i|i\in I]$, and call this ring the **polynomial ring in n variables $T_1,\dots,T_n$ over A**. In this case, we also write $T_1^{e_1}\cdots T_n^{e_n}$ for $\underline{T^e}$.

(3) For arbitrary A and I, we write T_i for the element $\underline{T^e}$ with $e_j=1$ if $j=i$ and $e_j=0$ if $j\neq i$. This realizes $\{T_i\}_{i\in I}$ as a subset of $A[T_i|i\in I]$, and we write $\iota_I: \{T_i\}_{i\in I}\to A[T_i|i\in I]$ for the inclusion. Elements of this form are called **indeterminates** or **variables**.

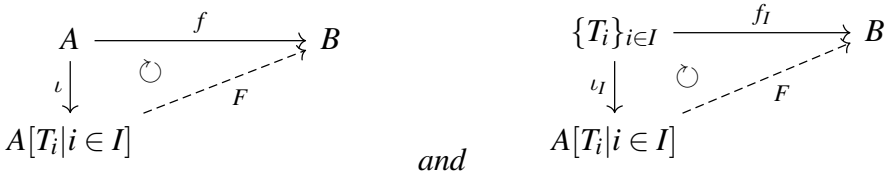
Further, we write $T_{i_1}^{e_{i_1}}\cdots T_{i_n}^{e_{i_n}}$ for an element $\underline{T^e}$ where $\{i_1,\dots,i_n\}$ is a finite subset of I that contains all $i\in I$ for which $e_i\neq 0$. Elements of this form are called **monomials**.

We also think of A as a subring of $A[T_i|i\in I]$, with respect to the inclusion ι from Lemma 1.9.2, and we write a for $\iota(a)$, by abuse of notation.

This notation extends to all elements of $A[T_i|i\in I]$, which are sums of monomials and which are called **polynomials**. Typical examples of polynomials in $\mathbb{R}[T_1,T_2,T_3]$ are $T_3^2-T_1T_2$ and T_2+2.

Sometimes, we use different symbols than T_i for the indeterminates, such as in X^2+Y^2-1, considered as an element of $\mathbb{R}[X,Y,Z]$.

Proposition 1.9.3. *Let A be a ring and $\{T_i\}_{i\in I}$ an indexed set. The polynomial ring $A[T_i|i\in I]$ together with the inclusions $\iota : A \to A[T_i|i\in I]$ and $\iota_I : \{T_i\}_{i\in I} \to A[T_i|i\in I]$ satisfies the following universal property: for every ring homomorphisms $f : A \to B$ and every map $f_I : \{T_i\}_{i\in I} \to B$, there is a unique ring homomorphism $F : A[T_i|i\in I] \to B$ such that $f = F\circ\iota$ and $f_I = F\circ\iota_I$, i.e. the diagrams*



commute.

Proof. If F exists, then $f = F\circ\iota$ and $f_I = F\circ\iota_I$ imply for every polynomial $\sum a_{\underline{e}}\underline{T}^{\underline{e}} \in A[T_i|i\in I]$ that

$$F\big(\sum a_{\underline{e}}\underline{T}^{\underline{e}}\big) \;=\; \sum F(a_{\underline{e}})\prod_{i\in I_0} F(T_i)^{e_i} \;=\; \sum f(a_{\underline{e}})\prod_{i\in I_0} f_I(T_i)^{e_i}$$

where I_0 is the finite set of all $i\in I$ for which $e_i \neq 0$. This shows that F is uniquely determined if it exists.

Assuming that the association $\sum a_{\underline{e}}\underline{T}^{\underline{e}} \mapsto \sum f(a_{\underline{e}})\prod_{i\in I_0} f_I(T_i)^{e_i}$ defines a ring homomorphism $F : A[T_i|i\in I] \to B$, it follows that $F\circ\iota(a) = F(a) = f(a)$ for all $a\in A$ and $F\circ\iota_I(T_i) = F(T_i) = f_I(T_i)$ for all $i\in I$, as desired.

We turn to the verification that F is indeed a ring homomorphism. Clearly, $F(1) = f(1) = 1$. Given two polynomials $\sum a_{\underline{e}}\underline{T}^{\underline{e}}$ and $\sum b_{\underline{e}}\underline{T}^{\underline{e}}$, we have

$$\begin{aligned}
F\Big(\sum a_{\underline{e}}\underline{T}^{\underline{e}} + \sum b_{\underline{e}}\underline{T}^{\underline{e}}\Big) \;&=\; \sum f(a_{\underline{e}}+b_{\underline{e}})\prod f_I(T_i)^{e_i} \\
&=\; \sum f(a_{\underline{e}})\prod f_I(T_i)^{e_i} + \sum f(b_{\underline{e}})\prod f_I(T_i)^{e_i} \\
&=\; F\big(\textstyle\sum a_{\underline{e}}\underline{T}^{\underline{e}}\big) + F\big(\sum b_{\underline{e}}\underline{T}^{\underline{e}}\big), \\
F\Big(\big(\textstyle\sum a_{\underline{e}}\underline{T}^{\underline{e}}\big)\cdot\big(\sum b_{\underline{e}}\underline{T}^{\underline{e}}\big)\Big) \;&=\; \sum_{\underline{e}}\Big(\sum_{\underline{f}+\underline{g}=\underline{e}} f(a_{\underline{f}})f(b_{\underline{g}})\Big)\prod f_I(T_i)^{e_i} \\
&=\; \Big(\sum f(a_{\underline{e}})\prod f_I(T_i)^{e_i}\Big)\cdot\Big(\sum f(b_{\underline{e}})\prod f_I(T_i)^{e_i}\Big) \\
&=\; F\big(\textstyle\sum a_{\underline{e}}\underline{T}^{\underline{e}}\big)\cdot F\big(\sum b_{\underline{e}}\underline{T}^{\underline{e}}\big),
\end{aligned}$$

which shows that F is a ring homomorphism and completes the proof. □

1.10 Field extensions

In this section, we study roots of polynomials. The theory that we have developed in the previous sections allows for a very powerful reinterpretation of the evaluation of a polynomial in an element, which we explain in the example of a polynomial $f = \sum c_i T^i$ with coefficients c_i in a ring A. The value of f in an element a is usually defined as

$$f(a) \;=\; \sum c_i a^i,$$

where we replace the symbol T in $\sum c_i T^i$ by the element a of A and interpret the resulting term $\sum c_i a^i$ as an element of A, using the addition and multiplication of A. If $\mathrm{ev}_a : A[T] \to A$ is the unique ring homomorphism that extends $\mathrm{id}_A : A \to A$ by $\mathrm{ev}_a(T) = a$, then

$$\mathrm{ev}_a(f) \;=\; \mathrm{ev}_a\Big(\sum c_i T^i\Big) \;=\; \sum c_i \mathrm{ev}_a(T)^i \;=\; \sum c_i a^i \;=\; f(a).$$

The reinterpretation of $f(a)$ as $\mathrm{ev}_a(f)$ is a powerful approach to study roots of polynomials and leads to the following generalization of the concept of values of polynomials.

Definition 1.10.1. Let $\alpha : A \to B$ be ring homomorphism, $f \in A[T]$ and $a \in B$. Let $\mathrm{ev}_{\alpha,a} : A[T] \to B$ be the ring homomorphism with $\mathrm{ev}_{\alpha,a}(T) = a$ and whose restriction to A is α. We define the **value of f in a (with respect to α)** as $f(a) = \mathrm{ev}_{\alpha,a}(f)$. We say that a **is a root of** f if $f(a) = 0$.

Notation. Often α will be the identity map $\mathrm{id}_A : A \to A$ or the embedding of a subring A into a ring B. Usually, we take the ring homomorphism α as given and suppress it from the notation, i.e. we write $\mathrm{ev}_a : A[T] \to B$ and say that $f(a) = \mathrm{ev}_a(f)$ is the value of f in a if α is apparent from the context.

Lemma 1.10.2. *Let $\alpha : A \to B$ and $\beta : B \to C$ be a ring homomorphisms, $f \in A[T]$ and $a \in B$ a root of f with respect to α. Then $\beta(a)$ is a root of f with respect to $\beta \circ \alpha : A \to C$.*

Proof. If a is a root of f, then $\mathrm{ev}_{\alpha,a} : A[T] \to B$ sends f to 0 by the definition of a root. By the universal property of polynomial rings (Proposition 1.3.12), $\mathrm{ev}_{\beta\circ\alpha,\beta(a)} : A[T] \to C$ is equal to $\beta \circ \mathrm{ev}_{\alpha,a}$. Thus we have $\mathrm{ev}_{\beta\circ\alpha,\beta(a)}(f) = \beta\big(\mathrm{ev}_{\alpha,a}(f)\big) = \beta(0) = 0$, which shows that $\beta(a)$ is a root of f, as claimed. □

Proposition 1.10.3. *Let K be a field, $f \in K[T]$ a polynomial of positive degree and $a \in K$ a root of f. Then there is a unique polynomial $g \in K[T]$ such that $f = (T-a)g$, and this polynomial has degree $\deg g = \deg f - 1$.*

Proof. The uniqueness of g is clear since $K[T]$ is an integral domain and thus the multiplication by $T - a$ is injective. It is also clear that $\deg f = \deg(T-a) + \deg g = 1 + \deg g$ if $f = (T-a)g$.

As our next step, we show that $\ker(\mathrm{ev}_a) = \langle T - a\rangle$. Since $\mathrm{ev}_a(T-a) = a - a = 0$, we have $\langle T-a\rangle \subset \ker(\mathrm{ev}_a)$. To prove the inverse inclusion, note that $T-a$ is irreducible in $K[T]$, cf. Example 1.7.4. Since $K[T]$ is a principal ideal domain, cf. Example 1.6.4 and Proposition 1.6.7, the ideal $\langle T-a\rangle$ generated by $T-a$ is maximal, cf. Exercise 1.24. Since $\mathrm{ev}_a(1) = 1$, the ideal $\ker(\mathrm{ev}_a)$ is proper, and thus $\langle T-a\rangle \subset \ker(\mathrm{ev}_a)$ implies that $\ker(\mathrm{ev}_a) = \langle T-a\rangle$.

By our assumptions, $\mathrm{ev}_a(f) = f(a) = 0$, which means that $f \in \ker(\mathrm{ev}_a) = \langle T-a\rangle$. Thus $f = (T-a)g$ for some $g \in K[T]$, as desired, which concludes the proof. □

Remark. In fact, the polynomial g with $f = (T-a)g$ in Proposition 1.10.3 can be found by using polynomial division.

Corollary 1.10.4. *Let A be an integral domain and $f \in A[T]$ a nonzero polynomial of degree n. Then f has at most n pairwise distinct roots in A.*

Proof. If f has has k pairwise distinct zeros $a_1, \dots, a_k \in A$, then by Lemma 1.10.2 also $\frac{a_1}{1}, \dots, \frac{a_k}{1} \in \operatorname{Frac} A$ are k pairwise distinct roots of f. Thus we can assume without loss of generality that $A = K$ is a field.

We prove the result by induction on $n = \deg f$. If $n = 0$, then $f = b$ is a nonzero constant, which has no root in K.

If $n > 0$ and a is a root of f, then Proposition 1.10.3 implies that $f = (T-a)g$ for some $g \in K[T]$ with $\deg g = \deg f - 1$. If b is a another root of f, then $\operatorname{ev}_b(T-a) \cdot \operatorname{ev}_b(g) = \operatorname{ev}_b(f) = 0$, and thus either $\operatorname{ev}_b(T-a) = 0$, i.e. $b = a$, or $\operatorname{ev}_b(g) = 0$, i.e. b is a root of g. By the inductive hypothesis, g has at most $n-1$ pairwise distinct roots in K, which implies that f has at most n distinct roots, as claimed. □

Definition 1.10.5. A **field extension** is a ring homomorphism $\alpha : K \to L$ of fields. The **degree of L over K** is the dimension $[L : K] = \dim_K L$ of L as a K-vector space. A field extension $\alpha : K \to L$ is **finite** if $[L : K]$ is finite.

Notation. We write L/K for a field extension if α is clear from the context. We also say that L is an *extension field of K*. Note that a field extension $\alpha : K \to L$ is injective, cf. Exercise 1.15.

Proposition 1.10.6. *Let K be a field and $f \in K[T]$ an irreducible polynomial. Then $L = K[T]/\langle f \rangle$ is a field and the canonical homomorphism $\alpha : K \to K[T] \to L$ is a field extension. The class $\overline{T}$ of T in L is a root of f. If $\beta : K \to A$ is a ring homomorphism and $a \in A$ is a root of f with respect to β, then there is a unique ring homomorphism $\bar{\beta} : L \to A$ with $\beta = \bar{\beta} \circ \alpha$ and $\bar{\beta}(\overline{T}) = a$.*

Proof. Since $K[T]$ is a principal domain, cf. Example 1.6.4 and Proposition 1.6.7, the irreducible polynomial f generates a maximal ideal $\langle f \rangle$, cf. Exercise 1.24. By Lemma 1.3.14, $L = K[T]/\langle f \rangle$ is a field, and thus $\alpha : K \to L$ a field extension. By definition of $L = K[T]/\langle f \rangle$, the quotient map $\operatorname{ev}_{\overline{T}} : K[T] \to L$ maps f to $\operatorname{ev}_{\overline{T}}(f) = \overline{f} = \overline{0}$, i.e. $\overline{T} = \operatorname{ev}_{\overline{T}}(T)$ is a root of f.

By the universal property of polynomial rings (Proposition 1.9.3), a ring homomorphism $\beta : K \to A$ and an element $a \in B$ determine a unique ring homomorphism $\hat{\beta} : K[T] \to B$ with $\hat{\beta}|_K = \beta$ and $\hat{\beta}(T) = a$. If a is a root of f, then $\langle f \rangle \subset \ker(ev_a)$. Thus the universal property of the quotient $\operatorname{ev}_{\overline{T}} : K[T] \to L$ (Proposition 1.3.8), implies that there is a unique morphism $\bar{\beta}$ such that $\hat{\beta} = \bar{\beta} \circ \operatorname{ev}_{\overline{T}}$, and thus $\beta = \bar{\beta} \circ \alpha$, and $\bar{\beta}(\overline{T}) = a$. □

Corollary 1.10.7. *Let K be a field and $f \in K[T]$ a polynomial of positive degree. Then there is a field extension L/K that contains a root $a \in L$ of f.*

Proof. Since $K[T]$ is a unique factorization domain, f has a factorization into irreducible elements. Let f_0 be an irreducible divisor of f and $L = K[T]/\langle f_0 \rangle$. By Proposition 1.10.6, L is a field extension of K and the class $\overline{T}$ of T in L is a root of f_0. Thus $f_0 \in \ker(\operatorname{ev}_{\overline{T}})$ and also $f \in \ker(\operatorname{ev}_{\overline{T}})$, which shows that $\overline{T} \in L$ is a root of f. □

Definition 1.10.8. A field K is **algebraically closed** if every polynomial $f \in K[T]$ of positive degree has a root in K.

Example 1.10.9. The fundamental theorem of algebra asserts that $\mathbb{C}$ is algebraically closed.

Proposition 1.10.10. *Let K be an algebraically closed field.*

(1) For every polynomial $f \in K[T]$ of degree n, there are elements $u, a_1, \dots, a_n \in K$ such that $f = u\prod_{i=1}^{n}(T - a_i)$.

(2) Every finite field extension $K \to L$ is an isomorphism.

Proof. We prove (1) the claim by induction in $n = \deg f$. If $n = 0$, then $f = u \in K$. If $n > 0$, then f has a root $a_n \in K$ since K is algebraically closed. By Proposition 1.10.6, $f = (T - a_n)g$ for a polynomial $g \in K[T]$ of degree $n-1$. By the inductive hypothesis, $g = u\prod_{i=1}^{n-1}(T - a_n)$, and thus $f = u\prod_{i=1}^{n}(T - a_n)$, as claimed.

To prove (2), consider a finite field extension $\alpha : K \to L$ and $a \in L$, which determines a morphism $\mathrm{ev}_a : K[T] \to L$ that extends α and maps T to a by the universal property of polynomial rings (Proposition 1.9.3). The subring $\mathrm{im}(\mathrm{ev}_a)$ of L is an integral domain and thus $\ker(\mathrm{ev}_a)$ a prime ideal of $K[T]$ by Lemma 1.3.14. Since $K[T] = \bigoplus_{i \geqslant 0} K \cdot T^i$ is a K-vector space of infinite dimension, but $\dim_K L$ is finite, ev_a cannot be injective. Since $K[T]$ is a principal ideal domain, the nonzero prime ideal $\ker(\mathrm{ev}_a)$ is generated by an irreducible polynomial $f \in K[T]$, cf. . By (1), $f = u(T - b)$ for some $u, b \in K$, i.e. $\mathrm{ev}_a(T - b) = 0$. We conclude that $\alpha(b) = \mathrm{ev}_a\big(T - (T - b)\big) = \mathrm{ev}_a(T) - \mathrm{ev}_a(T - b) = a$. This shows that the injective ring homomorphism $\alpha : K \to L$ is surjective and thus an isomorphism, as claimed. □

1.11 Gauss's lemma and polynomial rings over unique factorization domains

In this section, we study Gauss's lemma and apply it to prove that polynomial rings over unique factorization domains are again unique factorization domains. Throughout the section, A is a unique factorization domain and $K = \mathrm{Frac}A$ its field of fractions. We consider A as a subring of K and write a for $\frac{a}{1} \in K$.

Recall from Section 1.7, in particular see Corollary 1.7.14, that $\mathcal{P}(A)$ is the set of principal prime ideals of A and that we define for every $\mathfrak{p} = \langle f \rangle$ in $\mathcal{P}(A)$ the function $\mathrm{ord}_\mathfrak{p} : A - \{0\} \to \mathbb{N}$ with values $\mathrm{ord}_\mathfrak{p}(a) = \max\{n \geqslant 0 \,|\, f^n | a\}$, which we call the *order of a in $\mathfrak{p}$*.

Recall further that if $a = u\prod p_i$ is a factorization into prime elements $p_i \in A$, then $\mathrm{ord}_\mathfrak{p}(a) = \#\{i \mid \langle p_i \rangle = \mathfrak{p}\}$ is the number of prime factors p_i in this factorization that generate $\mathfrak{p}$. Since these numbers add up in products, we gain the formula

$$\mathrm{ord}_\mathfrak{p}(ab) \;=\; \mathrm{ord}_\mathfrak{p}(a) + \mathrm{ord}_\mathfrak{p}(b).$$

The main result of this section, Gauss's lemma, extends this formula to the content of polynomials over K.

As a first step, we generalize the concept of the order in $\mathfrak{p}$ to K and $K[T]$.

Definition 1.11.1. Let $\mathfrak{p} \in \mathcal{P}(A)$ and $\frac{a}{b} \in K^\times$. The **order of** $\frac{a}{b}$ **in** $\mathfrak{p}$ is

$$\operatorname{ord}_\mathfrak{p}\left(\frac{a}{b}\right) \;=\; \operatorname{ord}_\mathfrak{p}(a) - \operatorname{ord}_\mathfrak{p}(b).$$

Let $f = \sum c_i T^i$ be a nonzero polynomial in $K[T]$ of degree n. The **order of** f **in** $\mathfrak{p}$ is

$$\operatorname{ord}_\mathfrak{p}(f) \;=\; \min_{i=0,\dots,n} \big\{ \operatorname{ord}_\mathfrak{p}(c_i) \big\}.$$

Remark. The value $\operatorname{ord}_\mathfrak{p}(\frac{a}{b})$ is well-defined, as can be seen as follows. If $\frac{a}{b} = \frac{a'}{b'}$ in K, i.e. $ab' = a'b$ in A, then $\operatorname{ord}_\mathfrak{p}(a) + \operatorname{ord}_\mathfrak{p}(b') = \operatorname{ord}_\mathfrak{p}(a') + \operatorname{ord}_\mathfrak{p}(b)$ and thus

$$\operatorname{ord}_\mathfrak{p}(\tfrac{a}{b}) \;=\; \operatorname{ord}_\mathfrak{p}(a) - \operatorname{ord}_\mathfrak{p}(b) \;=\; \operatorname{ord}_\mathfrak{p}(a') - \operatorname{ord}_\mathfrak{p}(b') \;=\; \operatorname{ord}_\mathfrak{p}(\tfrac{a'}{b'}),$$

which shows that $\operatorname{ord}_\mathfrak{p}(\frac{a}{b})$ does not dependent on the representative for $\frac{a}{b}$.

Example 1.11.2. Let $A = \mathbb{Z}$ and $K = \mathbb{Q}$. Consider a prime number $p \in \mathbb{Z}$ and the prime ideal $\mathfrak{p} = \langle p \rangle$. Then we have $\operatorname{ord}_\mathfrak{p}(ap^i) = i$ for all $i \geqslant 0$ and $a \in \mathbb{Z}$ such that $p \nmid a$. Thus $\operatorname{ord}_\mathfrak{p}(\frac{a}{b}p^i) = i$ for all $i \in \mathbb{Z}$ and $a, b \in \mathbb{Z}$ with $p \nmid ab$.

As an example, we examine the orders of the polynomial $f = 6T^2 + \frac{3}{2}T - 9 \in \mathbb{Q}[T]$:

$$\operatorname{ord}_{\langle 2 \rangle}(f) \;=\; -1, \qquad \operatorname{ord}_{\langle 3 \rangle}(f) \;=\; 1, \qquad \text{and} \qquad \operatorname{ord}_{\langle p \rangle}(f) \;=\; 0$$

for all prime numbers $p \geqslant 5$.

Definition 1.11.3. A **principal fractional ideal of** A is a subset of K of the form $\langle a \rangle_A = \{ca \in K \mid c \in A\}$ for some $a \in K$. Given two principal fractional ideals I and J of A, we define their product as the principal fractional ideal

$$I \cdot J \;=\; \{ab \mid a \in I, b \in J\}.$$

Given a prime ideal $\mathfrak{p} = \langle p \rangle$ and $n \in \mathbb{Z}$, we define its n-th power as the principal fractional ideal $\mathfrak{p}^n = \langle p^n \rangle_A$.

Remark. Note that if $I = \langle a \rangle_A$ and $J = \langle b \rangle_A$ for some $a, b \in K$, then $I \cdot J = \langle ab \rangle_A$, which shows that $I \cdot J$ is indeed a principal fractional ideal. Note further that this coincides with the usual product of ideals if I and J are ideals of A. Similarly, p^n is the usual n-fold self-product of $\mathfrak{p}$ if n is positive. Finally we note that a principal fractional ideal I of A is an ideal of A if and only if $I \subset A$.

Example 1.11.4. Let $A = \mathbb{Z}$ and $K = \mathbb{Q}$. The principal fractional ideal generated by $\frac{a}{b} \in \mathbb{Q}^\times$ is

$$\langle \tfrac{a}{b} \rangle_\mathbb{Z} \;=\; \{\tfrac{na}{b} \in \mathbb{Q} \mid n \in \mathbb{Z}\},$$

which is the additive subgroup of $\mathbb{Q}$ generated by $\frac{a}{b}$. If $p \in \mathbb{Z}$ is prime, then $\langle p \rangle_\mathbb{Z}^{-1} = \langle \frac{1}{p} \rangle_\mathbb{Z}$. If

$$\frac{a}{b} \;=\; \frac{p_1 \cdots p_r}{q_1 \cdots q_s}, \qquad \text{then} \qquad \left\langle \frac{a}{b} \right\rangle_\mathbb{Z} \;=\; \langle p_1 \rangle_\mathbb{Z} \cdots \langle p_r \rangle_\mathbb{Z} \cdot \langle q_1 \rangle_\mathbb{Z}^{-1} \cdots \langle q_s \rangle_\mathbb{Z}^{-1}.$$

We finally are prepared to introduce the central concept of this section, which is the content of a polynomial.

Definition 1.11.5. Let $f \in K[T]$ be a nonzero polynomial. The **content of** f is the principal fractional ideal

$$\operatorname{cont}(f) = \prod_{\mathfrak{p}\in\mathcal{P}(A)} \mathfrak{p}^{\operatorname{ord}_{\mathfrak{p}}(f)},$$

and f is **primitive** if $\operatorname{cont}(f) = \langle 1\rangle_A$.

Remark. Note that the product $\prod \mathfrak{p}^{\operatorname{ord}_{\mathfrak{p}}(f)}$ in the definition of the content of a polynomial $f \in K[T]$ is finite since for every nonzero $a \in A$ with factorization $a = u\prod p_i$, there are only finitely many $\mathfrak{p} \in \mathcal{P}(A)$ such that $\operatorname{ord}_{\mathfrak{p}}(a) = \#\{p_i \mid \mathfrak{p} = \langle p_i\rangle\}$ is nonzero.

Choose a generator $p_{\mathfrak{p}}$ for every $\mathfrak{p} \in \mathcal{P}(A)$. Since $\langle a\rangle \cdot \langle b\rangle = \langle ab\rangle$, we have

$$\operatorname{cont}(f) = \left\langle \prod_{\mathfrak{p}\in\mathcal{P}(A)} p_{\mathfrak{p}}^{\operatorname{ord}_{\mathfrak{p}}(f)} \right\rangle_A.$$

In the literature, sometimes the content of a polynomial is defined as the element $\prod p_{\mathfrak{p}}^{\operatorname{ord}_{\mathfrak{p}}(f)}$, which is an element of K that is well-defined up to taking associates. This is particularly appealing if there a natural choices of generators of prime ideals, such as the positive prime numbers in $\mathbb{Z}$.

Example 1.11.6. Let $A = \mathbb{Z}$ and $K = \mathbb{Q}$. Consider the polynomial $f = 6T^2 + \frac{3}{2}T - 9$ from Example 1.11.2. Its content is

$$\operatorname{cont}(f) = \prod_{\mathfrak{p}\in\mathcal{P}(\mathbb{Z})} \mathfrak{p}^{\operatorname{ord}_{\mathfrak{p}}(f)} = \langle 2\rangle_{\mathbb{Z}}^{-1} \cdot \langle 3\rangle_{\mathbb{Z}} = \left\langle \frac{3}{2} \right\rangle_{\mathbb{Z}}.$$

An example of a primitive polynomial is $f = 6T^2 + 2T - 9$.

Lemma 1.11.7. *Consider $a \in K^\times$ and $f \in K[T]$. Then the following hold true:*

(1) $a \in A$ *if and only if* $\operatorname{ord}_{\mathfrak{p}}(a) \geqslant 0$ *for all* $\mathfrak{p} \in \mathcal{P}(A)$*;*

(2) $a \in A^\times$ *if and only if* $\operatorname{ord}_{\mathfrak{p}}(a) = 0$ *for all* $\mathfrak{p} \in \mathcal{P}(A)$*;*

(3) $f \in A[T]$ *if and only if* $\operatorname{cont}(f) \subset A$*;*

(4) $\operatorname{cont}(af) = \langle a\rangle_A \cdot \operatorname{cont}(f)$*;*

(5) $\operatorname{cont}(a) = \langle a\rangle_A$.

Proof. Let $a = \frac{b}{c}$ with $b, c \in A$. Then $\frac{b}{c} \in A$ if and only if $c|b$, which is equivalent with $\operatorname{ord}_{\mathfrak{p}}(a) = \operatorname{ord}_{\mathfrak{p}}(b) - \operatorname{ord}_{\mathfrak{p}}(c) \geqslant 0$ for all $\mathcal{P}(A)$ by Corollary 1.7.14. Thus (1). Since $\frac{b}{c} \in A^\times$ if and only if $b|c$ and $c|b$, this implies (2) at once.

If $f \in A[T]$, then $\operatorname{ord}_{\mathfrak{p}}(f) \geqslant 0$ for all $\mathfrak{p} \in \mathcal{P}(A)$ and thus $\operatorname{cont}(f) \subset A$. Conversely, assume that $\operatorname{cont}(f) \subset A$. As remarked before, $\operatorname{cont}(f) = \langle \prod p_{\mathfrak{p}}^{\operatorname{ord}_{\mathfrak{p}}(f)}\rangle$ where the product is taken over all $\mathfrak{p} \in \mathcal{P}(A)$ such that $\operatorname{ord}_{\mathfrak{p}}(f) \neq 0$ and $p_{\mathfrak{p}}$ is a chosen generator of each $\mathfrak{p}$. Thus we have $\operatorname{cont}(f) \subset A$ if and only if $\prod p_{\mathfrak{p}}^{\operatorname{ord}_{\mathfrak{p}}(f)} \in A$, which is the case if

and only if $\operatorname{ord}_\mathfrak{p}(f) \geqslant 0$ for all $\mathfrak{p} \in \mathcal{P}(A)$. This can only happen if all coefficients a_i of f have positive order $\operatorname{ord}_\mathfrak{p}(a_i)$ in all $\mathfrak{p} \in \mathcal{P}(A)$, i.e. $a_i \in A$. This shows that $f \in A[T]$. Thus (3).

Since $\operatorname{ord}_\mathfrak{p}(ac_i) = \operatorname{ord}_\mathfrak{p}(a) + \operatorname{ord}_\mathfrak{p}(c_i)$ for every coefficient c_i of $f = \sum c_i T^i$, we have

$$\begin{aligned}\operatorname{cont}(af) &= \Big\langle \prod_{\operatorname{ord}_\mathfrak{p}(af)\neq 0} p_\mathfrak{p}^{\operatorname{ord}_\mathfrak{p}(af)} \Big\rangle = \Big\langle \prod_{\operatorname{ord}_\mathfrak{p}(af)\neq 0} p_\mathfrak{p}^{\operatorname{ord}_\mathfrak{p}(a)+\operatorname{ord}_\mathfrak{p}(f)} \Big\rangle \\ &= \Big\langle \prod_{\operatorname{ord}_\mathfrak{p}(a)\neq 0} p_\mathfrak{p}^{\operatorname{ord}_\mathfrak{p}(a)} \cdot \prod_{\operatorname{ord}_\mathfrak{p}(f)\neq 0} p_\mathfrak{p}^{\operatorname{ord}_\mathfrak{p}(f)} \Big\rangle = \langle a\rangle_A \cdot \operatorname{cont}(f)\end{aligned}$$

where $p_\mathfrak{p}$ is a generator of $\mathfrak{p}$. This shows (4).

In particular, (4) implies $\operatorname{cont}(a) = \langle a\rangle_A \cdot \operatorname{cont}(1) = \langle a\rangle_A \cdot \langle 1\rangle = \langle a\rangle_A$, which shows (5) and concludes the proof. □

Theorem 1.11.8 (Gauss's lemma)**.** *Let $f, g \in K[T]$ be nonzero polynomials. Then*

$$\operatorname{cont}(fg) = \operatorname{cont}(f) \cdot \operatorname{cont}(g).$$

Proof. Let $\mathfrak{p} \in \mathcal{P}(A)$. We begin with the proof of the claim that if $f, g \in A[T]$ and $\operatorname{ord}_\mathfrak{p}(f) = \operatorname{ord}_\mathfrak{p}(g) = 0$, then $\operatorname{ord}_\mathfrak{p}(fg) = 0$. To this end, we consider the ring homomorphism $\alpha : A[T] \to (A/\mathfrak{p})[T]$ that sends a polynomial $f = \sum c_i T^i$ to $\bar{f} = \sum \bar{c}_i T^i$ where $\bar{c}_i \in A/\mathfrak{p}$ is image of c_i under the canonical surjection $A \to A/\mathfrak{p}$. Thus if $f = \sum c_i T^i$, then $\operatorname{ord}_\mathfrak{p}(\sum c_i T^i) = 0$ if and only if $c_i \notin \mathfrak{p}$ for some i, which is the case if and only if $\bar{c}_i \neq 0$ for some i, i.e. $\bar{f} \neq 0$ in $(A/\mathfrak{p})[T]$. Thus $\operatorname{ord}_\mathfrak{p}(f) = \operatorname{ord}_\mathfrak{p}(g) = 0$ implies that $\bar{f}$ and $\bar{g}$ are nonzero in $(A/\mathfrak{p})[T]$. Since $\mathfrak{p}$ is a prime ideal, $A/\mathfrak{p}$ is an integral domain, and so is $(A/\mathfrak{p})[T]$. Thus $\bar{f}\bar{g} \neq 0$, which means that $\operatorname{ord}_\mathfrak{p}(fg) = 0$, as claimed.

Consider arbitrary nonzero $f, g \in K[T]$ and let $c_f, c_g \in K^\times$ be generators of the content of f and g, respectively, i.e. $\operatorname{cont}(f) = \langle c_f\rangle_A$ and $\operatorname{cont}(g) = \langle c_g\rangle_A$. Then $f_0 = c_f^{-1} f$ and $g_0 = c_g^{-1} g$ have content $\operatorname{cont}(f_0) = \langle c_f\rangle_A^{-1} \operatorname{cont}(f) = \langle 1\rangle$ and $\operatorname{cont}(g_0) = \langle c_g\rangle_A^{-1} \operatorname{cont}(g) = \langle 1\rangle$. By Lemma 1.11.7, $f_0, g_0 \in A[T]$ and $\operatorname{ord}_\mathfrak{p}(f_0) = \operatorname{ord}_\mathfrak{p}(g_0) = 0$ for all $\mathfrak{p} \in \mathcal{P}(A)$. By what we have proven before, $\operatorname{ord}_\mathfrak{p}(f_0 g_0) = 0$ for all $\mathfrak{p} \in \mathcal{P}(A)$, and thus

$$\operatorname{cont}(fg) = \langle c_f c_g\rangle_A \cdot \underbrace{\operatorname{cont}(f_0 g_0)}_{=\langle 1\rangle} = \langle c_f\rangle_A \cdot \langle c_g\rangle_A = \operatorname{cont}(f) \cdot \operatorname{cont}(g),$$

which completes the proof. □

Theorem 1.11.9. *Let A be a unique factorization domain and $K = \operatorname{Frac} A$. Then $A[T]$ is a unique factorization domain whose prime elements are the prime elements of A and the primitive irreducible polynomials in $K[T]$.*

Proof. Since A is an integral domain, $\deg(fg) = \deg f + \deg g$ for all $f, g \in A[T]$. Thus, in particular, $A[T]^\times = A^\times$.

Let $p \in A$ be irreducible and $p = gh$ a factorization in $A[T]$. Then $g, h \in A$ and thus $g \in A^\times$ or $h \in A^\times$ since p is irreducible. This shows that p is irreducible in $A[T]$.

Let $f \in K[T]$ be a primitive irreducible polynomial and $f = gh$ a factorization with $g, h \in A[T]$. Since f is irreducible, $g \in K^\times$ or $h \in K^\times$. Without loss of generality, we may assume that $g \in K^\times$, i.e. in fact $g \in A - \{0\}$. By Lemma 1.11.7 (5), we have $\operatorname{cont}(g) = \langle g \rangle_A$. Since $\operatorname{cont}(g) \cdot \operatorname{cont}(h) = \operatorname{cont}(f) = \langle 1 \rangle$ by Gauss's lemma (Theorem 1.11.8) and both $\operatorname{cont}(g)$ and $\operatorname{cont}(h)$ are contained in A, we must have $\operatorname{cont}(g) = \operatorname{cont}(h) = \langle 1 \rangle$. Thus $\langle g \rangle = \langle 1 \rangle$, which shows that $g \in A^\times$. This shows that f is irreducible in $A[T]$.

Our next step is to show that every nonzero polynomial $f \in A[T]$ has a factorization into irreducible elements of the exhibited forms. Let $f = u \prod g_i$ be a factorization in $K[T]$, i.e. $u \in K^\times$ and $g_i \in K[T]$ are irreducible. We let $c_i \in K^\times$ be elements such that $\operatorname{cont}(g_i) = \langle c_i \rangle$ and define $g_{i,0} = c_i^{-1} g_i$, which are polynomials of content $\langle 1 \rangle$ and thus primitive and in $A[T]$. Since $g_{i,0}$ and g_i are associated in $K[T]$, $g_{i,0}$ is irreducible in $K[T]$, and since it is primitive, it is irreducible in $A[T]$, as shown above. If $u_0 = \prod c_i$, then $f = u_0 \prod g_{i,0}$. By Gauss's lemma (Theorem 1.11.8), we have

$$\langle u_0 \rangle \;=\; \langle u_0 \rangle \prod \underbrace{\operatorname{cont} g_{i,0}}_{=\langle 1 \rangle} \;=\; \operatorname{cont}\left(u_0 \prod g_{i,0}\right) \;=\; \operatorname{cont}(f) \;\subset\; A,$$

which shows that $u_0 \in A$. Let $u_0 = v \prod p_j$ be a factorization into irreducible elements p_i in A, which are irreducible in $A[T]$ by what we have shown above. This yields the factorization $f = v \prod p_i \prod g_{i,0}$ of f in $A[T]$. We conclude that every nonzero polynomial of $A[T]$ has a factorization into irreducible factors of the exhibited forms.

Moreover, we conclude that we have found all irreducible elements of $A[T]$. Indeed, let $f \in A[T]$ be irreducible and $f = u \prod g_i$ a factorization into irreducible elements of the exhibited forms. Since f is irreducible, this factorization contains precisely one irreducible factor g_1, which is either a prime in A or a primitive irreducible polynomial in $K[T]$. Thus $f = u g_1$ is of this form as well.

We conclude the proof by establishing the uniqueness of factorizations. Consider two factorizations

$$f \;=\; u \prod_{i=1}^{r} p_i \prod_{i=1}^{n} g_i \;=\; v \prod_{i=1}^{s} q_i \prod_{i=1}^{m} h_i$$

where $p_i, q_i \in A$ are prime and $g_i, h_i \in K[T]$ are primitive and irreducible. Then $\tilde{u} = u \prod p_i$ and $\tilde{v} = v \prod q_i$ are in $K^\times$ and thus $f = \tilde{u} \prod g_i = \tilde{v} \prod h_i$ are factorizations in $K[T]$. Since $K[T]$ is a unique factorization domain, there is a bijection $\sigma : \{1, \ldots, n\} \to \{1, \ldots, m\}$ such that $g_i \sim h_{\sigma(i)}$ for all $i = 1, \ldots, n$. Thus $\prod g_i = w \prod h_i$ for some $w \in K^\times$. By Gauss's lemma (Theorem 1.11.8) and Lemma 1.11.7 (5), we have $\prod \operatorname{cont}(g_i) = \langle w \rangle_A \cdot \prod \operatorname{cont}(h_i)$. Since $\operatorname{cont}(g_i) = \operatorname{cont}(h_i) = \langle 1 \rangle$, we have $\langle w \rangle_A = \langle 1 \rangle$ and thus $w \in A^\times$. Since $A[T]$ is an integral domain, we can cancel the factor $\prod h_i$ in the equality

$$u \prod_{i=1}^{r} p_i \left(w \prod_{i=1}^{n} h_i \right) \;=\; u \prod_{i=1}^{r} p_i \prod_{i=1}^{n} g_i \;=\; v \prod_{i=1}^{s} q_i \prod_{i=1}^{n} h_i,$$

which yields that equality $(uw) \prod p_i = v \prod q_i$. Since A is a unique factorization domain, there is a bijection $\tau : \{1, \ldots, r\} \to \{1, \ldots, s\}$ such that $p_i \sim q_{\tau(i)}$ for all $i = 1, \ldots, r$. This shows that f has a unique factorization and concludes the proof of the theorem. $\square$

Corollary 1.11.10. *Let A be a unique factorization domain and $n \geqslant 1$. Then $A[T_1, \dots, T_n]$ is a unique factorization domain.*

Proof. This follows immediately by induction over n from Theorem 1.11.9. □

1.12 Irreducibility criteria

In this section, we develop some criteria for the irreducibility of polynomials over fields and unique factorization domains.

Lemma 1.12.1. *Let A be a unique factorization domain, $K = \operatorname{Frac} A$ and $f \in K[T]$ a primitive polynomial. Then f is irreducible in $A[T]$ if and only if f is irreducible in $K[T]$.*

Proof. If f is irreducible in $K[T]$, then is is irreducible in $A[T]$ by Theorem 1.11.9. Conversely, assume that f is irreducible in $A[T]$ and consider a factorization $f = gh$ in $K[T]$. Choose $c_g, c_h \in K^\times$ with $\operatorname{cont}(g) = \langle c_g \rangle_A$ and $\operatorname{cont}(h) = \langle c_h \rangle_A$ and define $g_0 = c_g^{-1} g$ and $h_0 = c_h^{-1} h$. By Lemma 1.11.7 (5), we have $\operatorname{cont}(g_0) = \langle c_g^{-1} \rangle_A \cdot \operatorname{cont}(g) = \langle 1 \rangle$ and $\operatorname{cont}(h_0) = \langle c_h^{-1} \rangle_A \cdot \operatorname{cont}(h) = \langle 1 \rangle$, which shows that $g_0, h_0 \in A[T]$. Since

$$\langle c_g c_h \rangle_A = \langle c_g c_h \rangle_A \cdot \underbrace{\operatorname{cont}(g_0 h_0)}_{= \langle 1 \rangle} = \operatorname{cont}(c_g c_h g_0 h_0) = \operatorname{cont}(gh) = \operatorname{cont}(f) = 1,$$

we see that $c_g c_h \in A^\times$. Thus $f = (c_g c_h) g_0 h_0$ is a factorization in $A[T]$. Since f is irreducible, either $g_0 \in A^\times$ or $h_0 \in A^\times$, which means that either $g = c_g g_0 \in K^\times$ or $h = c_h h_0 \in K^\times$. This shows that f is irreducible in $K[T]$. □

Lemma 1.12.2. *Let A be a unique factorization domain and $f \in A[T]$ a polynomial of degree at least 2. If f has a root in A, then f is not irreducible in $A[T]$.*

Proof. If $a \in A$ is a root of f, then $f = (T - a)g$ for some $g \in K[T]$ by Proposition 1.10.3. Since $\operatorname{cont}(T - a) = \langle 1 \rangle$, we conclude that $\operatorname{cont}(g) = \operatorname{cont}(f)$ and thus $g \in A[T]$. Since $\deg g = \deg f - 1 \geqslant 1$, the polynomial g is not a unit in $A[T]$, which shows that f is not irreducible. □

Proposition 1.12.3. *Let A be a unique factorization domain, $K = \operatorname{Frac} A$ and $f \in A[T]$ a primitive polynomial of degree 2 or 3. If f has no root in K, then f is irreducible in $A[T]$.*

Proof. If $f = gh$ in $A[T]$, then $\operatorname{cont}(g) = \operatorname{cont}(h) = 1$. If $\deg g = 1$, i.e. $g = aT - b$ for $a, b \in A$ with $a \neq 0$, then $\frac{b}{a} \in K$ is a root of g and therefore of f. Thus $\deg g \neq 1$, and similarly $\deg h \neq 1$. Since $\deg g + \deg h = \deg f$ is 2 or 3, we conclude that $\deg g = 0$ or $\deg h = 0$. Since $\operatorname{cont}(g) = \operatorname{cont}(h) = \langle 1 \rangle$, we conclude that $g \in A^\times$ or $h \in A^\times$, which shows that f is irreducible. □

Proposition 1.12.4 (Eisenstein criterion)**.** *Let A be a unique factorization domain, $p \in A$ a prime element, $f = \sum a_i T^i$ a polynomial in $A[T]$ of degree $n \geqslant 1$ and $K = \operatorname{Frac} A$. If $p \mid a_i$ for $i = 0, \dots, n-1$, but $p \nmid a_n$ and $p^2 \nmid a_0$, then f is irreducible in $K[T]$.*

Proof. Consider $f = gh$ with $g,h \in K[T]$. We intend to show that one of g and h is a unit in $K[T]$, i.e. a nonzero constant polynomial. Since $f \neq 0$, it is clear that $g \neq 0$ and $h \neq 0$.

We have $g = c_g g_0$ and $h = c_h h_0$ for some primitive $g_0, h_0 \in A[T]$ and $c_g, c_h \in K$ with $\operatorname{cont}(g) = \langle c_g \rangle$ and $\operatorname{cont}(h) = \langle c_h \rangle$. Since $\operatorname{cont}(f) = \operatorname{cont}(gh) = \langle c_g c_h \rangle$, also $c = c_g c_h \in A$, and thus $f = c g_0 h_0$ is an equation in $A[T]$. If we can show that one of g_0 and h_0 is a constant polynomial, then one of g and h is constant. Thus we assume without loss of generality that $g,h \in A[T]$.

Let $g = \sum b_i T^i$ and $h = \sum c_i T^i$. Then by our assumptions, $p \mid a_0 = b_0 c_0$ and thus $p \mid b_0$ or $p \mid c_0$. By the symmetry of g and h, we can assume that $p \mid b_0$. Since $p^2 \nmid a_0 = b_0 c_0$, we conclude that $p \nmid c_0$. Since $p \nmid a_n = b_k c_l$, where $k = \deg g$ and $l = \deg h$, we have $p \nmid b_k$, and thus $m = \min\{i \in \mathbb{N} \mid p \nmid b_i\}$ exists. Then

$$a_m = \underbrace{b_m c_0}_{\notin \langle p \rangle} + \underbrace{b_{m-1} c_1 + \ldots + b_0 c_m}_{\in \langle p \rangle}$$

is not in $\langle p \rangle$, i.e. $p \nmid a_m$. Since $p \mid a_i$ for $i = 0, \ldots, n-1$, we conclude that $m = n$, which shows that $\deg h = n - m = 0$. This shows that h is invertible in $K[T]$, which concludes the proof that f is irreducible in $K[T]$. □

Example 1.12.5. We can apply the Eisenstein criterion (Proposition 1.12.4) to the polynomial $f = T^5 - 2T + 6 \in \mathbb{Z}[T]$ and the prime number 2 to deduce that f is irreducible in $\mathbb{Q}[T]$. Since f is primitive, it is also irreducible in $\mathbb{Z}[T]$. For the same reason, $3f = 3T^5 - 6T + 18$ is irreducible in $\mathbb{Q}[T]$, but it is not primitive and not irreducible in $\mathbb{Z}[T]$.

Proposition 1.12.6 (Reduction criterion)**.** *Let A be a unique factorization domain, B an integral domain, $K = \operatorname{Frac} A$ and $L = \operatorname{Frac} B$. Let $\alpha : A \to B$ be a ring homomorphism and $\hat{\alpha} : A[T] \to B[T]$ the extension of α with $\hat{\alpha}(T) = T$. Let $f \in A[T]$ and $\hat{f} = \hat{\alpha}(f)$. If $\deg \hat{f} = \deg f$ and if $\hat{f}$ is irreducible in $L[T]$, then f is irreducible in $K[T]$.*

Proof. Consider $f = gh \in K[T]$. After multiplying g and h with a suitable unit of K, we can assume that $g,h \in A[T]$. Let $\hat{g} = \hat{\alpha}(g)$ and $\hat{h} = \hat{\alpha}(h)$. Then $\hat{f} = \hat{g}\hat{h}$. Since $\deg \hat{g} + \deg \hat{h} = \deg \hat{f} = \deg f = \deg g + \deg h$, we have $\deg \hat{g} = \deg g$ and $\deg \hat{h} = \deg h$. Since $\hat{f}$ is irreducible in $L[T]$, one of $\hat{g}$ and $\hat{h}$ has degree 0. Thus one of g and h has degree 0, which shows that f is irreducible in $K[T]$. □

Example 1.12.7. Consider $f = 13T^3 + 15T + 7 \in \mathbb{Z}[T]$ and the quotient map $\alpha : \mathbb{Z} \to \mathbb{F}_2$. Then $\hat{f} = T^3 + T + 1$ has the same degreee as f. Since $\hat{f}$ does not have any root in $\mathbb{F}_2$, Proposition 1.12.3 implies that $\hat{f}$ is irreducible in $\mathbb{F}_2[T]$. Thus by the reduction criterion (Proposition 1.12.6), f is irreducible in $\mathbb{Q}[T]$. Since f is primitive, it is irreducible in $\mathbb{Z}[T]$ by Theorem 1.11.9.

1.13 Exercises

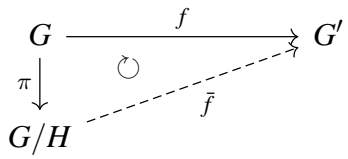
Exercise 1.1. Proof Lemma 1.1.5.

Exercise 1.2 (Group homomorphisms)**.** Let G and H be commutative groups. A **group homomorphism between G and H** is a map $f : G \to H$ such that $f(ab) = f(a)f(b)$ for all $a, b \in G$.

(1) Let $f : G \to H$ be a group homomorphism. Show that $f(e_G) = e_H$ and $f(a^{-1}) = f(a)^{-1}$ for all $a \in G$ where e_G is the neutral elements of G and e_H is the neutral element of H.

(2) Show that the identity map $\mathrm{id} : G \to G$ is a group homomorphism and that the composition $g \circ f : G \to H'$ of two group homomorphisms $f : G \to H$ and $g : H \to H'$ is a group homomorphism.

Exercise 1.3 (Universal property of quotient groups)**.** Let H be a subgroup of a commutative group G and G/H the quotient. Show that the association $a \mapsto [a]$ defines a group homomorphism $\pi : G \to G/H$ with $\pi(H) = \{0\}$. Show that for every group homomorphism $f : G \to G'$ with $f(H) = \{0\}$, there is a unique group homomorphism $\bar{f} : G/H \to G'$ such that $f = \bar{f} \circ \pi$, i.e. the diagram



commutes.

Exercise 1.4 (Cyclic groups)**.** Let G be a commutative group and $a \in G$. We define

$$a^n = \underbrace{a \cdots a}_{n\text{-times}} \quad \text{for } n > 0, \qquad a^0 = e, \quad \text{and} \qquad a^n = \underbrace{a^{-1} \cdots a^{-1}}_{-n\text{-times}} \quad \text{for } n < 0.$$

We call G a **cyclic group** if there is an element $a \in G$ such that every other element $b \in G$ is of the form $b = a^n$ for some $n \in \mathbb{Z}$.

(1) Show that there is a cyclic group C_n for every $n \geqslant 1$.

(2) Are there infinite cyclic groups?

Exercise 1.5. Let A be a ring and $B \subset A$ a subset. Prove that B is a subring if and only if addition α and multiplication μ of A restrict to maps $\alpha_B : B \times B \to B$ and $\mu_B : B \times B \to B$, respectively, such that (B, α_B, μ_B) is a ring and such that the inclusion map $B \to A$ is a ring homomorphism.

Exercise 1.6. Show that every finite integral domain is a field.

Exercise 1.7 (Gaussian integers)**.** Let $i \in \mathbb{C}$ be a square root of -1. Show that the subset $\mathbb{Z}[i] = \{a + bi \in \mathbb{C} | a, b \in \mathbb{Z}\}$ is a subring of $\mathbb{C}$. Is $\mathbb{Z}[i]$ an integral domain? Is it a field? Show that $\mathbb{Q}[i] = \{a + bi | a, b \in \mathbb{Q}\}$ is a subring of $\mathbb{C}$ that is a field.

Remark: $\mathbb{Z}[i]$ is called the **ring of Gaussian integers**.

Exercise 1.8. Proof Lemma 1.3.2.

Exercise 1.9. Show that the set $C^\infty(\mathbb{R})$ of all smooth functions $f : \mathbb{R} \to \mathbb{R}$ is a ring with respect to value-wise addition and multiplication, i.e. $(f+g)(x) := f(x)+g(x)$ and $(f \cdot g)(x) := f(x) \cdot g(x)$. Which of the following maps are ring homomorphisms?

(1) $\mathrm{ev}_a : \mathbb{C}^\infty(\mathbb{R}) \to \mathbb{R}$ with $\mathrm{ev}_a(f) := f(a)$ where $a \in \mathbb{R}$;

(2) $d : \mathbb{C}^\infty(\mathbb{R}) \to \mathbb{C}^\infty(\mathbb{R})$ with $d(f) := \frac{df}{dt}$.

Exercise 1.10. Let A be a ring. Show that $A[T]$ with the addition and multiplication as defined in Definition 1.3.11 is indeed a ring with zero is 0 and one 1. Show that the realization of elements $a \in A$ as constant polynomials defines an injective ring homomorphism $A \to A[T]$. Under which conditions on A is $A[T]$ an integral domain?

Exercise 1.11. Show that the set $\mathbb{F}_p[T] = \{\sum_{i=0}^n a_i T^i | n \geqslant 0, a_i \in \mathbb{F}_p\}$ of polynomials with coefficients in $\mathbb{F}_p$ forms a ring w.r.t. to the addition $\sum a_i T^i + \sum b_i T^i = \sum (a_i + b_i) T^i$ and the multiplication

$$\Big(\sum_{i=0}^{n} a_i T^i\Big) \cdot \Big(\sum_{j=0}^{m} b_j T^j\Big) := \sum_{k=0}^{n+m} \Big(\sum_{i+j=k} a_i b_j\Big) T^k.$$

Which of the following maps are ring homomorphisms?

(1) $\mathrm{ev}_c : \mathbb{F}_p[T] \to \mathbb{F}_p$ with $\mathrm{ev}_c(\sum a_i T^i) = \sum a_i c^i$ where $c \in \mathbb{F}_p$;

(2) $\mathrm{Frob} : \mathbb{F}_p[T] \to \mathbb{F}_p[T]$ with $\sum a_i T^i \to \sum a_i T^{pi}$.

Exercise 1.12. Show that the embedding $i : \mathbb{R} \to \mathbb{R}[T]$ of real numbers as constant polynomials is a ring homomorphism. Show that $\mathbb{R}[T]$ together with $i : \mathbb{R} \to \mathbb{R}[T]$ satisfies the following universal property. For every ring homomorphism $f : \mathbb{R} \to B$ and for every map $\tilde{f} : \{T\} \to B$, there is a unique ring homomorphism $F : \mathbb{R}[T] \to B$ such that $f = F \circ i$ and $F(T) = \tilde{f}(T)$.

Exercise 1.13. Let A be a ring and $\{I_i\}_{i \in I}$ be a family of ideals in A.

(1) Show that the intersection $\bigcap_{i \in I} I_i$ is an ideal of A.

(2) For finite I, show that the product $\prod_{i \in I} I_i = (\{\prod a_i | a_i \in I_i\})$ is an ideal of A that is contained in $\bigcap_{i \in I} I_i$. Under which assumption is $\prod I_i = \bigcap I_i$?

(3) For finite I, show that $\sum I_i$ is indeed an ideal.

(4) Let $f : A \to B$ be a ring homomorphism and I an ideal of B. Show that $f^{-1}(I)$ is an ideal of A. Show that $f^{-1}(I)$ is prime if I is prime. Is $f^{-1}(I)$ maximal if I is maximal? Is the image $f(J)$ of an ideal J of A an ideal of B?

Exercise 1.14. (1) Describe all ideals of $\mathbb{Z}$. Which of them are principal ideals, which of them are prime and which of them are maximal?

(2) Let $f \in \mathbb{R}[T]$ be of degree $\leqslant 2$. When is (f) a prime ideal, when is it a maximal ideal? When is the quotient ring isomorphic to $\mathbb{R}$? When is it isomorphic to $\mathbb{C}$?

Exercise 1.15. Let K be a field and $\alpha : K \to A$ a ring homomorphism. Show that α is injective unless A is the trivial ring.

Exercise 1.16. Let A be a ring and $A^\times$ its unit group.

(1) Show that the map

$$\begin{array}{ccc} A^\times \times A^\times & \longrightarrow & A^\times \\ (a,b) & \longmapsto & ab \end{array}$$

is well-defined and turns $A^\times$ into an abelian group.

(2) Let $f : A \to B$ be a ring homomorphism. Show that $f(A^\times) \subset B^\times$ and that the restriction $f|_{A^\times} : (A^\times, \cdot) \to (B^\times, \cdot)$ of f is a group homomorphism.

(3) Show that the map $A^\times \times A \to A$, defined by $(a,b) \mapsto ab$ is a group action of $(A^\times, \cdot)$ on A.

(4) Show that $(a) = A$ if and only if $a \in A^\times$.

(5) Let A be an integral domain. Consider the map $\Phi : A \to \{\text{ideals of } A\}$ that sends a to the principal ideal (a) of A. Show that $\Phi(a) = \Phi(b)$ if and only if a and b are contained in the same orbit of the action of $A^\times$ on A, i.e. $a \sim b$.

Exercise 1.17. Let $e_1, \ldots, e_n$ be pairwise coprime positive integers. Show that the underlying additive group of $\mathbb{Z}/e_1\mathbb{Z} \times \cdots \times \mathbb{Z}/e_n\mathbb{Z}$ is a cyclic group.

Exercise 1.18. Let A be an integral domain and $a, b, c, d, e \in A$.

(1) Show that if d is a greatest common divisor of b and c and e is a greatest common divisor of ab and ac, then $(e) = (ad)$. Conclude that $\gcd(ab, ac) = (a) \cdot \gcd(b, c)$.

(2) If A is a principal ideal domain, then d is a greatest common divisor of a and b if and only if $(a,b) = (d)$. Conclude that every two elements of a principal ideal domain have a greatest common divisor.

(3) Find an integral domain A with elements $a, b, d \in A$ such that d is a greatest common divisor of a and b, but $(a,b) \neq (d)$.

Exercise 1.19. Let A be a Euclidean domain and $a, b \in A$. Show that in general the sequence $r_0, \ldots, r_n, q_2, \ldots, q_{n+1} \in A$ nor its length n are uniquely determined. In particular, find an example of integers $a, b \in \mathbb{Z}$ and sequences $r_0, \ldots, r_n, q_2, \ldots, q_{n+1} \in \mathbb{Z}$ and $r_0, \ldots, r_m, q_2, \ldots, q_{m+1} \in \mathbb{Z}$ of different lengths $n \neq m$ that each satisfy the assumptions of the Euclidean algorithm.

Exercise 1.20 (Polynomial division)**.** Let K be a field. Use polynomial division to show that $K[T]$ is a Euclidean domain.

*__Exercise 1.21.__ Show that the ring $\mathbb{Z}[T]/\langle T^2 - T + 5\rangle$ is a principal ideal domain but not a Euclidean domain. This can be done along the following steps.

(1) Reason that $N(a + bT) = a^2 + ab + 5b^2$ defines a map $N : \mathbb{Z}[T]/\langle T^2 - T + 5\rangle \to \mathbb{N}$. Show that $N(0) = 0$, $N(1) = 1$ and $N(xy) = N(x)N(y)$. Use the map N to show that the units of $\mathbb{Z}[T]/\langle T^2 - T + 5\rangle$ are ± 1 and that 2 and 3 are irreducible.

(2) Show that every Euclidean domain A that is not a field contains an element $a \notin A^\times \cup \{0\}$ such for every $b \in A$ there is an element $u \in A^\times \cup \{0\}$ such that $a|(b-u)$. Conclude that $\mathbb{Z}[T]/\langle T^2 - T + 5\rangle$ is not a Euclidean domain, by considering the cases $a = 2$ and $a = T$.

(3) Show that N is a **Dedekind-Hasse norm**, i.e. for all nonzero $x, y \in \mathbb{Z}[T]/\langle T^2 - T + 5\rangle$, there are $p, q, r \in \mathbb{Z}[T]/\langle T^2 - T + 5\rangle$ such that $px = qy + r$ and $N(r) < N(y)$. Conclude that $\mathbb{Z}[T]/\langle T^2 - T + 5\rangle$ is a principal ideal domain.

Exercise 1.22 (Universal property of $\mathbb{Z}$)**.** Show that the ring $\mathbb{Z}$ satisfies the following universal property: for every ring A, there is a unique ring homomorphism $f : \mathbb{Z} \to A$. Use this and the universal property of the quotient map $\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ to show that for a ring A whose underlying additive group $(A, +)$ is isomorphic to $(\mathbb{Z}/n\mathbb{Z}, +)$ (as a group), there is a unique ring isomorphism $\mathbb{Z}/n\mathbb{Z} \to A$.

Remark: We say that $\mathbb{Z}$ is an **initial object in the category of rings**.

Exercise 1.23. Let A be a ring and I an ideal of A. Show that I is contained in a maximal ideal of A.

Hint: This is a consequence of Zorn's Lemma.

Exercise 1.24. Let A be a principal ideal domain and $a \in A$. Show that $\langle a\rangle$ is a maximal ideal if and only if a is irreducible.

Exercise 1.25. Let A be an integral domain and $a, b \in A$. Show that $\langle a\rangle = \langle b\rangle$ if and only if $a \mid b$. Is this still true for an arbitrary ring?

Exercise 1.26. Let $\mathbb{Z}[\sqrt{-5}]$ be the set of complex numbers of the form $z = a + b\sqrt{-5}$ with $a, b \in \mathbb{Z}$ and $\sqrt{-5} = i\sqrt{5}$.

(1) Show that $\mathbb{Z}[\sqrt{-5}]$ is a subring of $\mathbb{C}$.

(2) Show that the association $a + b\sqrt{-5} \mapsto a^2 + 5b^2$ defines a map $N : \mathbb{Z}[\sqrt{-5}] \to \mathbb{Z}$ with $N(zz') = N(z)N(z')$ and $N(1) = 1$.
Remark: $N(z)$ is the square of the usual absolute value of the complex number z.

(3) Conclude that $z \in \mathbb{Z}[\sqrt{-5}]^\times$ if and only if $N(z) \in \mathbb{Z}^\times$. Determine $\mathbb{Z}[\sqrt{-5}]^\times$.

(4) Show that 2, 3, $(1 + \sqrt{-5})$ and $(1 - \sqrt{-5})$ are irreducible, but not prime.

(5) Show that 6 and $2 + 2\sqrt{-5}$ do not have a greatest common divisor.

Exercise 1.27. (1) Determine all units, prime elements and irreducible elements of $\mathbb{Z}/6\mathbb{Z}$.

(2) Let $\mathbb{R}[T_1, T_2] = (\mathbb{R}[T_1])[T_2]$ be the polynomial ring over $\mathbb{R}$ in T_1 and T_2 and I the ideal generated by $T_1^2 + T_2^2$. Is the class $\overline{T}_1 = T_1 + I$ a prime element in the quotient ring $\mathbb{R}[T_1, T_2]/I$? Is $\overline{T}_1$ irreducible?

Exercise 1.28. Let A be a unique factorization domain.

(1) Show that every prime ideal of A is generated by a set of prime elements.

(2) Find an example of a unique factorization domain A and prime elements $p_1, \dots, p_n$ of A such that $I = (p_1, \dots, p_n)$ is **not** a prime ideal.

(3) Show that the ideal $I = (2, 1+\sqrt{-5})$ in $\mathbb{Z}[\sqrt{-5}]$ is prime and that it does not contain any prime element.

Exercise 1.29. Let A be an integral domain and (a) a nonzero principal ideal of A. A **factorization of** (a) **into principal prime ideals** is an equality of the form $(a) = \prod_{i=1}^{n}(p_i)$ where (p_i) are principal prime ideals of A.

(1) Show that a factorization in principal prime ideals is unique, i.e. if $(a) = \prod_{i=1}^{n}(p_i)$ and $(a) = \prod_{j=1}^{m}(q_j)$ are two such factorizations, then there exists a bijection $\sigma : \{1, \dots, n\} \to \{1, \dots, m\}$ such that $(p_i) = (q_{\sigma(i)})$ for all $i = 1, \dots, n$.

(2) Show that A is a unique factorization domain if and only if every principal ideal of A has a factorization into principal prime ideals.

Exercise 1.30. Let A be a ring. The **spectrum of** A is the set $\mathrm{Spec}A$ of all prime ideals of A. A **principal open subset of** $\mathrm{Spec}A$ is a subset of the form

$$U_a \;=\; U_{A,a} \;=\; \{\, \mathfrak{p} \in \mathrm{Spec}A \,|\, a \notin \mathfrak{p} \,\}$$

with $a \in A$.

(1) Show that $U_0 = \varnothing$, $U_1 = \mathrm{Spec}A$ and $U_a \cap U_b = U_{ab}$ for all $a, b \in A$.
Remark: This shows that the principal open subsets of $\mathrm{Spec}A$ form a basis for a topology on $\mathrm{Spec}A$, which is called the **Zariski topology**.

(2) Let $f : A \to B$ be a ring homomorphism. By Exercise 1.13, the association $\mathfrak{p} \mapsto f^{-1}(\mathfrak{p})$ defines a map $\varphi : \mathrm{Spec}B \to \mathrm{Spec}A$. Show that $\varphi^{-1}(U_{A,a}) = U_{B,f(a)}$ for every $a \in A$.
Remark: This shows that the map $\varphi : \mathrm{Spec}B \to \mathrm{Spec}A$ is a continuous map.

(3) Describe the spectrum of the following rings: a field K, the integers $\mathbb{Z}$, and their quotient $\mathbb{Z}/6\mathbb{Z}$. Describe the maps of spectra that are induced by the inclusion $\mathbb{Z} \to \mathbb{Q}$ and the surjection $\mathbb{Z} \to \mathbb{Z}/6\mathbb{Z}$.

*Exercise 1.31. Study the map $\varphi : \mathrm{Spec}\,\mathbb{Z}[i] \to \mathrm{Spec}\,\mathbb{Z}$ of spectra that is induced by the inclusion $\mathbb{Z} \to \mathbb{Z}[i]$ of $\mathbb{Z}$ into the Gaussian integers.

(1) Show that if $\varphi(\mathfrak{q}) = \langle p \rangle$ for a prime number $p \in \mathbb{Z}$, then $\mathfrak{q}$ is a maximal ideal of $\mathbb{Z}[i]$ that is generated by a single prime element of $\mathbb{Z}[i]$.

(2) Show that the fibres $\varphi^{-1}(\mathfrak{p})$ have either one or two elements for each fibre, and show that both cases occur.

Remark: If $\varphi^{-1}(\mathfrak{p})$ has two elements, then we say that $\mathfrak{p}$ **splits in the extension** $\mathbb{Z} \to \mathbb{Z}[i]$. *Hint:* A useful tool is the Euclidean norm $N(a+ib) = a^2 + b^2$, which has some convenient properties (e.g. it is multiplicative and its fibres are finite).

(3) Show that if $\mathfrak{p} = \langle p \rangle$ for some prime number $p \in \mathbb{Z}$ and if $\varphi^{-1}(\mathfrak{p}) = \{\mathfrak{q}\}$ consists of only one prime ideal $\mathfrak{q}$, then $\mathbb{Z}[i]/\mathfrak{q}$ is a field with p or p^2 elements. Show that both cases occur.

Remark: If $\mathbb{Z}[i]/\mathfrak{q}$ has p^2 elements, then we say that $\mathfrak{p}$ **is inert in the extension** $\mathbb{Z} \to \mathbb{Z}[i]$.

(4) Show that if $\mathfrak{p} = \langle p \rangle$ for some prime number $p \in \mathbb{Z}$ and if $\varphi^{-1}(\mathfrak{p}) = \{\mathfrak{q}\}$ such that $\mathbb{Z}[i]/\mathfrak{q} \simeq \mathbb{Z}/\mathfrak{p}$, then $\mathfrak{q}$ is generated by an element $q \in \mathbb{Z}[i]$ such that $q^2 = p$.

Remark: In this case, we say that $\mathfrak{p}$ **ramifies in the extension** $\mathbb{Z} \to \mathbb{Z}[i]$.

Exercise 1.32. Proof Lemma 1.8.4.

Exercise 1.33. Proof Lemma 1.9.2.

Exercise 1.34. Let A be a ring and S a multiplicative subset. Show the following assertions.

(1) The localization map $A \to S^{-1}A$ is injective if and only if for every $a \in S$, the multiplication $m_a : A \to A$ by a is an injective map.

(2) If A is an integral domain, a unique factorization domain, a principal ideal domain, a Euclidean domain or a field and $0 \notin S$, then $S^{-1}A$ is so, too.

(3) Let $A = A_1 \times A_2$ and $h = (1,0)$. Show that the association $\frac{(a,b)}{h^i} \mapsto (a,0)$ defines a ring isomorphism $A[h^{-1}] \simeq A_1$.

(4) Find an example of a local ring A and a multiplicative subset S with $0 \notin S$ such that $S^{-1}A$ is not local.

Exercise 1.35. Let A be a ring.

(1) Show that $A[T_1, T_2] \simeq (A[T_1])[T_2]$.

(2) Let $h \in A$. Show that $A[h^{-1}] \simeq A[T]/\langle hT - 1 \rangle$.

Exercise 1.36. Let A be a ring, S a multiplicative subset of A and $\iota_S : A \to S^{-1}A$ the localization map. Show the following.

(1) Given an ideal I of A, show that the ideal of $S^{-1}A$ generated by $\iota_S(I)$ equals

$$I \cdot S^{-1}A \;=\; \left\{ \tfrac{a}{s} \in S^{-1}A \,\middle|\, a \in I, s \in S \right\},$$

and that $I \cdot S^{-1}A$ is prime if I is prime and does not intersect S.

(2) Show that for every prime ideal $\mathfrak{q}$ of $S^{-1}A$, the inverse image $\iota_S^{-1}(\mathfrak{q})$ is a prime ideal of A that does not intersect S.

(3) Show that this defines mutually inverse bijections

$$\begin{array}{ccc} \{\text{prime ideals } \mathfrak{p} \text{ of } A \text{ with } \mathfrak{p} \cap S = \varnothing\} & \longleftrightarrow & \{\text{prime ideals of } S^{-1}A\} \\ \mathfrak{p} & \longmapsto & \mathfrak{p} \cdot S^{-1}A \\ \iota_S^{-1}(\mathfrak{q}) & \longmapsfrom & \mathfrak{q} \end{array}$$

Exercise 1.37. Let A be a ring and S a multiplicative subset. Show that the localization map $\iota_S : A \to S^{-1}A$ defines an injection $\varphi : \mathrm{Spec}A[h^{-1}] \to \mathrm{Spec}A$ that satisfies $\varphi(U_{A[h^{-1}],\frac{a}{s}}) = U_{A,ah}$ for every $a \in A$ and $s = h^i$ with $i \geqslant 0$.

Remark: This shows that $\varphi : \mathrm{Spec}(A[h^{-1}]) \to \mathrm{Spec}A$ is an open topological embedding with image U_h.

Exercise 1.38. Let A be a unique factorization domain. Show that the set G of principal fractional ideals is an abelian group with respect to their product. What is the neutral element of this group? What is the inverse of $\langle a \rangle_A$ where $a \in K^\times$? Show that the association $a \mapsto \langle a \rangle_A$ define a surjective group homomorphism $K^\times \to G$ with kernel $A^\times$.

Exercise 1.39. Let K be a field and $f \in K[T]$ a polynomial.

(1) Show for $\deg f = 2$ and $\deg f = 3$ that f is irreducible in $K[T]$ if and only if f does not have a root in K.

(2) Find a field K and a polynomial $f \in K[T]$ of degree 4 that is not irreducible and does not have a root in K.

(3) Show that there exists a field extension L/K such that f factorizes in $L[T]$ as

$$f = u\prod_{i=1}^{n}(T - a_i)$$

with $u, a_1, \ldots, a_n \in L$.

Exercise 1.40. Let A be a ring and let $n\mathbb{Z}$ be the kernel of the unique ring homomorphism $\mathbb{Z} \to A$ where $n \geqslant 0$. The number $\mathrm{char}\, A = n$ is called the **characteristic of** A.

(1) Show that if n is positive, then n is the smallest positive integer such that

$$n \cdot 1 \;=\; \underbrace{1 + \cdots + 1}_{n-\text{times}} \;=\; 0.$$

If $n = 0$, then $k \cdot 1 \neq 0$ for any $k \geqslant 0$.

(2) Show that n is zero or a prime number if A is an integral domain.

(3) Let L/K be a field extension. Show that K and L have the same characteristic.

(4) Let K be a field of characteristic 0. Show that there is a unique ring homomorphism $\mathbb{Q} \to K$.

(5) Let p be a prime number and $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ the field with p elements. Let K be a field of characteristic p. Show that there is a unique ring homomorphism $\mathbb{F}_p \to K$.

(6) Give an example of a ring homomorphism $A \to B$ where A and B have different characteristics.

Remark: The image of the unique homomorphism $\mathbb{Q} \to K$ (if $\mathrm{char}\, K = 0$) or $\mathbb{F}_p \to K$ (if $\mathrm{char}\, K = p > 0$) is called the **prime field of** K.

Exercise 1.41. Let $\mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z}$ be the field with two elements 0 and 1.

(1) Show that $f = T^2 + T + 1$ is an irreducible polynomial in $\mathbb{F}_2[T]$.

(2) Show that $\mathbb{F}_4 = \mathbb{F}_2[T]/(f)$ is a field with four elements.

(3) Show that $\mathbb{F}_4^\times$ is a cyclic group with 3 elements.

(4) Show that $T^4 - T = \prod_{a \in \mathbb{F}_4}(T - a)$ (as a polynomial in $\mathbb{F}_4[T]$).

(5) Find a factorization of $T^4 - T$ in $\mathbb{F}_2[T]$.

Exercise 1.42. Let G be an abelian group with n elements. We define the **exponent of** G as the smallest positive integer m such that $g^m = e$ for all $g \in G$.

(1) Show that G is cyclic if and only if its exponent is n.

(2) Let K be a field and U a finite subgroup of order n of the multiplicative group $K^\times$ of K. Show that U is cyclic.
Hint: If m is the exponent of U, then every element of U is a zero of $T^m - 1$.

Exercise 1.43. (1) Show that all irreducible polynomials in $\mathbb{R}[T]$ are of degree 1 or 2.

(2) Define two complex numbers z and z' as equivalent if $z' = z$ or $z' = \bar{z}$, the complex conjugate of z. Denote the corresponding equivalence relation by $\sim$ and the class of z in the quotient set $\mathbb{C}/\sim$ by $[z]$. Show that the map

$$\begin{array}{ccc} \mathbb{C}/\sim & \longrightarrow & \{\text{maximal ideals of } \mathbb{R}[T]\} \\ [z] & \longmapsto & (\prod_{z' \in [z]} (T - z')) \end{array}$$

is a bijection.

(3) Describe $\operatorname{Spec}\mathbb{C}[T]$, assuming the fundamental theorem of algebra (Exercise 1.44).

(4) Make a drawing of $\operatorname{Spec}\mathbb{R}[T]$ and of the map $f^* : \operatorname{Spec}\mathbb{C}[T] \to \operatorname{Spec}\mathbb{R}[T]$ that is induced by the inclusion $f : \mathbb{R}[T] \to \mathbb{C}[T]$.

Exercise 1.44.* Prove the **fundamental theorem of algebra: given a polynomial $f \in \mathbb{C}[T]$ of positive degree, then there exists a $z \in \mathbb{C}$ such that $f(z) = 0$.

Exercise 1.45. Let A be a principal ideal domain with only one prime element p (up to associates).

(1) Show that every element $a \in A - \{0\}$ can be written in the form $a = up^n$ for uniquely determined $u \in A^\times$ and $n \geqslant 0$.

(2) Show that A has a unique maximal ideal $\mathfrak{m}$.

(3) Show that every other ideal of A is either equal to (0) or (1) or of the form $\mathfrak{m}^i = \underbrace{\mathfrak{m} \cdots \mathfrak{m}}_{i\text{-times}}$ for some $i \geqslant 1$.

(4) The intersection of all $\mathfrak{m}^i$ is $\bigcap_{i \in \mathbb{N}} \mathfrak{m}^i = \{0\}$.

(5) Show that A is a Euclidean ring.

(6) Give an example of a ring with these properties (including a proof).

Remark: A ing with these properties is called a *discrete valuation ring*.

Chapter 2

Categories

2.1 Classes

Category theory relies on the notion of a **class**, which is a collection of "mathematical objects" that might be "too large" to be a set. But given a class C, we can say for every mathematical object A whether $A \in C$ or $A \notin C$.

The reason to introduce classes beyond sets is **Russell's paradox**:

> Let C be the class of all sets A with $A \notin A$. If C was a set, then $C \in C$ if and only if $C \notin C$, which is absurd.

It would lead, in particular, to a contradictory concept if we would attempt to consider the "set of all sets", which cannot exist for the alluded paradox—provided we assume axioms for set theory that prevent that a set can contain itself.

We avoid a digression into foundations of set theory and hope that the reader finds the description of a class as a collection of mathematical objects sufficiently intuitive to continue reading.

2.2 Categories

Definition 2.2.1. A **category** $\mathcal{C}$ consists of a class $\mathrm{Ob}(\mathcal{C})$ of objects, a morphism set $\mathrm{Hom}_{\mathcal{C}}(A,B)$ for every pair of objects $A,B \in \mathrm{Ob}(\mathcal{C})$ and a composition law

$$\begin{array}{rccc} \circ: & \mathrm{Hom}_{\mathcal{C}}(A,B) \times \mathrm{Hom}_{\mathcal{C}}(B,C) & \longrightarrow & \mathrm{Hom}_{\mathcal{C}}(A,C) \\ & (\alpha : A \to B, \beta : B \to C) & \longmapsto & \beta \circ \alpha : A \to C \end{array}$$

for any three objects $A,B,C \in \mathrm{Ob}(\mathcal{C})$ such that the following axioms hold:

(1) for every object $A \in \mathrm{Ob}(\mathcal{C})$, there is a morphism $\mathrm{id}_A : A \to A$ such that for all objects $B,C \in \mathrm{Ob}(\mathcal{C})$ and all morphisms $\alpha : A \to B$ in $\mathrm{Hom}_{\mathcal{C}}(A,B)$ and $\gamma : C \to A$ in $\mathrm{Hom}_{\mathcal{C}}(C,A)$, we have $\alpha \circ \mathrm{id}_A = \alpha$ and $\mathrm{id}_A \circ \gamma = \gamma$; *(identity)*

(2) for all objects $A,B,C,D \in \mathrm{Ob}(\mathcal{C})$ and all morphisms $\alpha : A \to B$ in $\mathrm{Hom}_{\mathcal{C}}(A,B)$, $\beta : B \to C$ in $\mathrm{Hom}_{\mathcal{C}}(B,C)$ and $\gamma : C \to D$ in $\mathrm{Hom}_{\mathcal{C}}(C,D)$, we have $\gamma \circ (\beta \circ \alpha) = (\gamma \circ \beta) \circ \alpha$. *(associativity)*

Notation. We say that A *is an object in* $\mathcal{C}$ if $A \in \mathrm{Ob}(\mathcal{C})$ and that $\alpha : A \to B$ *is a morphism in* $\mathcal{C}$ if $\alpha \in \mathrm{Hom}(A,B)$. We write $\mathrm{Hom}(A,B)$ for $\mathrm{Hom}_{\mathcal{C}}(A,B)$ if the category $\mathcal{C}$ is clear from the context.

The morphism $\mathrm{id}_A : A \to A$ from axiom (1) is called the *identity morphism of* A. Identity morphisms are uniquely determined by axiom (1): if $\iota : A \to A$ is a morphism such that $\alpha \circ \iota = \alpha$ for all morphisms $\alpha : A \to B$, then $\iota = \mathrm{id}_A \circ \iota = \mathrm{id}_A$.

Example 2.2.2. We are already familiar with many examples of categories, such as some of the following. In most of these examples, the objects are sets with some additional structure, morphisms are maps with some additional properties and the composition law is the usual composition of maps. Therefore the identity map $\mathrm{id}_A : A \to A$ that maps $a \in A$ to $\mathrm{id}_A(a) = a$ satisfies (1). The associativity law (2) follows from an elementwise evaluation of morphisms $\alpha : A \to B$, $\beta : B \to C$ and $\gamma : C \to D$ in elements $a \in A$:

$$\gamma \circ (\beta \circ \alpha)(a) \;=\; \gamma\big(\beta \circ \alpha(a)\big) \;=\; \gamma\big(\beta\big(\alpha(a)\big)\big) \;=\; \gamma \circ \beta\big(\alpha(a)\big) \;=\; (\gamma \circ \beta) \circ \alpha(a).$$

This reasons that the axioms (1) and (2) are satisfied in most of the following examples. The other cases are left as an exercise.

(0) The **trivial category** $\mathcal{C}$: $\mathrm{Ob}(\mathcal{C}) = \{A\}$ and $\mathrm{Hom}(A,A) = \{\mathrm{id}_A : A \to A\}$, with the tautological composition $\mathrm{id}_A \circ \mathrm{id}_A = \mathrm{id}_A$.

(1) The **category** Sets **of sets**: $\mathrm{Ob}(\mathrm{Sets})$ is the class of all sets; $\mathrm{Hom}(A,B)$ consist of all maps $\alpha : A \to B$ for every pair of sets A and B; the composition law $\circ$ is the usual composition of maps.

(2) The **category** Ab **of abelian groups**: $\mathrm{Ob}(\mathrm{Ab})$ is the class of all abelian groups; $\mathrm{Hom}(A,B)$ consists of all group homomorphisms $\alpha : A \to B$ for every pair of abelian groups A and B; the composition law $\circ$ is the usual composition of group homomorphisms.

(3) The **category** Rings **of rings**: $\mathrm{Ob}(\mathrm{Rings})$ is the class of all rings; $\mathrm{Hom}(A,B)$ consists of all ring homomorphisms $\alpha : A \to B$ for every pair of rings A and B; the composition law $\circ$ is the usual composition of ring homomorphisms.

(4) The **category** Fields **of fields**: $\mathrm{Ob}(\mathrm{Fields})$ is the class of all fields; $\mathrm{Hom}(A,B)$ consists of all ring homomorphisms $\alpha : A \to B$ for every pair of fields A and B; the composition law $\circ$ is the usual composition of ring homomorphisms.

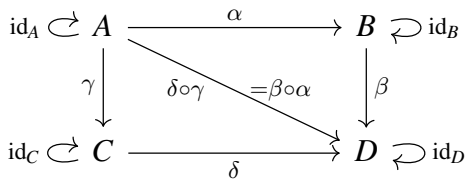
(5) The **category** Top **of topological spaces**: $\mathrm{Ob}(\mathrm{Top})$ is the class of all topological spaces; $\mathrm{Hom}(A,B)$ consists of all continuous maps $\alpha : A \to B$ for every pair of topological spaces A and B; the composition law $\circ$ is the usual composition of (continuous) maps; see section A.2 for definitions.

(6) The **category** Vect_K **of vector spaces over a field** K: $\mathrm{Ob}(\mathrm{Vect}_K)$ is the class of all K-vector spaces; $\mathrm{Hom}(A,B)$ consists of all K-linear maps $\alpha : A \to B$ for every pair of K-vector spaces A and B; the composition law $\circ$ is the usual composition of K-linear maps.

(7) The **category** Alg_A **of algebra over a ring** A: $\mathrm{Ob}(\mathrm{Alg}_A)$ is the class of all rings B together with a ring homomorphism $\iota_B : A \to B$ called the *structure map of* B; $\mathrm{Hom}(B,C)$ consists of all ring homomorphisms $\beta : B \to C$ such that $\iota_C = \beta \circ \iota_B$ for every pair of A-algebras B and C (where we suppress the structure maps $\iota_B : A \to B$ and $\iota_C : A \to C$ from the notation); the composition law $\circ$ is the usual composition of ring homomorphisms.

(8) The **opposite category** $\mathcal{C}^{\mathbf{op}}$ **of a category** $\mathcal{C}$: $\mathrm{Ob}(\mathcal{C}^{\mathrm{op}}) = \mathrm{Ob}(\mathcal{C})$ and $\mathrm{Hom}_{\mathcal{C}^{\mathrm{op}}}(B,A) = \mathrm{Hom}_{\mathcal{C}}(A,B)$ for all objects A and B of $\mathcal{C}$ where we denote the morphism in $\mathcal{C}^{\mathrm{op}}$ that corresponds to a morphism $\alpha : A \to B$ in $\mathcal{C}$ by $\alpha^{\mathrm{op}} : B \to A$. The composition is defined by $\alpha^{\mathrm{op}} \circ \beta^{\mathrm{op}} = (\beta \circ \alpha)^{\mathrm{op}}$ for $\alpha : A \to B$ and $\beta : B \to C$.

(9) One can visualize finite categories as graphs, for example:



(10) Categories of correspondences: a simple instance is the category $\mathcal{C}$ for which $\mathrm{Ob}(\mathcal{C})$ is the class of all sets and $\mathrm{Hom}(A,B)$ is the collection of all subsets S of $A \times B$, which we call *relations* or *correspondences*, and write $\alpha_S : A \to B$ for a subset S of $A \times B$. Given two correspondences $\alpha_S : A \to B$ and $\alpha_T : B \to C$, we define the composition $\alpha_T \circ \alpha_S : A \to C$ as the subset

$$\big\{(a,c) \in A \times C \,\big|\, (a,b) \in S \text{ and } (b,c) \in T \text{ for some } b \in B\big\}$$

of $A \times C$. We leave it as an exercise to verify that the subset $\{(a,a) \mid a \in A\}$ of $A \times A$ is the identity morphism $\mathrm{id}_A : A \to A$ and that the composition of correspondences is associative.

2.3 Monomorphisms, epimorphisms and isomorphisms

We have seen already examples of categories whose morphisms are not maps. In so far, it does not make sense to ask whether morphisms in an abstract category are injective, surjective or bijective. There is, however, a "categorical" characterization of injective, surjective and bijective maps, seen as morphisms in Sets, which is as follows. The precise relation is exhibited in Lemma 2.3.3.

Definition 2.3.1. Let $\mathcal{C}$ be a category and $\alpha : A \to B$ a morphism in $\mathcal{C}$. Then f is

- a **monomorphism** if $\alpha \circ \beta = \alpha \circ \gamma$ implies $\beta = \gamma$ for all $\beta : C \to A$ and $\gamma : C \to A$;
- an **epimorphism** if $\beta \circ \alpha = \gamma \circ \alpha$ implies $\beta = \gamma$ for all $\beta : B \to C$ and $\gamma : B \to C$;

- an **isomorphism** if there is a morphism $\beta : B \to A$ such that $\beta \circ \alpha = \mathrm{id}_A$ and $\alpha \circ \beta = \mathrm{id}_B$; we call β the **inverse of** α and denote it by α^{-1} if it exists.

Lemma 2.3.2. *Let $\mathcal{C}$ be a category.*

(1) Identity morphisms are isomorphisms.

(2) The inverse of an isomorphism is uniquely determined.

(3) Every isomorphism is both a monomorphism and an epimorphism.

Proof. Since $\mathrm{id}_A \circ \mathrm{id}_A = \mathrm{id}_A$, the identity morphism id_A is an isomorphism with inverse id_A. Thus (1).

Let $\alpha : A \to B$ be an isomorphism with inverse $\alpha^{-1} : B \to A$ and $\beta : B \to A$ a morphism such that $\beta \circ \alpha = \mathrm{id}_A$ and $\alpha \circ \beta = \mathrm{id}_B$. Then $\beta = \mathrm{id}_A \circ \beta = \alpha^{-1} \circ \alpha \circ \beta = \alpha^{-1}$, which shows (2).

Consider an isomorphism $\alpha : A \to B$ and morphisms $\beta : C \to A$ and $\gamma : C \to A$. Then $\beta = \alpha^{-1} \circ \alpha \circ \beta = \alpha^{-1} \circ \alpha \circ \gamma = \gamma$, which shows that α is a monomorphism. Similarly, we have for morphisms $\beta' : B \to C$ and $\gamma' : B \to C$ that $\beta' = \beta' \circ \alpha \circ \alpha^{-1} = \gamma' \circ \alpha \circ \alpha^{-1} = \gamma'$, which shows that α is an epimorphism. Thus (3), which completes the proof. □

Lemma 2.3.3. *Let $\alpha : A \to B$ be a map between sets. Considered as a morphism in* Sets*,*

(1) α is a monomorphism if and only if α is injective;

(2) α is an epimorphism if and only if α is surjective;

(3) α is an isomorphism if and only if α is bijective.

Proof. We begin with (1). Assume α is a monomorphism and consider $a, b \in A$ with $\alpha(a) = \alpha(b)$. We define two maps $\beta : \{c\} \to A$ and $\gamma : \{c\} \to A$ with $\beta(c) = a$ and $\gamma(c) = b$. Since $\alpha \circ \beta(c) = \alpha(a) = \alpha(b) = \alpha \circ \beta(c)$, we conclude that $\beta = \gamma$ since α is a monomorphism. Thus $a = b$, which shows that α is injective.

Conversely assume that α is injective and consider two maps $\beta : C \to A$ and $\gamma : C \to A$ with $\alpha \circ \beta = \alpha \circ \gamma$, i.e. $\alpha\big(\beta(c)\big) = \alpha\big(\gamma(c)\big)$ for all $c \in C$. Since f is injective, this implies that $\beta(c) = \gamma(c)$ for all $c \in C$, i.e. $\beta = \gamma$. Thus α is a monomorphism, which establishes (1).

We continue with (2). Assume α is an epimorphism and consider $b \in B$. Let $C = B \cup \{b'\}$ for some element $b' \notin B$. We define $\beta : B \to C$ and $\gamma : B \to C$ by $\beta(a) = \gamma(a) = a$ for all $a \in B - \{b\}$, $\beta(b) = b$ and $\gamma(b) = b'$. Then $\beta \neq \gamma$ and thus $\alpha \circ \beta \neq \alpha \circ \gamma$ since α is an epimorphism. Thus b must be in the image of α, which shows that α is surjective.

Conversely, assume that α is surjective and consider $\beta : B \to C$ and $\gamma : B \to C$ with $\beta \circ \alpha = \gamma \circ \alpha$, i.e. $\beta\big(\alpha(a)\big) = \gamma\big(\alpha(a)\big)$ for all $a \in A$. Since α is surjective, every $b \in B$ is of the form $b = \alpha(a)$ and thus $\beta(b) = \gamma(b)$ for all $b \in B$, i.e. $\beta = \gamma$. This shows that α is an epimorphism, which establishes (2).

We continue with (3). Assume that α is an isomorphism. By Lemma 2.3.2, α is both a monomorphism and an epimorphism. By (1) and (2), α is both injective and surjective, and therefore bijective.

Conversely assume that α is bijective. We define $\beta : B \to A$ by $\beta(\alpha(a)) = a$ which is well-defined since α is bijective. Then we have for all $a \in A$ and $b = \alpha(a)$ that $\beta \circ \alpha(a) = \beta(\alpha(a)) = a$ and $\alpha \circ \beta(b) = \alpha\big(\beta(b)\big) = b$. Thus β is the inverse of α, which shows that α is an isomorphism. □

Remark. We conclude that in Sets, every morphism that is both a monomorphism and an epimorphism, is an isomorphism. This is not true for all categories. For example, localizations of rings are epimorphisms in Rings, but not surjective in general. More specifically, the inclusion $\mathbb{Z} \to \mathbb{Q}$ is both a monomorphism and an epimorphism, but not an isomorphism, cf. Exercise 2.1.

The characterization of monomorphisms as injections, epimorphisms as surjections and isomorphisms as bijections holds in the categories Vect_K of K-vector spaces and Ab of abelian groups. In Rings monomorphisms coincide with injections, isomorphisms coincide with bijections, and surjections are epimorphisms, but not every epimorphism is surjective.

2.4 Initial and terminal objects, products and coproducts

Definition 2.4.1. Let $\mathcal{C}$ be a category. An object A in $\mathcal{C}$ is **initial** if for every object B in $\mathcal{C}$, there is a unique morphism $A \to B$. An object A in $\mathcal{C}$ is **terminal** if for every object B in $\mathcal{C}$, there is a unique morphism $B \to A$.

Remark. Initial and terminal objects do not have to exist in a category, but if they do, then they are unique up to a unique isomorphism. Indeed, if A and B are both initial or both terminal objects, then there are unique morphisms $\alpha : A \to B$ and $\beta : B \to A$ whose compositions $\beta \circ \alpha$ and $\alpha \circ \beta$ must be equal to the unique morphisms $A \to A$ and $B \to B$, which are the respective identities of A and B. Thus α is the unique isomorphism from A to B.

Example 2.4.2. We list initial and terminal objects for some categories.

(1) An initial object in Sets is the empty set $\varnothing$. A terminal object in Sets is a set $\{a\}$ with one element a. This makes clear that the terminal object is not general unique, but only unique up to unique isomorphism.

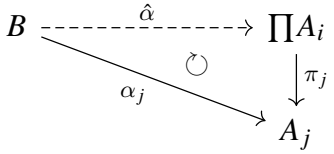
(2) Let K be a field. The trivial K-vector space $\{0\}$ is both initial and terminal in Vect_K. Similarly, the trivial group $\{e\}$ is both initial and terminal in Ab.

(3) An initial object in Rings is the ring of integers $\mathbb{Z}$, cf. Exercise 1.22. A terminal object in Rings is the trivial ring $\{0\}$.

(4) The category of fields has neither initial nor terminal objects since there are no morphisms between fields of different characteristics.

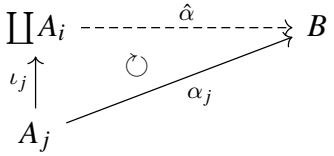
Definition 2.4.3. Let $\mathcal{C}$ be a category and $\{A_i\}_{i \in I}$ a family of objects in $\mathcal{C}$. A **product of** $\{A_i\}$ is an object $\prod_{i \in I} A_i$ in $\mathcal{C}$ together with a family $\{\pi_j : \prod A_i \to A_j\}_{j \in I}$ of morphisms

that satisfies the following universal property: for every object B and for every family of morphisms $\{\alpha_j : B \to A_j\}_{j\in I}$, there exists a unique morphism $\hat{\alpha} : B \to \prod A_i$ such that $\alpha_j = \pi_j \circ \hat{\alpha}$ for all $j \in I$, i.e. the diagram



commutes for every $j \in I$.

A **coproduct of** $\{A_i\}$ is an object $\coprod_{i\in I} A_i$ in $\mathcal{C}$ together with a family $\{\iota_j : A_j \to \coprod A_i\}_{j\in I}$ of morphisms that satisfies the following universal property: for every object B and for every family of morphisms $\{\alpha_j : A_j \to B\}_{j\in I}$, there exists a unique morphism $\hat{\alpha} : \coprod A_i \to B$ such that $\alpha_j = \hat{\alpha} \circ \iota_j$ for all $j \in I$, i.e. the diagram

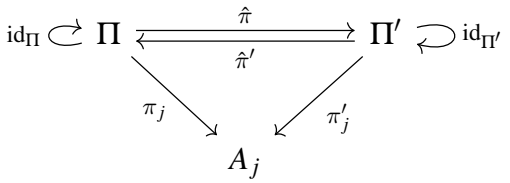


commutes for every $j \in I$.

Notation. The morphisms $\pi_j : \prod A_i \to A_j$ are called the *canonical projections* of the product $\prod A_i$ and the morphisms $\iota_j : A_j \to \coprod A_i$ are called the *canonical inclusions* of the coproduct $\coprod A_i$. Note that, in spite of its name, the canonical projections are in general neither surjective (as far as this makes sense) nor epimorphisms, and similarly, the canonical injections are neither injective (as far as this makes sense) nor monomorphisms.

Remark. Note that products and coproducts are unique up to a unique isomorphism that commutes with the canonical projections and inclusions, respectively. We verify this claim for products in the following, and leave the analogous case as an exercise.

Given two objects Π and Π' together with families $\{\pi_j : \Pi \to A_j\}_{j\in I}$ and $\{\pi'_j : \Pi' \to A_j\}_{j\in I}$ that satisfy each the universal property of the product, then the π'_j induce a unique morphism $\hat{\pi} : \Pi \to \Pi'$ such that $\pi_j = \pi'_j \circ \hat{\pi}$ for all $j \in I$, and the π_j induce a unique morphism $\hat{\pi}' : \Pi' \to \Pi$ such that $\pi'_j = \pi_j \circ \hat{\pi}'$ for all $j \in I$. Thus $\pi_j = \pi_j \circ \hat{\pi}' \circ \hat{\pi}$ and $\pi'_j = \pi'_j \circ \hat{\pi} \circ \hat{\pi}'$ for all $j \in I$. Since $\mathrm{id}_\Pi : \Pi \to \Pi$ is the unique morphism with $\pi_j = \pi_j \circ \mathrm{id}_\Pi$, we conclude that $\hat{\pi}' \circ \hat{\pi} = \mathrm{id}_\Pi$, and similarly, $\hat{\pi} \circ \hat{\pi}' = \mathrm{id}_{\Pi'}$, which verifies our claim. We illustrate these morphisms in the following diagram:



Example 2.4.4. We describe for the constructions of products and coproducts for some categories.

(1) Let $\{A_i\}_{i\in I}$ be a family of sets. Its product in Sets is the Cartesian product $\prod A_i$ together with the coordinate projections $\pi_j : \prod A_i \to A_j$. Its coproduct is the disjoint union

$$\coprod_{i\in I} A_i \;=\; \big\{(i,a)\,\big|\, i\in I \text{ and } a\in A_i\big\}$$

together with the inclusions $\iota_j : A_j \to \coprod A_i$ that are defined by $\iota_j(a) = (j,a)$.

(2) Let K be a field and $\{A_i\}_{i\in I}$ a family of K-vector spaces. Its product in Vect_K is the Cartesian product $\prod A_i$ together with the coordinate projections $\pi_j : \prod A_i \to A_j$. Its coproduct is the direct sum

$$\bigoplus_{i\in I} A_i \;=\; \Big\{(a_i)_{i\in I} \in \prod_{i\in I} A_i \,\Big|\, a_i = 0 \text{ for all but finitely many } i\in I\Big\}$$

together with the coordinate inclusions $\iota_j : A_j \to \bigoplus A_i$, i.e. $a\in A_j$ is sent to the element $(a_i)\in\bigoplus A_i$ with $a_i = a$ for $i = j$ and $a_i = 0$ for $i\neq j$. Note that if I is finite, then $\bigoplus A_i = \prod A_i$.

(3) The product and coproduct of a family of abelian groups $\{A_i\}_{i\in I}$ in Ab is similarly constructed: its product is the Cartesian product $\prod A_i$ together with the coordinate projections $\pi_j : \prod A_i \to A_j$ and its coproduct is the direct sum $\bigoplus A_i$ together with the coordinate inclusions $\iota_j : A_j \to \bigoplus A_i$.

(4) Let $\{A_i\}$ be a family of rings. Its product in Rings is the Cartesian product $\prod A_i$ together with the coordinate projections $\pi_j : \prod A_i \to A_j$. Its product is the tensor product $\bigotimes A_i$ together with the canonical inclusions $\iota_j : A_j \to \bigotimes A_i$, cf. Corollary 3.3.7.

We summarize Remark 2.3 and Examples 2.4.2 and 2.4.4 in Table 2.1.

Category	mono.	epi.	isom.	initial	terminal	product	coproduct
Sets	inj.	surj.	bij.	$\varnothing$	$\{a\}$	$\prod A_i$	$\coprod A_i$
Vect_K	inj.	surj.	bij.	$\{0\}$	$\{0\}$	$\prod A_i$	$\bigoplus A_i$
Ab	inj.	surj.	bij.	$\{0\}$	$\{0\}$	$\prod A_i$	$\bigoplus A_i$
Rings	inj.	?	bij.	$\mathbb{Z}$	$\{0\}$	$\prod A_i$	$\bigotimes A_i$

Table 2.1: Characterizations of certain morphisms and objects for some categories

2.5 Functors

A functor is an association between categories, similar to morphisms between objects in a category itself. Functors allow to connect different categories and to study properties of a category in its objects in terms of other categories. There are two fundamental variants of functors: those that preserve the direction of morphisms (covariant functors) and those the reverse the direction (contravariant functors).

Definition 2.5.1. Let $\mathcal{C}$ and $\mathcal{D}$ be categories. A **covariant functor** $\mathcal{F}:\mathcal{C}\to\mathcal{D}$ consists of an assignment of an object $\mathcal{F}(A)$ in $\mathcal{D}$ to every object A in $\mathcal{C}$ and a map

$$\begin{array}{ccc}\mathrm{Hom}_{\mathcal{C}}(A,B) & \longrightarrow & \mathrm{Hom}_{\mathcal{D}}\big(\mathcal{F}(A),\mathcal{F}(B)\big)\\ \alpha:A\to B & \longmapsto & \mathcal{F}(\alpha):\mathcal{F}(A)\to\mathcal{F}(B)\end{array}$$

for every pair of objects A and B in $\mathcal{C}$ such that $\mathcal{F}(\mathrm{id}_A)=\mathrm{id}_{\mathcal{F}(A)}$ for all objects A in $\mathcal{C}$ and such that $\mathcal{F}(\beta\circ\alpha)=\mathcal{F}(\beta)\circ\mathcal{F}(\alpha)$ for all morphisms $\alpha:A\to B$ and $\beta:B\to C$ in $\mathcal{C}$.

A **contravariant functor** $\mathcal{F}:\mathcal{C}\to\mathcal{D}$ consists of an assignment of an object $\mathcal{F}(A)$ in $\mathcal{D}$ to every object A in $\mathcal{C}$ and a map

$$\begin{array}{ccc}\mathrm{Hom}_{\mathcal{C}}(A,B) & \longrightarrow & \mathrm{Hom}_{\mathcal{D}}\big(\mathcal{F}(B),\mathcal{F}(A)\big)\\ \alpha:A\to B & \longmapsto & \mathcal{F}(\alpha):\mathcal{F}(B)\to\mathcal{F}(A)\end{array}$$

for every pair of objects A and B in $\mathcal{C}$ such that $\mathcal{F}(\mathrm{id}_A)=\mathrm{id}_{\mathcal{F}(A)}$ for all objects A in $\mathcal{C}$ and such that $\mathcal{F}(\beta\circ\alpha)=\mathcal{F}(\alpha)\circ\mathcal{F}(\beta)$ for all morphisms $\alpha:A\to B$ and $\beta:B\to C$ in $\mathcal{C}$.

Example 2.5.2. We provide a list of examples of functors.

(0) Let $\mathcal{C}$ be a category. The **identity functor** $\mathbf{id}_{\mathcal{C}}:\mathcal{C}\to\mathcal{C}$ is the covariant functor with $\mathrm{id}_{\mathcal{C}}(A)=A$ and $\mathrm{id}_{\mathcal{C}}(\alpha)=\alpha$ for all objects A and morphisms α in $\mathcal{C}$. There is a contravariant functor $\mathcal{C}\to\mathcal{C}^{\mathrm{op}}$ to the opposite category of $\mathcal{C}$ that is the identity on objects and that maps a morphism $\alpha:A\to B$ to its opposite morphism $\alpha^{\mathrm{op}}:B^{\mathrm{op}}\to A^{\mathrm{op}}$.

(1) Taking the unit groups of a ring defines a functor $(-)^\times:\mathrm{Rings}\to\mathrm{Ab}$ that assigns to each ring A its unit group $A^\times$ and that maps a ring homomorphism $\alpha:A\to B$ to its restriction $\alpha^\times:A^\times\to B^\times$, which is a group homomorphism.

(2) Let K be a field. Passing to the dual K-vector space defines a contravariant functor $(-)^*:\mathrm{Vect}_K\to\mathrm{Vect}_K$ that assigns to a K-vector space A its **dual space** $A^*=\mathrm{Hom}_K(A,K)$ and that maps a K-linear map $\alpha:A\to B$ to its **dual** (or **transpose**) $\alpha^*:B^*\to A^*$ that sends a K-linear map $\beta:B\to K$ to the composition $\alpha^*(\beta)=\beta\circ\alpha:A\to K$.

(3) For a category $\mathcal{C}$ whose objects are sets with additional structure, whose morphisms are maps that satisfying certain conditions and whose composition law is the composition of maps, there is a **forgetful functor** $\mathcal{F}:\mathcal{C}\to\mathrm{Sets}$ that is the covariant functor that assigns to each object A in $\mathcal{C}$ its underlying set $\underline{A}$ and to its morphism $\alpha:A\to B$ in $\mathcal{C}$ the underlying map $\underline{\alpha}:\underline{A}\to\underline{B}$.

This applies, in particular, to the categories Vect_K, Ab and Rings. For example, the forgetful functor $\mathcal{F}:\mathrm{Ab}\to\mathrm{Sets}$ assigns to each abelian group G its underlying set $\underline{G}=\{a\in G\}$ and to each group homomorphism $\alpha:G\to H$ the underlying map $a\mapsto\alpha(a)$ from $\underline{G}$ to $\underline{H}$.

(4) There are also **forgetful functors** into other categories. For instance for every ring A, there is a forgetful functor $\mathcal{F}:\mathrm{Alg}_A\to\mathrm{Rings}$ that sends an A-algebra B with structure map $\iota_B:A\to B$ to the ring B and an algebra morphism $\beta:B\to C$ to itself, forgetting that it commutes with the structure maps of B and C. Note that

since for every ring B, there is a unique morphism $\mathbb{Z} \to B$, the forgetful functor $\mathcal{F} : \mathrm{Alg}_{\mathbb{Z}} \to \mathrm{Rings}$ identifies the two categories.

(5) For many categories $\mathcal{C}$, there are **free functors** $\mathcal{F} : \mathrm{Sets} \to \mathcal{C}$. We explain this in the examples of Ab and Alg_A.

The free functor $\mathcal{F} : \mathrm{Sets} \to \mathrm{Ab}$ is the covariant functor that assigns to each set A the free abelian group $\mathcal{F}(A) = \bigoplus_{a \in A} \mathbb{Z}$ and to each map $\alpha : A \to B$ the group homomorphism $\mathcal{F}(\alpha) : \mathcal{F}(A) \to \mathcal{F}(B)$ that sends an elements $(n_a) \in \bigoplus_{a \in A} \mathbb{Z}$ to $(m_b) \in \bigoplus_{b \in B} \mathbb{Z}$ with $m_b = \sum_{\alpha(a)=b} n_a$, which is a finite sum since $n_a = 0$ for almost all $a \in A$.

The free functor $\mathcal{F} : \mathrm{Sets} \to \mathrm{Alg}_A$ is the covariant functor that assigns to each set B the polynomial ring $A[T_b | b \in B]$ together with the canonical inclusion $\iota_A : A \to A[T_b | b \in B]$ as the structure map, and that assigns to each map $\alpha : B \to C$ the ring homomorphism $\hat{\alpha} : A[T_b | b \in B] \to A[T_c | c \in C]$ that extends $\mathrm{id}_A : A \to A$ and that maps T_b to $T_{\alpha(b)}$ for every $b \in B$.

(6) Let $\mathcal{C}$ be a category and A an object in $\mathcal{C}$. Then $\mathrm{Hom}(A, -) : \mathcal{C} \to \mathrm{Sets}$ is the covariant functor that assigns to an object B in $\mathcal{C}$ the set $\mathrm{Hom}(A,B)$ of morphisms $\alpha : A \to B$ in $\mathcal{C}$ and that assigns to a morphism $\beta : B \to C$ the map $\beta_* : \mathrm{Hom}(A,B) \to \mathrm{Hom}(A,C)$ that sends $\alpha : A \to B$ to $\beta_*(\alpha) = \beta \circ \alpha : A \to C$. Similarly, $\mathrm{Hom}(-, A) : \mathcal{C} \to \mathrm{Sets}$ is the contravariant functor that assigns to an object B in $\mathcal{C}$ the set $\mathrm{Hom}(B,A)$ of morphisms $\alpha : B \to A$ in $\mathcal{C}$ and that assigns to a morphism $\beta : B \to C$ the map $\beta_* : \mathrm{Hom}(C,A) \to \mathrm{Hom}(B,A)$ that sends $\alpha : C \to A$ to $\beta^*(\alpha) = \alpha \circ \beta : B \to A$.

2.6 Adjoint functors

Definition 2.6.1. Let $\mathcal{C}$ and $\mathcal{D}$ be categories and $\mathcal{F} : \mathcal{C} \to \mathcal{D}$ and $\mathcal{G} : \mathcal{D} \to \mathcal{C}$ covariant functors. Then $\mathcal{F}$ is **left adjoint to** $\mathcal{G}$ and $\mathcal{G}$ is **right adjoint to** $\mathcal{F}$, which we denote as $\mathcal{F} \dashv \mathcal{G}$, if there is a bijection

$$\Psi_{A,B} : \quad \mathrm{Hom}_{\mathcal{C}}\big(A, \mathcal{G}(B)\big) \quad \longrightarrow \quad \mathrm{Hom}_{\mathcal{D}}\big(\mathcal{F}(A), B\big)$$

for every pair of an object A in $\mathcal{C}$ and an object B in $\mathcal{D}$ such that the diagram

$$\begin{array}{ccccc}
\gamma & \in & \mathrm{Hom}_{\mathcal{C}}\big(A', \mathcal{G}(B)\big) & \xrightarrow{\Psi_{A',B}} & \mathrm{Hom}_{\mathcal{D}}\big(\mathcal{F}(A'), B\big) \quad \ni \quad \delta \\
\big\downarrow & & \big\downarrow & & \big\downarrow \\
\mathcal{G}(\beta) \circ \gamma \circ \alpha & \in & \mathrm{Hom}_{\mathcal{C}}\big(A, \mathcal{G}(B')\big) & \xrightarrow{\Psi_{A,B'}} & \mathrm{Hom}_{\mathcal{D}}\big(\mathcal{F}(A), B'\big) \quad \ni \quad \beta \circ \delta \circ \mathcal{F}(\alpha)
\end{array}$$

commutes for every pair of a morphism $\alpha : A \to A'$ in $\mathcal{C}$ and a morphism $\beta : B \to B'$ in $\mathcal{D}$.

Example 2.6.2. We have already seen some examples of adjoint functors, which are the forgetful and free functors from Example 2.5.2. We explain this in detail in the following case.

Let A be a ring and $\mathcal{G} : \mathrm{Alg}_A \to \mathrm{Sets}$ the forgetful functor from Example 2.5.2.(4), which sends an A-algebra B to its underlying set $\underline{B}$ and a homomorphism $\alpha : B \to C$ of A-algebras to itself, considered as a map $\underline{\alpha} : \underline{B} \to \underline{C}$ between the respective underlying sets. Let $\mathcal{F} : \mathrm{Sets} \to \mathrm{Alg}_A$ be the free functor from Example 2.5.2.(5), which sends a set B to the polynomial ring $A[T_b | b \in B]$ and a map $\alpha : B \to C$ to the homomorphism $\bar{\alpha} : A[T_b | b \in B] \to A[T_c | c \in C]$ that maps T_b to $T_{\alpha(b)}$.

Let B be a set, C an A-algebra and $\alpha : B \to \underline{C}$ a map. The universal property of polynomial rings (Proposition 1.9.3) implies that there is a unique A-linear ring homomorphism $\hat{\alpha} : A[T_b | b \in B] \to C$ that sends T_b to $\alpha(b)$. This means that the association $\alpha \mapsto \hat{\alpha}$ defines a bijection

$$\Psi_{B,C} : \quad \mathrm{Hom}_{\mathrm{Sets}}\big(B, \underline{C}\big) \quad \longrightarrow \quad \mathrm{Hom}_A\big(A[T_b | b \in B], C\big)$$

for every set B and every A-algebra C.

Given a map $\beta : B \to B'$ and an A-linear ring homomorphism $\gamma : C \to C'$, we aim to show that the diagram

$$\begin{array}{ccc} \mathrm{Hom}_{\mathrm{Sets}}\big(B', \underline{C}\big) & \xrightarrow{\Psi_{B',C}} & \mathrm{Hom}_A\big(A[T_b | b \in B'], C\big) \\ \Big\downarrow{\scriptstyle \underline{\gamma} \circ - \circ \beta} & & \Big\downarrow{\scriptstyle \gamma \circ - \circ \bar{\beta}} \\ \mathrm{Hom}_{\mathrm{Sets}}\big(B, \underline{C'}\big) & \xrightarrow{\Psi_{B,C'}} & \mathrm{Hom}_A\big(A[T_b | b \in B], C'\big) \end{array}$$

commutes. Given a map $\delta : B' \to \underline{C}$, both $\Psi_{B,C'}(\underline{\gamma} \circ \delta \circ \beta)$ and $\gamma \circ \Psi_{B',C}(\delta) \circ \bar{\beta}$ are A-linear homomorphisms $A[T_b | b \in B] \to C'$, which are determined by the images of the indeterminates T_b. We have

$$\begin{aligned} \Big(\Psi_{B,C'}(\underline{\gamma} \circ \delta \circ \beta)\Big)(T_b) \;&=\; \underline{\gamma}\Big(\delta\big(\beta(b)\big)\Big) \\ &= \gamma\big(\hat{\delta}(T_{\beta(b)})\big) \;=\; \big(\gamma \circ \Psi_{B',C}(\delta)\big)(T_{\beta(b)}) \;=\; \big(\gamma \circ \Psi_{B',C}(\delta) \circ \bar{\beta}\big)(T_b), \end{aligned}$$

which shows that both maps are equal and that the diagram in question commutes. Thus the free functor $\mathcal{F} : \mathrm{Sets} \to \mathrm{Alg}_A$ is left adjoint to the forgetful functor $\mathcal{G} : \mathrm{Alg}_A \to \mathrm{Sets}$.

2.7 Exercises

Exercise 2.1. Let K be a field.

(1) Show that in the categories Ab, Vect_K and Rings, a morphism is an isomorphism if and only if it is a bijective map.

(2) Show that in the categories Ab, Vect_K and Rings, a morphism is a monomorphism if and only if it is an injective map.

(3) Show that in the categories Ab and Vect_K, a morphism is an epimorphism if and only if it is a surjective map.

(4) Let A be a ring and S a multiplicative subset. Show that the localization map $\iota_S : A \to S^{-1}A$ is an epimorphism in Rings. Conclude that the inclusion $\mathbb{Z} \to \mathbb{Q}$ is both a monomorphism and an epimorphisms, but not isomorphisms.

Exercise 2.2. (1) Let $\mathcal{F} : \mathcal{C} \to \mathcal{D}$ be a functor and $\alpha : A \to B$ an isomorphism in $\mathcal{C}$. Show that $\mathcal{F}(\alpha)$ is an isomorphism in $\mathcal{D}$.

(2) Give an example of a functor $\mathcal{F} : \mathcal{C} \to \mathcal{D}$ and an epimorphism α in $\mathcal{C}$ such that $\mathcal{F}(\alpha)$ is not an epimorphism.

(3) Give an example of a functor $\mathcal{F} : \mathcal{C} \to \mathcal{D}$ and a monomorphism α in $\mathcal{C}$ such that $\mathcal{F}(\alpha)$ is not a monomorphism.

Exercise 2.3. Let $\mathcal{C}$ be a category and $\{A_i\}_{i\in I}$ a family of objects in $\mathcal{C}$. Assume that $\mathcal{C}$ has a product $\prod_{i\in I} A_i$ and a coproduct $\coprod_{i\in I} A_i$. Let B be another object of $\mathcal{C}$.

(1) Show that there is a bijection $\mathrm{Hom}_{\mathcal{C}}\left(B, \prod_{i\in I} A_i\right) \longrightarrow \prod_{i\in I} \mathrm{Hom}_{\mathcal{C}}(B, A_i)$.

(2) Show that there is a bijection $\mathrm{Hom}_{\mathcal{C}}\left(\coprod_{i\in I} A_i, B\right) \longrightarrow \prod_{i\in I} \mathrm{Hom}_{\mathcal{C}}(A_i, B)$.

Exercise 2.4. Let $\mathcal{C}$ be a category and $\alpha : A \to B$ a morphism in $\mathcal{C}$. A **(categorical) image of** α is an object $\mathrm{im}(\alpha)$ in $\mathcal{C}$ together with a morphism $\pi : A \to \mathrm{im}(\alpha)$ and a monomorphism $\iota : \mathrm{im}(\alpha) \to B$ such that $\alpha = \iota \circ \pi$ that satisfies the following universal property: for every object C, morphism $\pi' : A \to C$ and monomorphism $\iota' : C \to B$ such that $\alpha = \iota' \circ \pi'$ there is a unique morphism $\beta : \mathrm{im}(\alpha) \to C$ such that $\pi' = \beta \circ \pi$ and $\iota = \iota' \circ \beta$.

(1) Draw a diagram taking all the above objects and morphisms into consideration.

(2) Assume that the image $\mathrm{im}(\alpha)$ of α exits. Show that if C together with $\pi' : A \to C$ and $\iota' : C \to B$ is an image of α, then there is a unique isomorphism $\beta : \mathrm{im}(\alpha) \to C$ such that$\pi' = \beta \circ \pi$ and $\iota = \iota' \circ \beta$.

(3) Let $\alpha : A \to B$ be a morphism in Sets. Consider the set

$$\mathrm{im}(\alpha) \;=\; \{\, b \in B \,|\, b = \alpha(a) \text{ for some } a \in A \,\},$$

and the maps $\pi : A \to \mathrm{im}(\alpha)$ with $\pi(a) = \alpha(a)$ and $\iota : \mathrm{im}(\alpha) \to B$ with $\iota(b) = b$. Show that $\mathrm{im}(\alpha)$ together with π and ι is a categorical image of α in Sets.

(4) Show that the analogous statements to (3) hold for Ab, Vect_K and Rings.

Exercise 2.5. Let $\mathcal{C}$ be a category. A **zero object of** $\mathcal{C}$ is an object $\mathbf{0}$ that is both initial and terminal. If $\mathcal{C}$ has a zero object $\mathbf{0}$, then we call for any two objects A and B of $\mathcal{C}$, the unique morphism $0 : A \to \mathbf{0} \to B$ from A to B the **zero morphism**.

(1) Show that the categories Ab and Vect_K have a zero object. Show that in both categories a morphism $\alpha : A \to B$ is a zero morphism if and only if $\alpha(a) = 0$ for all $a \in A$ (where 0 stays for the zero element of B).

(2) Show that the categories Sets and Rings do not have a zero object.

Assume that $\mathcal{C}$ has a zero object $\mathbf{0}$.

(3) Show that the composition of a morphism with a zero morphism (in any order) is a zero morphism.

A **(categorical) kernel of a morphism** $\alpha : A \to B$ is an object $\ker\alpha$ together with a morphism $\iota : \ker\alpha \to A$ such that $\alpha \circ \iota = 0$ that satisfies the following universal property: for every object C and every morphism $\iota' : C \to A$ such that $\alpha \circ \iota' = 0$, there is a unique morphism $\beta : C \to \ker\alpha$ such that $\iota' = \iota \circ \beta$.

(4) Draw a diagram taking all the above objects and morphisms into consideration.

(5) Let $\alpha : A \to B$ be a morphism of abelian groups. Show that $\ker\alpha = \{a \in A | \alpha(a) = 0\}$ together with the inclusion $\ker\alpha \to A$ as subgroup is a categorical kernel of α.

(6) What is the problem with categorical kernels in Rings?

Exercise 2.6. Show that the category Top of topological spaces has initial and terminal objects as well as products and coproducts.

Recall from Exercise 1.30 the definition of $\mathrm{Spec}A$ as the set of all prime ideals $\mathfrak{p}$ of A together with the topology generated by the principal open subsets U_a where a varies through all elements of A. For a ring homomorphism $f : A \to B$, we define $f^* = \mathrm{Spec}\, f$ as the map that sends a prime ideal $\mathfrak{p}$ of $\mathrm{Spec}\, B$ to $f^*(\mathfrak{p}) = f^{-1}(\mathfrak{p})$.

Show that this defines a contravariant functor $\mathrm{Spec} : \mathrm{Rings} \to \mathrm{Top}$. Show that the spectrum of the product of a finite number of rings, together with the induced morphisms of spectra defined by the canonical projections, is a coproduct of their spectra.

Chapter 3

Modules

3.1 Definitions

Throughout this section, we let A be a ring.

Definition 3.1.1. An **A-module** is an (additively written) commutative group M (with neutral element 0_M) together with an *action of A on M*, which is a map

$$\begin{array}{ccc} A \times M & \longrightarrow & M \\ (a,m) & \longmapsto & a.m \end{array}$$

that satisfies

(1) $1.m = m$;

(2) $(ab).m = a.(b.m)$;

(3) $(a+b).m = a.m + b.m$;

(4) $a.(m+n) = a.m + a.n$

for all $a,b \in A$ and $m,n \in M$ where $a.m + b.n$ must be read as $(a.m) + (b.n)$.

Let M and N be A-modules. A **homomorphism from M to N** (or **A-linear map**) is a group homomorphism $f : M \to N$ such that $f(a.m) = a.f(m)$ for all $a \in A$ and $m \in M$. Together with the usual composition of group homomorphisms, this defines the category Mod_A of A-modules. A homomorphism $f : M \to N$ of A-modules is an *isomorphism* if it is bijective.

A **submodule of M** is a subgroup N of M such that $a.n \in N$ for all $a \in A$ and all $n \in N$. Given a subset S of M, we define the **submodule generated by** S as the submodule

$$\langle S \rangle_A \;=\; \left\{ \sum_{i=1}^{n} a_i.s_i \in M \;\middle|\; n \geqslant 1, a_1, \ldots, a_n \in A, s_1, \ldots, s_n \in S \right\}$$

of M if S is not empty and $\langle S \rangle_A = \{0\}$ if $S = \varnothing$. A subset S of M **generates** M if $\langle S \rangle_A = M$. A module M is *finitely generated* if $M = \langle S \rangle_A$ for a finite subset S of M. A submodule N of M is **cyclic** if $N = \langle \{m\} \rangle_A$ for one element $m \in M$.

Remark. In the following, we explain some notational conventions and some first observations.

- If there is no danger of confusion, we denote both the neutral element of A and the neutral element of M by 0. The axioms of a module imply that $a.0 = 0.m = 0$ for all $a \in A$ and $m \in M$.
- An homomorphism $f : M \to N$ of A-modules is an isomorphism if and only if it is an isomorphism in the categorical sense, i.e. if and only if there is a homomorphism $g : N \to M$ (which is the inverse bijection) such that $g \circ f = \mathrm{id}_M$ and $f \circ g = \mathrm{id}_N$.
- A submodule N of M is itself an A-module with respect to the restriction of the action $A \times M \to M$ to $A \times N \to N$.
- The submodule generated by a subset S of M is the smallest submodule of M that contains S; in particular, it is indeed a submodule. Note that $\langle M \rangle_A = M$, thus every A-module has a generating set. If $S = \{s_1, \ldots, s_n\}$, then we also write $\langle s_1, \ldots, s_n \rangle_A$ for $\langle S \rangle_A$.

Example 3.1.2. We discuss some examples of modules, submodules and homomorphisms.

(0) The **trivial A-module** is the trivial group $\{0\}$ together with the tautological A-action given by $a.0 = 0$ for all $a \in A$.

(1) The underlying additive group of A together with the action of A on itself given by the multiplication $a.m = a \cdot m$ for all $a, m \in A$ is an A-module. A submodule of A is the same thing as an ideal of A.

(2) Let M be an A-module. The **trivial submodule of M** is $\{0\}$. The **improper submodule of M** is M.

(3) Let M and N be A-modules. The **zero map from M to N** is the homomorphism $\mathbf{0} : M \to N$ that sends every element $m \in M$ to 0.

(4) For a field K, it is apparent from the respective definitions that a K-module is the same thing as a K-vector space, and a homomorphism of K-modules is the same thing as a homomorphism of K-vector spaces. In so far, modules can be seen as a generalization of vector spaces from fields to rings.

(5) Let $f : A \to B$ be a ring homomorphism. Then B is an A-module with respect to the action of A on B given by $a.b = f(a) \cdot b$ for all $a \in A$ and $b \in B$. More general, every B-module M is an A-module with respect to the action $a.m = f(a).m$ where $a \in A$ and $m \in M$.

(6) Let A be a unique factorization domain and $K = \mathrm{Frac}A$ its field of fractions, which is an A-module with respect to the inclusion $A \to K$. A principal fractional ideal is the same thing as a cyclic submodule of the A-module K. This shows that the notation $\langle m \rangle_A$ is unambiguous.

(7) Every A-module is an abelian group. More precisely, there is a forgetful functor $\mathcal{F} : \mathrm{Mod}_A \to \mathrm{Ab}$ that sends an A-module to its underlying abelian group and an

A-linear map to its underlying group homomorphism. Conversely, every abelian group M is naturally a $\mathbb{Z}$-module with the action $\mathbb{Z} \times M \to M$ defined by

$$a.m \;=\; \underbrace{m+\cdots+m}_{a\text{-times}}$$

for $a > 0$ and $a.m = -\big((-a).m\big)$ for $a < 0$. Since this is the only possible action of $\mathbb{Z}$ on an abelian group, we conclude that a $\mathbb{Z}$-module is the same thing as an abelian group. More precisely, since every group homomorphisms between abelian groups is $\mathbb{Z}$-linear, that the forgetful functor $\mathcal{F} : \mathrm{Mod}_{\mathbb{Z}} \to \mathrm{Ab}$ is an identification of the two categories.

We can extend our list of examples in terms of the following constructions.

Definition 3.1.3. Let $f : M \to N$ be a homomorphism of A-modules. The **image of** f is the subset

$$\mathrm{im}\, f \;=\; \big\{ n \in N \,\big|\, n = f(m) \text{ for some } m \in M \big\}$$

of N and **the kernel of** f is the subset

$$\ker f \;=\; \big\{ m \in M \,\big|\, f(m) = 0 \big\}$$

of M.

Lemma 3.1.4. *Let $f : M \to N$ be a homomorphism of A-modules. Then $\mathrm{im}\, f$ is a submodule of N and $\ker f$ is a submodule of M. The A-linear map f is injective if and only if $\ker f = \{0\}$.*

Proof. We have to show that both $\mathrm{im}\, f$ and $\ker f$ are closed under subtraction and the A-action. We begin with $\mathrm{im}\, f$. Consider $a \in A$ and $n, n' \in \mathrm{im}\, f$, i.e. $n = f(m)$ and $n' = f(m')$ for some $m, m' \in M$. Then both

$$n - n' \;=\; f(m) - f(m') \;=\; f(m - m') \qquad \text{and} \qquad a.n \;=\; a.f(m) \;=\; f(a.m)$$

are in $\mathrm{im}\, f$, which shows that $\mathrm{im}\, f$ is a submodule of N. Consider $a \in A$ and $m, m' \in \ker f$, i.e. $f(m) = f(m') = 0$. Then

$$f(m - m') \;=\; f(m) - f(m') \;=\; 0 \qquad \text{and} \qquad f(a.m) \;=\; a.f(m) \;=\; a.0 \;=\; 0,$$

which shows that both $m - m'$ and $a.m$ are in $\ker f$. This shows that $\ker f$ is a submodule of M.

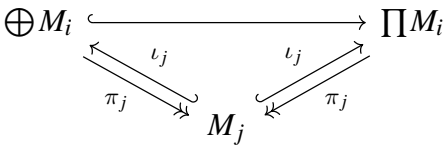
If f is injective, then $0 \in N$ has only one inverse image, which is $0 \in M$. Thus $\ker f = \{0\}$. If conversely, $\ker f = \{0\}$ and $f(m) = f(m')$, then $f(m - m') = f(m) - f(m') = 0$. Since $\ker f = \{0\}$, we conclude that $m - m' = 0$ and thus $m = m'$, which shows that f is injective. □

Definition 3.1.5. Let $\{M_i\}_{i\in I}$ be a family of A-modules. The **product of** $\{M_i\}$ is the Cartesian product $\prod_{i\in I} M_i$ together with the componentwise addition and the componentwise A-action, which is defined by $a.(m_i)_{i\in I} = (a.m_i)_{i\in I}$. The **direct sum of** $\{M_i\}$ is the submodule

$$\bigoplus_{i\in I} M_i \;=\; \left\{(m_i) \in \prod_{i\in I} M_i \,\middle|\, m_i = 0 \text{ for all but finitely many } i \in I\right\}$$

of the product $\prod M_i$.

Remark. Note that the inclusion $\bigoplus M_i \to \prod M_i$ is an isomorphism if I is finite. Both the product and the direct sum come equipped with the coordinate projections $\pi_j : \prod M_i \to M_j$ and $\pi_j : \bigoplus M_i \to M_j$ and with the coordinatewise injections $\iota_j : M_j \to \prod M_i$ and $\iota_j : M_j \to \bigoplus M_i$, which map an element $m \in M_j$ to the tuple $(m_i)_{i\in I}$ with $m_j = m$ and $m_i = 0$ for $i \neq j$. This homomorphisms together can be illustrated as follows:



The product $\prod M_i$ together with the family of *canonical projections* π_j is the categorical product of $\{M_i\}$ and the direct sum $\bigoplus M_i$ together with the family of *canonical inclusions* ι_j is the categorical coproduct of $\{M_i\}$. The proof is left as Exercise 3.2

3.2 Quotients

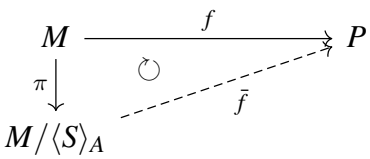
Let A be a ring.

Definition 3.2.1. Let M be an A-module and N a submodule. The **quotient of** M **by** N is the quotient M/N of abelian groups together with the A-action given by $a.[m] = [a.m]$ for $a \in A$ and the class $[m] \in M/N$ of $m \in M$.

Proposition 3.2.2 (Universal property of the quotient module)**.** *Let M be an A-module and N a submodule.*

(1) The A-action on the quotient M/N is well-defined and turns M/N into an A-module. The quotient map $\pi : M \to M/N$ with $\pi(m) = [m]$ for $m \in M$ is a homomorphism of A-modules.

(2) Let $S \subset N$ be a subset such that generates N, i.e. $N = \langle S\rangle_A$. Then $M/\langle S\rangle_A$ together with the quotient map $\pi : M \to M/\langle S\rangle_A$ satisfies the following universal property: for every A-module P and for every homomorphism $f : M \to P$ such that $f(S) \subset \{0\}$ there is a unique homomorphism $\bar{f} : M/\langle S\rangle_A \to P$ such that $f = \bar{f} \circ \pi$, i.e. the diagram



commutes.

Proof. We begin with the verification that the A-action on M is well-defined. Let $[m] = [m']$ in M/N, i.e. $n = m - m' \in N$. Then

$$a.[m] \;=\; [a.m] \;=\; [a.(m'+n)] \;=\; [a.m'+a.n] \;=\; [a.m'] + [a.n] \;=\; a.[m']$$

since $a.n \in N$ and thus $[a.n] = [0]$ in M/N. Thus the rule $a.[m] = [a.m]$ yields a well-defined map $A \times (M/N) \to (M/N)$.

We continue with the verification of the axioms of an A-module. By Proposition 1.1.7, the quotient M/N is a commutative group. Axioms (1)–(4) of a module hold since

(1) $1.[m] = [1.m] = [m]$,

(2) $(ab).[m] = [(ab).m] = [a.(b.m)] = a.[b.m] = a.(b.[m])$,

(3) $(a+b).[m] = [(a+b).m] = [a.m+b.m] = [a.m] + [b.m] = a.[m] + b.[m]$,

(4) $a.[m+n] = [a.(m+n)] = [a.m+a.n] = [a.m] + [a.n] = a.[m] + a.[n]$

for all $a, b \in A$ and $m, n \in M$. This shows that M/N is an A-module, as claimed.

By Exercise 1.3, the map $\pi : M \to M/N$ is a group homomorphism. For $a \in A$ and $m \in M$, we have $\pi(a.m) = [a.m] = a.[m] = a.\pi(m)$, which shows that π is A-linear. This completes the proof of (1).

We continue with the proof of (2). Given a homomorphism $f : M \to P$ with $f(S) \subset \{0\}$, we claim that the association $[m] \mapsto f(m)$ does not depend on the choice of representative $m \in [m]$ and defines a homomorphism $M/\langle S\rangle_A \to P$. Once we have proven this, it is clear that from the definition of $\bar{f}$ that $f = \bar{f} \circ \pi$. Note that $f = \bar{f} \circ \pi$ implies the uniqueness of $\bar{f}$ since it requires that $\bar{f}([m]) = \bar{f}(\pi(m)) = f(m)$.

In order to show that $\bar{f}$ is well-defined, consider $m, n \in M$ such that $[m] = [n]$. By Proposition 1.1.7, we have $m - n \in \langle S\rangle_A$, and thus $m - n = \sum a_i s_i$ for some $a_i \in A$ and $s_i \in S$. It follows that

$$f(m) \;=\; f(n + \textstyle\sum a_i s_i) \;=\; f(n) + \sum f(a_i) \underbrace{f(s_i)}_{=0} \;=\; f(n),$$

which shows that the value $\bar{f}([m]) = \bar{f}([n])$ does not depend on the choice of representative for $[m] = [n]$. The map $\bar{f}$ is a homomorphism since

$$\bar{f}([a.m]) \;=\; f(a.m) \;=\; a.f(m) \;=\; a.\bar{f}([m])$$

for all $a \in A$ and $m \in M$. This concludes the proof of the proposition. □

3.3 The tensor product

As usual, we let A be a ring. Let M and N be two A-modules. For every $(m,n) \in M \times N$, we let $A_{(m,n)}$ be a copy of A and consider the direct sum $\bigoplus_{(m,n)\in M\times N} A_{(m,n)}$. We write $a.[m,n]$ for the element of $\bigoplus A_{(m',n')}$ whose coefficient in $A_(m,n)$ is a and whose other coefficients are 0, i.e. $a.[m,n]$ is the image of a under the canonical inclusion of $A_{(m,n)}$ into $\bigoplus A_{(m',n')}$. We write $[m,n] = 1.[m,n]$.

Definition 3.3.1. The **tensor product of M and N** is the quotient

$$M \otimes_A N \;=\; \Big(\bigoplus_{(m,n)\in M\times N} A_{(m,n)} \Big) \,/\, \langle S \rangle$$

of $\bigoplus A_{(m,n)}$ by the submodule generated by the subset S of $\bigoplus A.(m,n)$ that consists of all elements of the forms

$$\begin{aligned} &[a.m,n]-a.[m,n], && [m+m',n]-[m,n]-[m',n],\\ &[m,a.n]-a.[m,n], && [m,n+n']-[m,n]-[m,n'] \end{aligned}$$

with $a \in A$, $m,m' \in M$ and $n,n' \in N$. We write $m \otimes n$ for the class of $[m,n]$ in $M \otimes_A N$, and call an element of this form in $M \otimes_A N$ a **pure tensor**.

Remark. If the ring A is clear from the context, we simply write $M \otimes N$ for $M \otimes_A N$. The defining relations of $M \otimes N$ imply at once the following rules for the tensor product:

$$\begin{aligned} (a.m)\otimes n \;&=\; a.(m\otimes n) \;=\; m\otimes(a.n),\\ (m+m')\otimes n \;&=\; m\otimes n + m'\otimes n,\\ m\otimes(n+n') \;&=\; m\otimes n + m\otimes n' \end{aligned}$$

for all $a \in A$, $m,m' \in M$ and $n,n' \in N$.

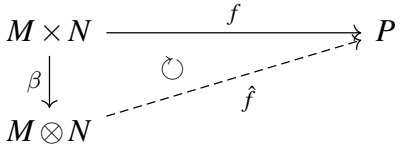
Since the elements (m,n) generate the A-module $\bigoplus A.(m,n)$, every element in the quotient $M \otimes N$ is a sum of pure tensors, i.e. of the form $\sum m_i \otimes n_i$ for some $m_i \in M$ and $n_i \in N$. Since $f(\sum m_i \otimes n_i) = \sum f(m_i \otimes n_i)$ for a homomorphism $f : M \otimes N \to P$, it suffices to determine the images of pure tensors to describe a homomorphism from $M \otimes N$ into another A-module.

Definition 3.3.2. Let M, N and P be A-modules. A **bilinear map from $M \times N$ to P** is a map $f : M \times N \to P$ such that the association $m \mapsto f(m,n)$ defines an A-linear map $f(-,n) : M \to P$ for every $n \in N$ and such that the association $n \mapsto f(m,n)$ defines an A-linear map $f(m,-) : N \to P$ for every $m \in M$.

Proposition 3.3.3 (Universal property of the tensor product)**.** *Let M and N be two A-modules.*

(1) The association $(m,n) \mapsto m \otimes n$ defines a bilinear map $\beta : M \times N \to M \otimes N$.

(2) The tensor product $M \otimes N$ together with the bilinear map $\beta : M \times N \to M \otimes N$ satisfies the following universal property: for every A-module P and every bilinear map $f : M \times N \to P$, there is a unique homomorphism $\hat{f} : M \otimes N \to P$ such that $f = \hat{f} \circ \beta$, i.e. the diagram



commutes.

Proof. We begin with (1). Given $n \in N$, the map $\beta(-,n) : M \to M \otimes N$ is A-linear since

$$\begin{aligned} \beta\big((m,n)+(m',n)\big) &= \beta(m+m',n) = (m+m')\otimes n \\ &= m\otimes n + m'\otimes n = \beta(m,n)+\beta(m',n), \\ \beta\big(a.(m,n)\big) &= \beta(a.m,n) = (a.m)\otimes n = a.(m\otimes n) = a.\beta(m,n) \end{aligned}$$

for all $a \in A$ and $m,m' \in M$. Similarly, the map $\beta(m,-) : N \to M \otimes N$ is A-linear for every $m \in M$. This shows that β is bilinear. Thus (1).

We continue with (2). Given a bilinear map $M \times N \to P$, we claim that the association $\sum m_i \otimes n_i \mapsto \sum f(m_i,n_i)$ defines a homomorphism $\hat{f} : M \otimes N \to P$. Once we have proven this, it is clear from the definition of $\hat{f}$ that $f = \hat{f} \circ \pi$. Note that $f = \hat{f} \circ \pi$ implies the uniqueness of $\hat{f}$ since it requires that $\hat{f}(\sum m_i \otimes n_i) = \sum \hat{f}(m_i \otimes n_i) = \sum f(m_i,n_i)$.

To show that $\hat{f}$ is indeed a well-defined homomorphism of A-modules, we first consider the homomorphism $\tilde{f} : \bigoplus_{(m,n)\in M\times N} A.(m,n) \to P$ that maps $\sum a_i.(m_i,n_i)$ to $\sum a_i f(m_i,n_i)$. We leave it as an exercise to see that this defines indeed a homomorphism, which follows, for instance, from the universal property of the coproduct $\bigoplus A.(m,n)$.

As our next step, we verify whether all defining relation of $M \otimes N$ are in the kernel of $\tilde{f}$. Let $a \in A$, $m,m' \in M$ and $n,n' \in N$. Since $f(a.m,n) = a.f(m,n)$, the element $(a.m,n) - a.(m,n)$ is in $\ker \tilde{f}$. Since $f(m+m',n) = f(m,n) + f(m',n)$, the element $(m+m',n) - (m,n) - (m',n)$ is in $\ker \tilde{f}$. Similarly, the elements $(m,a.n) - a.(m,n)$ and $(m,n+n') - (m,n) - (m,n')$ are in $\ker \tilde{f}$. Thus by the universal property of quotient modules (Proposition 3.2.2), there is a unique morphism $\hat{f} : M \otimes N \to P$ such that $\tilde{f} = \hat{f} \circ \pi$ where $\pi : \bigoplus A.(m,n) \to M \otimes N$ is the quotient map. We conclude that

$$\hat{f}(\textstyle\sum m_i \otimes n_i) = \sum \hat{f}(m_i \otimes n_i) = \sum \hat{f}\big(\pi(m_i,n_i)\big) = \sum \tilde{f}(m_i,n_i) = \sum f(m_i,n_i),$$

which shows that the constructed homomorphism $\hat{f}$ is as prescribed. Thus (2). □

Lemma 3.3.4. *Let M, N and P be A-modules. Then the following holds.*

(1) $M \otimes \{0\} = \{0\}$ *and* $M \otimes A \simeq M$;

(2) $M \otimes N \simeq N \otimes M$;

(3) $(M \otimes N) \otimes P \simeq M \otimes (N \otimes P)$;

(4) $M \otimes (N \oplus P) \simeq (M \otimes N) \oplus (M \otimes P)$.

Proof. We leave the proof as Exercise 3.1. □

Remark. Properties (2) and (3) imply that we can write unambiguously $\otimes_{i=1}^{r} M_i = M_1 \otimes \ldots \otimes M_r$ for A-modules $M_1, \ldots, M_r$. The tensor product $\otimes M_i$ together with the map $\prod M_i \to \otimes M_i$ that sends $(m_1, \ldots, m_r)$ to $m_1 \otimes \cdots \otimes m_r$ satisfies an analogous universal property to that of $M \otimes N$, which is based on the notion of multilinear maps in place of bilinear maps.

Lemma 3.3.5 (Extension of scalars)**.** *Let $f : A \to B$ be a ring homomorphism and M an A-module. Then the association*

$$\begin{array}{rccc} \theta: & B \times (B \otimes_A M) & \longrightarrow & B \otimes_A M \\ & (a, \sum b_i \otimes m_i) & \longmapsto & \sum (ab_i) \otimes m_i \end{array}$$

is an action of B on M that endows $B \otimes_A M$ with the structure of a B-module.

Proof. We begin with the proof that the action of B on $B \otimes_A M$ is well-defined as a map. We claim that the map $\theta_c : B \times M \to B \otimes_A M$ with $\theta_c(b,m) = (cb) \otimes m$ is bilinear for every $c \in C$. Indeed, the map $\theta_c(-,m) : B \to B \otimes_A M$ is A-linear for every $m \in M$ since

$$\begin{aligned} \theta_c(a.b,m) &= (cf(a)b) \otimes m = a.\big((cb) \otimes m\big) = a.\theta_c(b,m), \\ \theta_c(b+b',m) &= \big(c(b+b')\big) \otimes m = (cb) \otimes m + (cb') \otimes m = \theta_c(b,m) + \theta_c(b',m) \end{aligned}$$

for all $a \in A$ and $b, b' \in B$. Similarly, the map $\theta_c(b,-) : M \to B \otimes_A M$ is A-linear for every $b \in B$. Thus θ_c is bilinear, as claimed.

By the universal property of the tensor product (Proposition 3.3.3), there is a homomorphism $\hat{\theta}_c : B \otimes_A M \to B \otimes_A M$ with $\hat{\theta}_c(b \otimes m) = (cb \otimes m)$. We conclude that

$$f(a, \sum b_i \otimes m_i) = \hat{\theta}_a(\sum b_i \otimes m_i) = \sum \hat{\theta}_a(b_i \otimes m_i) = \sum (ab_i) \otimes m_i,$$

which shows that f is well-defined as a map. It also follows that $\theta(a,-) = \theta_a$ is an A-linear map $B \otimes_A M \to B \otimes_A M$, which implies axioms (4) of a module at once. Axioms (1)–(3) hold since

$$\begin{aligned} 1.\big(\sum c_i \otimes m_i\big) &= \sum (1 \cdot c_i) \otimes m_i = \sum c_i \otimes m_i, \\ (ab).\big(\sum c_i \otimes m_i\big) &= \sum (abc_i) \otimes m_i = a.\big(\sum b.(c_i \otimes m_i)\big), \\ (a+b).\big(\sum c_i \otimes m_i\big) &= \sum (ac_i + bc_i) \otimes m_i = a.\big(\sum c_i \otimes m_i\big) + b.\big(\sum c_i \otimes m_i\big) \end{aligned}$$

for all $a, b, c_i \in A$ and $m_i \in M$. This shows that $B \otimes_A M$ is a B-module, as claimed. □

Proposition 3.3.6 (Tensor product of rings)**.** *Let $\alpha_B : A \to B$ and $\alpha_C : A \to C$ be ring homomorphisms. Then the following holds true.*

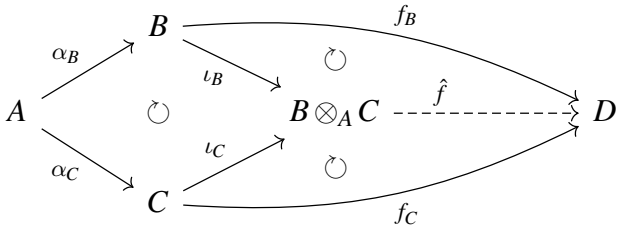
(1) The tensor product $B \otimes_A C$ is a ring with respect to the multiplication

$$\begin{array}{rccc} m: & (B \otimes_A C) \times (B \otimes_A C) & \longrightarrow & B \otimes_A C. \\ & (b \otimes c, b' \otimes c') & \longmapsto & (bb') \otimes (cc') \end{array}$$

(2) The associations $b \mapsto b \otimes 1$ and $c \mapsto 1 \otimes c$ define ring homomorphisms $\iota_B : B \to B \otimes_A C$ and $\iota_C : C \to B \otimes_A C$, respectively, which are called the canonical inclusions*. We have $\iota_B \circ \alpha_B = \iota_C \circ \alpha_C$.*

(3) The tensor product $B \otimes_A C$ together with the canonical inclusions $\iota_B : B \to B \otimes_A C$ and $\iota_C : C \to B \otimes_A C$ satisfies the following universal property: given ring homomorphisms $f_B : B \to D$ and $f_C : C \to D$ such that $f_B \circ \alpha_B = f_C \circ \alpha_C$, then

there is a unique ring homomorphism $\hat{f} : B \otimes_A C \to D$ such that $f_B = \hat{f} \circ \iota_B$ and $f_C = \hat{f} \circ \iota_C$, i.e. the diagram



commutes.

Proof. We begin with the verification that the multiplication m is well-defined as a map. Since the map

$$\begin{array}{rccc} \tilde{m}: & B \times C \times B \times C & \longrightarrow & B \otimes_A C \\ & (b,c,b',c') & \longmapsto & (bb') \otimes (cc') \end{array}$$

is linear in each of the four factors, we can use the universal property of the tensor product (Proposition 3.3.3) successively to gain a homomorphism

$$\hat{m}: \quad B \otimes_A C \otimes_A B \otimes_A C \quad \longrightarrow \quad B \otimes_A C$$

with $\hat{m}(b \otimes c \otimes b' \otimes c') = \tilde{m}(b,c,b',c') = (bb') \otimes (cc')$. Thus the prescribed association m is the composition of $\hat{m}$ with the bilinear map $\beta : (B \otimes_A C) \times (B \otimes_A C) \to B \times C \times B \times C$, which shows that m is well-defined as a map.

We continue with the verification that $B \otimes_A C$ is a ring. Since it is an A-module, it is, in particular, an abelian group with respect to addition. The neutral element for multiplication m is $1 \otimes 1$. The associativity and commutativity of m can easily be derived from the corresponding properties of B and C, which results in a proof that $B \otimes_A C$ is a multiplicative monoid. Distributivity follows directly from the construction of m as $\hat{m} \circ \beta$. Thus $B \otimes_A C$ is a ring, which establishes (1).

We continue with (2). The map $\iota_B : B \to B \otimes_A C$ is a ring homomorphism since $\iota_B(1) = 1 \otimes 1$ and

$$\begin{aligned} \iota_B(a+b) &= (a+b) \otimes 1 = a \otimes 1 + b \otimes 1 = \iota_B(a) + \iota_B(b), \\ \iota_B(ab) &= (ab) \otimes 1 = (a \otimes 1) \cdot (b \otimes 1) = \iota_B(a) \cdot \iota_B(b) \end{aligned}$$

for all $a, b \in B$. That ι_C is a ring homomorphism can be verified analogously. Since

$$\iota_B\big(\alpha_B(a)\big) = (a.1) \otimes 1 = 1 \otimes (a.1) = \iota_C\big(\alpha_C(a)\big),$$

for all $a \in A$, we have $\iota_B \circ \alpha_B = \iota_C \circ \alpha_C$. Thus (2).

We continue with (3). Given two ring homomorphisms $f_B : B \to D$ and $f_C : C \to D$ such that $f_B \circ \alpha_B = f_C \circ \alpha_C$, we claim that the association $b \otimes c \mapsto f_B(b) \cdot f_C(c)$ defines a ring homomorphism $\hat{f} : B \otimes_A C \to D$. Once we have proven this, it is clear from the definition of $\hat{f}$ that $f_B = \hat{f} \circ \iota_B$ and $f_C = \hat{f} \circ \iota_C$. Conversely, $f_B = \hat{f} \circ \iota_B$ and $f_C = \hat{f} \circ \iota_C$ imply that

$$\hat{f}(b \otimes c) = \hat{f}(b \otimes 1) \cdot \hat{f}(1 \otimes c) = \hat{f}\big(\iota_B(b)\big) \cdot \hat{f}\big(\iota_C(c)\big) = f_B(b) \cdot f_C(c),$$

which shows the uniqueness of $\hat{f}$.

To show that $\hat{f}$ is indeed a well-defined ring homomorphism, we consider the map $f : B \times C \to D$ that sends (b,c) to $f_B(b) \cdot f_C(c)$. If we consider D as an A-module with respect to the ring homomorphism $f_B \circ \alpha_B = f_C \circ \alpha_C : A \to D$, then f is bilinear: it is A-linear in the first argument since

$$\begin{aligned} f(b+b',c) &= f_B(b+b')f_C(c) = f_B(b)f_C(c) + f_B(b')f_C(c) = f(b,c) + g(b',c), \\ f(a.b,c) &= f_B\big(\alpha_B(a) \cdot b\big) \cdot f_C(c) = (f_B \circ \alpha_B)(a) \cdot f_B(b) \cdot f_C(c) = a.f(b,c), \end{aligned}$$

for all $a \in A$, $b.b' \in B$ and $c \in C$. Similarly, it can be verified that f is A-linear in the second argument. Thus the universal property of tensor products of A-modules (Proposition 3.3.3) establishes a homomorphism $\hat{f} : B \otimes_A C \to D$ of A-modules such that $\hat{f}(b \otimes c) = f(b,c) = f_B(b) \cdot f_C(c)$, as desired.

As an A-linear map, $\hat{f}$ is in particular additive. Since $\hat{f}(1 \otimes 1) = f_B(1)f_C(1) = 1$ and

$$\hat{f}\big((b \otimes c) \cdot (b' \otimes c')\big) = f_B(bb')f_C(cc') = f_B(b)f_C(c)f_B(b')f_C(c') = \hat{f}(b,c)\hat{f}(b',c'),$$

$\hat{f}$ is a ring homomorphism, which concludes the proof of (3). □

Recall from Exercise 1.22 that there is a unique morphism $\mathbb{Z} \to A$ from $\mathbb{Z}$ into any given ring A.

Corollary 3.3.7 (Coproduct of rings)**.** *Let B and C be rings and $\alpha_B : \mathbb{Z} \to B$ and $\alpha_C : \mathbb{Z} \to C$ the unique ring homomorphisms from $\mathbb{Z}$ to B and C, respectively. Then $B \otimes_{\mathbb{Z}} C$ together with the canonical inclusions $\iota_B : B \to B \otimes_{\mathbb{Z}} C$ and $\iota_C : C \to B \otimes_{\mathbb{Z}} C$ is a coproduct of B and C in the category* Rings.

Proof. We verify the universal property of the coproduct. Consider two ring homomorphisms $f_B : B \to D$ and $f_C : C \to D$. Then both $f_B \circ \alpha_B$ and $f_C \circ \alpha_C$ are equal to the unique morphism $\mathbb{Z} \to D$, i.e. $f_B \circ \alpha_B = f_C \circ \alpha_C$. Thus we can apply the universal property of the tensor product $B \otimes_{\mathbb{Z}} C$ (Proposition 3.3.6), which shows that there is a unique morphism $\hat{f} : B \otimes_{\mathbb{Z}} C \to D$ such that $f_B = \hat{f} \circ \iota_B$ and $f_C = \hat{f} \circ \iota_C$. This verifies the universal property of the coproduct. □

3.4 The isomorphism theorems for modules

Let A be a ring in this section.

Theorem 3.4.1 (First isomorphism theorem)**.** *Let $f : M \to N$ be a homomorphism of A-modules. Then*

$$\begin{array}{ccc} M/\ker f & \longrightarrow & \operatorname{im} f \\ [m] & \longmapsto & f(m) \end{array}$$

is an isomorphism.

Proof. Since $f(\ker f) = \{0\}$, the universal properties of quotient modules (Proposition 3.2.2) shows that f factors into the quotient map $\pi : M \to M/\ker f$ followed by a homomorphism $\bar{f} : M/\ker f \to N$ for which $\bar{f}([m]) = f(m)$. Tautologically, the image of $\bar{f}$ is $\operatorname{im} f$, i.e. we can restrict $\bar{f}$ to a surjective homomorphism $\hat{f} : M/\ker f \to \operatorname{im} f$, which is the association described in the theorem. This verifies in particular that $\hat{f}$ is well-defined.

To show that $\hat{f}$ is injective, it suffices to shows that $\ker \hat{f} = \{[0]\}$ by Lemma 3.1.4. This is the case since if $[m] \in \ker \hat{f}$, then $f(m) = \hat{f}([m]) = 0$. Thus $m \in \ker f$ and $[m] = [0]$. We conclude that $\hat{f}$ is an isomorphism, which concludes the proof of the theorem. □

Definition 3.4.2. Let M be an A-module and N, P submodules of M. The **internal sum of N and P** is the submodule

$$N + P \;=\; \langle \{N \cup P\} \rangle$$

of M.

Theorem 3.4.3 (Second isomorphism theorem)**.** *Let M be an A-module and N, P submodules of M. Then*

$$\begin{array}{ccc} N/(N \cap P) & \longrightarrow & (N+P)/P \\ [n] & \longmapsto & [n] \end{array}$$

is an isomorphism.

Proof. Consider the composition $f : N \to M/P$ of the injection $N \to M$ with the quotient map $M \to M/P$. Its kernel is $\ker f = N \cap P$. Its image consists of all cosets $n + P$ with $n \in N$, i.e. $\operatorname{im} f = (N+P)/P$. By the first isomorphism theorem (Theorem 3.4.1), we obtain an induced isomorphism $\hat{f} : N/\ker f \to \operatorname{im} f$, which is precisely the map of the theorem. □

Theorem 3.4.4 (Third isomorphism theorem)**.** *Let M be an A-module and N a submodule of M. Let $\pi : M \to M/N$ be the quotient map. Then*

$$\begin{array}{rccc} \Phi : & \{\text{submodules } P \subset M \text{ containing } N\} & \longrightarrow & \{\text{submodules } Q \text{ of } M/N\} \\ & P & \longmapsto & P/N = \pi(P) \end{array}$$

is an inclusion preserving bijection, and

$$\begin{array}{ccc} M/P & \longrightarrow & (M/N)/(P/N) \\ m+P & \longmapsto & (m+P)+(P/N) \end{array}$$

is a ring isomorphism for every submodule P of M containing N.

Proof. Note that $\pi(P) = \operatorname{im}(P \to M \to M/N)$ is a module as the image of a module, and therefore a submodule of M/N. Thus Φ is well-defined. The inverse bijection to Φ is given by sending a submodule Q of M/N to the inverse image $\pi^{-1}(Q)$, which is

a submodule of M since both $\pi(m-m') = \pi(m) - \pi(m')$ and $\pi(a.m) = a.\pi(m)$ are in Q for all $a \in A$ and $m, m' \in \pi^{-1}(Q)$. Clearly, $N \subset \pi^{-1}(Q)$. Since $\pi\big(\pi^{-1}(Q)\big) = Q$, the map Φ is surjective. Conversely, $\pi^{-1}\big(\pi(P)\big) = P+N$ for a submodule $P \subset M$. If $N \subset P$, then $P+N = P$, which shows that Φ is injective. Thus Φ is a bijection as claimed. It is clear that Φ is inclusion preserving.

Given a chain of submodules $N \subset P \subset M$, we consider the composition $f : M \to (M/P)/(P/N)$ of the quotient maps $M \to M/P$ and $M/P \to (M/P)/(P/N)$. As the composition of surjective maps it is surjective, i.e. $\operatorname{im} f = (M/P)/(P/N)$. Its kernel is $N+P = P$. Thus the first isomorphism theorem (Theorem 3.4.1) shows that f induces an isomorphism $\hat{f} : M/\ker f \to \operatorname{im} f$ with $\hat{f}([m]) = f(m)$, which is the map described in the theorem. □

3.5 Irreducible and indecomposable A-modules

Let A be a ring.

Definition 3.5.1. An A-module M is **irreducible** (or **simple**) if M is not trivial and the only submodules of M are $\{0\}$ and M. A **decomposition of** M is a family of submodules $\{N_i\}_{i\in I}$ of M such that the map

$$\begin{array}{ccc} \bigoplus_{i\in I} N_i & \longrightarrow & M \\ (n_i) & \longmapsto & \sum_{i\in I} n_i \end{array}$$

is an isomorphism. We write $M = \bigoplus N_i$ for a decomposition. The A-module M is indecomposable if for every decomposition $M = N_1 \oplus N_2$ either $N_1 = \{0\}$ or $N_2 = \{0\}$.

Remark. Obviously, every irreducible A-module is indecomposable. The converse is not true. For example, the $\mathbb{Z}$-module $\mathbb{Z}/4\mathbb{Z} = \{\bar{0}, \bar{1}, \bar{2}, \bar{3}\}$ is not irreducible since the subgroup $N_1 = \{\bar{0}, \bar{2}\}$ is a proper, nontrivial submodule. But it is indecomposable since there is no submodule N_2 such that $\mathbb{Z}/4\mathbb{Z}$ is isomorphic to $N_1 \oplus N_2$ and since $\mathbb{Z}/4\mathbb{Z}$ does not have any other nontrivial proper submodule.

Lemma 3.5.2 (Schur's lemma)**.** *Let M and N be irreducible A-modules and $f : M \to N$ a homomorphism. Then f is either an isomorphism or the zero map.*

Proof. Assume that f is not the zero map. Then its kernel $\ker f$ is not M and its image is not $\{0\}$. Since M and N are simple, we have $\ker f = \{0\}$ and $\operatorname{im} f = N$. Thus f is surjective and injective by Lemma 3.1.4, which shows that f is an isomorphism. □

3.6 Exact sequences

Let A be a ring.

Definition 3.6.1. Let $I \subset \mathbb{Z}$ be a set of consecutive integers, i.e. if $m < p < n$ for $m, n \in I$ and $p \in \mathbb{Z}$, then $p \in I$. Let $I^+ = \{i \in I \mid i-1 \in I\}$ and $I^\circ = \{i \in I \mid i-1, i+1 \in I\}$. A

sequence of A-modules (with index set I) is a family $\{M_i\}_{i\in I}$ of A-modules together with a family of homomorphisms $\{d_i : M_{i-1} \to M_i\}_{i\in I^+}$, which can be illustrated as

$$\ldots \xrightarrow{d_{i-1}} M_{i-1} \xrightarrow{d_i} M_i \xrightarrow{d_{i+1}} M_{i+1} \xrightarrow{d_{i+2}} \ldots$$

A sequence of A-modules $\big(\{M_i\},\{d_i\}\big)$ is **exact at** M_i for $i \in I^\circ$ if $\operatorname{im} d_i = \ker d_{i+1}$, and it is **exact** if it is exact at M_i for all $i \in I^\circ$. A **short exact sequence** is an exact sequence of the form

$$0 \longrightarrow N \xrightarrow{i} M \xrightarrow{p} Q \longrightarrow 0$$

i.e. $I = \{0,\ldots,4\}$, $N = M_1$, $M = M_2$, $Q = M_3$, $M_0 = M_4 = 0$ and $i = d_2$, $p = d_3$, $d_1 = d_4 = \mathbf{0}$ where 0 denotes the trivial A-module and $\mathbf{0}$ the zero map.

Remark. We can prolong any sequence by an infinite sequence of trivial modules and zero maps to get a sequence with index set $I = \mathbb{Z}$. In the case of a short exact sequence $0 \to N \to M \to Q \to 0$, this yields a "long" exact sequence

$$\ldots \longrightarrow 0 \longrightarrow 0 \longrightarrow N \longrightarrow M \longrightarrow Q \longrightarrow 0 \longrightarrow 0 \longrightarrow \ldots$$

with index set $\mathbb{Z}$.

Lemma 3.6.2. *A 5-term sequence $0 \to N \xrightarrow{i} M \xrightarrow{p} Q \to 0$ of A-modules is exact if and only if $i : N \to M$ is injective, $p : M \to Q$ is surjective and $\operatorname{im} i = \ker p$. In this case, $N \simeq \ker p$ and $Q \simeq M/\operatorname{im} i$.*

Proof. The sequence $0 \to N \xrightarrow{i} M \xrightarrow{p} Q \to 0$ is exact if and only if it is exact at N, M and Q. It is exact at N if and only if $\ker i = \operatorname{im}\mathbf{0} = \{0\}$, which is equivalent with i being injective by Lemma 3.1.4. It is exact at Q if and only if $\operatorname{im} p = \ker\mathbf{0} = Q$, which is equivalent with p being surjective. By definition, i is exact at M if and only if $\operatorname{im} i = \ker p$. This proves the first claim.

If the sequence is exact, then the injective homomorphism i defines an isomorphism $N \to \operatorname{im} i = \ker p$, which is our second claim. Since $\operatorname{im} p = Q$, the first isomorphism theorem (Theorem 3.4.1) shows that $Q \simeq M/\ker p = M/\operatorname{im} i$, which verifies our last claim. □

Example 3.6.3. We discuss some examples of short exact sequences.

(1) Let N be a submodule of M with inclusion map $i : N \to M$, and let $p : M \to M/N$ be the quotient map. Then Lemma 3.6.2 attests that

$$0 \longrightarrow N \xrightarrow{i} M \xrightarrow{p} M/N \longrightarrow 0$$

is a short exact sequence.

(2) The sequences

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \mathbb{Z}/2\mathbb{Z} & \xrightarrow{i} & \mathbb{Z}/4\mathbb{Z} & \xrightarrow{p} & \mathbb{Z}/2\mathbb{Z} & \longrightarrow & 0 \\ & & \bar{0} & \longmapsto & \bar{0}; \quad \bar{0},\bar{2} & \longmapsto & \bar{0} & & \\ & & \bar{1} & \longmapsto & \bar{2}; \quad \bar{1},\bar{3} & \longmapsto & \bar{1} & & \end{array}$$

and

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \mathbb{Z}/2\mathbb{Z} & \xrightarrow{i} & \mathbb{Z}/6\mathbb{Z} & \xrightarrow{p} & \mathbb{Z}/3\mathbb{Z} & \longrightarrow & 0 \\ & & \bar{0} & \longmapsto & \bar{0}; \quad \bar{0},\bar{3} & \longmapsto & \bar{0} & & \\ & & \bar{1} & \longmapsto & \bar{3}; \quad \bar{1},\bar{4} & \longmapsto & \bar{1} & & \\ & & & & \bar{2},\bar{5} & \longmapsto & \bar{2} & & \end{array}$$

of $\mathbb{Z}$-modules are short exact.

Definition 3.6.4. A short exact sequence $0 \to N \xrightarrow{i} M \xrightarrow{p} Q \to 0$ is **split** if there exists an isomorphism $f : M \to N \oplus Q$ such that the diagram

$$\begin{array}{ccccccccc} 0 & \longrightarrow & N & \xrightarrow{i} & M & \xrightarrow{p} & Q & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle \mathrm{id}_N} & & \downarrow{\scriptstyle f} & & \downarrow{\scriptstyle \mathrm{id}_Q} & & \\ 0 & \longrightarrow & N & \xrightarrow{\iota_N} & N \oplus Q & \xrightarrow{\pi_Q} & Q & \longrightarrow & 0 \end{array}$$

commutes.

Theorem 3.6.5. *Let $0 \to N \xrightarrow{i} M \xrightarrow{p} Q \to 0$ be a short exact sequence. Then the following are equivalent.*

(1) The sequence $0 \to N \xrightarrow{i} M \xrightarrow{p} Q \to 0$ is split.

(2) There is a retract *to i, which is a homomorphism $r : M \to N$ such that $r \circ i = \mathrm{id}_N$.*

(3) There is a section *to p, which is a homomorphism $s : Q \to M$ such that $p \circ s = \mathrm{id}_Q$.*

Proof. We will establish the circle of inclusion (1)⇒(2)⇒(3)⇒(1). We begin with (1)⇒(2). Given an isomorphism $f : M \to N \oplus Q$ that splits the short exact sequence, then we define $r = \mathrm{id}_N^{-1} \circ \pi_N \circ f$, which yields the diagram

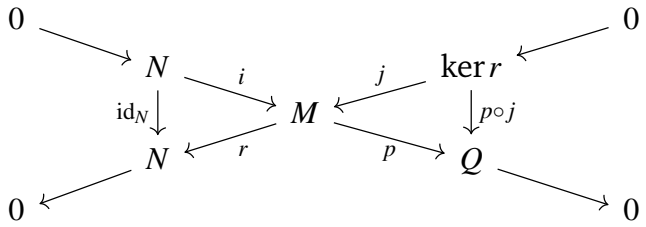
$$\begin{array}{ccccccccc} 0 & \longrightarrow & N & \underset{i}{\overset{r}{\rightleftarrows}} & M & \xrightarrow{p} & Q & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle \mathrm{id}_N} & & \downarrow{\scriptstyle f} & & \downarrow{\scriptstyle \mathrm{id}_Q} & & \\ 0 & \longrightarrow & N & \underset{\pi_N}{\overset{\iota_N}{\rightleftarrows}} & N \oplus Q & \xrightarrow{\pi_Q} & Q & \longrightarrow & 0 \end{array}$$

Since the squares of the diagram commute and since $\pi_N \circ \iota_N = \mathrm{id}_N$, we conclude that

$$r \circ i \;=\; \mathrm{id}_N^{-1} \circ \pi_N \circ f \circ i \;=\; \mathrm{id}_N^{-1} \circ \pi_N \circ \iota_N \circ \mathrm{id}_N \;=\; \mathrm{id}_N^{-1} \circ \mathrm{id}_N \circ \mathrm{id}_N \;=\; \mathrm{id}_N,$$

which shows that r is a retract to i. Thus (2).

We continue with (2)⇒(3). Let $r : M \to N$ be a retract to i. Then r is surjective. Thus by the first isomorphism theorem (Theorem 3.4.1), we have $N \sim M/\ker r$. This yields a commutative diagram



whose diagonals are short exact sequences. We claim that $p \circ j : \ker r \to Q$ is an isomorphism. Once we have proven this, we can define $s = j \circ (p \circ j)^{-1}$, which is a section to p since

$$p \circ s \;=\; (p \circ j) \circ (p \circ j)^{-1} \;=\; \mathrm{id}_Q.$$

We begin with the injectivity of $p \circ j$. If $m \in \ker(p \circ j)$, then $j(m) \in \ker p = \operatorname{im} i$. Thus $j(m) = i(n)$ for some $n \in N$ and

$$j(m) \;=\; i(n) \;=\; i \circ \underbrace{r \circ i}_{=\mathrm{id}_N}(n) \;=\; i \circ \underbrace{r \circ j}_{=0}(m) \;=\; 0.$$

Since j is injective, this means that $m = 0$, which shows that $\ker(p \circ j) = \{0\}$. By Lemma 3.1.4, we conclude that $p \circ j$ is injective.

We continue with the surjectivity of $p \circ j$. Given $q \in Q$, there is an $m \in M$ such that $p(m) = q$ since p is surjective. Thus

$$q \;=\; p(m) - i \circ \underbrace{r \circ p}_{=0}(m) \;=\; p\big(m - i \circ r(m)\big) \;=\; p(m')$$

where we define $m' = m - i \circ r(m)$. Since

$$r(m') \;=\; r(m) - \underbrace{r \circ i}_{=\mathrm{id}_N} \circ r(m) \;=\; r(m) - r(m) \;=\; 0,$$

$m' = j(m'')$ for some $m'' \in \ker r$. Thus $q = p(m') = (p \circ j)(m'')$, which shows that $p \circ j$ is surjective. This shows that $p \circ j$ is an isomorphism, and thus (3).

We continue with (3)⇒(1). Let $s : Q \to M$ be a section to p. Consider the homomorphism $g : N \oplus Q \to M$ that sends (n,q) to $i(n) + s(q)$. We claim that g is an isomorphism.

We begin with the injectivity of g. If $(n,q) \in \ker g$, i.e. $i(n) + s(q) = 0$, then $i(n) = -s(q) \in (\operatorname{im} i) \cap (\operatorname{im} s)$. Since $p \circ s = \mathrm{id}_Q$, the restriction of p to $\operatorname{im} s \to Q$ is injective. Thus we have

$$(\operatorname{im} i) \cap (\operatorname{im} s) \;=\; (\ker p) \cap (\operatorname{im} s) \;=\; \{m \in \operatorname{im} s \mid p(m) = 0\} \;=\; \{0\}.$$

This shows that $i(n) = s(q) = 0$. Since i and s are injective, we conclude that $(n,q) = (0,0)$, which shows that g is injective.

We continue with the surjectivity of g. By the first isomorphism theorem (Theorem 3.4.1), P induces an isomorphism $\bar{p} : M/\ker p \to Q$. Thus for $[m] \in M/\ker p$ with $m \in M$, we have

$$[m] \;=\; \bar{p}^{-1}\big(p(m)\big) \;=\; \bar{p}^{-1}\big(\underbrace{p \circ s}_{=\mathrm{id}_Q} \circ p(m)\big) \;=\; [s \circ p(m)],$$

and thus $m \in s(Q) + i(N) = g(N \oplus Q)$. This shows that g is surjective.

Thus g is an isomorphism and has an inverse $f : M \to N \oplus Q$. This yields the following diagram:

$$\begin{array}{ccccccccc}
0 & \longrightarrow & N & \xrightarrow{i} & M & \underset{s}{\overset{p}{\rightleftarrows}} & Q & \longrightarrow & 0 \\
 & & \downarrow{\scriptstyle \mathrm{id}_N} & & \uparrow{\scriptstyle g}\;\downarrow{\scriptstyle f} & & \downarrow{\scriptstyle \mathrm{id}_Q} & & \\
0 & \longrightarrow & N & \xrightarrow{\iota_N} & N \oplus Q & \xrightarrow{\pi_Q} & Q & \longrightarrow & 0
\end{array}$$

By the definition of g, we have $g \circ \iota_N = i \circ \mathrm{id}_N^{-1}$ and $p \circ g = \mathrm{id}_Q^{-1} \circ \pi_Q$. Thus

$$\iota_N \circ \mathrm{id}_N \;=\; f \circ g \circ \iota_N \circ \mathrm{id}_N \;=\; f \circ i \circ \mathrm{id}_N^{-1} \circ \mathrm{id}_N \;=\; f \circ i$$

and

$$\mathrm{id}_Q \circ p \;=\; \mathrm{id}_Q \circ p \circ g \circ f \;=\; \mathrm{id}_Q \circ \mathrm{id}_Q^{-1} \circ \pi_Q \circ f \;=\; \pi_Q \circ f.$$

This shows that the sequence $0 \to N \to M \to Q \to 0$ splits, which establishes (1) and concludes the proof of the theorem. □

Example 3.6.6. We discuss the concept of split short exact sequences in some examples.

(1) The prototype of a split exact sequence is of the form

$$0 \longrightarrow N \xrightarrow{\iota_N} N \oplus Q \xrightarrow{\pi_Q} Q \longrightarrow 0$$

where N and Q are two A-modules.

(2) Consider the two short exact sequences

$$0 \longrightarrow \mathbb{Z}/2\mathbb{Z} \xrightarrow{i} \mathbb{Z}/4\mathbb{Z} \xrightarrow{p} \mathbb{Z}/2\mathbb{Z} \longrightarrow 0$$

and

$$0 \longrightarrow \mathbb{Z}/2\mathbb{Z} \xrightarrow{i} \mathbb{Z}/6\mathbb{Z} \xrightarrow{p} \mathbb{Z}/3\mathbb{Z} \longrightarrow 0$$

of $\mathbb{Z}$-modules from Example 3.6.3. The first sequence is not split since $\mathbb{Z}/4\mathbb{Z}$ is not isomorphic to $(\mathbb{Z}/2\mathbb{Z}) \oplus (\mathbb{Z}/2\mathbb{Z})$. The second sequence is split since the surjection $p : \mathbb{Z}/6\mathbb{Z} \to \mathbb{Z}/3\mathbb{Z}$ has a section, which is the homomorphism $s : \mathbb{Z}/3\mathbb{Z} \to \mathbb{Z}/6\mathbb{Z}$ with $s(\bar{a}) = 2\bar{a}$ for $a \in \{\bar{0}, \bar{1}, \bar{2}\}$, using Theorem 3.6.5. This yields a new proof for the fact that $\mathbb{Z}/6\mathbb{Z} \simeq (\mathbb{Z}/2\mathbb{Z}) \oplus (\mathbb{Z}/3\mathbb{Z})$.

(3) As the following example shows, it is not enough to require that $M \simeq N \oplus Q$ for a short exact sequence $0 \to N \to M \to Q \to 0$ to split. The sequence

$$0 \longrightarrow \mathbb{Z} \xrightarrow{\ i\ } \mathbb{Z} \oplus \bigoplus_{i \geqslant 1} \mathbb{Z}/2\mathbb{Z} \xrightarrow{\ p\ } \bigoplus_{i \geqslant 0} \mathbb{Z}/2\mathbb{Z} \longrightarrow 0$$

with $i(a) = (2a, 0, 0, \dots)$ and $p(a_0, \bar{a}_1, \bar{a}_2, \dots) = (\bar{a}_0, \bar{a}_1, \bar{a}_2, \dots)$ is easily seen to be exact, and the middle term is isomorphic to the direct sum of the two outer terms. Still, the sequence is not split since a section s to p must, in particular, map the element $(\bar{1}, \bar{0}, \bar{0}, \dots)$ to an element $(a, \bar{0}, \bar{0}, \dots)$ with $a \in \mathbb{Z}$ odd. This would define a non-zero homomorphism $\mathbb{Z}/2\mathbb{Z} \to \mathbb{Z}$, which does not exist.

Definition 3.6.7. Let M be an A-module, and N a submodule of M. A **complement of N in M** is a submodule Q of M such that the homomorphism

$$\begin{array}{ccc} N \oplus Q & \longrightarrow & M \\ (n, q) & \longmapsto & n + q \end{array}$$

is an isomorphism.

Remark. Equivalently, a submodule Q of M is a complement of N in M if they intersect trivially, i.e. $N \cap Q = \{0\}$, and if they generate M, i.e. $N + Q = M$. We also write $M = N \oplus Q$ if these two conditions hold.

Corollary 3.6.8. *Let M be an A-module and N a submodule of M. Then N has a complement in M if and only if there is a retract $r : M \to N$ to the inclusion map $i : N \to M$, i.e. $r \circ i = \mathrm{id}_N$.*

Proof. Assume that N has a complement $Q \subset M$. Then $M = N \oplus Q$ and the inclusion $i : N \to M$ is equal to the canonical inclusion $\iota_N : N \to N \oplus Q = M$, for which that the canonical projection $\pi_N : M = N \oplus Q \to N$ is a retract.

Conversely, assume $r : M \to N$ is a retract to i. Let $\pi : M \to M/N$ be the quotient map. Then the short exact sequence

$$0 \longrightarrow N \underset{i}{\overset{r}{\leftrightarrows}} M \xrightarrow{\ \pi\ } M/N \longrightarrow 0$$

splits, and there is a section $s : M/N \to M$ to π by Theorem 3.6.5. Thus $M = N \oplus Q$ for $Q = \mathrm{im}\, s$, which shows that N has a complement in M. □

3.7 Exact functors

Throughout this section, we let A and B be rings. Let M and N be A-modules. Then the set $\mathrm{Hom}_A(M, N)$ of A-linear maps from M to N is an A-module with respect to valuewise addition and scalar multiplications, which are defined by the rules

$$(f + g)(m) \;=\; f(m) + g(m) \qquad \text{and} \qquad (a.f)(m) \;=\; a.\big(f(m)\big)$$

for $m \in M$, $a \in A$ and $f, g \in \mathrm{Hom}_A(M,N)$. The verification that these operations turn $\mathrm{Hom}_A(M,N)$ indeed into an A-module are left as Exercise 3.4.

Definition 3.7.1. A covariant functor $\mathcal{F} : \mathrm{Mod}_A \to \mathrm{Mod}_B$ is **additive** if the map

$$\mathrm{Hom}_A(M,N) \quad \longrightarrow \quad \mathrm{Hom}_B\big(\mathcal{F}(M), \mathcal{F}(N)\big)$$

is a group homomorphism for all A-modules M and N. A contravariant functor $\mathcal{F} : \mathrm{Mod}_A \to \mathrm{Mod}_B$ is **additive** if the map

$$\mathrm{Hom}_A(M,N) \quad \longrightarrow \quad \mathrm{Hom}_B\big(\mathcal{F}(N), \mathcal{F}(M)\big)$$

is a group homomorphism for all A-modules M and N.

In the following, we sometimes do not specify whether a functor is covariant or contravariant if we do not have to consider the direction of the image morphisms explicitly. In these cases, we consider both variants of a functor.

Lemma 3.7.2. *Let $\mathcal{F} : \mathrm{Mod}_A \to \mathrm{Mod}_B$ be an additive functor. Then the following holds true.*

(1) Let 0 be the trivial A-module. Then $\mathcal{F}(0)$ is the trivial B-module.

(2) Let $\mathbf{0} : M \to N$ be the zero map between two A-modules M and N. Then $\mathcal{F}(\mathbf{0})$ is the zero map.

Proof. We explain the proof in the case of a covariant functor. The case of a contravariant functor is analogous. Claim (2) is clear since $\mathrm{Hom}_A(M,N) \to \mathrm{Hom}_B\big(\mathcal{F}(M), \mathcal{F}(N)\big)$ is a group homomorphism.

Since the identity map $\mathrm{id}_0 : 0 \to 0$ is equal to the zero map, claim (2) implies that the identity map $\mathrm{id}_{\mathcal{F}(0)} : \mathcal{F}(0) \to \mathcal{F}(0)$ is the zero map. This can only be the case if $\mathcal{F}(0)$ is the trivial B-module. □

Example 3.7.3. We discuss some examples of additive functors.

(0) The identity functor $\mathrm{id} : \mathrm{Mod}_A \to \mathrm{Mod}_A$ that sends every A-module and every A-linear map to itself and the zero functor $\mathbf{0} : \mathrm{Mod}_A \to \mathrm{Mod}_B$ that sends every A-module to the trivial B-module 0 and every A-linear map $f : M \to N$ to the zero map $\mathbf{0} : 0 \to 0$ are additive.

(1) Let P be an A-module. Let $\mathrm{Hom}_A(P,-) : \mathrm{Mod}_A \to \mathrm{Mod}_A$ be the covariant functor that sends an A-module M to $\mathrm{Hom}_A(P,M)$ and an A-linear map $f : M \to N$ to the map $f_* : \mathrm{Hom}_A(P,M) \to \mathrm{Hom}_A(P,N)$ that sends $h : P \to M$ to $f \circ h : P \to N$. We leave the proof that this defines indeed a functor as Exercise 3.5. This functor is additive since $(f+g) \circ h = (f \circ h) + (g \circ h)$ for all A-linear maps $f, g : M \to N$ and $h \in \mathrm{Hom}_A(P,M)$. The proof of this equality is left as Exercise 3.4.

Note that $\mathrm{Hom}_A(A,-)$ is "essentially equal" to the identity functor $\mathrm{id} : \mathrm{Mod}_A \to \mathrm{Mod}_A$, by which we mean that $\mathrm{Hom}_A(A,M)$ is canonically isomorphic to M. The functor $\mathrm{Hom}_A(\mathbf{0},-)$ is essentially equal to the zero functor $\mathbf{0} : \mathrm{Mod}_A \to \mathrm{Mod}_A$.

(2) Let P be an A-module. Similar to $\mathrm{Hom}_A(P,-)$, we have a contravariant functor $\mathrm{Hom}_A(-,P):\mathrm{Mod}_A\to\mathrm{Mod}_A$ that sends an A-module M to $\mathrm{Hom}_A(M,P)$ and an A-linear map $f:M\to N$ to the map $f^*:\mathrm{Hom}_A(N,P)\to\mathrm{Hom}_A(M,P)$ that sends $h:N\to P$ to $h\circ f:M\to P$. Again, we leave the proof that this defines indeed a functor as Exercise 3.5. This functor is additive since $h\circ(f+g)=(h\circ f)+(h\circ g)$ for all A-linear maps $f,g:M\to N$ and $h\in\mathrm{Hom}_A(N,P)$. The proof is left as Exercise 3.4.

Note that if $A=K$ is a field, then the functor $\mathrm{Hom}_K(-,K)$ takes a K-vector space V to its dual space V^* and a K-linear map $f:V\to W$ to its adjoint $f^*:V^*\to W^*$.

(3) Let P be an A-module and $P\otimes_A-:\mathrm{Mod}_A\to\mathrm{Mod}_A$ the covariant functor that sends an A-module M to $P\otimes_A M$ and an A-linear map $f:M\to N$ to the map $f_P:P\otimes_A M\to P\otimes_A N$ that maps $p\otimes m$ to $p\otimes f(m)$ for all $p\in P$ and $m\in M$. We leave the proof that this is indeed a functor as Exercise 3.5. It is additive since

$$\begin{aligned}(f+g)_P(p\otimes m) &= p\otimes\big((f+g)(m)\big)\\ &= \big(p\otimes f(m)\big)+\big(p\otimes g(m)\big)\\ &= \big(f_P\big)+\big(g_P\big)\big)(p\otimes m)\end{aligned}$$

for all $p\in P$, $m\in M$ and $f,g\in\mathrm{Hom}_A(M,N)$.

(4) Let $f:A\to B$ be a ring homomorphism. The *restriction of scalars* is the covariant functor $\mathrm{Res}_{B/A}:\mathrm{Mod}_B\to\mathrm{Mod}_A$ that sends a B-module M to itself, but considered as an A-module with respect to the A-action given $a.m=f(a).m$ for $a\in A$ and $m\in M$, and that sends a B-linear map $f:M\to N$ to itself. We leave the proof that this defines indeed a functor as Exercise 3.6. This defines an inclusion $\mathrm{Hom}_B(M,N)\to\mathrm{Hom}_A(\mathrm{Res}_{B/A}(M),\mathrm{Res}_{B/A}(N))$, which is tautologically a group homomorphism since $\mathrm{Res}_{B/A}(M)=M$ and $\mathrm{Res}_{B/A}(N)=N$ as commutative groups. Thus the restriction of scalars is an additive functor.

(5) Let $f:A\to B$ be a ring homomorphism. The *extension of scalars* is the covariant functor $B\otimes_A-:\mathrm{Mod}_A\to\mathrm{Mod}_B$ that sends an A-module M to the B-module $B\otimes_A M$ whose action is given by $a.(b\otimes m)=(ab)\otimes m$ for $a,b\in B$ and $m\in M$ and that sends an A-linear map $f:M\to N$ to the map $f_B:B\otimes_A M\to B\otimes_A N$ that sends $a\otimes m$ to $a\otimes f(m)$ for $a\in B$ and $m\in M$. We leave the proof that this is indeed a functor as Exercise 3.6. The functor $B\otimes_A-$ is additive for the same reason as $P\otimes_A-$ from example (3) is additive.

Definition 3.7.4. An additive covariant functor $\mathcal{F}:\mathrm{Mod}_A\to\mathrm{Mod}_B$ is **left exact** (**right exact**) if for every short exact sequence $0\to N\xrightarrow{i}M\xrightarrow{p}Q\to 0$ of A-modules, the induced sequence

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \mathcal{F}(N) & \xrightarrow{\mathcal{F}(i)} & \mathcal{F}(M) & \xrightarrow{\mathcal{F}(p)} & \mathcal{F}(Q) & & \\ \Big(& & \mathcal{F}(N) & \xrightarrow{\mathcal{F}(i)} & \mathcal{F}(M) & \xrightarrow{\mathcal{F}(p)} & \mathcal{F}(Q) & \longrightarrow & 0 \quad \Big) \end{array}$$

is exact. An additive contravariant functor $\mathcal{F}: \mathrm{Mod}_A \to \mathrm{Mod}_B$ is **left exact** (**right exact**) if for every short exact sequence $0 \to N \xrightarrow{i} M \xrightarrow{p} Q \to 0$ of A-modules, the induced sequence

$$\begin{pmatrix} 0 \longrightarrow \mathcal{F}(Q) \xrightarrow{\mathcal{F}(p)} \mathcal{F}(M) \xrightarrow{\mathcal{F}(i)} \mathcal{F}(N) & \\ \mathcal{F}(Q) \xrightarrow{\mathcal{F}(p)} \mathcal{F}(M) \xrightarrow{\mathcal{F}(i)} \mathcal{F}(N) \longrightarrow 0 & \end{pmatrix}$$

is exact. A functor $\mathcal{F}: \mathrm{Mod}_A \to \mathrm{Mod}_B$ is exact if it is both left and right exact.

Example 3.7.5. Both the identity $\mathrm{id}: \mathrm{Mod}_A \to \mathrm{Mod}_A$ and the zero functor $\mathbf{0}: \mathrm{Mod}_A \to \mathrm{Mod}_B$ are tautologically exact functors.

Note that there are additive functors that are neither left nor right exact. For example the covariant functor $\mathcal{F}: \mathrm{Mod}_{\mathbb{Z}} \to \mathrm{Mod}_{\mathbb{Z}}$ that sends a $\mathbb{Z}$-module M to $2M = \{2m \in M \mid m \in M\}$ and a homomorphism $f: M \to N$ to its restriction $f|_{2M}: 2M \to 2N$ is additive. However, $\mathcal{F}$ sends the short exact sequence $0 \to \mathbb{Z}/2\mathbb{Z} \to \mathbb{Z}/4\mathbb{Z} \to \mathbb{Z}/2\mathbb{Z} \to 0$ from Example 3.6.3 to the sequence $0 \to 0 \to \mathbb{Z}/2\mathbb{Z} \to 0 \to 0$, which is not exact at the middle term $\mathbb{Z}/2\mathbb{Z}$.

Proposition 3.7.6. *Let P be an A-module. Then both $\mathrm{Hom}_A(P,-)$ and $\mathrm{Hom}_A(-,P)$ are left exact.*

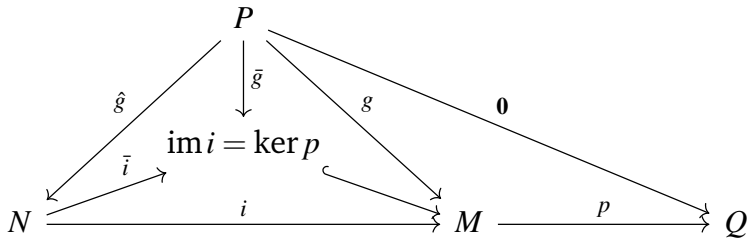
Proof. Let $0 \to N \xrightarrow{i} M \xrightarrow{p} Q \to 0$ be a short exact sequence of A-modules. We first consider the sequence

$$0 \longrightarrow \mathrm{Hom}_A(P,N) \xrightarrow{i_*} \mathrm{Hom}_A(P,M) \xrightarrow{p_*} \mathrm{Hom}_A(P,Q),$$

which is exact if i_* is injective and $\mathrm{im}\, i_* = \ker p_*$.

We begin with the injectivity of i_*. Consider two homomorphisms $f, g \in \mathrm{Hom}_A(P,N)$ such that $i_*(f) = i_*(g)$. Since i is a monomorphism, the equality $i \circ f = i \circ g$ implies that $f = g$. Thus i_* is injective.

We continue with showing that $\mathrm{im}\, i_* = \ker p_*$. Since $p \circ i = \mathbf{0}$ and zero maps are preserved under additive functors by Lemma 3.7.2, we conclude that $p_* \circ i_* = \mathbf{0}$, which shows that $\mathrm{im}\, i_*$ is contained in $\ker p_*$. In order to show the converse inclusion, we consider a homomorphism $g: P \to M$ in the kernel of p_*, i.e. $p \circ g = p_*(g) = \mathbf{0}$. This means that $\mathrm{im}\, g$ is contained in $\ker p$ and restricts to a homomorphism $\bar{g}: P \to \ker p$. Similarly, the injection $i: N \to M$ restricts to an isomorphism $\bar{i}: N \to \mathrm{im}\, i$. This yields the commutative diagram



where $\hat{g} = \bar{i}^{-1} \circ \bar{g}$. This shows that $g = i \circ \hat{g} = i_*(\hat{g})$ is in the image of i_*. Thus $\operatorname{im} i_* = \ker p_*$, which concludes the proof that $\operatorname{Hom}_A(P,-)$ is left exact.

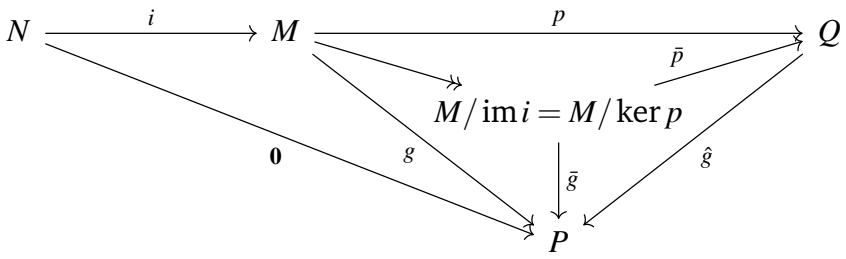
We continue to consider the sequence

$$0 \longrightarrow \operatorname{Hom}_A(Q,P) \xrightarrow{p^*} \operatorname{Hom}_A(M,P) \xrightarrow{i^*} \operatorname{Hom}_A(N,P),$$

which is exact if p^* is injective and $\operatorname{im} p^* = \ker i^*$.

We begin with the injectivity of p^*. Consider two homomorphisms $f, g \in \operatorname{Hom}_A(Q,P)$ such that $p^*(f) = p^*(g)$. Since p is an epimorphism, the equality $f \circ p = g \circ p$ implies that $f = g$. Thus p^* is injective.

We continue with showing that $\operatorname{im} p^* = \ker i^*$. Since $p \circ i = \mathbf{0}$ and zero maps are preserved under additive functors by Lemma 3.7.2, we conclude that $i^* \circ p^* = \mathbf{0}$, which shows that $\operatorname{im} p^*$ is contained in $\ker i^*$. In order to show the converse inclusion, we consider a homomorphism $g : M \to P$ in the kernel of i^*, i.e. $g \circ i = i^*(g) = \mathbf{0}$. This means that $\operatorname{im} i$ is contained in $\ker g$, and thus g factors into the projection $M \to M/\operatorname{im} i$, followed by a homomorphism $\bar{g} : M/N \to P$. By the first isomorphism theorem (Theorem 3.4.1), the surjection $p : M \to Q$ restricts to an isomorphism $\bar{p} : M/\ker p \to Q$. This yields the commutative diagram



where $\hat{g} = \bar{g} \circ \bar{p}^{-1}$. This shows that $g = \hat{g} \circ p = p^*(\hat{g})$ is in the image of p^*. Thus $\operatorname{im} p^* = \ker i^*$, which concludes the proof that $\operatorname{Hom}_A(-,P)$ is left exact. □

Remark. The functors $\operatorname{Hom}_A(P,-)$ and $\operatorname{Hom}_A(-,P)$ are in general not right exact. For instance, let $A = \mathbb{Z}$ and $P = \mathbb{Z}/n\mathbb{Z}$. Consider the short exact sequence

$$0 \longrightarrow \mathbb{Z} \xrightarrow{i} \mathbb{Z} \xrightarrow{p} \mathbb{Z}/n\mathbb{Z} \longrightarrow 0$$

of $\mathbb{Z}$-modules where $n \geqslant 2$ is an integer and where $i(a) = na$ and $p(a) = \bar{a}$ for $a \in \mathbb{Z}$. If we apply $\operatorname{Hom}_{\mathbb{Z}}(\mathbb{Z}/n\mathbb{Z},-)$ to p, we get

$$p_* : \operatorname{Hom}_{\mathbb{Z}}(\mathbb{Z}/n\mathbb{Z},\mathbb{Z}) \longrightarrow \operatorname{Hom}_{\mathbb{Z}}(\mathbb{Z}/n\mathbb{Z},\mathbb{Z}/n\mathbb{Z}),$$

which is not surjective since the only homomorphism $\mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}$ is the zero map, but there are homomorphisms $\mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ that are not zero. This shows that $\operatorname{Hom}_{\mathbb{Z}}(\mathbb{Z}/n\mathbb{Z},-)$ is not exact.

If we apply $\operatorname{Hom}_{\mathbb{Z}}(-,\mathbb{Z}/n\mathbb{Z})$ to i, the map

$$i^* : \operatorname{Hom}_{\mathbb{Z}}(\mathbb{Z},\mathbb{Z}/n\mathbb{Z}) \longrightarrow \operatorname{Hom}_{\mathbb{Z}}(\mathbb{Z},\mathbb{Z}/n\mathbb{Z}),$$

maps every homomorphism $f : \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ to $f \circ i$, which is the zero map since for all $a \in \mathbb{Z}$, we have $f \circ i(a) = f(n.a) = n.f(a) = 0$ in $\mathbb{Z}/n\mathbb{Z}$. This shows that i^* is not surjective and thus $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}/n\mathbb{Z}, -)$ not exact.

If A is a ring and P an A-module such that $\mathrm{Hom}_A(P, -)$ is exact, then P is called *projective*. If $\mathrm{Hom}_A(-, P)$ is exact, then P is called *injective*.

Lemma 3.7.7. *Let $i : N \to M$ and $p : M \to Q$ be homomorphisms of A-modules.*

(1) If the sequence

$$0 \longrightarrow \mathrm{Hom}_A(P,N) \xrightarrow{p_*} \mathrm{Hom}_A(P,M) \xrightarrow{i_*} \mathrm{Hom}_A(P,Q)$$

is exact for every A-module P, then $0 \to N \xrightarrow{i} M \xrightarrow{p} Q$ is exact.

(2) If the sequence

$$0 \longrightarrow \mathrm{Hom}_A(Q,P) \xrightarrow{p^*} \mathrm{Hom}_A(M,P) \xrightarrow{i^*} \mathrm{Hom}_A(N,P)$$

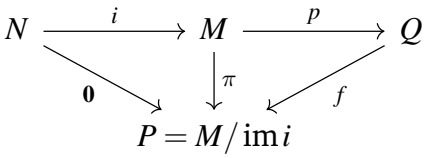
is exact for every A-module P, then $N \xrightarrow{i} M \xrightarrow{p} Q \to 0$ is exact.

Proof. We only prove (2). The proof of (1) is analogous and left as an exercise. We begin with the surjectivity of p. Let $\pi : Q \to Q/\operatorname{im} p$ be the quotient map and consider the hypothesis of (2) for $P = Q/\operatorname{im} p$. Then $p^*(\pi) = \pi \circ p : M \to Q \to P$ is the zero map. Since p^* is injective, we conclude that π is the zero map. Since π is surjective, this means that $P = Q/\operatorname{im} p = 0$ and thus $\operatorname{im} p = Q$. This shows that p is surjective.

We continue with $\operatorname{im} i = \ker p$. Using the hypothesis of (2) for $P = Q$ yields

$$p \circ i \;=\; \mathrm{id}_Q \circ p \circ i \;=\; i^*(\mathrm{id}_Q \circ p) \;=\; \underbrace{i^* \circ p^*}_{=\mathbf{0}}(\mathrm{id}_Q) \;=\; \mathbf{0},$$

which shows that $\operatorname{im} i \subset \ker p$. To establish the converse inclusion, we consider the quotient map $\pi : M \to M/\operatorname{im} i$ and use the hypothesis of (2) for $P = M/\operatorname{im} i$. Then $i^*(\pi) = \pi \circ i : N \to P$ is the zero map, i.e. $\pi \in \ker i^* = \operatorname{im} p^*$. Thus $\pi = p^*(f)$ for some homomorphism $f : Q \to P$, which can be illustrated follows:



We conclude that $\ker p \subset \ker \pi = \operatorname{im} i$, which completes the proof. □

Proposition 3.7.8. *Let P be an A-module. The functor $P \otimes_A - : \mathrm{Mod}_A \to \mathrm{Mod}_A$ is right exact.*

Proof. Let M and N be A-modules. Then map

$$\Psi_{M,N}: \quad \mathrm{Hom}_A\big(M, \mathrm{Hom}_A(P,N)\big) \quad \longrightarrow \quad \mathrm{Hom}_A(P \otimes_A M, N)$$

that sends a homomorphism $f : M \to \mathrm{Hom}_A(P,N)$ to the homomorphism $\Psi_{M,N}(f) : P \otimes_A M \to N$ that maps $p \otimes m$ to $\big(f(m)\big)(p)$ for $p \in P$ and $m \in M$ is a well-defined isomorphism of A-modules. Moreover, the functor $\mathrm{Hom}_A(P,-) : \mathrm{Mod}_A \to \mathrm{Mod}_A$ is right adjoint to $P \otimes_A -$. We leave the details as Exercise 3.15.

This means that given a short exact sequence $0 \to N \xrightarrow{i} M \xrightarrow{p} Q \to 0$ of A-modules, we obtain a commutative diagram

$$\begin{array}{ccccccc}
0 & \longrightarrow & \mathrm{Hom}_A(Q, \mathcal{G}(P')) & \xrightarrow{p^*} & \mathrm{Hom}_A(M, \mathcal{G}(P')) & \xrightarrow{i^*} & \mathrm{Hom}_A(N, \mathcal{G}(P')) \\
 & & \downarrow{\scriptstyle \Psi_{Q,\mathcal{G}(P')}} & & \downarrow{\scriptstyle \Psi_{M,\mathcal{G}(P')}} & & \downarrow{\scriptstyle \Psi_{N,\mathcal{G}(P')}} \\
0 & \longrightarrow & \mathrm{Hom}_A(P \otimes_A Q, P') & \xrightarrow{(p_P)^*} & \mathrm{Hom}_A(P \otimes_A M, P') & \xrightarrow{(i_P)^*} & \mathrm{Hom}_A(P \otimes_A N, P')
\end{array}$$

for every A-module P' and $\mathcal{G}(P') = \mathrm{Hom}_A(P, P')$.

By Proposition 3.7.6, the functor $\mathrm{Hom}(-, \mathcal{G}(P'))$ is left exact, which shows that the upper row of the diagram is exact. Since all the vertical maps are isomorphisms, this implies that the lower row is exact. Since this holds for every A-module P', Lemma 3.7.7 implies that the sequence

$$P \otimes_A N \xrightarrow{i_P} P \otimes_A M \xrightarrow{p_P} P \otimes_A Q \longrightarrow 0$$

is exact. This shows that $P \otimes_A -$ is right exact and concludes the proof. □

Remark. The functor $P \otimes_A -$ is in general not exact since it does not preserve injectivity. Indeed consider $A = \mathbb{Z}$ and the injective homomorphism $i : \mathbb{Z} \to \mathbb{Z}$ that maps a to na for some fixed integer $n \geqslant 2$. Then the induced homomorphism $i_{\mathbb{Z}/n\mathbb{Z}} : (\mathbb{Z}/n\mathbb{Z}) \otimes_{\mathbb{Z}} \mathbb{Z} \to (\mathbb{Z}/n\mathbb{Z}) \otimes_{\mathbb{Z}} \mathbb{Z}$ maps the nonzero element $\bar{1} \otimes 1$ to

$$\bar{1} \otimes n \;=\; \bar{n} \otimes 1 \;=\; \bar{0} \otimes 1 \;=\; \bar{0} \otimes 0,$$

which is the zero in $(\mathbb{Z}/n\mathbb{Z}) \otimes_{\mathbb{Z}} \mathbb{Z}$. Thus $i_{\mathbb{Z}/n\mathbb{Z}}$ is not injective.

The following property can be deduced from the fact that $P \otimes_A -$ is right exact.

Proposition 3.7.9. *Let I be an ideal of A and M an A-module. Define*

$$IM \;=\; \langle a.m | a \in I, m \in M \rangle_A \;=\; \Big\{ \sum a_i.m_i \,\Big|\, a_i \in I, m_i \in M \Big\},$$

which is a submodule of M. Then the map

$$\begin{array}{rccc}
f: & M \otimes_A (A/I) & \longrightarrow & M/IM \\
 & m \otimes [a] & \longmapsto & [a.m]
\end{array}$$

is an isomorphism of A-modules.

Proof. The submodule I of A defines a short exact sequences $0 \to I \xrightarrow{i} A \xrightarrow{p} A/I \to 0$ of A-modules, and similarly the submodule IM of M defines a short exact sequences $0 \to IM \xrightarrow{j} M \xrightarrow{q} M/IM \to 0$. By Proposition 3.7.8, applying $M \otimes_A -$ to the first of these short exact sequences yields an exact sequence

$$M \otimes_A I \xrightarrow{i_M} M \otimes_A A \xrightarrow{p_M} M \otimes_A (A/I) \longrightarrow 0.$$

The map $M \times I \to IM$ with $(m,a) \mapsto a.m$ is bilinear and defines thus a homomorphism $h : M \otimes_A I \to IM$ with $h(m \otimes a) = a.m$, which is surjective. The association $m \otimes a \mapsto a.m$ extends to the canonical isomorphism $g : M \otimes_A A \to M$. Since

$$(q \circ g) \circ i_M = \underbrace{q \circ j}_{=\mathbf{0}} \circ h = \mathbf{0},$$

we have $(q \circ g)(\operatorname{im} i_M) = \{0\}$. Since

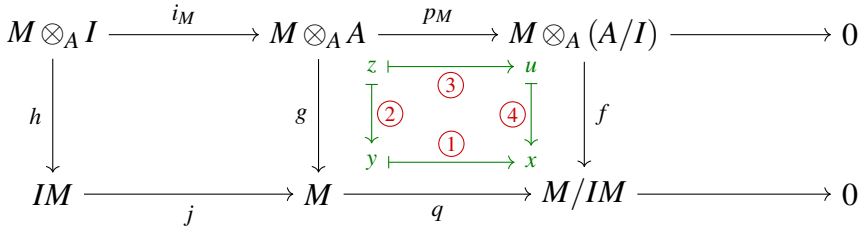
$$M \otimes_A (A/I) = \operatorname{im} p_M \simeq (M \otimes_A A)/\ker p_M = (M \otimes_A A)/\operatorname{im} i_M$$

by the first isomorphism theorem (Theorem 3.4.1), the universal property of quotients (Proposition 3.2.2) implies that the map $q \circ g : M \otimes_A A \to M/IM$ factors into p_M composed with the morphism $f : M \otimes_A (A/I) \to M/IM$ that maps $m \otimes [a]$ to $[a.m]$. This shows that f is well-defined and yields the commutative diagram

$$\begin{array}{ccccccc} M \otimes_A I & \xrightarrow{i_M} & M \otimes_A A & \xrightarrow{p_M} & M \otimes_A (A/I) & \longrightarrow & 0 \\ \downarrow{\scriptstyle h} & & \downarrow{\scriptstyle g} & & \downarrow{\scriptstyle f} & & \\ IM & \xrightarrow{j} & M & \xrightarrow{q} & M/IM & \longrightarrow & 0 \end{array}$$

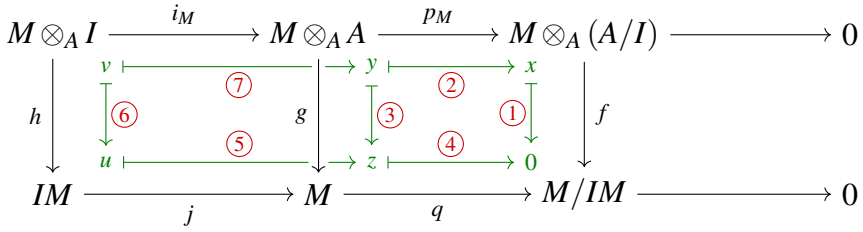
with exact rows. That f is an isomorphism can be proven by a "diagram chase", which we will demonstrate in the following. We will refer to the individual steps of the proof by red circled numbers in the illustrations.

We begin with the surjectivity of f, which can be proven along the following steps.



Let x be an element of M/IM. ① Since q is surjective, there is an element $y \in M$ such that $x = q(y)$. ② Since g is surjective, $y = g(z)$ for some $z \in M \otimes_A A$. ③ Let $u = p_M(z)$. ④ Then $f(u) = f \circ p_M(z) = q \circ g(z) = x$, which shows that f is surjective.

We continue with the injectivity of f, which can be proven along the following steps.



Let x be an element of the kernel of f. ① Then $f(x) = 0$. ② Since p_M is surjective, there is an element y in $M \otimes_A A$ with $p_M(y) = x$. ③ Let $z = g(y)$. ④ Since $q(z) = g \circ q(y) = f \circ p_M(y) = f(x) = 0$, we have $z \in \ker q$. ⑤ Since $\ker q = \operatorname{im} j$, we have $z = j(u)$ for some element $u \in IM$. ⑥ Since h is surjective, $u = h(v)$ for some $v \in M \otimes_A I$. ⑦ Since g is an isomorphism, we have $y = g^{-1}(z) = g^{-1} \circ j \circ h(v) = g^{-1} \circ g \circ i_M(v) = i_M(v)$. We conclude that $y \in \operatorname{im} i_M = \ker p_M$ and thus $x = p_M(y) = 0$. This shows that $\ker f = \{0\}$, which completes the proof of the injectivity of f. □

3.8 Free modules and torsion modules

Let A be a nontrivial ring, i.e. we assume $0 \neq 1$ throughout the section.

Definition 3.8.1. Let M be an A-module. A **basis of** M is a subset $\mathcal{B}$ of M such that the homomorphism

$$\begin{array}{ccc} \bigoplus_{s \in \mathcal{B}} A & \longrightarrow & M \\ (a_v)_{v \in \mathcal{B}} & \longmapsto & \sum_{v \in \mathcal{B}} a_v . v \end{array}$$

is an isomorphism, i.e. for every $m \in M$, there is a unique $(a_v) \in \bigoplus_{v \in \mathcal{B}} A$ such that $m = \sum a_v . v$. An A-module M is **free** if it has a basis. If M is a free A-module with basis $\mathcal{B}$, then the **rank of** M is defined as the cardinality $\operatorname{rk} M$ of $\mathcal{B}$.

Note that a basis is, in particular, a generating set. That the rank of a free A-module is well-defined follows from the following fact.

Proposition 3.8.2. *Let M be a free A-module. Then any two bases of M have the same cardinality.*

Proof. Let $\mathcal{B}$ be a basis of M and $f : \bigoplus_{v \in \mathcal{B}} A \to M$ the corresponding isomorphism. Let $\mathfrak{m}$ a maximal ideal of A and $K = A/\mathfrak{m}$ the residue field. Let $\pi_M : M \to M/\mathfrak{m}M$ be the quotient map where $\mathfrak{m}M = \langle a.m \mid a \in \mathfrak{m}, m \in M \rangle_A$ and $f_v : A \to M/\mathfrak{m}M$ the map with $f_v(a) = [a.v]$ for $v \in \mathcal{B}$. Since $f_v(a) = 0$ if $a \in \mathfrak{m}$, the universal property of quotient modules (Proposition 3.2.2) implies that f_v factors into the quotient map $\pi_A : A \to K$, followed by the homomorphism $\bar{f}_v : K \to M/\mathfrak{m}M$ that sends $[a]$ to $[a.v]$.

By Proposition 3.7.9, $M/\mathfrak{m}M \simeq M \otimes_A K$ is an K-module with respect to the action $[a].[m] = [a.m]$ for $[a] \in K$ and $[m] \in M/\mathfrak{m}M$. This means that $\bar{f}_v : K \to M/\mathfrak{m}M$ is, in fact, a K-linear map.

Let $\overline{\mathcal{B}} = \{[v] \mid v \in \mathcal{B}\}$. Adding the maps $\bar{f}_v$ for all $v \in \mathcal{B}$ yields the K-linear map $\bar{f} : \bigoplus_{[v]\in\overline{\mathcal{B}}} K \to M/\mathfrak{m}M$ that maps $([a_v])$ to $\sum[a_v.m]$, which is clearly surjective. Its injectivity can be shown as follows. Assume that $\sum[a_v.m] = \bar{f}([a_v]) = 0$, i.e. $\sum a_v.m = \sum \tilde{a}_v.v$ for some $\tilde{a}_v \in \mathfrak{m}$. Then $\tilde{a}_v = a_v$ since $\mathcal{B}$ is a basis. Thus $([a_v]) = ([0])$ in $\bigoplus_{[v]\in\overline{\mathcal{B}}} K$, which shows that $\bar{f}$ is injective. We conclude that $\bar{f} : \bigoplus_{[v]\in\overline{\mathcal{B}}} K \to M/\mathfrak{m}M$ is an isomorphism of K-vector spaces.

Since $v \mapsto [v]$ defines a bijection $\mathcal{B} \to \overline{\mathcal{B}}$, our claim follows from the fact that any two basis for the K-vector space $M/\mathfrak{m}M$ have the same cardinality. □

Example 3.8.3. We discuss some examples.

(1) Let $\mathcal{B}$ be a set. Then the A-module $\bigoplus_{v\in\mathcal{B}} A$ is tautologically free.

(2) Let K be a field. Then a basis of a K-vector space is the same thing as a basis in the sense of Definition 3.8.1. Since every K-vector space has a basis, every K-module is free.

(3) The polynomial ring $A[T]$, considered as an A-module, is free with bases $\mathcal{B} = \{1, T, T^2, \ldots, \}$.

(4) If $A \neq \{0\}$ is not a field, then A has a proper nonzero ideal I. In this case, the A-module $M = A/I$ is not free. Indeed if $M = \langle S\rangle_A$, then we note first that $S \neq \varnothing$ since I is proper, i.e. S contains an element $[s]$, which is the residue class of an element $s \in A$. By our assumptions, I contains a nonzero element a. Then $as \in I$ and thus $a.[s] = [as] = [0 \cdot s] = 0.[s]$, which shows that M is not free.

Definition 3.8.4. An element $a \in A$ is **regular** if it is not a zero divisor. Let M be an A-module. The **torsion submodule of** M is the subset

$$T(M) \;=\; \{m \in M \mid a.m = 0 \text{ for a regular element } a \in A\}$$

of M. An A-module M is **torsion-free** if $T(M) = \{0\}$ and M is a **torsion module** if $T(M) = M$.

Remark. Note that the product ab of regular elements a and b is regular since if ab is a zero divisor, then so is one of a and b. If A is an integral domain, then every nonzero element is regular.

Lemma 3.8.5. *Let M be an A-module and $T(M)$ its torsion submodule. Then the following holds true.*

(1) The subset $T(M)$ is a submodule of M.

(2) The quotient $M/T(M)$ is torsion-free.

(3) If M is free, then M is torsion-free.

Proof. We begin with (1). Let $m, n \in T(M)$ and $a, b \in A$ regular such that $a.m = b.n = 0$. Then ab is regular, and $(ab).(m+n) = (ab).m + (ab).n = 0$ shows that $m+n \in T(M)$. For $c \in A$, we have $(ca).m = c.(a.m) = 0$. This shows that $T(M)$ is a submodule. Thus (1).

We continue with (2). Let $[m] \in M/T(M)$ be the residue class of $m \in M$ and $a \in A$ regular. If $a.[m] = 0$, then $a.m \in T(M)$. Thus there is a regular $b \in A$ such that $(ba).m = b.(a.m) = 0$. Since ba is regular, this shows that $m \in T(M)$ i.e. $[m] = [0]$. We conclude that $M/T(M)$ is torsion-free. Thus (2).

We continue with (3). Consider $b.m = 0$ for $m \in M$ and $b \in A$ regular. Let $\mathcal{B}$ be a basis of M. Then $m = \sum a_v.v$ for a unique $(a_v) \in \bigoplus_{v\in\mathcal{B}} A$. Since $0 = b.m = \sum(ba_v).m$, we have $ba_v = 0$ for all $v \in \mathcal{B}$. Since b is regular, $a_v = 0$ for all $v \in \mathcal{B}$, which shows that $m = 0$. Thus M is torsion-free, which shows (3). □

Example 3.8.6. We discuss some examples.

(1) If I is an ideal of A that contains a regular element a, then as an A-module, A/I is a torsion module since $a.[m] = [a.m] = [0]$ for all $[m] \in A/I$. However, as an A/I-module, A/I is free and therefore torsion-free.

(2) The torsion submodule of a $\mathbb{Z}$-module M is the subset $T(M) = \{m \in M \mid a.m = 0 \text{ for some integer } m > 0\}$ of M.

(3) The additive group of $\mathbb{Q}$ is a torsion-free $\mathbb{Z}$-module that is not free. The proof is left as Exercise 3.18.

(4) Let K be a field, M a $K[T]$-module that is finite dimensional as a K-vector space and $m \in M$. Then the kernel of the homomorphism $f_m : K[T] \to M$ that sends a to $a.m$ is of infinite dimension over K and contains a regular element a. Thus $a.m = f_m(a) = 0$ shows that $m \in T(M)$. This shows that M is a torsion module.

3.9 Modules over principal ideal domains

In this section, we prove some profound theorems about modules over principal ideal domains: the elementary divisor theorem, the Smith normal form and the structure theorem for finitely generated modules over principal ideal domains. These three theorems are closely related and can be deduced from each other easily.

To begin with, we establish some auxiliary results. Throughout the rest of this chapter, we let A be a principal ideal domain.

Lemma 3.9.1. *Let M be a free A-module of finite rank r and N a submodule of M. Then N is free of rank $s \leqslant r$.*

Proof. We prove the claim by induction on r. If $r = 0$, there is nothing to prove.

Let $r > 0$. Let $\mathcal{B} = \{v_1, \ldots, v_r\}$ be a basis of M and $M' = \langle v_1, \ldots, v_{r-1}\rangle$, which is free of rank $r-1$. By the inductive hypothesis, $N' = N \cap M'$ is a submodule of M' that is free of rank $s' \leqslant r-1$. We have

$$\overline{N} \;=\; N/N' \;=\; N/(N\cap M') \;\xrightarrow{\sim}\; (N+M')/M' \;\subset\; M/M' \;\simeq\; A$$

by the second isomorphism theorem (Theorem 3.4.3). Under these inclusions and isomorphisms, $\overline{N}$ corresponds to an ideal $I = \langle a\rangle$ of A.

If $a = 0$, then $\overline{N} = \{0\}$, and $N = N'$ is free of rank $s = s' \leqslant r-1$, as desired. If $a \neq 0$, then multiplication by a defines an isomorphism $m_a : A \to \langle a \rangle$ of A-modules, which shows that $\overline{N} \simeq A$ as A-modules.

By Exercise 3.14, the free A-module $\overline{N}$ is projective and the short exact sequence $0 \to N' \to N \to \overline{N} \to 0$ splits, which shows that $N \simeq N' \oplus \overline{N}$. Thus N is free of rank $r = r' + 1 \leqslant (r-1) + 1 = r$, which concludes the proof of the lemma. □

Lemma 3.9.2. *Let $s \geqslant 0$ be an integer and $d_1, \ldots, d_s \in A$ such that $\langle 0 \rangle \neq \langle d_1 \rangle \subset \ldots \subset \langle d_s \rangle \neq \langle 1 \rangle$. Then s and the ideals $\langle d_1 \rangle, \ldots, \langle d_s \rangle$ are uniquely determined by the A-module $\prod_{i=1}^{s} A/\langle d_i \rangle$.*

Proof. Every prime ideal $\mathfrak{p}$ of A is principal, i.e. $\mathfrak{p} = \langle p \rangle$ for a prime element $p \in A$. Thus $\mathfrak{p}^k = \langle p^k \rangle$, which allows us to express the order of an element $d \in A$ in a prime ideal $\mathfrak{p}$ as $\mathrm{ord}_{\mathfrak{p}}(d) = \max\{k \mid d \in \mathfrak{p}^k\}$ and to write $\langle d \rangle = \prod \mathfrak{p}^{\mathrm{ord}_{\mathfrak{p}}(d)}$ where the product is taken over all prime ideals $\mathfrak{p}$ for which $\mathrm{ord}_{\mathfrak{p}}(d) \neq 0$; cf. section 1.7. Therefore the uniqueness of the ideals $\langle d_1 \rangle, \ldots, \langle d_s \rangle$ follows if we can express the orders $\mathrm{ord}_{\mathfrak{p}}(d_i)$ in terms of the A-module $M = \prod_{i=1}^{s} A/\langle d_i \rangle$ for every prime ideal $\mathfrak{p}$. We will achieve this as follows.

Let $p \in A$ be prime, $\mathfrak{p} = \langle p \rangle$ and $k \geqslant 0$ an integer. Then $\mathfrak{p}^k N = \{p^k.n \mid n \in N\}$ for an A-module N, and thus the A-linear map

$$\begin{array}{rccc} f_{p,k}: & A/\langle d \rangle & \longrightarrow & \mathfrak{p}^k(A/\langle d, p^{k+1} \rangle) \\ & \bar{a} & \longmapsto & p^k.\bar{a} \end{array}$$

is surjective. By Lemma 1.6.8, $\langle d, p^{k+1} \rangle = \gcd(d, p^{k+1})$, and thus $f_{p,k}(\bar{a}) = \bar{0}$ if and only if $p^k a \in \gcd(d, p^{k+1})$. Thus $\ker f_{p,k} = A/\langle d \rangle$ if $\mathrm{ord}_{\mathfrak{p}}(d) \leqslant k$ and $\ker f_{p,k} = \langle p \rangle + \langle d \rangle = \mathfrak{p}/\langle d \rangle$ if $\mathrm{ord}_{\mathfrak{p}}(d) > k$ where we use the notation $\mathfrak{p}/\langle d \rangle$ of the third isomorphism theorem (Theorem 3.4.4). Thus if $\mathrm{ord}_{\mathfrak{p}}(d) \leqslant k$, then $\mathfrak{p}^k(A/\langle d, p^{k+1} \rangle) = 0$, and otherwise

$$\mathfrak{p}^k(A/\langle d, p^{k+1} \rangle) \;\simeq\; (A/\langle d \rangle)/(\mathfrak{p}/\langle d \rangle) \;\simeq\; A/\mathfrak{p}.$$

Note that $k(\mathfrak{p}) = A/\mathfrak{p}$ is a field since $\mathfrak{p} = \langle p \rangle$ is a maximal ideal as a nonzero prime ideal in a principal ideal domain. Using Proposition 3.7.9 and Lemma 3.3.4.(4), we conclude that

$$\mathfrak{p}^k\big(M/\mathfrak{p}^{k+1}M\big) \;\simeq\; \mathfrak{p}^k\big(M \otimes_A (A/\mathfrak{p}^{k+1})\big) \;\simeq\; \prod_{i=1}^{s} \mathfrak{p}^k\Big(\big(A/\langle d_i \rangle\big) \otimes_A \big(A/\mathfrak{p}^{k+1}\big)\Big)$$
$$\simeq \prod_{i=1}^{s} \mathfrak{p}^k\big(A/\langle d_i, p^{k+1} \rangle\big)$$

is a $k(\mathfrak{p})$-vector space whose dimension equals the number of d_i's with $\mathrm{ord}_{\mathfrak{p}}(d_i) \geqslant k+1$. Since $\langle d_1 \rangle \subset \ldots \subset \langle d_s \rangle$, we have $\mathrm{ord}_{\mathfrak{p}}(d_s) \leqslant \cdots \leqslant \mathrm{ord}_{\mathfrak{p}}(d_1)$, and thus

$$v_{\mathfrak{p}}(d_i) \;=\; \min\Big\{k \in \mathbb{N} \;\Big|\; \dim_{k(\mathfrak{p})} \mathfrak{p}^k\big(M/\mathfrak{p}^{k+1}M\big) < i\Big\}.$$

This formula shows that $v_{\mathfrak{p}}(d_i)$ is determined by $M = \prod_{i=1}^{s} A/\langle d_i \rangle$. Note that since $\langle d_1 \rangle \neq \langle 1 \rangle$, the element d_1 is divisible by a prime element $p \in A$. Thus for $\mathfrak{p} = \langle p \rangle$,

$$s \;=\; \dim_{k(\mathfrak{p})}(M/\mathfrak{p}M)$$

determines s, which completes the proof of the lemma. □

The elementary divisor theorem

Theorem 3.9.3 (Elementary divisor theorem)**.** *Let M be a free A-module of finite rank r and N a submodule of M. Then there is a basis $\mathcal{B} = \{v_1, \dots, v_n\}$ of M, an $s \leqslant r$ and elements $d_1, \dots, d_s \in A$ such that $\{d_1.v_1, \dots, d_s.v_s\}$ is a basis for N and such that $\langle 0\rangle \neq \langle d_1\rangle \subset \dots \subset \langle d_s\rangle$. Moreover, the integer s and the ideals $\langle d_1\rangle, \dots, \langle d_s\rangle$ are uniquely determined by the submodule N of M.*

Remark. The elementary divisor theorem was first proven by Schering who calls the elements $d_1, \dots, d_s$ the elementary divisors of N. The term 'elementary divisor' was introduced by Weierstrass who used it, in a slightly different context, for prime powers as they appear in the structure theorem for finitely generated A-modules (Theorem 3.9.5). Throughout the literature, there is a certain inconsistency in the usage of the term 'elementary divisor', but nowadays it more common to use it for the alluded prime powers, and to call the elements $d_1, \dots, d_s$ the *invariants of* N. We shall follow this latter convention in our lecture notes.

Proof. We prove the existence claim by induction on r. The claim is trivial for $r = 0$.

Let $r > 0$. If $N = \{0\}$, the claim is trivial. Thus we may assume that N is not trivial. A homomorphism $f : M \to A$ of A-modules maps N to a submodule I_f of A, which is an ideal of A. Let $g : M \to A$ be a homomorphism such that I_g is maximal in $\{I_f \mid f : M \to A\}$ with respect to inclusion. Since A is a principal ideal domain, $I_g = \langle d_g\rangle$ for some $d_g \in I_g$, i.e. $d_g = g(v_g)$ for some $v_g \in N$.

We claim that $f(v_g) \in I_g$ for every A-linear map $f : M \to A$. Indeed, let $f : M \to A$ be an A-linear map, $a = f(v_g)$ and d a greatest common divisor of a and d_g. Then $\langle d\rangle = \gcd(a, d_g)$ contains both $\langle a\rangle$ and $I_g = \langle d_g\rangle$. By Lemma 1.6.8, we have $d = ba + cd_g$ for some $b, c \in A$. Thus $f'(v_g) = ba + cd_g = d$ for $f' = bf + cg$, which shows that $\langle d\rangle \subset I_{f'}$. Thus $I_g \subset \langle d\rangle \subset I_{f'}$, which must be equalities by the maximality of I_g. We conclude that $\langle a\rangle \subset \langle d\rangle = I_g$, which verifies our claim.

Next we claim that $I_g \neq \langle 0\rangle$. Indeed, choose a basis $\tilde{\mathcal{B}}$ for M and consider the A-linear maps

$$\begin{array}{rccc} f_w : & M & \longrightarrow & A \\ & \sum\limits_{v \in \tilde{\mathcal{B}}} a_v.v & \longmapsto & a_w \end{array}$$

for $w \in \tilde{\mathcal{B}}$. Since $N \neq \{0\}$ by assumption, it contains a nonzero element $n = \sum_{v \in \tilde{\mathcal{B}}} c_v.v$, i.e. $c_w \neq 0$ for some $w \in \mathcal{B}$. Then $f_w(n) = c_w \neq 0$, which shows that I_{f_w} properly contains $\langle 0\rangle$. Since I_g is maximal, it cannot be equal to $\langle 0\rangle$, which verifies our claim.

Write $v_g = \sum_{v \in \tilde{\mathcal{B}}} d_v.v$. Then $d_w = f_w(v_g) \in I_g = \langle d_g\rangle$, i.e. d_v is divisible by d_g for all $v \in \mathcal{B}$. Thus there is an element $w_g \in M$ such that $v_g = d_g.w_g$. This means that $d_g = g(v_g) = d_g \cdot g(w_g)$ and thus $g(w_g) = 1$ since $\langle d_g\rangle = I_g \neq \langle 0\rangle$.

Next we claim that the A-linear map

$$\begin{array}{rccc} h : & \langle w_g\rangle_A \oplus \ker g & \longrightarrow & M \\ & (a.w_g, m) & \longmapsto & a.w_g + m \end{array}$$

is an isomorphism. Indeed, h is injective since $g(a.w_g) = a \cdot g(w_g) = a$ and thus $\langle w_g\rangle_A \cap \ker g = \{a.w_g \in M \mid a = 0\} = \{0\}$. To show that surjectivity of h, we consider $m \in M$

and define $m' = m - g(m).w_g$. Then $g(m') = g(m) - g(m) \cdot g(w_g) = 0$, i.e. $m' \in \ker g$. Thus $m = h\big(g(m).w_g, m'\big)$, which shows that h is surjective and verifies our claim.

By Lemma 3.9.1, $M' = \ker g$ is a free submodule of M. Since $\langle w_g\rangle_A \cap N = \langle v_g\rangle_A$, by the maximality of g, we have $N = \langle v_g\rangle_A \oplus N'$ for the submodule $N' = \ker g \cap N$ of M'.

Since $r = \operatorname{rk} M$ equals $\operatorname{rk}(\langle w_g\rangle_A \oplus \ker g) = 1 + \operatorname{rk} M'$, we have $\operatorname{rk} M' = r - 1$. Thus by the inductive hypothesis, there is a basis $\mathcal{B}' = \{v_1, \ldots, v_{r-1}\}$ of M', an $s \leqslant r$ and elements $d_1, \ldots, d_{s-1} \in A$ such that $\{d_1.v_1, \ldots, d_{s-1}.v_{s-1}\}$ is a basis for N and such that $\langle d_1\rangle \subset \ldots \subset \langle d_{s-1}\rangle$. Since $d_1.v_1$ is contained in a basis, we have $d_1 \neq 0$ and thus $\langle 0\rangle \neq \langle d_1\rangle$. Let $v_s = w_g$ and $d_s = d_g$. Then $\mathcal{B} = \{v_1, \ldots, v_s\}$ is a basis of M and $\{d_1.v_1, \ldots, d_s.v_s\}$ is a basis of N with $s \leqslant r$, as required.

What is left to prove is that $\langle d_{s-1}\rangle \subset \langle d_s\rangle$. Let $f : M \to A$ be the A-linear map with $f(\sum c_i.v_i) = \sum c_i$ and let $d \in A$ be a greatest common divisor of d_{s-1} and d_s. Then $\langle d\rangle = \gcd\{d_{s-1}, d_s\}$ contains both $\langle d_{s-1}\rangle$ and $\langle d_s\rangle$. By Lemma 1.6.8, we have $d = c_{s-1}d_{s-1} + c_s d_s$ for some $c_{s-1}, c_s \in A$. Thus $f\big((c_{s-1}d_{s-1}).v_{s-1} + (c_s d_s).v_s\big) = c_{s-1}d_{s-1} + c_s d_s = d$, and thus $I_g \subset \langle d\rangle \subset I_f$. By the maximality of $I_g = \langle d_s\rangle$, these inclusions are equalities and thus $\langle d_{s-1}\rangle \subset \langle d\rangle = \langle d_s\rangle$, which verifies our claim. This completes the proof of the existence claim of the theorem.

We turn to the uniqueness of s and the ideals $\langle d_1\rangle, \ldots, \langle d_s\rangle$. By Proposition 3.8.2, the rank s of the free A-module N is uniquely determined. Let $\mathcal{B} = \{v_1, \ldots, v_r\}$ and $d_1, \ldots, d_s$ be as described in the theorem. Then

$$M/N \;\simeq\; A^{r-s} \times \prod_{i=1}^{s} A/\langle d_i\rangle \qquad \text{and} \qquad T(M/N) \;\simeq\; \prod_{i=1}^{s} A/\langle d_i\rangle.$$

Note that $d_1 \neq 0$ since $d_1.v_1$ belongs to a basis of N. Let $t \leqslant s$ be the integer such that $\langle d_t\rangle \neq \langle d_{t+1}\rangle = \ldots = \langle d_s\rangle = \langle 1\rangle$. Then $T(M/N) \simeq \prod_{i=1}^{t} A/\langle d_i\rangle$ and $0 \neq \langle d_1\rangle \subset \ldots \langle d_t\rangle \neq \langle 1\rangle$. By Lemma 3.9.2, the integer t and the ideals $\langle d_1\rangle, \ldots, \langle d_t\rangle$ are uniquely determined by $T(M/N)$. From this and the knowledge of s, we necessarily obtain $\langle d_{t+1}\rangle = \ldots = \langle d_s\rangle = \langle 1\rangle$, which proves that s and the ideals $\langle d_1\rangle, \ldots, \langle d_s\rangle$ are uniquely determined by the submodule N of M. This concludes the proof of the theorem. □

The Smith normal form

Theorem 3.9.4 (Smith normal form)**.** *Let M and N be free A-modules of finite rank r and s, respectively, and $f : M \to N$ a homomorphism. Then there are bases $\{v_1, \ldots, v_r\}$ of M and $\{w_1, \ldots, w_s\}$ of N, an integer t with $0 \leqslant t \leqslant \min\{r, s\}$ and elements $d_1, \ldots, d_t \in A$ with $\langle 0\rangle \neq \langle d_1\rangle \subset \ldots \subset \langle d_t\rangle$ such that*

$$f\Big(\sum_{i=1}^{r} a_i.v_i\Big) \;=\; \sum_{i=1}^{t} (d_i a_i).w_i.$$

Moreover, the integer t and the ideals $\langle d_1\rangle, \ldots, \langle d_t\rangle$ are uniquely determined by $f : M \to N$.

Remark. The ideals $\langle d_1\rangle, \ldots, \langle d_t\rangle$ are called the *invariants of f*. The central claim of Theorem 3.9.4 can be expressed by saying that there are ordered bases for M and N such

that f can be represented by a matrix of the form

$$\left(\begin{array}{ccc|c} d_1 & & & \\ & \ddots & & 0 \\ & & d_t & \\ \hline & 0 & & 0 \end{array}\right)$$

where the zeros stay for matrix blocks of appropriate sizes whose coefficients are all zero.

Proof. By the elementary divisor theorem (Theorem 3.9.3), there is a basis $\{w_1,\dots,w_s\}$ of N, an integer $t \leqslant s$ and elements $d_1,\dots,d_t \in A$ with $\langle 0\rangle \neq \langle d_1\rangle \subset \dots \subset \langle d_t\rangle$ such that $\{d_1.w_1,\dots,d_t.w_t\}$ is a basis for the submodule $\operatorname{im} f$ of N. In particular, $\operatorname{im} f$ is free. By Exercise 3.14, $\operatorname{im} f$ is projective and the short exact sequence $0 \to \ker f \to M \to \operatorname{im} f \to 0$ splits, i.e. there is a section $s : \operatorname{im} f \to M$ to $f : M \to \operatorname{im} f$ and $M = \ker f \oplus s(\operatorname{im} f)$. As a consequence, the elements $v_i = s(d_i.w_i)$ for $i = 1,\dots,t$ form a basis of the submodule $s(\operatorname{im} f)$ of M.

By the elementary divisor theorem (Theorem 3.9.3), the submodule $\ker f$ of M is free. Since $\operatorname{rk} M = \operatorname{rk}(\ker f) + \operatorname{rk}(\operatorname{im} f)$, we have $\operatorname{rk} \ker f = r - t$. Let $\{v_{t+1},\dots,v_r\}$ be a basis of $\ker f$. Then $\{v_1,\dots,v_r\}$ is a basis for M. By construction, we have $f(\sum_{i=1}^r a_i.v_i) = \sum_{i=1}^t (d_i a_i).w_i$.

Since t and $\langle d_1\rangle,\dots,\langle d_t\rangle$ are determined by the submodule $\operatorname{im} f$ of N, their uniqueness follows at once from the elementary divisor theorem (Theorem 3.9.3). □

The structure theorem for finitely generated modules

Recall from Definition 3.1.1 that an A-module M is finitely generated if $M = \langle S\rangle_A$ for a finite subset S of M.

Theorem 3.9.5 (Structure theorem for finitely generated modules)**.** *Let M be a finitely generated A-module. Then there are integers $r,s,t,e_1,\dots,e_t \geqslant 0$, elements $d_1,\dots,d_s \in A$ and prime elements $p_1,\dots,p_t \in A$ such that $\langle 0\rangle \neq \langle d_1\rangle \subset \dots \subset \langle d_s\rangle \neq \langle 1\rangle$ and*

$$M \quad \simeq \quad A^r \times \prod_{i=1}^{s} A/\langle d_i\rangle \quad \simeq \quad A^r \times \prod_{i=1}^{t} A/\langle p_i^{e_i}\rangle.$$

Moreover, r,s,t and the ideals $\langle d_1\rangle,\dots,\langle d_s\rangle$ and $\langle p_1^{e_1}\rangle,\dots,\langle p_t^{e_t}\rangle$ are uniquely determined, up to a simultaneous permutation of the indices of the p_i and e_i.

Remark. The ideals $\langle d_1\rangle,\dots,\langle d_s\rangle$ are called the *invariants of M* and the prime powers $p_1^{e_1},\dots,p_t^{e_t}$ are called the *elementary divisors of M*.

Proof. Let $\{m_1,\dots,m_{\tilde{r}}\}$ be a set of generators of M. Then the homomorphism

$$\begin{array}{rccc} f: & A^{\tilde{r}} & \longrightarrow & M \\ & (a_i) & \longmapsto & \sum a_i.m_i \end{array}$$

is surjective. By the elementary divisor theorem (Theorem 3.9.3), there is a basis $\{v_1,\dots,v_{\tilde r}\}$ of $A^{\tilde r}$, an integer $\tilde s \geqslant 0$ and $d_1,\dots,d_{\tilde s} \in A$ with $\langle 0\rangle \neq \langle d_1\rangle \subset \dots \subset \langle d_{\tilde s}\rangle$ such that $\{d_1.w_1,\dots,d_{\tilde s}.w_{\tilde s}\}$ is a basis of $\ker f$. By the first isomorphism theorem (Theorem 3.4.1), we have

$$M \simeq A^{\tilde r}/\ker f \simeq A^{\tilde r-\tilde s} \times \prod_{i=1}^{\tilde s} A/\langle d_i\rangle.$$

Let $r = \tilde r - \tilde s$ and let s be the integer for which $\langle d_s\rangle \neq \langle d_{s+1}\rangle = \dots = \langle d_{\tilde s}\rangle = 1$. Then $\langle 0\rangle \neq \langle d_1\rangle \subset \dots \subset \langle d_s\rangle \neq \langle 1\rangle$. Since $A/\langle d_i\rangle = \{0\}$ for $i = s+1,\dots,\tilde s$, we gain the first isomorphism

$$M \simeq A^r \times \prod_{i=1}^{s} A/\langle d_i\rangle$$

of the theorem. Let d be an integer with factorization $d = p_1^{e_1}\cdots p_t^{e_t}$ into prime powers, i.e. $p_1,\dots,p_t \in A$ are pairwise distinct prime elements. Then $\langle p_i^{e_i}\rangle$ and $\langle p_j^{e_j}\rangle$ are coprime for $i \neq j$ since $\langle p_i^{e_i}\rangle + \langle p_j^{e_j}\rangle = \gcd(p_i^{e_i}, p_j^{e_j}) = \langle 1\rangle$, and thus the Chinese remainder theorem (Theorem 1.5.5) implies that $A/\langle d\rangle \simeq \prod A/\langle p_i^{e_i}\rangle$. Applying this to all factors $A/\langle d_i\rangle$ in the expression $M \simeq A^r \times \prod_{i=1}^s A/\langle d_i\rangle$ yields the second isomorphism of the theorem.

We turn to the uniqueness claims. The integer r is uniquely determined by the rank of $M/T(M) \simeq A^r$. The ideals $\langle d_1\rangle,\dots,\langle d_s\rangle$ are uniquely determined by $T(M) \simeq \prod_{i=1}^s A/\langle d_i\rangle$, as shown in Lemma 3.9.2.

In the following, we show that the ideals $\langle p_1^{e_1}\rangle,\dots,\langle p_t^{e_t}\rangle$ determine uniquely the integer s and the ideals $\langle d_1\rangle,\dots,\langle d_s\rangle$, which implies the uniqueness of t and the ideals $\langle p_1^{e_1}\rangle,\dots,\langle p_t^{e_t}\rangle$.

To explain the idea: since $v_{\mathfrak p}(d_1) \geqslant \dots \geqslant v_{\mathfrak p}(d_s)$ for all prime ideals $\mathfrak p$, the ideal $\langle d_1\rangle$ is determined as the product of the ideals $\langle p_i^{e_i}\rangle$ with the highest exponents e_i for every association class $[p_i] = \{p_j | p_j \sim p_i\}$ of prime elements in $\{p_1,\dots,p_t\}$. Successively, we define $\langle d_2\rangle$ as the product of the ideals $\langle p_i^{e_i}\rangle$ with the second highest exponents e_i, and so forth. We make this precise as follows.

Assume that $\prod_{i=1}^s A/\langle d_i\rangle \simeq \prod_{i=1}^t A/\langle p_i^{e_i}\rangle$ for elements $d_1,\dots,d_s \in A$ with $\langle 0\rangle \neq \langle d_1\rangle \subset \dots \subset \langle d_s\rangle \neq \langle 1\rangle$. Let $\mathcal P = \{\langle p_1\rangle,\dots,\langle p_t\rangle\}$ be the set of prime ideals generated by the prime elements $p_1,\dots,p_t$. Let

$$\mu_{\mathfrak p} = \#\big\{i \in \{1,\dots,t\} \,\big|\, \langle p_i\rangle = \mathfrak p\big\}$$

be the multiplicity of the occurrence of a prime ideal $\mathfrak p$ in $(\langle p_1\rangle,\dots,\langle p_t\rangle)$ and $s' = \max\{\mu_{\mathfrak p} \mid \mathfrak p \in \mathcal P\}$ be the maximal multiplicity that occurs. Since $A/\langle p_i^{e_i}\rangle \times A/\langle p_j^{e_j}\rangle$ is not cyclic if $\langle p_i\rangle = \langle p_j\rangle$, we conclude that $s \geqslant s'$. On the other hand, if $\langle d_s\rangle \subset \langle p_i^{e_i}\rangle = \mathfrak p^{e_i}$, then $\langle d_i\rangle \subset \langle d_s\rangle \subset \mathfrak p^{e_i}$ for all $i = 1,\dots,s$. Since $\langle d_1\rangle \neq \langle 1\rangle$, we must have $s \leqslant \mu_{\mathfrak p} \leqslant s'$ for some $\mathfrak p \in \mathcal P$, which shows that $s = s'$. Thus s is determined by the ideals $\langle p_1^{e_1}\rangle,\dots,\langle p_t^{e_t}\rangle$.

In order to determine the ideals $\langle d_1\rangle,\dots,\langle d_s\rangle$, we define for $\mathfrak p \in \mathcal P$ and $k = 1,\dots,s$ the integer

$$e_{\mathfrak p,k} = \max\Big\{e \in \mathbb N \,\Big|\, e = 0 \text{ or } \#\big\{i \in \{1,\dots,t\} \,\big|\, \langle p_i\rangle = \mathfrak p, e_i \geqslant e\big\} \geqslant k\Big\},$$

which is the k-th largest exponent e_i occurring among those ideals $\langle p_i^{e_i}\rangle$ for which $\langle p_i\rangle = \mathfrak{p}$ unless there are less than k such ideals, in which case $e_{\mathfrak{p},k} = 0$. This notation allows us to reorder the ideals $\langle p_1^{e_1}\rangle, \ldots, \langle p_t^{e_t}\rangle$ into increasing chains $\mathfrak{p}^{e_{\mathfrak{p},1}} \subset \ldots \subset \mathfrak{p}^{e_{\mathfrak{p},s}}$ for every $\mathfrak{p} \in \mathcal{P}$ such that the proper ideals in these chains are precisely the ideals $\langle p_1^{e_1}\rangle, \ldots, \langle p_t^{e_t}\rangle$. Since for $k \leqslant l$, the inclusion $\langle d_l\rangle \subset \mathfrak{p}^{e_i}$ implies that $\langle d_k\rangle \subset \mathfrak{p}^{e_j}$ for some $e_j \geqslant e_i$, we conclude that we must have

$$\langle d_i\rangle \;=\; \prod_{\mathfrak{p}\in\mathcal{P}} \mathfrak{p}^{e_{\mathfrak{p},i}}$$

for $i = 1, \ldots, s$. This shows that the ideals $\langle d_1\rangle, \ldots, \langle d_s\rangle$ and the ideals $\langle p_1^{e_1}\rangle, \ldots, \langle p_t^{e_t}\rangle$ determine each other. Thus the uniqueness of t and $\langle p_1^{e_1}\rangle, \ldots, \langle p_t^{e_t}\rangle$ follows from the uniqueness of s and $\langle d_1\rangle, \ldots, \langle d_s\rangle$, which concludes the proof of the theorem. □

Applications

Recall from Definition 3.1.1 that an A-module is cyclic if it is generated by a single element.

Corollary 3.9.6. *Let M be a finitely generated A-module. Then M is the direct sum of finitely many cyclic submodules.*

Proof. Let M be a finitely generated A-module. By the structure theorem of finitely generated modules (Theorem 3.9.5), there is an isomorphism $M \simeq \prod_{i=1}^{s} A/\langle d_i\rangle$ for some $r, s \in \mathbb{Z}$ and some nonunits $d_1, \ldots, d_s \in A$, which we allow to be 0, which accounts for the factors $A = A/\langle 0\rangle$. Composing its inverse with the canonical isomorphism from the finite direct sum with the finite product yields an isomorphism

$$f : \bigoplus_{i=1}^{s} A/\langle d_i\rangle \quad\xrightarrow{\sim}\quad \prod_{i=1}^{s} A/\langle d_i\rangle \quad\xrightarrow{\sim}\quad M$$

Since $A/\langle d_j\rangle$ is generated by $\bar{1}$, its image $N_j = f \circ \iota_j(A/\langle d_j\rangle)$ in M is generated by $f \circ \iota_j(\bar{1})$ where $\iota_j : A/\langle d_j\rangle \to \bigoplus_{i=1}^{s} A/\langle d_i\rangle$ is the canonical inclusion. Thus M is the direct sum $\bigoplus_{i=1}^{s} N_i$ of the cyclic submodules N_i of M. □

Corollary 3.9.7. *A finitely generated A-module is free if and only if it is torsion-free.*

Proof. By Lemma 3.8.5, free modules are torsion-free. Conversely, assume that M is torsion-free. By the structure theorem for finitely generated A-modules (Theorem 3.9.5), $M = A^r \times \prod A/\langle d_r\rangle$ for certain nonzero elements $d_1, \ldots, d_s \in A$. Since $T(M) \simeq \prod A/\langle d_i\rangle$ is trivial, we conclude that $M \simeq A^r$ is free. □

Corollary 3.9.8. *Let M be a finitely generated A-module and N a submodule of M. Then N is finitely generated.*

Proof. Let $\{v_1,\dots,v_r\}$ be a set of generators for M. Then the A-linear map

$$\begin{array}{rccc} f: & \bigoplus_{i=1}^{r} A & \longrightarrow & M \\ & (a_i) & \longmapsto & \sum a_i.v_i \end{array}$$

is surjective. By Lemma 3.9.1, the submodule $N' = f^{-1}(N)$ of the free A-module $\bigoplus_{i=1}^{r} A$ is free of rank $s \leqslant r$. Let $\{w_1,\dots,w_s\}$ be a basis of N'. Then $\{f(w_1),\dots,f(w_s)\}$ is a set of generators for N, which shows that N is finitely generated. □

The structure theorem for finitely generated abelian groups

Definition 3.9.9. An abelian group G is *finitely generated* if it is finitely generated as a $\mathbb{Z}$-module.

Theorem 3.9.10 (Structure theorem for finitely generated abelian groups)**.** *Let G be a finitely generated abelian group. Then there exist integers $r,s,t,d_1,\dots,d_s,e_1,\dots,e_t \geqslant 0$ and prime numbers $p_1,\dots,p_t \geqslant 0$ such that $\langle d_1\rangle \subset \dots \subset \langle d_s\rangle$ and*

$$G \quad \simeq \quad A^r \times \prod_{i=1}^{s} \mathbb{Z}/\langle d_i\rangle \quad \simeq \quad A^r \times \prod_{i=1}^{t} \mathbb{Z}/\langle p_i^{e_i}\rangle.$$

Moreover, the integers $r,s,t,d_1,\dots,d_s,e_1,\dots,e_t$ and the prime numbers $p_1,\dots,p_t$ are uniquely determined up to a simultaneous permutation of the indices of p_i and e_i.

Proof. Once we have observed that every ideal of $\mathbb{Z}$ has a unique nonnegative generator, this follows immediately from the structure theorem of finitely generated A-modules (Theorem 3.9.5), applied to $A = \mathbb{Z}$. □

The Jordan normal form

In order to explain how the theorem of the Jordan normal form follows from the structure theorem of finitely generated A-modules, we recall some facts from linear algebra. For the rest of this section, let K be a field.

Let M be a finitely dimensional K-vector space of dimension r and $\varphi : M \to M$ a K-linear endomorphism. Let $\mathcal{B} = \{v_1,\dots,v_r\}$ be a basis of M over K. Then $\varphi\big(\sum a_i.v_i\big) = U \cdot (a_i)$ for some $r \times r$–matrix U with coefficients in K where we consider (a_i) as a row vector. The *characteristic polynomial of* φ is the monic polynomial $\mathrm{Char}_\varphi = \det(\mathrm{id}_M \cdot T - U)$ of degree r, which is independent from the choice of basis $\mathcal{B}$ and thus a well-defined invariant of the automorphism φ of M.

For $i \geqslant 0$, we define the K-linear map $\varphi^i : M \to M$ as $\varphi^0 = \mathrm{id}_M$ if $i = 0$ and

$$\varphi^i = \underbrace{\varphi \circ \dots \circ \varphi}_{i\text{-times}}$$

if $i > 0$. For a polynomial $f = \sum c_i T^i$ in $K[T]$, we define the K-linear map $f(\varphi) : M \to M$ by $\big(f(\varphi)\big)(m) = \sum c_i.\varphi^i(m)$. The action

$$\begin{array}{ccc} K[T] \times M & \longrightarrow & M \\ (f,m) & \longmapsto & \big(f(\varphi)\big)(m) \end{array}$$

of $K[T]$ on M endows M with the structure of a $K[T]$-module. We leave the verification of the axioms of an $K[T]$-module as an exercise. Since M is finite dimensional as a K-vector space, M is finitely generated as a $K[T]$-module.

Note that by the very definition of the $K[T]$-action, $\varphi(m) = T.m$ for all $m \in M$. This means that a $K[T]$-module M is essentially the same thing as a K-vector space M together with an endomorphism $\varphi : M \to M$.

Theorem 3.9.11 (Cayley-Hamilton theorem)**.** *Let M be a finite dimensional K-vector space and $\varphi : M \to M$ be a K-linear endomorphism. Then* $\mathrm{Char}_\varphi(\varphi) = 0$.

Proof. For the sake of completeness, we recall the idea of the proof. Let $\mathcal{B} = \{v_1, \ldots, v_r\}$ be a basis for M and U be the matrix that represents φ in the basis $\mathcal{B}$, i.e. $\varphi(v_i) = \sum U_{i,j}.v_j$. Let V be the $r \times r$–matrix of endomorphisms of M with coefficients $V_{i,j} = \delta_{i,j}\varphi - U_{i,j}$ where $\delta_{i,j}$ is the Kronecker symbol. Then $V^\# V = \det(V) = \mathrm{Char}_\varphi(\varphi)$ as endomorphisms of M where $V^\#$ is the adjoint matrix of V. Since $V : v_i \mapsto \sum_{j=1}^r \big(\delta_{i,j}\varphi(v_j) - U_{i,j}(v_j)\big) = 0$ for all $i = 1, \ldots, r$, we conclude that $V^\# V$ is the trivial map, and thus $\mathrm{Char}_\varphi(\varphi) = 0$. □

Let $K[\varphi] = \{f(\varphi) \mid f \in K[T]\}$ be the K-algebra of endomorphism of M that is generated by φ. Let $\pi : K[T] \to K[\varphi]$ be the homomorphism that sends T to φ. By the Cayley -Hamilton theorem (Theorem 3.9.11), the characteristic polynomial of φ is in the kernel of π. Thus $\ker \pi$ is not trivial and $K[\varphi]$ is finite dimensional over K.

Every nonzero ideal I of the principal ideal domain $K[T]$ is generated by a unique monic polynomial since the unit group $K[T]^\times = K^\times$ coincides with the choice of the leading coefficient of a generator for I. The *minimal polynomial of* φ is defined as the unique monic generator Min_φ of $\ker \pi$. Since $\mathrm{Min}_\varphi \in \ker \pi$, we have $\mathrm{Min}_\varphi(\varphi) = 0$. Since $\mathrm{Char}_\varphi \in \ker \pi = \langle \mathrm{Min}_\varphi \rangle$, the minimal polynomial Min_φ divides the characteristic polynomial Char_φ.

Theorem 3.9.12 (Structure theorem for finite dimensional $K[T]$-modules)**.** *Let M be a finite dimensional K-vector space and $\varphi : M \to M$ be a K-linear endomorphism. Then there are integers $s, t, e_1, \ldots, e_t \geqslant 0$, monic polynomials $f_1, \ldots, f_s$ with $\langle 0 \rangle \neq \langle f_1 \rangle \subset \ldots \subset \langle f_s \rangle \neq \langle 1 \rangle$ and monic irreducible polynomials $g_1, \ldots, g_t$ such that*

$$M \simeq \prod_{i=1}^{s} K[T]/\langle f_i \rangle \simeq \prod_{i=1}^{t} K[T]/\langle g_i^{e_i} \rangle$$

as $K[T]$-modules. Moreover, the integers s, t and the polynomials $f_1, \ldots, f_s, g_1, \ldots, g_t$ are uniquely determined by M, up to a simultaneous permutation of the indices of the g_i and e_i. The characteristic polynomial of φ is equal to $\mathrm{Char}_\varphi = \prod_{i=1}^s f_i$ *and the minimal polynomial of φ is equal to* $\mathrm{Min}_\varphi = f_1$.

Proof. Since a finite dimensional $K[T]$-module is a torsion module, cf. Example 3.8.6, and since every nonzero ideal of $K[T]$ is generated by a unique monic polynomial, everything follows at once from the structure theorem of finitely generated A-modules (Theorem 3.9.5) applied to $A = K[T]$, but for the claims that $\mathrm{Char}_\varphi = \prod_{i=1}^s f_i$ and $\mathrm{Min}_\varphi = f_1$. We will verify the latter claim in the following and leave the former claim as Exercise 3.25.

By what we have proven, we can assume that $M = \prod_{i=1}^s K[T]/\langle f_i\rangle$. Since both f_1 and Min_φ are monic, it is enough to show that $\langle \mathrm{Min}_\varphi\rangle = \langle f_1\rangle$. Since $f_1 \in \langle f_i\rangle$ for all $i = 1,\dots,s$, we have for all $m = \big([h_1],\dots,[h_s]\big) \in M$ that

$$f_1(\varphi).m \;=\; \big([f_1h_1],\dots,[f_1h_s]\big) \;=\; \big([0],\dots,[0]\big) \;=\; 0,$$

and thus $f_1 \in \langle \mathrm{Min}_\varphi\rangle$. Conversely, we have

$$\big([\mathrm{Min}_\varphi],\dots,[\mathrm{Min}_\varphi]\big) \;=\; \mathrm{Min}_\varphi(\varphi).1 \;=\; 0 \;=\; \big([0],\dots,[0]\big)$$

as elements of $\prod_{i=1}^s K[T]/\langle f_i\rangle$, and thus $\mathrm{Min}_\varphi \in \langle f_1\rangle$. Thus $\langle \mathrm{Min}_\varphi\rangle = \langle f_1\rangle$, which completes the proof. □

For $\lambda \in K$, we define the *Jordan block (of size e)* as the $e \times e$–matrix

$$\begin{pmatrix} \lambda & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & \lambda \end{pmatrix}$$

whose coefficients below and above the two diagonals with entries are all zero.

Theorem 3.9.13 (Jordan normal form)**.** *Let M be a K-vector space of finite dimension r and $\varphi : M \to M$ be a K-linear endomorphism whose minimal polynomial Min_φ factors in $K[T]$ into linear factors, i.e. $\mathrm{Min}_\varphi = \prod_{i=1}^s (T - a_i)$ for some $a_1,\dots,a_s \in K$ where $s = \deg(\mathrm{Min}_\varphi)$. Then there is a basis $\mathcal{B}$ of M, integers $t, e_1,\dots,e_t \geqslant 0$ and $\lambda_1,\dots,\lambda_t \in \{a_1,\dots,a_s\}$ such that φ is represented in the basis $\mathcal{B}$ by the matrix*

$$\begin{pmatrix} J_{e_1}(\lambda_1) & & \\ & \ddots & \\ & & J_{e_t}(\lambda_t) \end{pmatrix}$$

that has Jordan blocks $J_{e_i}(\lambda_i)$ on its diagonals and zero coefficients outside the diagonal blocks. Moreover, the integers $s, e_1,\dots,e_t \geqslant 0$ and $\lambda_1,\dots,\lambda_t \in K$ are uniquely determined by φ up to a simultaneous permutation of the indices of the e_i and λ_i.

Proof. By the structure theorem for finite dimensional $K[T]$-modules (Theorem 3.9.12), there are integers $s, e_1,\dots,e_t$ and irreducible monic polynomials $f_1,\dots,f_t \in K[T]$ such that $M \simeq \prod_{i=1}^s K[T]/\langle f_i^{e_i}\rangle$ as $K[T]$-modules, which allows us to assume that

$M = \prod_{i=1}^{s} K[T]/\langle f_i^{e_i}\rangle$ where φ acts as T, i.e. $\varphi\big([g_1],\dots,[g_t]\big) = \big([T\cdot g_1],\dots,[T\cdot g_t]\big)$. Since $\mathrm{Min}_\varphi(\varphi)$ acts trivial on M by the very definition of Min_φ, we have

$$\big([\mathrm{Min}_\varphi],\dots,[\mathrm{Min}_\varphi]\big) \;=\; \big(\mathrm{Min}_\varphi(\varphi)\big).\big([1],\dots,[1]\big) \;=\; \big([0],\dots,[0]\big)$$

as elements of $\prod_{i=1}^{s} K[T]/\langle f_i^{e_i}\rangle$, which shows that $\mathrm{Min}_\varphi \in \langle f_i^{e_i}\rangle$ for $i = 1,\dots,s$. Thus for every $i \in \{1,\dots,s\}$, the polynomial f_i divides $\mathrm{Min}_\varphi = \prod_{j=1}^{r} T - a_j$. Since f_i is irreducible, we conclude that $f_i = T - \lambda_i$ for some $\lambda_i \in \{a_1,\dots,a_s\}$.

Since φ leaves the submodules $M_i = K[T]/\langle (T-\lambda_i)^{e_i}\rangle$ of M invariant, we can concentrate on finding appropriate matrix representations U_i for the restrictions $\varphi_i : M_i \to M_i$ of φ to the submodules M_i. Putting the bases of the M_i together to a basis of M yields a matrix representation for φ whose diagonal blocks are the matrices U_i and whose other coefficients are all 0.

Consider the basis $\mathcal{B}_i = \big\{[1],[T-\lambda_i],\dots,[(T-\lambda_i)^{e_i-1}]\big\}$ of M_i. Since for $k = 0,\dots,e_i-1$,

$$\begin{aligned}\varphi_i.\big[(T-\lambda_i)^k\big] \;=\; \big[T(T-\lambda_i)^k\big] \;&=\; \big[(T-\lambda_i)(T-\lambda_i)^k + \lambda_i(T-\lambda_i)^k\big]\\ &=\; 1.\big[(T-\lambda_i)^{k+1}\big] + \lambda_i.\big[(T-\lambda_i)^k\big]\end{aligned}$$

and since $\big[(T-\lambda_i)^{e_i}\big] = [0]$, the endomorphism φ_i is represented in the basis $\mathcal{B}_i$ by the Jordan block $U_i = J_{e_i}(\lambda_i)$. Thus φ is represented in the basis $\mathcal{B} = \bigcup_{i=1}^{s}\mathcal{B}_i$ of $M = \bigoplus_{i=1}^{s} M_i$ by the matrix as described in the theorem.

Since the integers $s, e_1,\dots,e_t \geqslant 0$ and $\lambda_1,\dots,\lambda_t \in K$ are determined by the $K[T]$-module M, the uniqueness claim follows at once from the structure theorem for finite dimensional $K[T]$-modules (Theorem 3.9.12). □

Remark. If K is an algebraically closed field, then every polynomial is a product of linear factors by Proposition 1.10.10 and thus every endomorphism $\varphi : M \to M$ of a finite dimensional K-vector space M has a Jordan normal form.

This is not true if K is not algebraically closed. An example is the rotation $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ of the Euclidean plan $\mathbb{R}^2$ by 90°, which is represented by the matrix $\left(\begin{smallmatrix}0 & -1\\ 1 & 0\end{smallmatrix}\right)$ in the standard unit basis. If φ had a Jordan normal form with respect to some basis $\mathcal{B} = \{v_1, v_2\}$ of $\mathbb{R}^2$, then it would have an eigenvector, namely v_1. However, the rotation φ does not have an eigenvector, and thus cannot have a Jordan normal form.

In agreement with the lack of a Jordan normal form, we find that the minimal polynomial Min_φ is in this case equal to the characteristic polynomial $\mathrm{Char}_\varphi = T^2 + 1$ of φ since T^2+1 is irreducible in $\mathbb{R}[T]$ and thus has no nontrivial divisors. In particular, Min_φ is not a product of linear polynomials in $\mathbb{R}[T]$.

3.10 Exercises

Exercise 3.1. Proof Lemma 3.3.4.

Exercise 3.2. Let $\{M_i\}_{i\in I}$ be a family of A-modules.

(1) Show that $\prod_{i\in I} M_i$ together with the projections $\pi_j : \prod M_i \to M_j$ is the categorical product in Mod_A.

(2) Show that $\bigoplus_{i\in I} M_i$ together with the inclusions $\iota_j : M_j \to \bigoplus M_i$ is the categorical coproduct in Mod_A.

Exercise 3.3. Verify the following assertions.

(1) $\mathbb{R}^n \otimes_{\mathbb{R}} \mathbb{R}^m$ and $\mathbb{R}^{m\cdot n}$.

(2) $A[T_1] \otimes_A A[T_2] \simeq A[T_1, T_2]$ for any ring A.

(3) $(\mathbb{Z}/n\mathbb{Z}) \otimes_{\mathbb{Z}} (\mathbb{Z}/m\mathbb{Z}) \simeq \mathbb{Z}/d\mathbb{Z}$ where d is a greatest common divisor of the natural numbers m and n.

(4) $K \otimes_{\mathbb{Z}} L = \{0\}$ if K and L are fields of different characteristics.

(5) $\mathbb{Q} \otimes_{\mathbb{Z}} \mathbb{Q} \simeq \mathbb{Q}$.

Exercise 3.4. Let M and N be A-modules. Show that $\mathrm{Hom}_A(M,N)$ is an A-module with respect to the operations $f+g : m \mapsto f(m)+g(m)$ and $a.f : m \mapsto a.f(m)$ for $a \in A$ and $f, g \in \mathrm{Hom}_A(M,N)$. Show that

$$\begin{array}{ccc} \mathrm{Hom}_A(M,N) \times \mathrm{Hom}_A(N,P) & \longrightarrow & \mathrm{Hom}_A(M,P) \\ (f,g) & \longmapsto & g \circ f \end{array}$$

is an A-bilinear homomorphism. Conclude that the association $f \otimes g \mapsto g \circ f$ describes a homomorphism $\mathrm{Hom}_A(M,N) \otimes \mathrm{Hom}_A(N,P) \to \mathrm{Hom}_A(M,P)$ of A-modules.

Exercise 3.5. Let M, N, N' and P be A-modules and $f : N \to N'$ an A-linear homomorphism. Consider the associations

$$\begin{array}{rccc} f_M : & M \otimes_A N & \longrightarrow & M \otimes_A N', \\ & m \otimes n & \longmapsto & m \otimes f(n) \end{array} \qquad \begin{array}{rccc} f^* : & \mathrm{Hom}(N',P) & \longrightarrow & \mathrm{Hom}(N,P) \\ & g & \longmapsto & g \circ f \end{array}$$

$$\text{and} \quad \begin{array}{rccc} f_* : & \mathrm{Hom}(M,N) & \to & \mathrm{Hom}(M,N'). \\ & h & \longmapsto & f \circ h \end{array}$$

(1) Show that f_M is well-defined as a map and that all three maps are homomorphisms of A-modules.

(2) Conclude that $M \otimes_A (-)$, $\mathrm{Hom}(-,P)$ and $\mathrm{Hom}(M,-)$ are functors from Mod_A to Mod_A. Which of them are covariant, which of them are contravariant?

Exercise 3.6. Let $f : A \to B$ be a ring homomorphism.

(1) Show that sending an A-module M to $B \otimes_A M$ and sending an A-linear map $\alpha : M \to M'$ to the B-linear map $\alpha_B : B \otimes_A M \to B \otimes_A M'$ that is defined by $\alpha_B(b \otimes m) = b \otimes \alpha(m)$ defines a functor $B \otimes_A - : \mathrm{Mod}_A \to \mathrm{Mod}_B$.

(2) Show that a B-module N is A-module with respect to the action defined by $a.n = f(a).n$ for $a \in A$ and $n \in N$. Show that a B-linear map $\alpha : N \to N'$ is A-linear with respect to this action. Conclude that this defines a functor $\mathcal{F} : \mathrm{Mod}_B \to \mathrm{Mod}_A$.

(3) Show that the association $b \otimes n \mapsto b.n$ defines a B-linear map $\eta_N : B \otimes_A \mathcal{F}(N) \to N$.

(4) Let M be an A-module and N a B-module. Show that the association

$$\begin{array}{rccc} \Psi_{M,N}: & \mathrm{Hom}_A\left(M, \mathcal{F}(N)\right) & \longrightarrow & \mathrm{Hom}_B(B \otimes_A M, N) \\ & \gamma : M \to \mathcal{F}(N) & \longmapsto & \eta_N \circ \gamma_B : B \otimes_A M \to N \end{array}$$

is a well-defined bijection.

(5) Let $\alpha : M \to M'$ be an A-linear map and $\beta : N \to N'$ a B-linear map. Show that the diagram

$$\begin{array}{ccccccc} \gamma & \in & \mathrm{Hom}_A\left(M', \mathcal{F}(N)\right) & \xrightarrow{\Psi_{M',N}} & \mathrm{Hom}_B(B \otimes_A M', N) & \ni & \delta \\ \downarrow & & \downarrow & & \downarrow & & \downarrow \\ \mathcal{F}(\beta) \circ \gamma \circ \alpha & \in & \mathrm{Hom}_A\left(M, \mathcal{F}(N')\right) & \xrightarrow{\Psi_{M,N'}} & \mathrm{Hom}_B(B \otimes_A M, N') & \ni & \beta \circ \delta \circ \alpha_B \end{array}$$

commutes.

Remark: The functor $B \otimes_A - : \mathrm{Mod}_A \to \mathrm{Mod}_B$ is called the *extension of scalars from A to B* and the functor $\mathcal{F} : \mathrm{Mod}_B \to \mathrm{Mod}_A$ is usually called *the restriction of scalars from B to A*. The properties (4) and (5) say that $\mathcal{F}$ is *right-adjoint* to $B \otimes_A -$.

Exercise 3.7 (Schur's lemma for algebras over algebraically closed fields)**.** Let K be an algebraically closed field, A a K-algebra and V an irreducible A-module that is finite dimensional as a K-vector space. Show that every A-linear map $\phi : V \to V$ is of the form $\phi(v) = a.v$ for some $a \in K$.

Exercise 3.8. Let K be a field and $M = N = K^2$, considered as additive groups. Define a map $K[T] \times M \to M$ by

$$\left(\sum a_i T^i\right).(m,n) \quad = \quad \left(\sum (a_i m), \sum a_i n\right)$$

and a map $K[T] \times N \to N$ by

$$\left(\sum a_i T^i\right).(m,n) \quad = \quad \left(\sum (a_i m + i a_i n), \sum a_i n\right)$$

where $a_i, m, n \in k$. Show that M and N are $K[T]$-modules with respect to these maps. Show that neither M nor N is simple, but that N is indecomposable while M is not.
Hint: T acts on M as the matrix $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, and it acts on N as the matrix $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$.

Exercise 3.9. Let K be a field and $M = N = k^2$ the $K[T]$-modules from Exercise 3.8. Let $P = K$.

(1) Show that the map $K[T] \times P \to P$ with $\left(\sum a_i T^i\right).(m) = \sum a_i.m$ turns P into a $K[T]$-module.

(2) Show that the inclusion $a \mapsto (a,0)$ into the first coordinate defines injective $K[T]$-linear maps $i : P \to M$ and $j : P \to N$.

(3) Show that there are short exact sequences of the form

$$0 \longrightarrow P \xrightarrow{i} M \xrightarrow{p} P \longrightarrow 0 \qquad \text{and} \qquad 0 \longrightarrow P \xrightarrow{j} N \xrightarrow{q} P \longrightarrow 0$$

for some $K[T]$-linear maps p and q.

(4) Which of these sequences are split?

Exercise 3.10. Let K be a field and $0 \to V_1 \to \cdots \to V_n \to 0$ be an exact sequence of K-vector spaces. Show that $\sum(-1)^i \dim_K V_i = 0$.

Exercise 3.11.
Let A be a ring and $f : M \to N$ a homomorphism of A-modules that has a section $g : N \to M$, i.e. $f \circ g = \mathrm{id}_N$. Show that $M \simeq \ker f \oplus \mathrm{im}\, g$.

Exercise 3.12 (Short 5-lemma)**.** Given a ring A and a commutative diagram

$$\begin{array}{ccccccccc} 0 & \longrightarrow & N & \xrightarrow{i} & M & \xrightarrow{p} & Q & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle f_N} & & \downarrow{\scriptstyle f_M} & & \downarrow{\scriptstyle f_Q} & & \\ 0 & \longrightarrow & N' & \xrightarrow{i'} & M' & \xrightarrow{p'} & Q' & \longrightarrow & 0 \end{array}$$

of A-modules with exact rows, show that

(1) f_M is a monomorphism if f_N and f_Q are monomorphisms,

(2) f_M is an epimorphism if f_N and f_Q are epimorphisms, and

(3) f_M is an isomorphism if f_N and f_Q are isomorphisms.

Exercise 3.13. Let A be a ring and $f : M \to N$ a homomorphism of A-modules. Show that

(1) f is a monomorphism if and only if it is injective;

(2) f is an epimorphism if and only if it is surjective;

(3) f is an isomorphism (in the sense of category theory, cf. Chapter 2) if and only if it is bijective.

Exercise 3.14. Show that the following properties for an A-module P are equivalent.

(1) The functor $\mathrm{Hom}(P, -)$ is exact.

(2) There is an A-module Q such that $P \oplus Q$ is free.

(3) Every short exact sequence of A-modules of the form $0 \to N \to M \to P \to 0$ splits.

(4) For every epimorphism $p : M \to Q$ of A-modules and every homomorphism $f : P \to Q$, there is a homomorphism $g : P \to M$ such that $f = p \circ g$.

An A-module P with these properties is called *projective*. Conclude that every free A-module is projective. Show that $\mathbb{Z}/2\mathbb{Z}$ is a projective $\mathbb{Z}/6\mathbb{Z}$-module that is not free.

Remark: An A-module I is *injective* if $\mathrm{Hom}(-,I)$ is exact. It can be shown that there are analogous characterizations as in (3) and (4) for injective modules. However, there is no direct analogue to (2). For $A=\mathbb{Z}$, one can show that a $\mathbb{Z}$-module I is injective if and only if it is *divisible*, i.e. for every $m\in I$ and every integer $l>0$ there exists an $n\in I$ such that $l.n=m$.

Exercise 3.15. Let A be a ring and P an A-module.

(1) Let M and N be A-modules and $f: P\otimes_A M\to N$ a homomorphism. Show that $\Psi_{M,N}(f): m\mapsto \big[p\mapsto f(m\otimes p)\big]$ defines an isomorphism

$$\Psi_{M,N}: \quad \mathrm{Hom}_A(P\otimes_A M,N) \quad\longrightarrow\quad \mathrm{Hom}_A\big(M,\mathrm{Hom}_A(P,N)\big)$$

of A-modules whose inverse sends a homomorphism $g: M\to \mathrm{Hom}_A(P,N)$ to the homomorphism $P\otimes_A M\to N$ with $p\otimes m\mapsto \big(g(m)\big)(p)$.

(2) Let $\alpha: M\to M'$ and $\beta: N\to N'$ be homomorphisms. Show that the diagram

$$\begin{array}{ccc} \mathrm{Hom}_A\big(M',\mathrm{Hom}_A(P,N)\big) & \xrightarrow{\Psi_{M',N}} & \mathrm{Hom}_A(P\otimes_A M',N) \\ \Big\downarrow{\scriptstyle \beta_*\circ-\circ\alpha} & & \Big\downarrow{\scriptstyle \beta\circ-\circ\alpha_P} \\ \mathrm{Hom}_A\big(M,\mathrm{Hom}_A(P,N')\big) & \xrightarrow{\Psi_{M,N'}} & \mathrm{Hom}_A(P\otimes_A M,N') \end{array}$$

commutes where $\alpha_P: P\otimes_A M\to P\otimes_A M'$ and $\beta_*: \mathrm{Hom}_A(P,N)\to\mathrm{Hom}_A(P,N')$ are the homomorphisms that are induced by α and β, respectively.

Exercise 3.16. Let A be a ring and M_1 and M_2 A-modules.

(1) Show that the canonical injections $\iota_k: M_k\to M_1\oplus M_2$ and the canonical projections $\pi_k: M_1\oplus M_2\to M_k$ (for $k=1,2$) satisfy the relations

$$\iota_1\circ\pi_1+\iota_2\circ\pi_2 \;=\; \mathrm{id}_{M_1\oplus M_2} \qquad\text{and}\qquad \pi_k\circ\iota_l \;=\; \begin{cases}\mathrm{id}_{M_k} & \text{if } k=l,\\ \mathbf{0} & \text{if } k\neq l\end{cases}$$

for all $k,l=1,2$.

(2) Let P be an A-module and $i_k: M_k\to P$ and $p_k: P\to M_k$ homomorphisms for $k=1,2$ that satisfy the relations

$$i_1\circ p_1+i_2\circ p_2 \;=\; \mathrm{id}_P \qquad\text{and}\qquad p_k\circ i_l \;=\; \begin{cases}\mathrm{id}_{M_k} & \text{if } k=l,\\ \mathbf{0} & \text{if } k\neq l\end{cases}$$

for $k,l=1,2$. Show that the homomorphism $M_1\oplus M_2\to P$ that is induced by $\{i_k: M_k\to P\}_{k=1,2}$ is an isomorphism.

(3) Let B be a ring and $\mathcal{F} : \mathrm{Mod}_A \to \mathrm{Mod}_B$ an additive covariant functor. Show that the homomorphism $\mathcal{F}(M_1) \oplus \mathcal{F}(M_2) \to \mathcal{F}(M_1 \oplus M_2)$ that is induced by $\{\mathcal{F}(\iota_k) : \mathcal{F}(M_i) \to \mathcal{F}(M_1 \oplus M_2)\}_{k=1,2}$ is an isomorphism.

(4) Let $\mathcal{F} : \mathrm{Mod}_A \to \mathrm{Mod}_B$ be as before and $0 \to N \to M \to Q \to 0$ a split short exact sequence. Show that $0 \to \mathcal{F}(N) \to \mathcal{F}(M) \to \mathcal{F}(Q) \to 0$ is a split short exact sequence.

Exercise 3.17. Let A and B be rings.

(1) Let M and N be A-modules and $f, g : M \to N$ homomorphisms. Show that the homomorphism

$$\begin{array}{ccccccc} M & \xrightarrow{\Delta} & M \oplus M & \xrightarrow{(f,g)} & N \oplus N & \xrightarrow{\Sigma} & N \\ m & \longmapsto & (m,m) & & (m,n) & \longmapsto & m+n \\ & & (m,n) & \longmapsto & f(m)+g(n) & & \end{array}$$

is equal to $f + g : M \to N$.

(2) Let $\mathcal{F} : \mathrm{Mod}_A \to \mathrm{Mod}_B$ be a covariant functor such that for all A-modules M_1 and M_2, the homomorphism $\mathcal{F}(M_1) \oplus \mathcal{F}(M_2) \to \mathcal{F}(M_1 \oplus M_2)$ that is induced by $\{\mathcal{F}(\iota_k) : \mathcal{F}(M_i) \to \mathcal{F}(M_1 \oplus M_2)\}_{k=1,2}$ is an isomorphism where $\iota_k : M_k \to M_1 \oplus M_2$ are the canonical inclusions. Show that $\mathcal{F}$ is additive.

Exercise 3.18. Show that the additive group of $\mathbb{Q}$ is a torsion-free $\mathbb{Z}$-module. Show that every free submodule of $\mathbb{Q}$ is cyclic, and show that the same is true for finitely generated submodules of $\mathbb{Q}$. Give an example of a proper submodule $N \subsetneq \mathbb{Q}$ that is not cyclic.

Exercise 3.19. Let A be an integral domain.

(1) Show that $T(M \times N) \simeq T(M) \times T(N)$. Conclude that for $r \geqslant 0$, nonzero ideals $I_1, \ldots, I_s$ of A and $M = A^r \times \prod_{i=1}^s A/I_i$, we have $T(M) \simeq \prod_{i=1}^s A/I_i$ and $M/T(M) \simeq A^r$.

(2) Show that a homomorphism $f : M \to N$ of A-modules restricts to a homomorphism $T(M) \to T(N)$ between their respective torsion modules. Show that this defines a left exact functor $T : \mathrm{Mod}_A \to \mathrm{Mod}_A$.

Exercise 3.20. (1) Let $M = \mathbb{Z}^3$ and N the submodule generated by $(1,1,6)$ and $(1,-1,6)$. Determine the invariants of the submodule N. Determine the invariants and the elementary divisors of M/N. What is the rank of $(M/N)/T(M/N)$?

(2) Let $f : \mathbb{Z}^3 \to \mathbb{Z}^3$ be the $\mathbb{Z}$-linear map given by multiplication of row vectors in $\mathbb{Z}^3$ with the matrix

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}.$$

Determine the Smith normal form and the invariants of f. Determine the invariants and the elementary divisors of $\mathbb{Z}^3/\operatorname{im} f$. What is the rank of $\mathbb{Z}^3/\operatorname{im} f$ divided by its torsion submodule?

(3) Let N be the submodule of $\mathbb{Z}^4$ that is generated by

$$(1,1,1,0), \quad (1,1,0,1), \quad (1,0,1,1), \text{ and } \quad (0,1,1,1).$$

Find a basis $\{v_1,\dots,v_4\}$ of $\mathbb{Z}^4$ and integers $a_1,\dots,a_4$ such that $\{a_1v_1,\dots,a_4v_4\}$ is a basis of N.

Exercise 3.21. An A-module M is *flat* if $-\otimes_A M$ is exact.

(1) Show that every free A-module is flat. Conclude that every projective A-module is flat.

(2) Let I be an ideal of A. Show that $I\otimes_A M \simeq IM$ if M is flat.

Hint: For (1), Exercise 3.14 is useful. For (2), the proof of Proposition 3.7.9 is helpful.

*__Exercise 3.22.__ Let A be a principal ideal domain and $r,s \geqslant 0$ integers.

(1) Show that every A-linear map $f : A^r \to A^s$ is of the form $f(a_i) = U\cdot(a_i)$ for some $r\times s$-matrix U with coefficients in A where $U\cdot(a_i)$ denotes the usual multiplication of a matrix with a vector.

(2) Let $\mathcal{B} = \{v_1,\dots,v_r\}$ be a subset of A^r and $v_{i,j}$ the j-th coordinate of v_i for $i,j = 1,\dots,r$. Show that $\mathcal{B}$ is a basis of A^r if and only if the $r\times r$-matrix U with coefficients $U_{i,j} = v_{i,j}$ is invertible.

(3) Show that there are for every $r\times s$-matrix U an $r\times r$-matrix V and an $s\times s$-matrix W such that $D = WUV$ is in *Smith normal form*, i.e. there is an integer t with $0 \leqslant t \leqslant \min\{r,s\}$ and elements $d_1,\dots,d_t \in A$ with $0 \neq \langle d_1\rangle \subset \dots \subset \langle d_t\rangle$ such that $D_{i,i} = d_i$ for $i = 1,\dots,t$ and $D_{i,j} = 0$ if $i \neq j$ or $i = j > t$.

(4) Exhibit invertible matrices V such that multiplying U with V from the right (from the left) results in *(a)* multiplying a column (row) by a unit of A; *(b)* an exchange of columns (rows); *(c)* adding a multiple of a column (row) to another. Such matrices V are called *elementary matrices*.

(5) Let A be a Euclidean domain with Euclidean norm $N : A \to \mathbb{N}$. Develop an algorithm using elementary column and row operations to bring U into Smith normal form.

Hint: One can refine the Gaussian algorithm appropriately using the Euclidean norm for the pivot search. If a pivot does not divide all coefficients of a given column and row, then one can produce a new pivot of smaller norm with the help of the Euclidean algorithm.

Remark: An algorithm as in part (5) does not exist for principal ideal domains in general since there are examples of invertible matrices that are not products of elementary matrices.

Exercise 3.23. Let K be a field and $A = K[T]$ and consider $M = K^n$ as an A-module by letting T act as a complex $n\times n$-matrix U. Show that M is a cyclic A-module if U has a

Jordan normal form with only one Jordan block, i.e. if U is conjugated to a matrix of the form

$$\begin{pmatrix} \lambda & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & \lambda \end{pmatrix}$$

for some $\lambda \in K$.

Exercise 3.24. Consider the $\mathbb{C}[T]$-module $M = \mathbb{C}^3$ where T acts as one of the matrices

$$(1)\quad T = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} \qquad (2)\quad T = \begin{pmatrix} \lambda & 0 & 0 \\ 1 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} \qquad (3)\quad T = \begin{pmatrix} \lambda & 0 & 0 \\ 1 & \lambda & 0 \\ 0 & 1 & \lambda \end{pmatrix}$$

$$(4)\quad T = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \mu \end{pmatrix} \qquad (5)\quad T = \begin{pmatrix} \lambda & 0 & 0 \\ 1 & \lambda & 0 \\ 0 & 0 & \mu \end{pmatrix} \qquad (6)\quad T = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \nu \end{pmatrix}$$

and where λ, μ and ν are pairwise distinct complex numbers. Determine in each case the characteristic polynomial and the minimal polynomial of T, as well as the elementary divisors and the invariant factors of M.

Exercise 3.25. Let K be a field, M a finite dimensional K-vector space and $\varphi : M \to M$ a K-linear map. Let $I_1 = (f_1), \ldots, I_s = (f_s)$ be the invariant factors of M as $K[T]$-module where T acts as φ and where $f_1, \ldots, f_s$ are monic polynomials. Show that $\prod_{i=1}^{s} f_i$ is the characteristic polynomial of φ.

Hint: Reduce the situation to the case where M is cyclic and use that in this case, the characteristic polynomial equals the minimal polynomial.

Exercise 3.26. Let A be a ring and $\mathrm{Mat}_{n\times n}(A)$ the set of $n \times n$-matrices with coefficients in A.

(1) Show that $\mathrm{Mat}_{n\times n}(A)$ is a noncommutative ring with respect to matrix addition and matrix multiplication. What are 0 and 1?

(2) Show that the inclusion $f : A \to \mathrm{Mat}_{n\times n}(A)$ as diagonal matrices is a homomorphism of (noncommutative) rings, i.e. $f(a+b) = f(a) + f(b)$, $f(a \cdot b) = f(a) \cdot f(b)$ and $f(1) = 1$.

(3) The *determinant* is the map $\det : \mathrm{Mat}_{n\times n}(A) \to A$ that sends a matrix $T = (a_{i,j})_{i,j=1,\ldots,n}$ to the element

$$\det(T) \quad = \quad \sum_{\sigma \in S_n} \mathrm{sign}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)}$$

of A. Show that $\det$ is multiplicative, i.e. $\det(T \cdot T') = \det(T) \cdot \det(T')$ and $\det(1) = 1$.

(4) Show that a matrix T is a unit in $\mathrm{Mat}_{n\times n}(A)$, i.e. $TT' = 1$ for some matrix T', if and only if $\det(T)$ is a unit in A.

Exercise 3.27. Let K be a field and M a finite dimensional K-vector space. A K-linear map $\varphi : M \to M$ is called *diagonalizable* if it acts as a diagonal matrix with respect to some basis of M. Show that φ is diagonalizable if and only if the minimal polynomial is of the form

$$\mathrm{Min}_\varphi \;=\; \prod_{i=1}^{n}(T-\alpha_i)$$

for pairwise distinct $\alpha_1,\ldots,\alpha_n \in K$. Is the $\mathbb{C}$-linear map $\varphi : \mathbb{C}^2 \to \mathbb{C}^2$ given by the matrix $\left(\begin{smallmatrix} 1 & -1 \\ 1 & 1 \end{smallmatrix}\right)$ for the standard basis of $\mathbb{C}^2$ diagonalizable?

Exercise 3.28. Let K be a field, M and N finite dimensional K-vector spaces, and $\varphi : M \to M$ and $\psi : N \to N$ K-linear maps. Assume that their respective characteristic polynomials factor as

$$\mathrm{Char}_\varphi \;=\; \prod_{i=1}^{m}(T-\alpha_i), \text{ and } \quad \mathrm{Char}_\psi \;=\; \prod_{j=1}^{n}(T-\beta_j).$$

Show that the formula $(\varphi\otimes\psi)(m\otimes n) = \varphi(m)\otimes\psi(n)$ defines a K-linear homomorphism $\varphi\otimes\psi : M\otimes_K N \to M\otimes_K N$, whose characteristic polynomial is

$$\mathrm{Char}_{\varphi\otimes\psi} \;=\; \prod_{i,j}(T-\alpha_i\beta_j).$$

Chapter 4

Multilinear algebra

In this chapter, we introduce the tensor algebra, the symmetric algebra and the exterior algebra of an A-module, and study their basic properties. Since many statements in this chapter are similar in nature to previous results, and can be proven by similar techniques as we have seen them already, we pass through this chapter with a faster pace and omit several proofs.

4.1 Graded algebras

Let A be a (commutative) ring. In this chapter, we allow A-algebras to have a noncommutative multiplication. More precisely, we use it in the following sense.

Definition 4.1.1. An **A-algebra** is a not necessarily commutative ring B together with a *ring homomorphism* $\iota_B : A \to B$, which is a map such that $\iota_B(1) = 1$, $\iota_B(a+b) = \iota_B(a) + \iota_B(b)$ and $\iota_B(ab) = \iota(a)\iota_B(b)$ for all $a, b \in A$.

As usual, we suppress the ring homomorphism $\iota_B : A \to B$ from the notation and simply say that B is an A-algebra. Note that we require that the additive group of a (not necessarily commutative) A-algebra B is commutative. In particular, this implies that B is an A-module with respect to the A-action $a.b = \iota_B(a)b$ for $a \in A$ and $b \in B$.

Definition 4.1.2. A **graded A-algebra** is an A-algebra B together with a family $\{B_i\}_{i\in\mathbb{N}}$ of A-submodules B_i of B such that $B = \bigoplus_{i\in\mathbb{N}} B_i$, such that $\iota_B(A) \subset B_0$ and such that $ab \in B_{i+j}$ for all $i, j \in \mathbb{N}$, $a \in B_i$ and $b \in B_j$. Let B and C be graded A-algebras. A **graded homomorphism from A to B** is a map $f : B \to C$ such that $f(a+b) = f(a) + f(b)$ and $f(ab) = f(a)f(b)$ for all $a, b \in B$, such that $f(B_i) \subset C_i$ for all $i \in \mathbb{N}$ and such that $f \circ \iota_B = \iota_C$. This defines the category GrAlg_A of graded A-algebras.

An element $a \in B$ is **homogeneous** (**of degree** i) if $a \in B_i$ for some $i \in \mathbb{N}$ (and if $a \neq 0$). A **homogeneous ideal of** B is an A-submodule I of B such that $ab, ba \in I$ for all $a \in I$ and $b \in B$ and such that I is generated by homogeneous elements, i.e. $I = \bigoplus_{i\in\mathbb{N}} I_i$ for $I_i = I \cap B_i$.

Example 4.1.3. Every A-algebra B can be seen as a *trivially graded* A-algebra with $B_0 = B$ and $B_i = \{0\}$ for $i > 0$. A nontrivial example is the polynomial algebra $B = A[T]$

over A, which is graded by the A-submodules $B_i = \{aT^i \mid a \in A\}$. More generally, the polynomial algebra $B = A[T_1, \dots, T_n]$ in n indeterminates $T_1, \dots, T_n$ is graded by the A-submodules

$$B_i \;=\; \left\langle T_1^{e_1} \cdots T_n^{e_n} \,\middle|\, e_1, \dots, e_n \in \mathbb{N} \text{ with } e_1 + \cdots + e_n = i \right\rangle$$

of B.

Lemma 4.1.4. *Let $B = \bigoplus B_i$ be a graded A-algebra and $I = \bigoplus I_i$ a homogeneous ideal of B where $I_i = I \cap B_i$. Then the quotient $\overline{B} = B/I$ together with the submodules $\overline{B}_i = B_i/I_i$ is a graded A-algebra with respect to the multiplication $[a] \cdot [b] = [ab]$ for $a, b \in B$ and the composition $\iota_{\overline{B}} = \pi_I \circ \iota_B : A \to \overline{B}$ of $\iota_B : A \to B$ with the quotient map $\pi_I : B \to B/I$, which is a graded homomorphism.*

Proof. We begin with showing that $\overline{B} = \bigoplus \overline{B}_i$. The A-linear map $\bar{f} : \bigoplus \overline{B}_i \to \overline{B}$ with $\bar{f}\big((\bar{b}_i)\big) = \sum \bar{b}_i$ is surjective since the map $f : \bigoplus B_i \to B$ with $f\big((b_i)\big) = \sum b_i$ is surjective. It is injective for the following reason. Consider $(b_i) \in \bigoplus B_i$ such that $\big[\sum b_i\big] = \bar{f}\big((\bar{b}_i)\big) = \bar{0}$, i.e. $\sum b_i \in I$. Since $I = \bigoplus I_i$ and $I_i = I \cap B_i$, this means that $b_i \in I_i$ and thus $\bar{b}_i = \bar{0}$ for all $i \in \mathbb{N}$. This shows that f is injective and thus $\overline{B} = \bigoplus \overline{B}_i$.

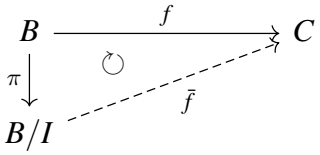
We continue to verify that the multiplication is well-defined on $\overline{B} = B/I$. Given $a, a', b, b' \in B$ with $[a] = [a']$ and $[b] = [b']$, i.e. both $c = a - a'$ and $d = b - b'$ are in I, we have

$$[a] \cdot [b] \;=\; [ab] \;=\; \big[(a'+c)(b'+d)\big] \;=\; \big[a'b' + \underbrace{a'd + cb' + cd}_{\in I}\big] \;=\; [a'b'] \;=\; [a'] \cdot [b'],$$

which shows that the product $[a] \cdot [b] = [ab]$ is well-defined. It is additive on the degrees of homogeneous elements $[a] \in \overline{B}_i$ and $[b] \in \overline{B}_j$ since $a \in B_i$ and $b \in B_j$ implies $ab \in B_{i+j}$ and thus $[a] \cdot [b] = [ab] \in \overline{B}_{i+j}$.

The quotient map $\pi_I : B \to \overline{B}$ is tautologically a graded homomorphism. In consequence, $\iota_{\overline{B}} = \pi_I \circ \iota_B : A \to \overline{B}$ is a ring homomorphism with $\iota_{\overline{B}}(A) \subset \overline{B}_0$. This concludes the proof of the lemma. □

Proposition 4.1.5. *Let $B = \bigoplus B_i$ be a graded A-algebra and $I = \bigoplus I_i$ a homogeneous ideal of B where $I_i = I \cap B_i$. Then the quotient B/I together with the quotient map $\pi : B \to B/I$ satisfies the following universal property: for every graded A-algebra C and every graded homomorphism $f : B \to C$ such that $f(I) = \{0\}$, there is a unique graded homomorphism $\bar{f} : B/I \to C$ such that $f = \bar{f} \circ \pi$, i.e. the diagram*



commutes.

Proof. We leave the proof as Exercise 4.1. □

4.2 The tensor algebra

Let A be a ring.

Definition 4.2.1. Let M be an A-module. The **tensor algebra of** M is the A-module

$$\mathrm{T}(M) \;=\; \bigoplus_{i\in\mathbb{N}} \mathrm{T}^i(M) \quad \text{where} \quad \mathrm{T}^0(M) \;=\; A \quad \text{and} \quad \mathrm{T}^i(M) \;=\; \underbrace{M\otimes_A\cdots\otimes_A M}_{i\text{-times}}.$$

Please note the calligraphic difference in the notation $\mathrm{T}(M)$ for the tensor algebra and the notation $T(M)$ for the torsion submodule of M.

Lemma 4.2.2. *Let M be an A-module. Then the* canonical inclusion $\iota_{\mathrm{T}(M)} : A = \mathrm{T}^0(M) \to \mathrm{T}(M)$ *and the multiplication map given by*

$$\begin{array}{rccc} m_{i,j}: & \mathrm{T}^i(M)\times\mathrm{T}^j(M) & \longrightarrow & \mathrm{T}^{i+j}(M) \\ & \big(m_1\otimes\cdots\otimes m_i, n_1\otimes\ldots\otimes n_j\big) & \longmapsto & m_1\otimes\ldots\otimes m_i\otimes n_1\otimes\ldots\otimes n_j \end{array}$$

on the homogeneous parts of $\mathrm{T}(M)$ *turn* $\mathrm{T}(M)=\bigoplus\mathrm{T}^i(M)$ *into a graded A-algebra.*

Proof. By definition, $\iota_{\mathrm{T}(M)}$ is a ring homomorphism with $\iota_{\mathrm{T}(M)}(A)\subset\mathrm{T}^0(M)$, and $\mathrm{T}(M)$ is an A-module with respect to the grading $\mathrm{T}(M)=\bigoplus\mathrm{T}^i(M)$. We leave it as an exercise to verify that the maps $m_{i,j}$ define a multiplication $m:\mathrm{T}(M)\times\mathrm{T}(M)\to\mathrm{T}(M)$ that turns $\mathrm{T}(M)$ into an A-algebra. □

Remark. Typically the tensor algebra $\mathrm{T}(M)$ is non-commutative. For example if $M=A^2$ with basis $\{v_1,v_2\}$, then $v_1\cdot v_2 = v_1\otimes v_2$ is not equal to $v_2\cdot v_1 = v_2\otimes v_2$. Note that if M is a free A-module (of rank r), then $\mathrm{T}^i(M)$ is a free A-module (of rank r^i), and thus $\mathrm{T}(M)$ is a free A-module.

Lemma 4.2.3. *Let M and N be A-modules and $f:M\to N$ a homomorphism. Then the map* $\mathrm{T}(f):\mathrm{T}(M)\to\mathrm{T}(N)$ *that maps a homogeneous element $m_1\otimes\cdots\otimes m_i\in\mathrm{T}^i(M)$ to $f(m_1)\otimes\cdots\otimes f(m_i)\in\mathrm{T}^i(N)$ is a graded homomorphism. This defines a covariant functor* $\mathrm{T}:\mathrm{Mod}_A\to\mathrm{GrAlg}_A$.

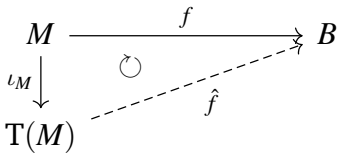
Proof. Let $g=\mathrm{T}(f)$ and $a,b\in\mathrm{T}(M)$. We have $g(a+b)=g(a)+g(b)$ by the multilinearity of the tensor product and $g(ab)=\big(g(a)\big)\cdot\big(g(b)\big)$ by the definition of the product of the tensor algebra and $\mathrm{T}(f)$. Also the remaining properties $g\circ\iota_{\mathrm{T}(M)}=\iota_{\mathrm{T}(N)}$ and $g\big(\mathrm{T}^i(M)\big)\subset\mathrm{T}^i(N)$ follow at once from the definition of $\mathrm{T}(f)$. Thus $\mathrm{T}(f):\mathrm{T}(M)\to\mathrm{T}(N)$ is a graded homomorphism.

We continue with verifying that this defines a covariant functor $\mathrm{T}:\mathrm{Mod}_A\to\mathrm{GrAlg}_A$. Clearly, $\mathrm{T}(\mathrm{id}_M):\mathrm{T}(M)\to\mathrm{T}(M)$ is the identity map. Given homomorphisms $f:M\to N$ and $g:N\to P$ of A-modules, we have

$$\begin{aligned}\big(\mathrm{T}(g\circ f)\big)\big(m_1\otimes\cdots\otimes m_i\big) \;&=\; (g\circ f)(m_1)\otimes\cdots\otimes(g\circ f)(m_1)\\ &=\; \big(\mathrm{T}(g)\big)\big(f(m_1)\otimes\cdots\otimes f(m_1)\big) \;=\; \big(\mathrm{T}(g)\circ\mathrm{T}(f)\big)\big(m_1\otimes\cdots\otimes m_i\big),\end{aligned}$$

which shows that $\mathrm{T}:\mathrm{Mod}_A\to\mathrm{GrAlg}_A$ is indeed a covariant functor. □

Proposition 4.2.4. *Let M be an A-module. Then the tensor algebra $\mathrm{T}(M)$ of M together with the canonical inclusion $\iota_M : M = \mathrm{T}^1(M) \to \mathrm{T}(M)$ satisfies the following universal property: for every graded A-algebra $B = \bigoplus B_i$ and every A-linear map $f : M \to B_1$, there is a unique graded homomorphism $\hat{f} : \mathrm{T}(M) \to B$ of graded A-algebras such that $f = \hat{f} \circ \iota_M$, i.e. the diagram*



commutes.

Proof. We leave the proof as Exercise 4.2. □

4.3 The symmetric algebra

Let A be a (commutative) ring. Let S_i be the symmetric group on $\{1,\dots,i\}$.

Definition 4.3.1. Let M be an A-module and $\mathrm{T}(M)$ the tensor algebra of M. Define $I_0 = \{0\}$ and for all $i > 0$ the submodules

$$I_i \;=\; \left\langle m_1 \otimes \cdots \otimes m_i - m_{\sigma(1)} \otimes \cdots \otimes m_{\sigma(i)} \in \mathrm{T}^i(M) \,\middle|\, m_1 \otimes \cdots \otimes m_i \in \mathrm{T}^i(M), \sigma \in S_i \right\rangle$$

of $\mathrm{T}^i(M)$. Let $I = \bigoplus_{i\in\mathbb{N}} I_i$. The **symmetric algebra of** M is the A-module $\mathrm{Sym}(M) = T(M)/I$.

Lemma 4.3.2. *Let M be an A-module, $\mathrm{T}(M)$ its tensor algebra and $I = \bigoplus I_i$ the submodule of $\mathrm{T}(M)$ from Definition 4.3.1. Then I is a graded ideal of $\mathrm{T}(M)$ and $\mathrm{Sym}(M) = \bigoplus \mathrm{Sym}^i(M)$ for $\mathrm{Sym}^i(M) = \mathrm{T}^i(M)/I_i$. The symmetric algebra $\mathrm{Sym}(M)$ is a commutative graded A-algebra.*

Proof. We begin with the verification that I is a graded ideal of $\mathrm{T}(M)$. Clearly, $I_i \subset \mathrm{T}^i(M)$. Given a permutation $\sigma \in S_i$ and elements $m_1 \otimes \cdots \otimes m_i \in \mathrm{T}^i(M)$ and $n_1 \otimes \cdots \otimes n_j \in \mathrm{T}^j(M)$, we have

$$\begin{aligned}\left(m_1 \otimes \cdots \otimes m_i - m_{\sigma(1)} \otimes \cdots \otimes m_{\sigma(i)}\right) \cdot \left(n_1 \otimes \cdots \otimes n_j\right) \\ = \; m_1 \otimes \cdots \otimes m_i \otimes n_1 \otimes \cdots \otimes n_j - m_{\sigma(1)} \otimes \cdots \otimes m_{\sigma(i)} \otimes n_1 \otimes \cdots \otimes n_j,\end{aligned}$$

which is an element of I^{i+j}. By linear extension to sums of homogeneous elements, this shows that $I\,\mathrm{T}(M) = I$ and, similarly, $\mathrm{T}(M)I = I$, which completes the proof that I is a graded ideal of $\mathrm{T}(M)$.

By Lemma 4.1.4, the quotient $\mathrm{Sym}(M) = \mathrm{T}(M)/I$ is a graded A-algebra. That $\mathrm{Sym}(M)$ is commutative can be verified on generators of the forms $\bar{m} = [m_1 \otimes \cdots \otimes m_i]$ and $\bar{n} = [n_1 \otimes \cdots \otimes n_j]$ of $\mathrm{Sym}(M)$. Since

$$m_1 \otimes \cdots \otimes m_i \otimes n_1 \otimes \cdots \otimes n_j - n_1 \otimes \cdots \otimes n_j \otimes m_1 \otimes \cdots \otimes m_i$$

is in I^{i+j}, we have $\bar{m}\bar{n} = \bar{n}\bar{m}$ in $\mathrm{Sym}(M)$, which shows that $\mathrm{Sym}(M)$ is commutative and completes the proof. □

Lemma 4.3.3. *Let M and N be A-modules and $f : M \to N$ an A-linear map. Then the association*

$$\begin{array}{rccc} \mathrm{Sym}(f): & \mathrm{Sym}(M) & \longrightarrow & \mathrm{Sym}(N) \\ & [m_1 \otimes \cdots \otimes m_i] & \longmapsto & [f(m_1) \otimes \cdots \otimes f(m_i)] \end{array}$$

defines a graded homomorphism of A-algebras. This yields a covariant functor $\mathrm{Sym} : \mathrm{Mod}_A \to \mathrm{GrAlg}_A$.

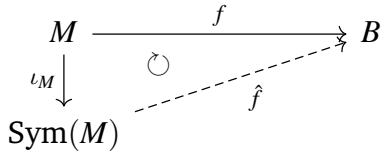
Proof. Let I_M and I_N be the graded ideals that define the respective symmetric algebras $\mathrm{Sym}(M) = \mathrm{T}(M)/I_M$ and $\mathrm{Sym}(N) = \mathrm{T}(N)/I_N$. From the definition of the graded homomorphism $\mathrm{T}(f) : \mathrm{T}(M) \to \mathrm{T}(N)$, it is evident that it maps I_M to I_N. Thus by the universal property of graded quotients (Proposition 4.2.4), the graded homomorphism $\mathrm{T}(M) \to \mathrm{T}(N) \to \mathrm{Sym}(N)$ induces a morphism $\mathrm{Sym}(f) : \mathrm{Sym}(M) \to \mathrm{Sym}(N)$ that maps $[m] \in \mathrm{Sym}(M) = \mathrm{T}(M)/I_M$ to $\big(\mathrm{T}(f)\big)(m)$.

Clearly, $\mathrm{Sym}(\mathrm{id}_M) = \mathrm{id}_{\mathrm{Sym}(M)}$. Given two homomorphisms $f : M \to N$ and $g : N \to P$ of A-modules, it is immediately verified on generators that $\mathrm{Sym}(g \circ f) = \mathrm{Sym}(g) \circ \mathrm{Sym}(f)$. Thus $\mathrm{Sym} : \mathrm{Mod}_A \to \mathrm{GrAlg}_A$ is a covariant functor. □

Proposition 4.3.4. *Let M be an A-module, $\mathrm{T}(M)$ its tensor algebra and $\mathrm{Sym}(M)$ its symmetric algebra.*

(1) The composition $\iota_M : M \to \mathrm{T}(M) \to \mathrm{Sym}(M)$ of the canonical inclusion $M = \mathrm{T}^1(M) \hookrightarrow \mathrm{T}(M)$ with the quotient map $\mathrm{T}(M) \to \mathrm{Sym}(M)$ is an injective homomorphism of A-modules with image $\mathrm{Sym}^1(M)$.

(2) The symmetric algebra $\mathrm{Sym}(M)$ together with the the inclusion $\iota_M : M \to \mathrm{Sym}(M)$ satisfies the following universal property: for every commutative graded A-algebra $B = \bigoplus B_i$ and every A-linear map $f : M \to B_1$, there is a unique graded homomorphism $\hat{f} : \mathrm{Sym}(M) \to B$ of graded A-algebras such that $f = \hat{f} \circ \iota_M$, i.e. the diagram



commutes.

Proof. We leave the proof as Exercise 4.3. □

Recall from Example 4.1.3 that the polynomial ring $B = A[T_1, \ldots, T_n]$ is a graded A-algebra with respect to the A-submodules

$$B_i \;=\; \big\{ aT_1^{e_1} \cdots T_n^{e_n} \,\big|\, a \in A, e_1, \ldots, e_n \in \mathbb{N} \text{ with } e_1 + \cdots + e_n = i \big\}.$$

Lemma 4.3.5. *Let M be a free A-module of finite rank r with basis $\mathcal{B} = \{v_1, \ldots, v_r\}$. Then the association*

$$\begin{array}{ccc} A[T_1, \ldots, T_r] & \longrightarrow & \mathrm{Sym}(M) \\ T_i & \longmapsto & v_i \end{array}$$

defines an graded isomorphism of graded A-algebras, and $\mathrm{Sym}^i(M)$ is free of rank

$$\binom{r-1+i}{i} = \frac{(r-1+i)!}{i!(r-1)!}.$$

Proof. We leave the proof as Exercise 4.4. □

4.4 The exterior algebra

Let A be a commutative ring.

Definition 4.4.1. Let M be an A-module and $\mathrm{T}(M)$ its tensor algebra. Define $I_0 = \{0\}$ and for $i > 0$ the submodules

$$I_i = \left\langle m_1 \otimes \ldots \otimes m_i \in \mathrm{T}^i(M) \,\middle|\, m_k = m_l \text{ for some } k \neq l \right\}$$

of $\mathrm{T}^i(M)$. Let $I = \bigoplus_{i \in \mathbb{N}} I_i$. The **exterior algebra of** M is the quotient $\Lambda(M) = \mathrm{T}(M)/I$. The **$i$-th exterior power of** M is the quotient $\Lambda^i(M) = \mathrm{T}^i(M)/I_i$. We write $m_1 \wedge \ldots \wedge m_i$ for the class of $m_1 \otimes \ldots \otimes m_i$ in $\Lambda(M)$.

Lemma 4.4.2. *Let M be an A-module, $\mathrm{T}(M)$ its tensor algebra and $I = \bigoplus_{i \in \mathbb{N}} I_i$ the submodule of $\mathrm{T}(M)$ from Definition 4.4.1. Then I is a graded ideal of $\mathrm{T}(M)$ and $\Lambda(M) = \bigoplus \Lambda^i(M)$ is a graded A-algebra.*

Proof. By definition, we have $I_i \subset \mathrm{T}^i(M)$ and $I = \bigoplus I_i$. Consider elements $m_1 \otimes \cdots \otimes m_i \in I_i$, i.e. $m_k = m_l$ for some $k \neq l$, and $n_1 \otimes \cdots \otimes n_j \in \mathrm{T}^j(M)$. Then their product

$$m_1 \otimes \cdots \otimes m_i \otimes n_1 \otimes \cdots \otimes n_j$$

still satisfies $m_k = m_l$ and is thus contained in I_{i+j}. Thus $I\Lambda(M) = I$, and similarly, $\Lambda(M)I = I$. This shows that $I = \bigoplus_{i \in \mathbb{N}} I_i$ is a graded ideal. By Proposition 4.1.5, the quotient $\Lambda(M) = \bigoplus \Lambda^i(M)$ is a graded A-algebra. □

Lemma 4.4.3. *Let M and N be A-modules. Then the association*

$$\begin{array}{rccc} \Lambda(f): & \Lambda(M) & \longrightarrow & \Lambda(N) \\ & m_1 \wedge \cdots \wedge m_i & \longmapsto & f(m_1) \wedge \cdots \wedge f(m_i) \end{array}$$

defines a graded homomorphism of A-algebras. This yields a covariant functor $\Lambda : \mathrm{Mod}_A \to \mathrm{GrAlg}_A$.

Proof. Let I_M and I_n be the respective defining ideals of $\Lambda(M) = \mathrm{T}(M)/I_M$ and $\Lambda(N) = \mathrm{T}(N)/I_N$. The graded homomorphism $\mathrm{T}(f) : \mathrm{T}(M) \to \mathrm{T}(N)$ maps I_M to I_N, and thus the composition $\mathrm{T}(M) \to \mathrm{T}(N) \to \Lambda(N)$ with the quotient map $\pi : \mathrm{T}(N) \to \Lambda(N)$ maps all elements of I_M to 0. Thus by the universal property of quotients of graded algebras (Proposition 4.1.5), there is a unique graded homomorphism $\Lambda(f) : \Lambda(M) \to \Lambda(N)$ that maps $m_1 \wedge \ldots \wedge m_i$ to $f(m_1) \wedge \ldots \wedge f(m_i)$, which completes the proof. □

Proposition 4.4.4. *Let M be an A-module and $\Lambda(M)$ its exterior algebra. Then the following holds.*

(1) For all $m, n \in M$, we have $m \wedge n = -n \wedge m$.

(2) The defining ideal I of $\Lambda(M) = \mathrm{T}(M)/I$ is the smallest graded ideal of $\mathrm{T}(M)$ that contains the submodule

$$I_2 \;=\; \big\langle m \otimes m \,\big|\, m \in M \big\rangle_A .$$

(3) If M is generated by r elements, then $\Lambda^i(M) = 0$ for all $i > r$.

Proof. Claim (1) follows since

$$0 \;=\; (m+n) \wedge (m+n) \;=\; \underbrace{m \wedge m}_{=0} + m \wedge n + n \wedge m + \underbrace{n \wedge n}_{=0}$$

implies $m \wedge n = -n \wedge m$, as desired.

We continue with (2). Given $m_1 \otimes \ldots \otimes m_i$ with $m_k = m_l$ for some $k < l$, we can use (1) repeatedly to gain the equality

$$m_1 \otimes \ldots \otimes m_i \;=\; \pm \underbrace{(m_k \otimes m_l)}_{\in I_2} \cdot \big(m_1 \otimes \ldots \widehat{m_k} \ldots \widehat{m_l} \ldots \otimes m_i\big),$$

which shows that every ideal containing I_2 contains I. Thus (2).

We continue with (3). If M is generated by $v_1, \ldots, v_r$, then every element of $\mathrm{T}^i(M)$ can be written as a linear combination

$$\sum a_\sigma v_{\sigma(1)} \otimes \ldots \otimes v_{\sigma(i)}$$

for certain $a_\sigma \in A$ where σ varies through all maps $\sigma : \{1, \ldots, i\} \to \{1, \ldots, r\}$. If $i > r$, then every map $\sigma : \{1, \ldots, i\} \to \{1, \ldots, r\}$ fails to be injective, and thus every element $v_{\sigma(1)} \otimes \ldots \otimes v_{\sigma(i)}$ in this expression is in I_i. This shows that $\sum a_\sigma v_{\sigma(1)} \wedge \ldots \wedge v_{\sigma(i)} = 0$ for $i > r$. Thus (3). □

Definition 4.4.5. Let M and N be A-modules and $i \geqslant 0$. A map $f : M^i \to N$ is *multi-linear* if for all $k \in \{1, \ldots, i\}$, for all $a \in A$ and for all elements $m = (m_1, \ldots, m_i)$ and $n = (n_1, \ldots, n_i)$ of M^i with $m_l = n_l$ for $l \neq k$, we have

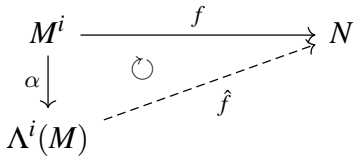
$$\begin{aligned} f\big((m_1, \ldots, m_{k-1}, a.m_k, m_{k+1}, \ldots, m_i)\big) &= a.f(m), \\ f\big((m_1, \ldots, m_{k-1}, m_k + n_k, m_{k+1}, \ldots, m_i)\big) &= f(m) + f(n). \end{aligned}$$

A map $M^i \to N$ is **alternating** if it is multi-linear and if $f\big((m_1, \ldots, m_i)\big) = 0$ for every $(m_1, \ldots, m_i) \in M^i$ for which $m_k = m_l$ for some $k \neq l$.

Proposition 4.4.6. *Let M be an A-module, $i \geqslant 0$ and $\Lambda^i(M)$ its i-th exterior power. Then the following holds.*

(1) The map $\alpha : M^i \to \Lambda^i(M)$ that sends $(m_1, \ldots, m_i)$ to $m_1 \wedge \ldots \wedge m_i$ is alternating.

(2) The i-th exterior power $\Lambda^i(M)$ of M together with the alternating map $\alpha : M^i \to \Lambda^i(M)$ satisfies the following universal property: for every A-module N and every alternating map $f : M^i \to N$, there is a unique A-linear map $\hat{f} : \Lambda^i(M) \to N$ such that $f = \hat{f} \circ \alpha$, i.e. the diagram



commutes.

Proof. The map α is the composition of the multilinear map $\beta : M^i \to \mathrm{T}^i(M)$ to the i-th tensor power of M followed by the quotient map $\pi : \mathrm{T}^i(M) \to \Lambda^i(M)$, and therefore multilinear itself. By the definition of $\Lambda^i(M)$, we have $\alpha\big((m_1, \ldots, m_i)\big) = 0$ if $m_k = m_l$ for some $k \neq l$. Thus α is alternating, as claimed in (1).

Consider an alternating map $f : M^i \to N$. By the universal property of the tensor product (Proposition 3.3.3), there is a unique A-linear map $\tilde{f} : \mathrm{T}^i(M) \to N$ such that $f = \tilde{f} \circ \beta$. For $m_1 \otimes \ldots \otimes m_i \in \mathrm{T}^i(M)$ with $m_k = m_l$ for some $k \neq l$, we have by our assumptions

$$\tilde{f}(m_1 \otimes \ldots \otimes m_i) \;=\; f\big((m_1, \ldots, m_i)\big) \;=\; 0.$$

Since the defining submodule I_i of $\Lambda^i(M) = \mathrm{T}^i(M)/I_i$ is generated by elements $m_1 \otimes \ldots \otimes m_i$ with $m_k = m_l$ for some $k \neq l$, the universal property of quotient modules (Propositions 3.2.2) implies that there is a unique morphism $\hat{f} : \Lambda^i(M) \to N$ such that $\tilde{f} = \hat{f} \circ \pi$. Thus $f = \tilde{f} \circ \alpha = \hat{h} \circ \pi \circ \alpha = \hat{f} \circ \beta$, as desired. The uniqueness of $\hat{f}$ follows from the construction. □

Example 4.4.7. The following is a key example of an alternating map. Let M be a free A-module with basis $\{v_1, \ldots, v_r\}$. Then the determinant $\det : M^r \to A$, which is defined by

$$\det\Big(\sum_{i=1}^{r} a_{i,1} v_i, \ldots, \sum_{i=1}^{r} a_{i,r} v_i\Big) \;=\; \sum_{\sigma \in S_i} \Big(\mathrm{sign}(\sigma) \prod_{i=1}^{r} a_{i,\sigma(i)}\Big),$$

is alternating. By the universal property of the r-th exterior power of M, this yields an A-linear map $\overline{\det} : \Lambda^r(M) \to A$ with $\det = \overline{\det} \circ \alpha$ where $\alpha : M^r \to \Lambda^r(M)$ is the alternating map that sends $(m_1, \ldots, m_r)$ to $m_1 \wedge \ldots \wedge m_r$. Note that since $\det\big((v_1, \ldots, v_r)\big) = 1$, the A-linear map $\overline{\det} : \Lambda^r(M) \to A$ is surjective.

Let $i, r \in \mathbb{N}$. A map $\sigma : \{1, \ldots, i\} \to \{1, \ldots, r\}$ is *strictly order preserving* if $\sigma(k) < \sigma(l)$ for all $k, l \in \{1, \ldots, i\}$ with $k < l$.

Theorem 4.4.8. *Let $i, r \in \mathbb{N}$ with $i \leqslant r$ and M be a free A-module with basis $\{v_1, \ldots, v_r\}$. Then $\Lambda^i(M)$ is a free A-module of rank $\binom{r}{i}$ with basis*

$$\mathcal{B}_i \;=\; \Big\{ v_{\sigma(1)} \wedge \ldots \wedge v_{\sigma(i)} \,\Big|\, \sigma : \{1,\ldots,i\} \to \{1,\ldots,r\} \text{ strictly order preserving} \Big\}$$

where we apply the convention that the empty wedge product is 1, i.e. $\mathcal{B}_0 = \{1\}$.

Proof. By Proposition 4.4.4.(1), we have

$$v_{k_{\sigma(1)}} \wedge \ldots \wedge v_{k_{\sigma(i)}} \;=\; \operatorname{sign}(\sigma).\big(v_{k_1} \wedge \ldots \wedge v_{k_i}\big)$$

for all $k_1, \ldots, k_i \in \{1,\ldots,r\}$ and permutations $\sigma \in S_i$, and $v_{k_1} \wedge \ldots \wedge v_{k_i} = 0$ if $k_r = k_s$ for some $r \neq s$. This shows that $\mathcal{B}_i$ generates M, and thus the natural map $\Phi : \bigoplus_{v \in \mathcal{B}_i} A.v \to \Lambda^i M$ is surjective.

We continue with showing that Φ is injective and thus an isomorphism. If $i = 0$, then $\Lambda^0(M) = \mathrm{T}^0(M) = A$ by definition, and thus $\mathcal{B}_0 = \{1\}$ is a basis. For $i = r$, the association $a \mapsto a.(v_1 \wedge \ldots \wedge v_r)$ defines an A-linear map $A \to \Lambda^r(M)$ that is inverse to the A-linear map $\underline{\det} : \Lambda(M)^r \to A$ from Example 4.4.7. Thus $\Lambda^r(M) \simeq A$ is free of rank 1 with basis $\mathcal{B}_r = \{v_1 \wedge \ldots \wedge v_r\}$.

For $1 \leqslant i \leqslant r-1$, we consider a relation

$$\sum_{\substack{\sigma:\{1,\ldots,i\}\to\{1,\ldots,r\} \\ \text{strictly order preserving}}} a_\sigma.\big(v_{\sigma(1)} \wedge \ldots \wedge v_{\sigma(i)}\big) \;=\; 0 \wedge \ldots \wedge 0.$$

with $a_\sigma \in A$. We need to prove that $a_\tau = 0$ for every strictly order preserving map $\tau : \{1,\ldots,i\} \to \{1,\ldots,r\}$. Since τ is injective, we can extend it to a bijection $\hat{\tau} : \{1,\ldots,r\} \to \{1,\ldots,r\}$. This allows us to define the A-linear map $f : \Lambda^i(M) \to \Lambda^r(M)$ with

$$f\big(v_{\sigma(1)} \wedge \ldots \wedge v_{\sigma(i)}\big) \;=\; v_{\sigma(1)} \wedge \ldots \wedge v_{\sigma(i)} \wedge v_{\hat{\tau}(i+1)} \wedge \ldots \wedge v_{\hat{\tau}(r)}$$

If $\sigma : \{1,\ldots,i\} \to \{1,\ldots,r\}$ is another order preserving map with $\operatorname{im}\sigma = \operatorname{im}\tau$, then necessarily $\sigma = \tau$. Thus if $\sigma \neq \tau$, then $\operatorname{im}\sigma$ contains an element of $\big\{\hat{\tau}(i+1), \ldots, \hat{\tau}(r)\big\}$. Therefore we can deduce that

$$\begin{aligned}
0 \;&=\; 0 \wedge \ldots \wedge 0 \wedge v_{\hat{\tau}(i+1)} \wedge \ldots \wedge v_{\hat{\tau}(r)} \\
&=\; \sum_{\substack{\sigma:\{1,\ldots,i\}\to\{1,\ldots,r\} \\ \text{strictly order preserving}}} a_\sigma.\big(v_{\sigma(1)} \wedge \ldots \wedge v_{\sigma(i)} \wedge v_{\hat{\tau}(i+1)} \wedge \ldots \wedge v_{\hat{\tau}(r)}\big) \\
&=\; a_\tau.\big(v_{\hat{\tau}(1)} \wedge \ldots \wedge v_{\hat{\tau}(i)} \wedge v_{\hat{\tau}(i+1)} \wedge \ldots \wedge v_{\hat{\tau}(r)}\big) \\
&=\; \big(\operatorname{sign}(\hat{\tau}) a_\tau\big).\big(v_1 \wedge \ldots \wedge v_i\big).
\end{aligned}$$

Since $\Lambda^r(M) \simeq A$, this implies that $a_\tau = 0$. Thus $\Phi : \bigoplus_{v \in \mathcal{B}_i} A.v \to \Lambda^i M$ is an isomorphism and $\mathcal{B}_i$ is a basis for $\Lambda^i(M)$. The rank of $\Lambda^i(M)$ is equal to

$$\#\mathcal{B}_i \;=\; \big\{ i\text{-subset of } \{1,\ldots,r\} \big\} \;=\; \tbinom{r}{j},$$

which completes the proof of the theorem. □

4.5 Exercises

Exercise 4.1. Prove Proposition 4.1.5.

Exercise 4.2. Prove Proposition 4.2.4.

Exercise 4.3. Prove Proposition 4.3.4.

Exercise 4.4. Prove Lemma 4.3.5.

Exercise 4.5. Let M be a free A-module of rank $r > 0$ with basis $\{v_1, \ldots, v_r\}$. Show that the map

$$\begin{array}{ccc} A[T_1, \ldots, T_r] & \longrightarrow & \mathrm{Sym}(M) \\ T_i & \longmapsto & v_i \end{array}$$

is a graded isomorphism of graded A-algebras. Show that $T^i(M)$ is free of rank r^i and that $\mathrm{Sym}^i(M)$ is free of rank $\binom{r+i-1}{i}$.

Exercise 4.6. Let A be a $\mathbb{Q}$-algebra and M a finitely generated A-module. The *exponential map* $\exp : \Lambda(M) \to \Lambda(M)$ is defined by the formula

$$\exp(x) \;=\; 1 + \sum_{k \geqslant 1} \frac{1}{k!} \underbrace{(x \wedge \ldots \wedge x)}_{k-\text{times}}.$$

Show that $\exp(x)$ is equal to a finite sum, and therefore well-defined as an element of $\Lambda(M)$. Calculate the expressions $\exp(m)$ and $\exp(m \wedge n + o \wedge p)$ where $m, n, o, p \in M$. Does the formula $\exp(x+y) = \exp(x) \wedge \exp(y)$ hold for any $x, y \in \Lambda(M)$?

Exercise 4.7. Let M be an A-module. Consider the ideals

$$I \;=\; \langle m \otimes m \,|\, m \in M \rangle \quad \text{and} \quad J \;=\; \langle m \otimes n + n \otimes m \,|\, m, n \in M \rangle$$

of $T(M)$. Show that $I = J$ if 2 is invertible in A. Give an example for A and M where $I \neq J$.

Exercise 4.8. Let $l \leqslant r$ be positive integers and M a free A-module with basis $\{v_1, \ldots, v_r\}$. For $i = 1, \ldots, r$ and $j = 1, \ldots, l$, let $a_{i,j} \in A$ and define the elements

$$m_j \;=\; \sum_{i=1}^{r} a_{i,j} v_i$$

of M.

(1) Show that there is a unique element $\delta_\sigma \in A$ for every strictly order preserving maps $\sigma : \{1, \ldots, l\} \to \{1, \ldots, n\}$ such that

$$m_1 \wedge \ldots \wedge m_l \;=\; \sum \delta_\sigma . \big(v_{\sigma(1)} \wedge \ldots \wedge v_{\sigma(l)}\big)$$

as elements of $\Lambda^l(M)$ where σ ranges through all strictly order preserving maps $\sigma : \{1, \ldots, l\} \to \{1, \ldots, n\}$. Show that

$$\delta_\sigma \;=\; \det(a_{i,j})_{\substack{i \in \mathrm{im}\,\sigma \\ j = 1, \ldots, l}}.$$

(2) Let $f : M \to M$ be an endomorphism and $\Lambda^r(f) : \Lambda^r(M) \to \Lambda^r(M)$ the induced linear map. Let $a_{i,j} \in A$ such that $f(v_i) = \sum_{j=1}^r a_{i,j}.v_j$ for $i = 1,\ldots,r$. Conclude that

$$\delta \;=\; \det(a_{i,j})_{i,j=1,\ldots,r}$$

is the unique element of A such that

$$\big(\Lambda^r(f)\big)\big(v_1 \wedge \ldots \wedge v_r\big) \;=\; \delta.\big(v_1 \wedge \ldots \wedge v_r\big).$$

Exercise 4.9. This is a continuation of Exercise 4.8. However, we assume that $A = k$ is field for this exercise. Consequently, M is a k-vector space.

(1) Show that $m_1 \wedge \ldots \wedge m_l \neq 0$ if and only if $\{m_1,\ldots,m_l\}$ is linearly independent.

(2) Assume that $\{m_1,\ldots,m_l\}$ and $\{m'_1,\ldots,m'_l\}$ are linearly independent subsets of M. Show that there is a $\lambda \in k^\times$ such that

$$m'_1 \wedge \ldots \wedge m'_l \;=\; \lambda \cdot m_1 \wedge \ldots \wedge m_l.$$

if and only if $\{m_1,\ldots,m_l\}$ and $\{m'_1,\ldots,m'_l\}$ span the same l-dimensional subvector space N of M.
Hint: If they span the same subvector space N, then one can find a $l \times l$-base change matrix. What is the effect of this matrix on the coefficients δ_σ from Exercise 4?

(3) Define $\mathbb{P}\big(\Lambda^l(M)\big) = \big(\Lambda^l(M) - \{0\}\big)/k^\times$ as the set of equivalent classes of nonzero elements of $\Lambda^l(M)$ modulo scalar multiplication by nonzero $\lambda \in k^\times$. Conclude from the previous part of the exercise that there is a well-defined inclusion

$$\big\{l\text{-dimensional subvector spaces of } M\big\} \quad\longrightarrow\quad \mathbb{P}\big(\Lambda^l(M)\big).$$

Remark: The set $\mathbb{P}\big(\Lambda^l(M)\big)$ is called the *projective space of* $\Lambda^l(M)$, the above inclusion is called the *Plücker embedding* and its image is called the *Grassmann variety* $\mathrm{Gr}(l,n)$ *of* l*-subspaces in* n*-space*.

Chapter 5

Groups

5.1 Basic definitions

Definition 5.1.1. A **group** is a set G together with a map

$$\begin{array}{rccc} m: & G\times G & \longrightarrow & G \\ & (a,b) & \longmapsto & a\cdot b = ab \end{array}$$

such that

(1) $(ab)c = a(bc)$ for all $a,b,c \in G$, *(associativity)*

(2) there is an $e \in G$ such that $ae = a$ for all $a \in G$, *(neutral element)*

(3) for every $a \in G$, there is a $b \in G$ such that $ab = e$ *(inverses)*

for all $a,b,c \in G$ where $(ab)c = m(m(a,b),c)$ and $a(bc) = m(a,m(b,c))$. We call the map m the *multiplication of* G.

Note that the associativity allows us to write products $a_1 \cdots a_n$ without ambiguity in which order we multiply the elements $a_1,\ldots,a_n \in G$.

Lemma 5.1.2. *Let G be a group and $e,a,b \in G$ such that $ce = c$ for all $c \in G$ and $ab = e$.*

(1) If $c \in G$ satisfies $c\cdot c = c$, then $c = e$.

(2) We have $ec = c$ for all $c \in G$. If an element $e' \in G$ satisfies $ce' = c$ for all $c \in G$, then $e' = e$.

(3) We have $ba = e$. If an element $b' \in G$ satisfies $ab' = e$, then $b' = b$.

Proof. Let $c \in G$ be an element with $c\cdot c = c$. By axiom (3), there is an element $d \in G$ such that $cd = e$. Thus $c = ce = ccd = cd = e$, which establishes (1).

Using this observation, we use all axioms of a group to conclude that

$$(ba)(ba) \;=\; b(ab)a \;=\; bea \;=\; (be)a \;=\; ba,$$

which implies that $ba = e$ by applying (1) to $c = ba$. This establishes the first claim of (3). Consequently, we have for every $c \in G$ that

$$ec \;=\; (cd)c \;=\; c(dc) \;=\; ce \;=\; c$$

where $d \in G$ satisfies $cd = e$ and thus also $dc = e$. This establishes the first claim of (2).

Given an element $e' \in G$ such that $ce' = c$ for all $c \in G$, then $e' = ee' = e$, which establishes the second claim of (2). Given an element $b' \in G$ such that $ab' = e$, then

$$b' \;=\; eb' \;=\; (ba)b' \;=\; b(ab') \;=\; be \;=\; b,$$

which establishes the second claim of (3) and completes the proof of the lemma. □

We call the element e the *neutral element of* G. We write a^{-1} for the element b with $ab = e$ and call it the *inverse of* a. The association $a \mapsto a^{-1}$ defines the *inversion* $i : G \to G$. For an integer $i > 0$ and $a \in G$, we define

$$a^i \;=\; \underbrace{a \cdots a}_{i\text{-times}}, \qquad a^0 \;=\; e, \qquad a^{-i} \;=\; \underbrace{a^{-1} \cdots a^{-1}}_{i\text{-times}}.$$

Note that $(ab)^{-1} = b^{-1}a^{-1}$ since $(b^{-1}a^{-1})(ab) = e$.

Definition 5.1.3. Let G and G' be groups. A **group homomorphism from G to G'** is a map $f : G \to G'$ such that $f(ab) = f(a)f(b)$ for all $a, b \in G$. This defines the category Groups of groups.

Lemma 5.1.4. *Let $f : G \to H$ be a group homomorphism. Then $f(e) = e$ and $f(a^i) = f(a)^i$ for all $a \in G$ and $i \in \mathbb{Z}$.*

Proof. Since $ee = e$, we have $f(e)f(e) = f(e)$, and thus Lemma 5.1.2,(1) implies that $f(e) = e$ is the neutral element of H. Thus the first claim of the lemma and $f(a^0) = f(a)^0$. For $i > 0$, we have $f(a^i) = f(a)^i$ by a repeated application of the defining property of a group homomorphism.

For $a \in G$ with inverse a^{-1}, we have $f(a)f(a^{-1}) = f(aa^{-1}) = f(e) = e$, and thus $f(a^{-1}) = f(a)^{-1}$ is the unique inverse of $f(a)$ in H. Using that $a^i = (a^{-1})^{-i}$, we conclude that $f(a^i) = f(a)^i$ for $i < 0$. □

Definition 5.1.5. A **subgroup of** G is a nonempty subset H of G such that $ab^{-1} \in H$ for all $a, b \in H$. We write $H < G$ to denote a subgroup H of G.

Let S be a subset of G. The **subgroup generated by** S is the intersection $\langle S \rangle$ of all subgroups of G that contain S. We write $\langle a_1, \ldots, a_n \rangle$ for $\langle \{a_1, \ldots, a_n\} \rangle$. A group G is **cyclic** if $G = \langle a \rangle$ for some $a \in G$.

Let I be a set and $\{G_i\}_{i \in I}$ be a family of groups. The **product of** $\{G_i\}$ is the group

$$\prod_{i \in I} G_i \;=\; \big\{ (a_i)_{i \in I} \,\big|\, a_i \in G_i \big\},$$

together with the coordinatewise multiplication given by $(a_i) \cdot (b_i) = (a_i b_i)$. The **direct sum of** $\{G_i\}$ is the subgroup

$$\bigoplus_{i \in I} G_i \;=\; \Big\{ (a_i)_{i \in I} \in \prod_{i \in I} G_i \,\Big|\, a_i = e \text{ for all but finitely many } i \in I \Big\}.$$

Remark. We include some remarks on these definitions. A subgroup of a group G is the same thing as a subset H of G such that the multiplication m of G restricts to a map $m_H : H \times H \to H$ and such that H is a group with respect to m_H. Since H is nonempty, it contains an element a and thus $e = aa^{-1}$, as well as $a^{-1} = ea^{-1}$.

Since the intersection of subgroups of G is a subgroup, $\langle S \rangle$ is indeed a subgroup of G. An isomorphism in Groups (in the sense of Definition 2.3.1) is a bijective group homomorphism. We leave the verification of these facts as an exercise, as well as the claims that the product and the direct sum of groups is a group.

5.2 Cosets

Definition 5.2.1. Let G be a group and H a subgroup of G. A **left coset of** H is a subset of the form $aH = \{ah \mid h \in H\}$ of G for some $a \in G$. A **right coset of** H is a subset of the form $Ha = \{ha \mid h \in H\}$ of G for some $a \in G$. We write

$$G/H \;=\; \{aH \mid a \in G\} \qquad \text{and} \qquad H\backslash G \;=\; \{Ha \mid a \in G\}$$

for the families of right and left cosets of G, respectively.

Lemma 5.2.2. *Let G be a group and H a subgroup of G. Then we have for all $a, b \in H$,*

(1) $aH = H$ if and only if $a \in H$;

(2) $b \in aH$ if and only if $aH = bH$, which is the case if and only if $a^{-1}b \in H$;

(3) $aH = bH$ or $aH \cap bH = \varnothing$;

(4) aH and bH have the same (possibly infinite) cardinality.

Proof. We begin with (1). If $aH = H$, then $a = ae \in aH = H$. Conversely, assume that $a \in H$. Since H is a subgroup of G, we have $ah \in H$ for all $h \in H$ and thus $aH \subset H$. Since $a^{-1} \in H$, we have $a^{-1}h \in H$ for all $h \in H$ and thus $h = a(a^{-1}h) \in aH$, which shows that $H \subset aH$. Thus $aH = H$ as claimed, which establishes (1).

We continue with (2). We have $b \in aH$ if and only if $b = ah$ for some $h \in H$, which is the case if and only if $bH = ahH = aH$ since $hH = H$ by (1). Multiplying $aH = bH$ with a^{-1} from the left yields $a^{-1}bH = H$, which is equivalent with $a^{-1}b = h \in H$ by (1). Thus (2).

We continue with (3). If there is an element $c \in aH \cap bH$, then $aH = cH = bH$ by (2). Thus (3).

We continue with (3). The map

$$\begin{array}{ccc} aH & \longrightarrow & bH \\ ah & \longmapsto & (ba^{-1})ah \end{array}$$

is a bijection whose inverse sends an element $bh \in bH$ to $(ab^{-1})bh$. Thus aH and bH have the same cardinality, which establishes (4). □

Definition 5.2.3. Let G be a group and $a \in G$. The **order of** G is the cardinality $\operatorname{ord}(G)$ of G. The **order of** a is the cardinality $\operatorname{ord}(a)$ of the cyclic subgroup $\langle a \rangle$ generated by a. Let H be a subgroup of G. The **index of** H **in** G is the cardinality $(G : H)$ of G/H.

Remark. If the order of $a \in G$ is finite, then it is the smallest positive integer i such that $a^i = e$. If the order of a is infinite, then $a^i \neq e$ for all positive integers i.

Theorem 5.2.4 (Lagrange's theorem)**.** *Let G be a finite group and H a subgroup of G. Then* $\operatorname{ord}(G) = (G : H) \cdot \operatorname{ord}(H)$. *In particular, both the order of H and the index of H in G divide the order of G.*

Proof. By Lemma 5.2.2.(3), the left cosets of H are disjoint, and thus

$$G = \coprod_{aH \in G/H} aH.$$

By Lemma 5.2.2.(4), all cosets have the same cardinality $\#H$, and thus $\#G = \#(G/H) \cdot \#H$, which proves the theorem. □

Corollary 5.2.5. *Let G be a finite group and $a \in G$. Then* $\operatorname{ord}(a)$ *divides* $\operatorname{ord}(G)$.

Proof. This follows at once from Lagrange's theorem (Theorem 5.2.4) applied to the cyclic subgroup $H = \langle a \rangle$. □

5.3 Normal subgroups and quotients

Definition 5.3.1. Let G be a group. A subgroup N of G is **normal** if $aN = Na$ for all $a \in G$. We write $N \lhd G$ for normal subgroups N of G. If N is a normal subgroup of G, then we call G/N the **quotient of** G **by** N.

Remark. Note that we can rewrite the condition $aN = Na$ as $aNa^{-1} = N$. This means that a subgroup N of G is normal if and only if $aNa^{-1} = N$ for all $a \in G$.

Proposition 5.3.2. *Let G be a group and N a normal subgroup of G. Then G/N is a group with respect to the multiplication*

$$\begin{array}{rccc} \bar{m}: & (G/N) \times (G/N) & \longrightarrow & G/N \\ & \big([a],[b]\big) & \longmapsto & [ab] \end{array}$$

where we write $[a]$ for aN. The quotient map $\pi_N : G \to G/N$ *with $\pi(a) = [a]$ is a surjective group homomorphism.*

Proof. We begin with the verification that $\bar{m}$ is well-defined. Consider $a, a', b, b' \in G$ with $[a] = [a']$ and $[b] = [b']$. Then

$$\bar{m}([a],[b]) = abN = ab'N = aNb' = a'Nb' = a'b'N = \bar{m}([a'],[b'])$$

since $b'N = Nb'$. Thus $\bar{m}$ is well-defined as a map.

We verify that $\bar{m}$ turns G/N indeed into a group. The multiplication $\bar{m}$ is associative since for all $a,b,c \in G$,

$$\big([a][b]\big)[c] \;=\; \big[(ab)c\big] \;=\; \big[a(bc)\big] \;=\; [a]\big([b][c]\big).$$

The class $[e]$ of e is neutral for G/N since $[a][e] = [ae] = [e]$ for all $a \in G$. Given $a \in G$, the class $[a^{-1}]$ is an inverse of $[a]$ in G/N, since $[a][a^{-1}] = [aa^{-1}] = [e]$.

The map π_N is clearly surjective. It is a group homomorphism since $\pi_N(ab) = [ab] = [a][b] = \pi_N(a)\pi_N(b)$. This completes the proof. □

Definition 5.3.3. Let $f : G \to H$ be a group homomorphism. The **kernel of** f is the subset $\ker f = \{a \in G \mid f(a) = e\}$ of G.

Lemma 5.3.4. *Let $f : G \to H$ be a group homomorphism. Then its kernel $\ker f$ is a normal subgroup of G.*

Proof. Since $f(e) = e$, the kernel $\ker f$ is not empty. Given $a,b \in \ker f$, we have $f(ab^{-1}) = f(a)f(b)^{-1} = e$, and thus $ab^{-1} \in \ker f$. Thus $\ker f$ is a subgroup of G. Let $a \in G$ and $b \in \ker f$. Then by Lemma 5.1.4,

$$f(aba^{-1}) \;=\; f(a)\underbrace{f(b)}_{=e}f(a)^{-1} \;=\; e,$$

which shows that $aba^{-1} \in \ker f$. Thus $a(\ker f)a^{-1} \subset \ker f$ for all $a \in G$. Multiplying with a^{-1} from the left and with a from the right yields $\ker f = a^{-1}\big(a(\ker f)a^{-1}\big)a \subset a^{-1}(\ker f)a$ for all $a \in G$. Replacing a by a^{-1} in this last inequality shows that $a(\ker f)a^{-1} = \ker f$ for all $a \in G$. Thus $\ker f$ is a normal subgroup of G. □

Remark. Note that a normal subgroup N equals the kernel $\ker \pi_N$ of the quotient map $\pi_N : G \to G/N$. Thus the subsets of G that are kernels of morphisms into other groups are precisely the normal subgroups of G.

Lemma 5.3.5. *A group homomorphism $f : G \to H$ is injective if and only if $\ker f = \{e\}$.*

Proof. Assume that f is injective. Since $f(e) = e$, the kernel $\ker f$ contains $e \in G$. Since f is injective, $e \in G$ is the only element that is mapped to $e \in H$ and thus $\ker f = \{e\}$.

Conversely assume that $\ker f = \{e\}$ and consider $a,b \in G$ with $f(a) = f(b)$. Then $f(a^{-1}b) = f(a)^{-1}f(b) = e$ and thus $a^{-1}b \in \ker f$. By our assumption, $a^{-1}b = e$ and thus $a = b$, as desired. □

Lemma 5.3.6. *Let $f : G \to H$ be a group homomorphism. Then its image*

$$\operatorname{im} f \;=\; \{a \in H \mid a = f(b) \text{ for some } b \in G\}$$

is a subgroup of H.

Proof. Since $f(e) = e$, the image $\operatorname{im} f$ is not empty. Given $a,b \in \operatorname{im} f$, i.e. $a = f(c)$ and $b = f(d)$ for some $c,d \in G$, we have $ab^{-1} = f(c)f(d)^{-1} = f(cd^{-1})$, which shows that $ab^{-1} \in \operatorname{im} f$. This shows that $\operatorname{im} f$ is a subgroup of H. □

5.4 The isomorphism theorems

Theorem 5.4.1 (First isomorphism theorem)**.** *Let $f : G \to H$ be a group homomorphism. Then the association*

$$\begin{array}{rccc} \bar{f}: & G/\ker f & \longrightarrow & \operatorname{im} f \\ & [a] & \longmapsto & f(a) \end{array}$$

is an isomorphism of groups.

Proof. We begin to verify that $\bar{f}$ is well-defined as a map. Given $a, b \in G$ with $[a] = [b]$, i.e. $a = bh$ for some $h \in \ker f$, then

$$f(a) \;=\; f(bh) \;=\; f(b)f(h) \;=\; f(b)e \;=\; f(b),$$

which shows that the value $\bar{f}\big([a]\big) = f(a)$ does not depend on the choice of representative a for $[a]$.

We continue with showing that $\bar{f}$ is a group homomorphism. Given $a, b \in G$, we have

$$\bar{f}\big([a][b]\big) \;=\; \bar{f}\big([ab]\big) \;=\; f(ab) \;=\; f(a)f(b) \;=\; \bar{f}\big([a]\big)\bar{f}\big([b]\big),$$

as desired.

The group homomorphism $\bar{f}$ is surjective by the definition of its codomain as the image $\operatorname{im} f$ of f. Since $\bar{f}\big([a]\big) = e$ if and only if $a \in \ker f$, the kernel of $\bar{f}$ consists of a unique element, which is $[e] = \ker f$. Thus by Lemma 5.3.5, $\bar{f}$ is injective. This shows that $\bar{f}$ is an isomorphism of groups. □

Theorem 5.4.2 (Second isomorphism theorem)**.** *Let G be a group, H a subgroup and N a normal subgroup of G. Let $HN = \{hn \in G \mid h \in H, n \in N\}$. Then*

(1) HN is a subgroup of G;

(2) $H \cap N$ is a normal subgroup of H;

(3) the map

$$\begin{array}{ccc} H/(H\cap N) & \longrightarrow & (HN)/N \\ a(H\cap N) & \longmapsto & aN \end{array}$$

is an isomorphism of groups.

Proof. We begin with (1). Clearly, HN is not empty. Consider $hn, h'n' \in HN$ where $h, h' \in H$ and $n, n' \in N$. Then $n'' = (h^{-1}h')^{-1}n^{-1}(h^{-1}h')$ is in N since N is a normal subgroup and thus

$$(hn)^{-1}h'n' \;=\; n^{-1}h^{-1}h'n' \;=\; h^{-1}h'n''n'$$

is an element of HN. This shows that HN is a subgroup of G. Thus (1).

We continue with (2). By Lemma 5.3.4, the kernel $H \cap N = \ker f$ of the restriction $f : H \to G/N$ of the quotient map $G \to G/N$ to H is a normal subgroup of H. Thus (2).

We continue with (3). By the first isomorphism theorem (Theorem 5.4.1), f induces an isomorphism

$$\bar{f}: H/(H\cap N) \;=\; H/\ker f \longrightarrow \operatorname{im} f \;=\; HN/N$$

that maps $a(H \cap N)$ to aN, which verifies (3). □

Theorem 5.4.3 (Third isomorphism theorem)**.** *Let G be a group and N a normal subgroup. Let $\pi : G \to G/N$ be the quotient map. Then*

$$\begin{array}{rccc} \Phi: & \{\textit{subgroups of } G \textit{ containing } N\} & \longrightarrow & \{\textit{subgroups of } G/N\} \\ & H & \longmapsto & H/N = \pi(H) \end{array}$$

is an inclusion preserving bijection. A subgroup H of G that contains N is a normal subgroup of G if and only if H/N is a normal subgroup of G/N, and in this case, the map

$$\begin{array}{ccc} G/H & \longrightarrow & (G/N)/(H/N) \\ aH & \longmapsto & (aH)(H/N) \end{array}$$

is a group isomorphism for every subgroup H of G containing N.

Proof. We begin with the first claim of the theorem. By Lemma 5.3.6, the image $\pi(H)$ of H is a subgroup of G/N, which shows that Φ is well-defined as a map. Conversely, the inverse image $\pi^{-1}(H')$ of a subgroup H' of G/N contains N and is a subgroup of G since for $a,b \in G$ with $\pi(a),\pi(b) \in H'$, we have $\pi(a^{-1}b) = \pi(a)^{-1}\pi(b)$. Since $H' = \pi\big(\pi^{-1}(H')\big)$, we conclude that Φ is surjective. If H is a subgroup of G containing N, then $H = HN = H/N$ as subsets of G and thus $H = \pi^{-1}(H/N)$, which shows that Φ is injective. It is clear that Φ is inclusion preserving. This verifies the first claim of the theorem.

We continue with the second claim. Let H be a subgroup of G containing N. Then $NH = H$ and thus $[a]H = aH$ for the class $[a] = aN$ of an element $a \in G$, and similarly $H[a] = Ha$. Thus we have $aH = Ha$ if and only if $[a]H = H[a]$, which shows that H is normal in G if and only if H/N is normal in G/N. Thus the second claim.

We continue with the last claim. Let H be a normal subgroup of G containing N. Then the group homomorphism $f : G \to (G/N)/(H/N)$ sending a to $aH(H/N)$ is surjective with kernel $HN = H$. Thus by the first isomorphism theorem (Theorem 5.4.1), this yields a group isomorphism

$$\bar{f}: G/H = G/\ker f \longrightarrow \operatorname{im} f = (G/N)/(H/N)$$

that sends aH to $f(a) = aH(H/N)$, which establishes the last claim and concludes the proof. □

5.5 Group actions

Definition 5.5.1. Let G be a group and X a set. A **(left) action of G on X** is a map

$$\begin{array}{ccc} G \times X & \longrightarrow & X \\ (a,x) & \longmapsto & a.x \end{array}$$

such that

$$e.x = x \qquad \text{and} \qquad (ab).x = a.(b.x)$$

for all $a,b \in G$ and $x \in X$.

One often refers to a left action of G on X by the notation $G \circlearrowright X$. Since we will not consider right actions in this text, we will simply refer to a left action by an action of G on X. Often, we suppress the map $G \times X \to X$ from the notation, and simply say that G acts on X, and write $a.x$ for the image of (a,x) under the action.

Example 5.5.2. Let G be a group and H a subgroup of G. Then the action

$$\begin{array}{ccc} H \times G & \longrightarrow & G \\ (a,x) & \longmapsto & ax \end{array}$$

of H on G is called the *left translation by H*, and the action

$$\begin{array}{ccc} H \times G & \longrightarrow & G \\ (a,x) & \longmapsto & axa^{-1} \end{array}$$

of H on G is called the *conjugation by H*. We leave it as an exercise to verify that both maps define indeed actions of H on G.

Definition 5.5.3. Let G be a group that acts on a set X and $x \in X$. The **orbit of** x is the subset $\mathcal{O}(x) = \{a.x \mid a \in G\}$ of X, and x is a **fixed point** if $\mathcal{O}(x) = \{x\}$. We write $G\backslash X = \{\mathcal{O}(x) \mid x \in X\}$ for the collection of all subset of X that are orbits of X. The **stabilizer of** x is the subset $\mathsf{Stab}_G(x) = \{a \in G \mid a.x = x\}$ of G.

Lemma 5.5.4. *Let G be a group that acts on a set X. Then $y \in \mathcal{O}(x)$ if and only if $\mathcal{O}(x) = \mathcal{O}(y)$ for all $x, y \in X$, and $\mathcal{O}(x) \cap \mathcal{O}(y) = \varnothing$ if not. Consequently,*

$$X \;=\; \coprod_{\mathcal{O} \in G\backslash X} \mathcal{O}.$$

Proof. We have $y \in \mathcal{O}(x)$, i.e. $y = a.x$ for some $a \in G$, if and only if

$$\mathcal{O}(x) \;=\; \big\{b.x \,\big|\, b \in G\big\} \;=\; \big\{(ca).x \,\big|\, c \in G\big\} \;=\; \big\{c.y \,\big|\, c \in G\big\} \;=\; \mathcal{O}(y),$$

which establishes the first claim. If $\mathcal{O}(x) \cap \mathcal{O}(y)$ contains an element z, then this implies that $\mathcal{O}(x) = \mathcal{O}(z) = \mathcal{O}(y)$, which verifies the second claim. Thus X decomposes into a disjoint union of orbits of the action of G on X. This concludes the proof of the lemma. □

Lemma 5.5.5. *Let G be a group acting on X, $a \in G$ and $x \in X$. Then $\mathsf{Stab}_G(x)$ is a subgroup of G and $\mathsf{Stab}_G(a.x) = a\big(\mathsf{Stab}_G(x)\big)a^{-1}$.*

Proof. Since $e.x = x$, the subset $\mathsf{Stab}_G(x)$ is not empty. For $b, c \in \mathsf{Stab}_G(x)$, we have

$$(b^{-1}c).x \;=\; b^{-1}.(c.x) \;=\; b^{-1}.x \;=\; b^{-1}.(b.x) \;=\; (b^{-1}b).x \;=\; e.x \;=\; x,$$

which shows that $\mathsf{Stab}_G(x)$ is a subgroup of G.

We continue to verify that $\mathsf{Stab}_G(a.x) = a\big(\mathsf{Stab}_G(x)\big)a^{-1}$. For $b \in \mathsf{Stab}_G(x)$, we have

$$(aba^{-1}).(a.x) \;=\; (aba^{-1}a).x \;=\; (ab).x \;=\; a.(b.x) \;=\; a.x,$$

which shows that $a^{-1}ba \in \mathrm{Stab}_G(a.x)$. For $c \in \mathrm{Stab}_G(a.x)$, we have

$$(a^{-1}ca).x \;=\; a^{-1}.\big(c.(a.x)\big) \;=\; a^{-1}.(a.x) \;=\; (a^{-1}a).x \;=\; e.x \;=\; x,$$

which shows that $c \in a\,\mathrm{Stab}_G(x)a^{-1}$. Thus $\mathrm{Stab}_G(a.x) = a\big(\mathrm{Stab}_G(x)\big)a^{-1}$. □

Lemma 5.5.6. *Let G be a group acting on a set X. Then $\#\mathcal{O}(x) = \big(G : \mathrm{Stab}_G(x)\big)$ for every $x \in X$.*

Proof. Consider the association

$$\begin{array}{rccc} \Phi: & G/\mathrm{Stab}_G(x) & \longrightarrow & \mathcal{O}(x), \\ & a\,\mathrm{Stab}_G(x) & \longmapsto & a.x \end{array}$$

which is well-defined as a map since for $b \in \mathrm{Stab}_G(x)$, we have $(ab).x = a(b.x) = a.x$. By the definition of $\mathcal{O}(x)$, the map Φ is surjective. It is injective since $a.x = b.x$ implies that $x = (a^{-1}b).x$ and thus $a^{-1}b \in \mathrm{Stab}_G(x)$, which means that $a\,\mathrm{Stab}_G(x) = b\,\mathrm{Stab}_G(x)$. Thus Φ is a bijection and the cardinality of $\mathcal{O}(x)$ equals $\big(G : \mathrm{Stab}_G(x)\big)$. □

5.6 Centralizer and normalizer

Definition 5.6.1. Let G be a group, $a \in G$ and H a subgroup of G. The **center of** G is the subset $Z(G) = \{a \in G \mid ab = ba \text{ for all } b \in G\}$ of G. The **centralizer of** a **in** G is the subset $C_G(a) = \{b \in G \mid ab = ba\}$ of G. The **normalizer of** H **in** G is the subset $\mathrm{Norm}_G(H) = \{a \in G \mid aH = Ha\}$ of G.

Remark. Let $G \times G \to G$ be the action of G on itself by conjugation, i.e. $a.x = axa^{-1}$ for $a, x \in G$. Then the center $Z(G)$ is the set of fixed points in G and the centralizer of an element a equals $C_G(a) = \mathrm{Stab}_G(a)$. By Lemma 5.5.5, this implies that the centralizer of a is a subgroup of G. It is evident from the definition that the center of G is a normal and commutative subgroup of G.

Let X be the collection of all subgroups of G and $G \times X \to X$ the conjugation action, i.e. $a.H = aHa^{-1}$. Then the normalizer of a subgroup H equals the stabilizer $\mathrm{Norm}_G(H) = \mathrm{Stab}_G(H)$ of H with respect to this action. By Lemma 5.5.5, this implies that the normalizer of H is a subgroup of G. By its very definition, $\mathrm{Norm}_G(H)$ is the largest subgroup of G that contains H as a normal subgroup. By Lemma 5.5.6, we have

$$\#\{aHa^{-1} \mid a \in G\} \;=\; \#\mathcal{O}(H) \;=\; \big(G : \mathrm{Norm}_G(H)\big).$$

Proposition 5.6.2 (Class equation)**.** *Let G be a finite group that acts on itself by conjugation. Let $S \subset G$ a set of representatives of the* conjugation classes $\mathcal{O}(x) = \{axa^{-1} \mid a \in G\}$ *of elements $x \in G$. Let $S' \subset S$ be the subset of all $X \in S$ such that $\#\mathcal{O}(x) > 1$. Then*

$$\mathrm{ord}(G) \quad = \quad \sum_{x \in S} \big(G : C_G(x)\big) \quad = \quad \mathrm{ord}\,\big(Z(G)\big) + \sum_{x \in S'} \big(G : C_G(x)\big).$$

Proof. By Lemma 5.5.5, $\#\mathcal{O}(x) = \big(G : C_G(x)\big)$ for every $x \in S$. Thus Lemma 5.5.4 yields the first equality

$$G \;=\; \coprod_{x\in S} \mathcal{O}(x) \;=\; \sum_{x\in S} \big(G : C_G(x)\big).$$

The second equality follows from the observation that $\big(G : C_G(x)\big) = \#\mathcal{O}(x) = 1$ if and only if $x \in Z(G)$. □

Theorem 5.6.3 (Cauchy's theorem)**.** *Let G be a finite group and p a prime number that divides the order of G. Then there exists an element $a \in G$ of order p.*

Proof. We first prove the theorem for the case that G is commutative. We proceed by induction on $n = \operatorname{ord}(G)$. If $n = 1$, then there is nothing to prove.

Assume that $n > 1$. Choose an element $a \in G$. If $\operatorname{ord}(a) = kp$ for a positive integer k, then $\operatorname{ord}(a^k) = p$, and we are done. If p does not divide $\operatorname{ord}(a)$, then $\operatorname{ord}(G/\langle a\rangle)$ is divisible by p and contains an element $[b] = b\langle a\rangle$ of order p. Since G is abelian, $[b]^i = (b\langle a\rangle)^i = b^i\langle a\rangle = [b^i]$. Thus $b^i \in \langle a\rangle$ if and only if i is divisible by p. Since $e \in \langle a\rangle$, this implies that $\operatorname{ord}(b) = kp$ for some positive integer k. Thus b^k has order p as desired. This proves that every abelian group contains an element of order p.

We turn to the case of an arbitrary group G, which we also prove by induction on $n = \operatorname{ord}(G)$. If $n = 1$, then there is nothing to do.

Assume that $n > 1$. If there is an element $a \in G$ that is not in $Z(G)$ and such that p does not divide $\big(G : C_G(a)\big)$, then p divides $\operatorname{ord}\big(C_G(a)\big) = \operatorname{ord}(G)\big/\big(G : C_G(a)\big)$. Since $a \notin Z(G)$, the subgroup $C_G(a)$ has less elements than G, and thus we find an element of order p in $C_G(a)$ by the inductive hypothesis.

We are left with the case that p divides $\big(G : C_G(a)\big)$ for all $a \in G - Z(G)$. Then the class equation (Proposition 5.6.2)

$$\underbrace{\operatorname{ord}(G)}_{\text{divisible by } p} \;=\; \operatorname{ord}\big(Z(G)\big) + \sum_{x\in S'} \underbrace{\big(G : C_G(x)\big)}_{\text{divisible by } p}$$

shows that $\operatorname{ord}\big(Z(G)\big)$ is divisible by p. Since $Z(G)$ is commutative, it contains an element of order p, as we have proven before. □

5.7 Sylow subgroups

Definition 5.7.1. Let G be a finite group and p a prime number. A ***p*-group** is a finite group H whose order is a power of p. A ***p*-subgroup of** G is a subgroup that is a p-group. A ***p*-Sylow subgroup of** G is a p-subgroup P of G such that p does not divide $(G : P)$.

Lemma 5.7.2. *Let G be a p-group that acts on a finite set X. Then*

$$\#\big\{x \in X \,\big|\, \mathcal{O}(x) = \{x\}\big\} \;\equiv\; \#X \pmod{p}.$$

Proof. By Lemma 5.5.4, we have

$$\#X \quad = \quad \#\big\{x \in X \,\big|\, \mathcal{O}(x) = \{x\}\big\} \; + \sum_{\substack{\mathcal{O}\in G\backslash X \\ \#\mathcal{O}>1}} \#\mathcal{O}.$$

By Lemma 5.5.6, the cardinality $\#\mathcal{O}(x) = \big(G : \mathrm{Stab}_G(x)\big)$ of an orbit $\mathcal{O}(x)$ is a divisor of $\mathrm{ord}(G)$ and therefore a power of p since G is a p-group. Thus if $\#\mathcal{O}(x) > 1$, then $\#\mathcal{O}(x)$ is divisible by p, which proves the claim of the lemma. □

Theorem 5.7.3 (Sylow theorems)**.** *Let G be a finite group and p a prime number. Then the following holds.*

(1) Every p-subgroup H of G is contained in a p-Sylow subgroup. In particular, G contains a p-Sylow subgroup.

(2) All p-Sylow subgroups are conjugate to each other.

(3) Let n_p be the number of p-Sylow subgroups of G. Then $n_p \equiv 1 \pmod{p}$ and $n_p = \big(G : \mathrm{Norm}_G(P)\big)$ for any p-Sylow subgroup P of G. In particular, n_p divides $(G : P)$.

Proof. To begin with, we show by induction on $n = \mathrm{ord}(G)$ that G has a p-Sylow subgroup. If $n = 1$, the result is trivial.

Assume $n > 1$. If G has a proper subgroup such that p does not divide $(G : H)$, then H contains a p-Sylow subgroup P by the inductive hypothesis. Since p does neither divide $(G : H)$ nor $(H : P)$, it does not divide $(G : P) = (G : H)(H : P)$, which shows that P is a p-Sylow subgroup of G.

If p divides $(G : H)$ for all proper subgroups H of G, then the class equation (Proposition 5.6.2)

$$\underbrace{\mathrm{ord}(G)}_{\text{divisible by } p} \quad = \quad \mathrm{ord}\big(Z(G)\big) + \sum_{x\in S'} \underbrace{\big(G : C_G(x)\big)}_{\text{divisible by } p}$$

shows that $\mathrm{ord}\big(Z(G)\big)$ is divisible by p. By Cauchy's theorem (Theorem 5.6.3), $Z(G)$ contains an element a of order p. Thus the subgroup $N = \langle a \rangle$ of G has order p and is normal in G as a subgroup of $Z(G)$. Let $\pi : G \to G/N$ be the quotient map.

By the inductive hypothesis, G/N contains a p-Sylow subgroup P', i.e. $\mathrm{ord}(P') = p^k$ for some $k \geqslant 0$ and $m = \big((G/N) : P'\big)$ is not divisible by p. Thus $P = \pi^{-1}(P')$ is a p-group with $\mathrm{ord}(P') \cdot \mathrm{ord}(N) = p^{k+1}$ elements. Since $\mathrm{ord}(G) = \mathrm{ord}(N) \cdot \mathrm{ord}(G/N) = mp^{k+1}$, the index $(G : P) = m$ is not divisible by p. Thus P is a p-Sylow subgroup of G, which concludes the proof that G has a p-Sylow subgroup.

Let P be a p-Sylow subgroup and $X = \{aPa^{-1} \mid a \in G\}$. Consider the action of G on X by conjugation. Since $X = \mathcal{O}(P)$ and since $\mathrm{Norm}_G(P)$ is the stabilizer $\mathrm{Stab}_G(P)$ of P with respect to this action, we have $\#X = \big(G : \mathrm{Norm}_G(P)\big)$ by Lemma 5.5.4. Since $\mathrm{Norm}_G(P)$ contains the p-Sylow subgroup P, this implies that p does not divide $\#X$.

Let H be a p-subgroup of G and let us consider the action of H on X by conjugation. By Lemma 5.7.2, the number of fixed points is congruent to $\#X$ modulo p. Since p does

not divide $\#X$, there is a fixed point, i.e. a conjugate $P' = aPa^{-1}$ of P such that H is contained in $\mathrm{Norm}_G(P')$. Since P' has the same cardinality as P, it is also a p-Sylow subgroup of G.

Since $\mathrm{Norm}_G(P')$ is the largest subgroup of G that contains P' as a normal subgroup, P' is normal in HP'. By the second isomorphism theorem (Theorem 5.4.2), we have

$$(HP')/P' \;\simeq\; H/(H \cap P'),$$

which shows that $(HP' : P')$ is a divisor of $\mathrm{ord}(H)$ and therefore a power of p. We conclude that $\mathrm{ord}(HP') = (HP' : P') \cdot \mathrm{ord}(P')$ is a power of p, and thus HP' is a p-subgroup G that contains P'. Since P' is a p-Sylow subgroup, it is a maximal p-subgroup and thus $HP' = P'$. This shows that H is contained in the p-Sylow subgroup P', which establishes (1).

If H is a p-Sylow subgroup itself, then it has the same cardinality as P' and thus $H = P' = aPa^{-1}$ is a conjugate of P. This establishes (2).

Moreover, this shows that the action of the p-Sylow subgroup H on X by conjugation has only one fixed point, namely $H = aPa^{-1}$ itself. Therefore Lemma 5.7.2 implies that $n_p \equiv 1 \pmod{p}$. Since $\mathrm{Norm}_G(P)$ is the stabilizer of P under the action of G on X by conjugation, Lemma 5.5.6 implies that $n_p = \#X = \big(G : \mathrm{Norm}_G(P)\big)$. In particular, n_P divides $(G : P) = \big(G : \mathrm{Norm}_G(P)\big) \cdot \big(\mathrm{Norm}_G(P) : P\big)$. This completes the proof of (3) and the theorem. □

5.8 Exercises

In the following exercises, let G be a group with multiplication $m : G \times G \to G$, inversion $i : G \to G$ and neutral element e.

Exercise 5.1 (Isomorphisms, monomorphisms and epimorphisms)**.** Let $f : G \to H$ be a group homomorphism. Show that f is an isomorphism in Groups (in the sense of Definition 2.3.1) if and only if f is bijective. Show that f is a monomorphism if and only if f is injective. Show that f is an epimorphism if and only if f is surjective.

Exercise 5.2 (Subgroups)**.** Let H be a subset of G. Show that H is a subgroup of G if and only if $e \in H$, $m(H \times H) \subset H$ and $i(H) \subset H$. In other words, H is a subgroup if and only if it is a group with respect to the restrictions of m and i to H.

Exercise 5.3 (The center)**.** Show that the *center of* G

$$Z(G) \;=\; \{\, a \in G \,|\, ab = ba \text{ for all } b \in G \,\}$$

is a subgroup of G. Show that $Z(G)$ is commutative. Show that every subgroup of $Z(G)$ is normal in G. Is every commutative subgroup of G normal?

Exercise 5.4 (The subgroup generated by a subset)**.** (1) Let $\{H_i\}_{i \in I}$ be a family of subgroups of G. Show that the intersection $\bigcap_{i \in I} H_i$ is a subgroup of G.

(2) Let $S \subset G$ be a subset. Show that

$$\bigcap_{H<G \text{ with } S\subset H} H \quad = \quad \{a_1 a_2^{-1} \cdots a_{2n-1} a_{2n}^{-1} \,|\, n \geqslant 1 \text{ and } a_1, \ldots, a_n \in S \cup \{e\}\}$$

and conclude that there is a unique smallest subgroup $\langle S \rangle$ of G that contains S.

Exercise 5.5 (Orders of elements in commutative groups)**.** Let G be a commutative group and $a, b \in G$. Show that $\mathrm{ord}(ab)$ divides $\mathrm{ord}(a) \cdot \mathrm{ord}(b)$. Is this also true if G is not commutative?

Exercise 5.6 (Cyclic groups and the Klein four-group)**.** (1) Classify all cyclic groups up to isomorphism. Which of them are commutative?

(2) Show that a cyclic group of order n has a unique subgroup of order d for each divisor d of n.

(3) Is the *Klein four-group* $V = (\mathbb{Z}/2\mathbb{Z}) \times (\mathbb{Z}/2\mathbb{Z})$ cyclic? Is it commutative?

Exercise 5.7 (Dihedral groups)**.** Let D_n be the group of symmetries of a regular polygon with n sides. Show that $D_n = \langle r, s \rangle$ where r is a rotation around the center of the polygon by an angle of $2\pi/n$ and s is the reflection at a line passing through the center of the polygon and one of its vertices. What is the number of elements of D_n? Show that $D_3 \simeq S_3$, and that for $n \geqslant 4$, the dihedral group D_n is not isomorphic to a symmetric group.

Exercise 5.8 (Symmetric groups)**.** The *symmetric group* S_n is the group of permutations of the numbers $1, \ldots, n$, together with composition as multiplication, i.e. $\sigma \cdot \tau = \sigma \circ \tau$. An element σ of S_n is called a *cycle (of length l)* if $\mathrm{ord}(\sigma) = l$ and if there is an $i \in \{1, \ldots, n\}$ such that $\sigma(j) = j$ if $j \notin \{i, \sigma(i), \ldots, \sigma^{l-1}(i)\}$; we write $\sigma = (i, \sigma(i), \ldots, \sigma^{l-1}(i))$ in this case.

(1) Show that $(i, \ldots, \sigma^{l-1}(i)) = (j, \ldots, \sigma^{l-1}(j))$ if $j = \sigma^n(i)$ for some $n \geqslant 0$.

(2) Two cycles $\sigma = (i, \ldots, \sigma^{l-1}(i))$ and $\tau = (j, \ldots, \tau^{k-1}(j))$ are called *disjoint* if the sets $\{i, \ldots, \sigma^{l-1}(i)\}$ and $\{j, \ldots, \tau^{k-1}(j)\}$ are disjoint. Show that σ and τ are disjoint if and only if $\sigma\tau = \tau\sigma$.

(3) Show that every element of S_n can be written as a product of disjoint cycles.

(4) A *transposition* is a cycle (i, j) of length 2. Show that every element of S_n can be written as a product of transpositions.

Exercise 5.9 (The sign)**.** Let σ be an element of S_n and $\sigma = \tau_n \circ \cdots \circ \tau_1$ and $\sigma = \tau'_m \circ \cdots \circ \tau'_1$ two representations of σ as a product of transpositions $\tau_1, \ldots, \tau_n$ and $\tau'_1, \ldots, \tau'_m$.

(1) Show that $n - m$ is even. Conclude that the map $\mathrm{sign} : S_n \to \{\pm 1\}$ that sends σ to $(-1)^n$ is well-defined.

(2) Show that sign is a group homomorphism.

Exercise 5.10 (Theorem of Cayley)**.** Let $G = \{a_1, \ldots, a_n\}$ be of finite order n. Define the map $f : G \to S_n$ that sends a_l to the permutation σ_l with $\sigma_l(i) = j$ such that $a_l a_i = a_j$. Show that f is an injective group homomorphism. Conclude that every finite group is isomorphic to a subgroup of a symmetric group.

Exercise 5.11 (The alternating group)**.** The *alternating group* A_n is defined as the kernel of sign : $S_n \to \{\pm 1\}$. A group G is called *simple* if $G \neq \{e\}$ and if the only normal subgroups of G are $\{e\}$ and G.

(1) Show that a cyclic group G of order n is simple if and only if n is a prime number.

(2) Show that A_3 is simple. Show that A_4 is not simple. What about A_1 and A_2?

(3) Show that A_n is simple for $n \geqslant 5$.[1]

Exercise 5.12 (Quaternion group)**.** The quaternion group Q consists of the elements $\{\pm 1, \pm i, \pm j, \pm k\}$, and the multiplication is determined by the following rules: 1 is the neutral element, $(-1)^2 = 1$ and

$$i^2 = j^2 = k^2 = -1, \quad (-1)i = -i, \quad (-1)j = -j, \quad (-1)k = -k, \quad ij = k = -ji.$$

(1) Is Q commutative?

(2) Describe all subgroups of Q.

(3) Which subgroups are normal? What are the respective quotient groups?

Exercise 5.13. Classify all groups with 6 elements and all groups with 8 elements up to isomorphism.

Exercise 5.14 (Transitivity of index)**.** Let H be a subgroup of G and K a subgroup of H. Show that $(G : K) = (G : H)(H : K)$.

Exercise 5.15 (Quotients by non-normal subgroups)**.** Let H be subgroup of G. Show that the association $([a], [b]) \mapsto [ab]$ is not well-defined on cosets $[a], [b] \in G/H$ if H is not normal in G.

Exercise 5.16 (Alternative characterization of normal subgroups)**.** A subgroup H of G is normal if and only if $gHg^{-1} \subset H$ for every $g \in G$.

Exercise 5.17 (Exercises on normal subgroups)**.** Show the following statements.

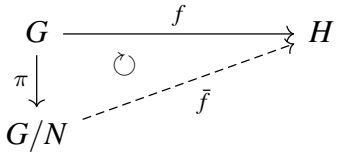
(1) Every subgroup of index 2 is normal.

(2) Every subgroup of a commutative group is normal. Is there a non-commutative group G such that every subgroup H of G is normal?

(3) The intersection of two normal subgroups is a normal subgroup. If both normal subgroups have finite index, then their intersection has also finite index.

[1] This exercise is more difficult than others, but solutions can be found in the literature.

Exercise 5.18 (Universal property of the quotient)**.** Let N be a normal subgroup of G. Show that the quotient map $\pi : G \to G/N$ satisfies the following universal property: for every group homomorphism $f : G \to H$ with $f(a) = e$ for $a \in N$ there exists a unique group homomorphism $\bar{f} : G/N \to H$ such that $f = \bar{f} \circ \pi$, i.e. the diagram

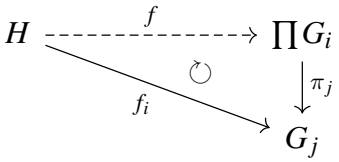


commutes.

Exercise 5.19 (Universal property of the product)**.** Let $\{G_i\}_{i\in I}$ be a family of groups and $G = \prod G_i$ their product.

(1) Show that the map $\pi_i : G \to G_i$ that sends $(g_i)_{i\in I}$ to g_i is a surjective group homomorphism for every $i \in I$. These maps are called the *canonical projections*.

(2) Show that the product together with the canonical projections satisfies the following universal property: for every family of group homomorphisms $\{f_i : H \to G_i\}_{i\in I}$, there is a unique group homomorphism $f : H \to \prod G_i$ such that $f_j = \pi_j \circ f$ for every $j \in I$, i.e. the diagram

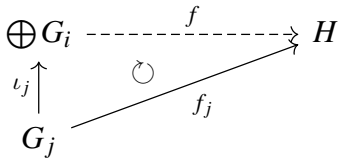


commutes for every $j \in I$.

Exercise 5.20 (Universal property of the direct sum)**.** Let $\{G_i\}_{i\in I}$ be a family of commutative groups and $G = \bigoplus G_i$ their direct sum.

(1) Show that the map $\iota_i : G_i \to G$ that sends g to $(g_j)_{j\in I}$ with $g_i = g$ and $g_j = e_j$ for $j \neq i$ is an injective group homomorphism for every $i \in I$. These maps are called the *canonical injections*.

(2) Show that the direct sum together with the canonical injections satisfies the following universal property: for every family of group homomorphisms $\{f_i : G_i \to H\}_{i\in I}$ of commutative groups, there is a unique group homomorphism $f : \bigoplus G_i \to H$ such that $f_j = f \circ \iota_j$ for every $j \in I$, i.e. the diagram



commutes for every $j \in I$.

(3) Is the same true if H is a non-commutative group?

Exercise 5.21 (Some group actions)**.** Show that the following maps are group actions:

(1) $S_n \times \{1,\dots,n\} \to \{1,\dots,n\}$, with $\sigma.i = \sigma(i)$;

(2) $\mathrm{GL}_n(\mathbb{R}) \times \mathbb{R}^n \to \mathbb{R}^n$, with $g.v = g \cdot v$ (usual matrix multiplication);

(3) $\mathbb{R}^\times \times \mathbb{R}^n \to \mathbb{R}^n$, with $a.v = a \cdot v$ (scalar multiplication);

(4) the permutation of the vertices of a regular n-gon by elements of the dihedral group D_n.

Exercise 5.22 (Center and centralizer)**.** Consider the action of G on itself by conjugation.

(1) Show that

$$\big\{x \in G \,\big|\, \mathcal{O}(x) = \{x\}\big\} \;=\; \big\{a \in G \,\big|\, ab = ba \text{ for all } b \in G\big\}.$$

(2) Show that $C_G(x) = \{a \in G \mid ax = xa\}$.

(3) Show that

$$Z(G) \;=\; \bigcap_{x \in G} C_G(x).$$

Exercise 5.23 (Normalizer)**.** Let H be a subgroup of G. Show that its normalizer $\mathrm{Norm}_G(H)$ is the largest subgroup of G containing H such that H is a normal subgroup of $\mathrm{Norm}_G(H)$. Show further that the following properties are equivalent:

(1) H is normal in G;

(2) $\mathrm{Norm}_G(H) = G$;

(3) H is a fixed point for the action of G on the set of all subgroups of G by conjugation.

Exercise 5.24 (Short exact sequences)**.** A short exact sequence of groups is a sequence

$$\{e\} \xrightarrow{f_1} N \xrightarrow{f_2} G \xrightarrow{f_3} Q \xrightarrow{f_4} \{e\}$$

of groups and group homomorphism such that $\operatorname{im} f_i = \ker f_{i+1}$ for $i = 1,2,3$.

(1) Show that $\operatorname{im} f_i = \ker f_{i+1}$ for $i = 1,2,3$ holds if and only if f_2 is injective, if $\operatorname{im} f_2 = \ker f_3$ and if f_3 is surjective.

(2) Show that N is isomorphic to $N' = \operatorname{im} f_1$, that N' is a normal subgroup of G and that $G/N' \simeq Q$ in case of a short exact sequence.

Exercise 5.25. Calculate all orbits and stabilizers for the action of D_4 on itself by conjugation.

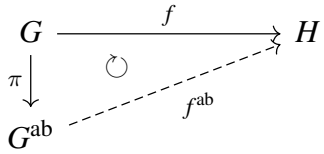
Exercise 5.26 (Commutator subgroup)**.** The commutator of two elements $a,b \in G$ is $[a,b] = aba^{-1}b^{-1}$. The commutator subgroup of G is the subgroup $[G,G]$ generated by the commutators $[a,b]$ of all pairs of elements a and b of G.

(1) Show that $[a,b] = e$ if and only if $ab = ba$. Conclude that $[G,G] = \{e\}$ if and only if G is commutative.

(2) Show that $c[a,b]c^{-1} = [cac^{-1}, cbc^{-1}]$ and conclude that $[G,G]$ is a normal subgroup of G.

(3) Show that the quotient group $G^{\mathrm{ab}} = G/[G,G]$ is commutative.

(4) Show that G^{ab} together with the projection $\pi : G \to G^{\mathrm{ab}}$ satisfies the following universal property: for every group homomorphism $f : G \to H$ into a commutative group H, there exists a unique group homomorphism $f^{\mathrm{ab}} : G^{\mathrm{ab}} \to H$ such that $f = f^{\mathrm{ab}} \circ \pi$:



Exercise 5.27. Determine all p-Sylow subgroups of S_4 for $p \in \{2,3\}$.

Exercise 5.28. Let $\mathrm{ord}(G) = 6$ and n_p the number of p-Sylow subgroups of G. Find all possibilities for n_2 and n_3, using the Sylow theorems. Find examples of groups with 6 elements that realize these possibilities.

Exercise 5.29. Let $\mathrm{ord}(G) = pq$ for prime numbers p and q. Show that G is not simple.

Hint: If $p = q$, then use the class equation. If $p \neq q$, then use the Sylow theorems.

Chapter 6

Outlook to algebraic geometry

In this chapter, we introduce some concepts from algebraic geometry and use them to study curves in the affine plane. We will discuss some central theorems like Hilbert's Basissatz, Hilbert's Nullstellensatz, Nakayama's lemma and Krull's principal ideal theorem. Even though the proofs of these theorems are fairly elementary and accessible within the framework of this course, we omit them for the purpose of a compact presentation of this chapter.

6.1 Hilbert's Basissatz

Definition 6.1.1. A ring A is **Noetherian** if every ideal of A is finitely generated.

Theorem 6.1.2 (Hilbert's Basissatz)**.** *Let A be a Noetherian ring. Then $A[T]$ is Noetherian.*

We do not prove this theorem in these notes.

Corollary 6.1.3. *Let K be a field, n a positive integer and I an ideal in $K[T_1,\dots,T_n]$. Then $K[T_1,\dots,T_n]/I$ is Noetherian.*

Proof. By Hilbert's Basissatz (Theorem 6.1.2), $K[T_1,\dots,T_n] = (\cdots(K[T_1])[T_2]\cdots)[T_n]$ is Noetherian. Let J be an ideal of $K[T_1,\dots,T_n]/I$ and $\pi : K[T_1,\dots,T_n] \to K[T_1,\dots,T_n]/I$ the quotient map. Since $K[T_1,\dots,T_n]$ is Noetherian, $J' = \pi^{-1}(J)$ is finitely generated, i.e. $J' = \langle f_1,\dots,f_r\rangle$ for some $f_1,\dots,f_r \in K[T_1,\dots,T_n]$. Then $J = \langle \pi(f_1),\dots,\pi(f_r)\rangle$ is also finitely generated, which concludes the proof. □

Another useful fact is the following.

Lemma 6.1.4. *Let A be a Noetherian ring and S a multiplicative set in A. Then $S^{-1}A$ is Noetherian.*

Proof. Let $\iota : A \to S^{-1}A$ be the canonical map and consider an ideal I of $S^{-1}A$. Since A is Noetherian, $\iota^{-1}(I)$ is generated by finitely many elements $a_1,\dots,a_r \in A$. Then

$b_i = \iota(a_i) \in I$ fir $i = 1, \dots, r$. Consider an element $\frac{b}{s} \in I$. Then $\frac{b}{1} = s\frac{b}{s} \in I$ and $\frac{b}{1} = \iota(a)$ for some $a \in \iota^{-1}(I)$. Thus $a = \sum c_i a_i$ for some $c_1, \dots, c_r \in A$ and

$$\frac{b}{s} \;=\; \frac{1}{s}\iota(a) \;=\; \frac{1}{s}\sum_{i=1}^{r} c_i \iota(a_i) \;=\; \sum_{i=1}^{r} \frac{c_i}{s} b_i,$$

which shows that I is generated by $b_1, \dots, b_r$. This concludes our proof that every ideal of $S^{-1}A$ is finitely generated. □

6.2 Affine varieties

For the rest of this chapter, we fix an algebraically closed field K and $n \in \mathbb{N}$. The reader might assume safely that $K = \mathbb{C}$, but, in fact, everything is valid for an arbitrary algebraically closed field.

Definition 6.2.1. The **affine n-space over** K is the set K^n.

We recall the multi-index notation for polynomials in several variables from section 1.9. For $\underline{e} = (e_1, \dots, e_n) \in \mathbb{N}^n$ and $a = (a_1, \dots, a_n) \in K^n$, we write $T^{\underline{e}}$ for the monomial $T_1^{e_1} \cdots T_n^{e_n}$ in $K[T_1, \dots, T_n]$ and $a^{\underline{e}}$ for the element $a_1^{e_1} \cdots a_n^{e_n}$ of K. A polynomial $f = \sum c_{\underline{e}} T^{\underline{e}}$ in $K[T_1, \dots, T_n]$ defines a function

$$\begin{array}{rccc} f: & K^n & \longrightarrow & K, \\ & a & \longmapsto & f(a) = \sum c_{\underline{e}} a^{\underline{e}} \end{array}$$

which we denote by the same symbol f.

Definition 6.2.2. Let S be a subset of $K[T_1, \dots, T_n]$. The **vanishing set of** S is the subset

$$\mathcal{V}(S) \;=\; \big\{ a \in K^n \,\big|\, f(a) = 0 \text{ for all } f \in S \big\}$$

of K^n. We write $\mathcal{V}(f_1, \dots, f_r)$ for $\mathcal{V}(\{f_1, \dots, f_r\})$.

Lemma 6.2.3. *The vanishing sets in K^n satisfy the following properties:*

(1) $\mathcal{V}(S) = \mathcal{V}(I)$ for every subset S of $K[T_1, \dots, T_n]$ and the ideal I generated by S;

(2) $\mathcal{V}(0) = K^n$ and $\mathcal{V}(1) = \varnothing$;

(3) $\mathcal{V}(T_1 - a_1, \dots, T_n - a_n) = \{(a_1, \dots, a_n)\}$ for all $a \in K^n$;

(4) $\mathcal{V}(S) = \bigcap_{f \in S} \mathcal{V}(f)$ for all subsets S of $K[T_1, \dots, T_n]$;

(5) $\mathcal{V}(f \cdot g) = \mathcal{V}(f) \cup \mathcal{V}(g)$ for all $f, g \in K[T_1, \dots, T_n]$.

Proof. We begin with (1). Since $S \subset I$, we have $\mathcal{V}(I) \subset \mathcal{V}(S)$. Conversely, I consists of elements of the form $\sum c_i f_i$ with $c_i \in A$ and $f_i \in S$. For such an element $\sum c_i f_i$ and $a \in \mathcal{V}(S)$, we have

$$\big(\sum c_i f_i\big)(a) \;=\; \sum c_i \underbrace{f_i(a)}_{=0} \;=\; 0,$$

which shows that $a \in \mathcal{V}(I)$. Thus (1).

Part (2) follows since the zero polynomial 0 maps every $a \in K^n$ to 0 and the constant polynomial 1 maps a to 1. Part (3) follows since

$$\mathcal{V}(T_1 - a_1, \ldots, T_n - a_n) \;=\; \big\{b \in K^n \,\big|\, b_1 - a_1 = 0, \ldots, b_n - a_n = 0\big\} \;=\; \big\{(a_1, \ldots, a_n)\big\}.$$

Part (4) follows since

$$\mathcal{V}(S) \;=\; \big\{a \in K^n \,\big|\, f(a) = 0 \text{ for all } f \in S\big\} \;=\; \bigcap_{f \in S} \big\{a \in K^n \,\big|\, f(a) = 0\big\} \;=\; \bigcap_{f \in S} \mathcal{V}(f).$$

Part (5) follows since

$$\begin{aligned} \mathcal{V}(f \cdot g) \;&=\; \big\{a \in K^n \,\big|\, (f \cdot g)(a) = 0\big\} \\ &=\; \big\{a \in K^n \,\big|\, f(a) = 0\big\} \cup \big\{a \in K^n \,\big|\, g(a) = 0\big\} \;=\; \mathcal{V}(f) \cup \mathcal{V}(g). \quad \square \end{aligned}$$

Remark. Note that by Hilbert's Basissatz (Theorem 6.1.2), every ideal I of $K[T_1, \ldots, T_n]$ is finitely generated. Thus for every subset S of $K[T_1, \ldots, T_n]$, the ideal $I = \langle S \rangle$ is finitely generated, i.e. $I = \langle f_1, \ldots, f_r \rangle$ for some $f_1, \ldots, f_r \in K[T_1, \ldots, T_n]$, and thus $\mathcal{V}(S) = \mathcal{V}(I) = \mathcal{V}(f_1, \ldots, f_r)$.

As a consequence, properties (2), (4) and (5) show that the vanishing sets form the closed subsets of a topology for K^n. This topology is called the *Zariski topology*.

Definition 6.2.4. An **affine K-variety** is a subset V of K^n of the form $V = \mathcal{V}(I)$ for some ideal I of $K[T_1, \ldots, T_n]$. A **point in** K^n is an element $(a_1, \ldots, a_n)$ of K^n.

Theorem 6.2.5 (Hilbert's Nullstellensatz, weak form)**.** *Let I be a proper ideal of $K[T_1, \ldots, T_n]$. Then $\mathcal{V}(I)$ is not empty.*

We do not prove this theorem in these notes.

Corollary 6.2.6. *The map*

$$\begin{array}{rccc} \Phi: & K^n & \longrightarrow & \big\{\textit{maximal ideals of } K[T_1, \ldots, T_n]\big\} \\ & (a_1, \ldots, a_n) & \longmapsto & \langle T_1 - a_1, \ldots, T_n - a_n \rangle \end{array}$$

is a bijection.

Proof. Note that $\langle T_1 - a_1, \ldots, T_n - a_n \rangle$ is a maximal ideal since $K[T_1, \ldots, T_n]/\langle T_1 - a_1, \ldots, T_n - a_n \rangle$ is isomorphic to K, which is a field. Thus Φ is well-defined.

We establish the injectivity of Φ by contradiction. If Φ was not injective, then there was an $(a_1, \ldots, a_n) \in K^n$ and $b_i \neq a_i$ such that $T_i - b_i \in \langle T_1 - a_1, \ldots, T_n - a_n \rangle$. But then

$$1 \;=\; \underbrace{(a_i - b_i)}_{\neq 0}^{-1} \big((T_i - a_i) - (T_i - b_i)\big)$$

was an element of $\langle T_1 - a_1, \ldots, T_n - a_n \rangle$, which is not a case and thus a delivers the desired contradiction. This shows that Φ is injective.

To show the surjectivity of Φ, consider a maximal ideal $\mathfrak{m}$ of $K[T_1,\dots,T_n]$. By Hilbert's Nullstellensatz (Theorem 6.2.5), $\mathcal{V}(\mathfrak{m})$ is not empty and thus contains a point $a=(a_1,\dots,a_n)$ of K^n. Since the polynomials $f_i = T_i - a_i$ vanish in a for all $i=1,\dots,n$, the point a is contained in $\mathcal{V}(\mathfrak{m}\cup\{f_1,\dots,f_n\})$. This implies that $\mathfrak{m}' = \langle \mathfrak{m}\cup\{f_1,\dots,f_n\}\rangle$ is a proper ideal that contains $\mathfrak{m}$. Since $\mathfrak{m}$ is maximal, we conclude that $\mathfrak{m}=\mathfrak{m}'$ and thus $f_1,\dots,f_n\in\mathfrak{m}$. Since $\langle f_1,\dots,f_n\rangle$ is itself a maximal ideal, we see that $\mathfrak{m}=\langle f_1,\dots,f_n\rangle = \Phi(a_1,\dots,a_n)$ is in the image of Φ. Thus Φ is surjective, which concludes the proof. □

6.3 Regular functions

Definition 6.3.1. Let V be an affine variety in K^n. The **vanishing ideal of** V is the ideal

$$\mathcal{I}(V) \;=\; \big\{ f\in K[T_1,\dots,T_n] \,\big|\, f(a)=0 \text{ for all } a\in V \big\}$$

of $K[T_1,\dots,T_n]$. The **ring of regular functions on** V if $\mathcal{O}(V) = K[T_1,\dots,T_n]/\mathcal{I}(V)$.

Remark. The set $\mathcal{I}(V)$ is indeed an ideal: for $f,g\in\mathcal{I}(V)$ and $h\in K[T_1,\dots,T_n]$, we have

$$(f+g)(a) \;=\; f(a)+g(a) \;=\; 0 \qquad \text{and} \qquad (hf)(a) \;=\; h(a)f(a) \;=\; 0.$$

Lemma 6.3.2. *Let V be an affine variety in K^n and $f,g\in K[T_1,\dots,T_n]$. Then $f|_V = g|_V$ as functions $V\to K$ if and only if $[f]=[g]$ as elements of $\mathcal{O}(V)$.*

Proof. Define $h=f-g$. Then $f(a)=g(a)$ for all $a\in V$ if and only if $h(a)=0$ for all $a\in V$, i.e. $h\in\mathcal{I}(V)$. This means exactly that $[f]=[g+h]=[g]$ in $\mathcal{O}(V)$. □

Definition 6.3.3. Let A be a ring and I an ideal of A. The radical of I is the subset

$$\sqrt{I} \;=\; \{a\in A \mid a^i\in I \text{ for some } i>0\}$$

of A. An ideal I of A is a **radical ideal** if $\sqrt{I}=I$.

Lemma 6.3.4. *Let A be a ring and I an ideal of A. Then $\sqrt{I}$ is the intersection of all prime ideals of A containing I.*

Proof. Let $\mathfrak{p}$ be a prime ideal that contains I and consider $a\in\sqrt{I}$, i.e. $a^i\in I$ for some $i>0$. Then $a^i\in\mathfrak{p}$ and thus $a\in\mathfrak{p}$ since $\mathfrak{p}$ is prime. Thus $\sqrt{I}$ is contained in the intersection of all prime ideals $\mathfrak{p}$ of A with $I\subset\mathfrak{p}$.

Conversely consider an element $a\in A$ that is not contained in $\sqrt{I}$ and define $S=\{a^i \mid i\in\mathbb{N}\}$. Since $S\cap\sqrt{I}=\varnothing$, the ideal $S^{-1}\sqrt{I}$ is a proper ideal of $S^{-1}A$ and thus contained in a maximal ideal $\mathfrak{m}$ of $S^{-1}A$. Thus the prime ideal $\mathfrak{p}=\iota_S^{-1}(\mathfrak{m})$ contains $\sqrt{I}$ where $\iota_S: A\to S^{-1}A$ is the canonical homomorphism. Since $\mathfrak{p}\cap S=\varnothing$, the prime ideal $\mathfrak{p}$ does not contain a. This shows that the intersection of all prime ideals $\mathfrak{p}$ of A with $I\subset\mathfrak{p}$ is contained in $\sqrt{I}$, which completes the proof. □

Remark. As an immediate consequence of Lemma 6.3.4, we see that the radical ideal is indeed an ideal as the intersection of ideals. Moreover, we conclude that the radical of an ideal is a radical ideal, and that all prime ideals are radical ideals.

Theorem 6.3.5 (Hilbert's Nullstellensatz, strong form)**.** *Let I be an ideal of $K[T_1,\dots,T_n]$. Then $\mathcal{I}\big(\mathcal{V}(I)\big)=\sqrt{I}$ and $\mathcal{V}(\sqrt{I})=\mathcal{V}(I)$.*

It is not very hard to deduce the strong form of Hilbert's Nullstellensatz from the weak form. We will omit this proof however.

Corollary 6.3.6. *Let V be an affine variety in K^n. The map*

$$\begin{array}{rccc}\Phi_V: & V & \longrightarrow & \big\{\text{maximal ideals of } \mathcal{O}(V)\big\}\\ & (a_1,\dots,a_n) & \longmapsto & \langle[T_1-a_1],\dots,[T_n-a_n]\rangle\end{array}$$

is a bijection.

Proof. Let $\pi: K[T_1,\dots,T_n]\to\mathcal{O}(V)$ be the quotient map and $\mathfrak{m}$ a maximal ideal of $\mathcal{O}(V)$. Then $\pi^{-1}(\mathfrak{m})$ is a maximal ideal of $K[T_1,\dots,T_n]$ since it is the kernel of the surjection $K[T_1,\dots,T_n]\to\mathcal{O}(V)/\mathfrak{m}$ whose image is a field. Since π is surjective, two ideals $\mathfrak{m}$ and $\mathfrak{m}'$ of $\mathcal{O}(V)$ coincide if $\pi^{-1}(\mathfrak{m})=\pi^{-1}(\mathfrak{m}')$. This defines an embedding π^* of the set of maximal ideals of $\mathcal{O}(V)$ into the set of maximal ideals of $K[T_1,\dots,T_n]$.

Let $\mathfrak{n}$ be a maximal ideal of $K[T_1,\dots,T_n]$. By Corollary 6.2.6, there is a unique $a=(a_1,\dots,a_n)$ in K^n such that $\mathfrak{n}=\langle T_1-a_1,\dots,T_n-a_n\rangle$. By the third isomorphism theorem for rings (Theorem 1.4.3), there is a unique maximal ideal $\mathfrak{m}$ of $\mathcal{O}(V)$ with $\mathfrak{n}=\pi^{-1}(\mathfrak{m})$ if and only if $\mathcal{I}(V)\subset\mathfrak{n}$. Note that $\mathfrak{n}=\sqrt{\mathfrak{n}}=\mathcal{I}\big(\mathcal{V}(\mathfrak{n})\big)=\mathcal{I}\big(\{a\}\big)$ by Hilbert's Nullstellensatz (Theorem 6.3.5). Thus if $a\in V$, i.e. $\{a\}\subset V$, then $\mathcal{I}(V)\subset\mathcal{I}\big(\{a\}\big)=\mathfrak{n}$ implies that $\Phi_K(a)$ is indeed a maximal ideal of $\mathcal{O}(V)$. This shows that Φ_V is well-defined.

The injectivity of Φ implies that Φ_V is injective. The map Φ_V is surjective since if the inverse image $\mathfrak{n}=\pi^{-1}(\mathfrak{m})$ of a maximal ideal $\mathfrak{m}$ of $\mathcal{O}(V)$ contains $\mathcal{I}(V)$, and thus $\Phi^{-1}(\mathfrak{n})\subset V$, which shows that $\Phi_V^{-1}(\mathfrak{m})=\Phi^{-1}(\mathfrak{n})\cap V$ is not empty. □

Definition 6.3.7. Let V be an affine variety in K^n with ring of regular functions $\mathcal{O}(V)$. Let $a\in V$ and $\mathfrak{m}=\Phi_V(a)$. The **stalk of** $\mathcal{O}(V)$ **in** a is $\mathcal{O}_{V,a}=\mathcal{O}(V)_{\mathfrak{m}}$. We denote the unique maximal ideal $\mathfrak{m}\mathcal{O}_{V,a}$ of $\mathcal{O}_{V,a}$ by $\mathfrak{m}_a$.

6.4 Plane curves

From this section on, we will concentrate our study to plane curves. For simplifying our notation, we use $X=T_1$ and $Y=T_2$.

Definition 6.4.1. A **plane curve** is the vanishing set $C=\mathcal{V}(f)$ of a polynomial $f\in K[X,Y]$ in K^2. A plane curve C is *irreducible* if $C=\mathcal{V}(f)$ for an irreducible polynomial $f\in K[X,Y]$.

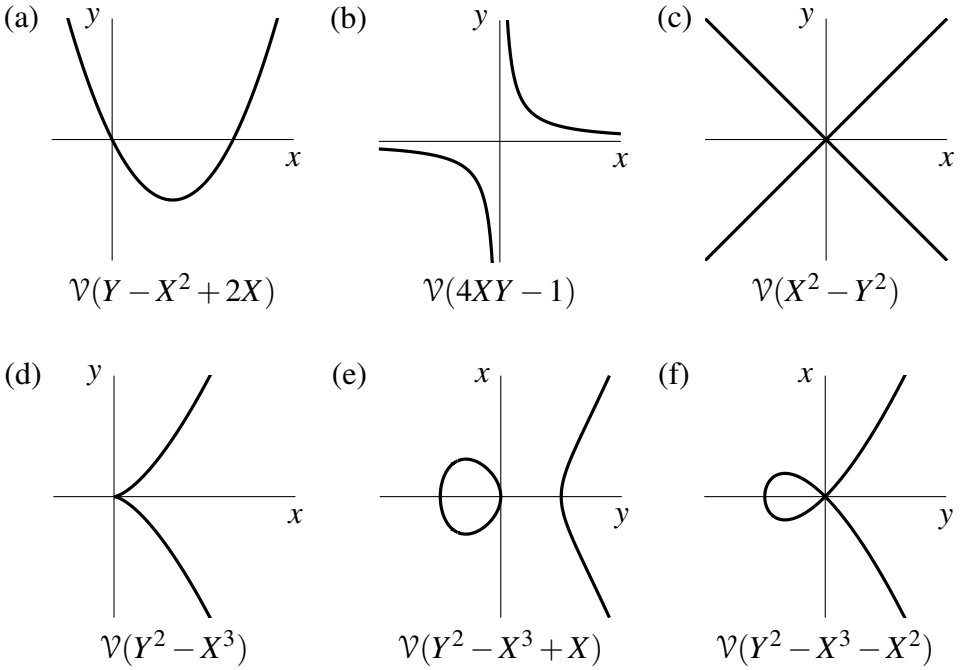
Remark. Since $\mathcal{V}(fg)=\mathcal{V}(f)\cup\mathcal{V}(g)$, every plane curve is a finite union of irreducible plane curves.

Note that, in particular, $\mathcal{V}(f^2)=\mathcal{V}(f)\cup\mathcal{V}(f)=\mathcal{V}(f)$. This shows that f does not need to be irreducible for $\mathcal{V}(f)$ to be irreducible.

Lemma 6.4.2. *Let C be a plane curve. If C is irreducible, then $\mathcal{O}(V)$ is an integral domain.*

Proof. Let $C = V(f)$ for an irreducible $f \in K[X,Y]$. Then $\langle f \rangle$ is a prime ideal and thus $\mathcal{I}(C) = \sqrt{\langle f \rangle} = \langle f \rangle$. This shows that $\mathcal{O}(V) = K[X,Y]/\mathcal{I}(C)$ is an integral domain. □

Example 6.4.3. We illustrate some example of complex plane curves $C = \mathcal{V}(f)$, i.e. for the case $K = \mathbb{C}$ and $f \in \mathbb{C}[X,Y]$. The illustrations capture the respective *real parts* $C \cap \mathbb{R}^2$ of the complex curves $C \subset \mathbb{C}^2$ where we denote the coordinates of $\mathbb{R}^2$ by x and y.



6.5 Singular points

Definition 6.5.1. Let $f = \sum a_{i,j} X^i Y^j \in K[X,Y]$. The **formal partial derivatives of** f are the polynomials

$$\frac{\partial f}{\partial X} = \sum_{i,j \in \mathbb{N}} (i+1) a_{i+1,j} X^i Y^j \qquad \text{and} \qquad \frac{\partial f}{\partial Y} = \sum_{i,j \in \mathbb{N}} (j+1) a_{i,j+1} X^i Y^j.$$

The **Jacobian matrix of** f is the matrix $J_f = \left(\frac{\partial f}{\partial X}, \frac{\partial f}{\partial Y}\right)$, which defines a function

$$\begin{array}{rccc} J_f: & K^2 & \longrightarrow & K^2 \\ & a & \longmapsto & \left(\frac{\partial f}{\partial X}(a), \frac{\partial f}{\partial Y}(a)\right) \end{array}$$

Definition 6.5.2. Let $f = \sum a_{i,j} X^i Y^j \in K[X,Y]$ be an irreducible polynomial and $C = \mathcal{V}(f)$ a plane curve. A point $a \in C$ is **singular** if $J_f(a) = (0,0)$. Otherwise, a is **nonsingular**.

Remark. In the complex case $K = \mathbb{C}$, a point a of a plane curve $C = \mathcal{V}(f)$ is nonsingular iff and only if there is an open neighbourhood U of a in $\mathbb{C}^2$ (in the usual topology of $\mathbb{C}^2$) such that $C \cap U$ is a complex submanifold of U. Naively speaking, this is the case if and only if there is only one tangent direction to C at a.

To explain, if $J_f(a) \neq 0$, then the *tangent line at* $a = (x,y)$ is given as

$$\mathcal{T}_a(f) \;=\; \mathcal{V}\Big(\big(\tfrac{\partial f}{\partial X}(a)\big)(X-x) + \big(\tfrac{\partial f}{\partial Y}(a)\big)(Y-y) \Big).$$

Note that a rigorous definition of the "number of tangent directions" is somewhat subtle, since the point $(0,0)$ of the curve $C = \mathcal{V}(Y^2 - X^3)$ is a singular point even though geometrically, $\mathcal{V}(Y)$ seems to be the only tangent line at $(0,0)$; cf. Examples 6.4.3.(d) and 6.5.7.

Example 6.5.3. Let $f = \sum c_{i,j} X^i Y^j$ be a polynomial in $K[X,Y]$. The criterion $J_f(a) = (0,0)$ assumes a particularly simple shape for the origin $\mathbf{o} = (0,0)$ of K^2. First of all notice that $\mathbf{o} \in \mathcal{V}(f)$ if and only if $f(\mathbf{o}) = c_{0,0} = 0$. Thus we can assume that f has a trivial constant coefficient.

Since $\frac{\partial f}{\partial X} = \sum (i+1) c_{i+1,j} X^i Y^j$ and $\frac{\partial f}{\partial Y} = \sum (j+1) c_{i,j+1} X^i Y^j$, we have $J_f(\mathbf{o}) = (c_{1,0}, c_{0,1})$. Thus $\mathbf{o}$ is a singular point of $\mathcal{V}(f)$ if and only if $c_{0,0} = c_{1,0} = c_{0,1} = 0$.

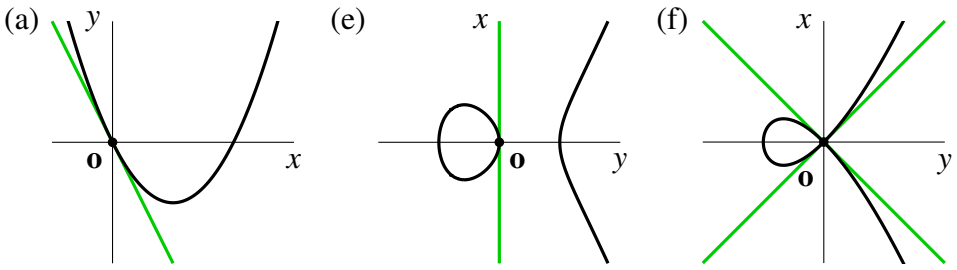
The origin $\mathbf{o}$ is a point of the plane complex curve $C = \mathcal{V}(f)$ in the cases (a), (e) and (f) of Example 6.4.3. We inspect in each case whether $\mathbf{o}$ is a singular point.

The curve in (a) is defined by $f = Y - X^2 + 2X$, and thus $J_f = (-2X+2, 1)$. Since $J_f(\mathbf{o}) = (2,1)$ is nonzero, $\mathbf{o}$ is a nonsingular point of C.

The curve in (e) is defined by $f = Y^2 - X^3 + X$, and thus $J_f = (-3X^2+1, 2Y^2)$. Since $J_f(\mathbf{o}) = (1,0)$ is nonzero, $\mathbf{o}$ is a nonsingular point of C.

The curve in (f) is defined by $f = Y^2 - X^3 - X^2$, and thus $J_f = (-3X^2 - 2X, 2Y^2)$. Since $J_f(\mathbf{o}) = (0,0)$ is zero, $\mathbf{o}$ is a singular point of C.

We illustrate the three curves, including the tangent lines $\mathcal{T}_{\mathbf{o}}(f)$ in the first two cases. In the third case, there are two "tangent directions" as indicated in the illustration.



$\mathcal{T}_{\mathbf{o}}(Y - X^2 + 2X)$ $\qquad$ $\mathcal{T}_{\mathbf{o}}(Y^2 - X^3 + X)$

Definition 6.5.4. Let $C = \mathcal{V}(f)$ be a plane curve and $a \in C$. Let $\mathcal{O}_{V,a}$ be the stalk at a and $\mathfrak{m}_a$ its maximal ideal. The **residue field of** C **at** a is the quotient $k(a) = \mathcal{O}_{V,a}/\mathfrak{m}_a$. The **cotangent space of** C **at** a is the quotient $T_a^*(C) = \mathfrak{m}_a/\mathfrak{m}_a^2$.

Lemma 6.5.5. *Let $C = \mathcal{V}(f)$ be a plane curve and $a \in C$. Let $k(a)$ the residue field and $T_a^*(C)$ the cotangent space of C at a. Then the canonical map $\iota : K \to \mathcal{O}_{V,a} \to k(a)$ is an*

isomorphism and $\mathfrak{m}_a/\mathfrak{m}_a^2$ *is a* $k(a)$*-vector space with respect to the action* $[b].[c] = [bc]$ *for* $b \in A$ *and* $c \in \mathfrak{m}_a$.

Proof. Let $\mathfrak{m} = \Phi(a) = \langle X-x, Y-y\rangle$ be the maximal ideal associated with $a = (x,y)$, cf. Corollary 6.2.6. Then

$$k(a) \;=\; \mathcal{O}_{C,a}/\mathfrak{m}_a \;=\; \big(K[X,Y]/\langle f\rangle\big)_{\mathfrak{m}} \big/ \langle [X-x],[Y-y]\rangle \\ \simeq\; K[X,Y]/\langle X-x, Y-y\rangle \;\simeq\; K$$

as K-algebras, which shows that the canonical morphism $K \to k(a)$ is an isomorphism. This establishes the first claim.

Since for every $c \in \mathfrak{m}_a$, the kernel of the $\mathcal{O}_{C,a}$-linear map $m_c : \mathcal{O}_{C,a} \to \mathfrak{m}_a/\mathfrak{m}_a^2$ that sends $b \in \mathcal{O}_{C,a}$ to $[bc]$ is $\ker m_c = \mathfrak{m}$, we obtain a well-defined map $k(a) \times (\mathfrak{m}_a/\mathfrak{m}_a^2) \to \mathfrak{m}_a/\mathfrak{m}_a^2$. The axioms of a $k(a)$-action follows at once from the corresponding properties of the $\mathcal{O}_{C,a}$-action on $\mathfrak{m}$, which establishes the second claim. □

Theorem 6.5.6. *Let* $C = \mathcal{V}(f)$ *be a plane curve,* $a \in C$ *and* $T_a^*(C)$ *the cotangent space of* C *at* a. *If* a *is singular, then* $\dim_K\big(T_a^*(C)\big) = 2$ *and if* a *is nonsingular, then* $\dim_K\big(T_a^*(C)\big) = 1$.

Proof. Let $a = (x,y)$. The map $(X,Y) \mapsto (X-x, Y-y)$ defines an automorphism $K[X,Y] \to K[X,Y]$ that induces a variable transformation $K^2 \to K^2$ that sends a point $a' = (x',y')$ to $(x'-x, y'-y)$. In particular, it sends a to the origin $\mathbf{o} = (0,0)$. Thus we can assume without loss of generality that $a = \mathbf{o}$. Since $(0,0) \in C = \mathcal{V}(f)$, the constant term $c_{0,0}$ of $f = \sum c_{i,j}X^iY^j$ is zero.

Let $\mathfrak{m} = \langle X,Y\rangle$ be the maximal ideal of $K[X,Y]$ that corresponds to $\mathbf{o}$. Consider the map

$$\begin{array}{rccc} \theta: & \mathfrak{m}/\mathfrak{m}^2 & \longrightarrow & K^2, \\ & [g] & \longmapsto & J_g(\mathbf{o}) \end{array}$$

which is well-defined since every $g \in \mathfrak{m}^2 = \langle X^2, XY, Y^2\rangle$ is without linear terms and thus $J_g(\mathbf{o}) = \big(\frac{\partial g}{\partial X}, \frac{\partial g}{\partial Y}\big)(0,0) = (0,0)$. Conversely, $J_g(\mathbf{o}) = (0,0)$ implies that $g \in \mathfrak{m}$ has neither a constant term nor linear terms and is thus in $\mathfrak{m}^2$. This shows that θ is injective.

Given $(c,d) \in K^2$, the polynomial $g = cX + dY$ has Jacobian matrix $J_g = (c,d)$ and thus $\theta\big([g]\big) = J_g(\mathbf{o}) = (c,d)$. Thus θ is surjective. This shows that $\theta : \mathfrak{m}/\mathfrak{m}^2 \to K^2$ is an isomorphism of K-vector spaces where we use the identification of K with $k(\mathbf{o})$ from Lemma 6.5.5.

Let $\pi : K[X,Y] \to K[X,Y]/\langle f\rangle = \mathcal{O}(C)$ be the quotient map and $\iota : \mathcal{O}(C) \to \mathcal{O}_{C,o}$ the canonical map to the localization. Let $\bar{\mathfrak{m}} = \pi(\mathfrak{m})$, which is a maximal ideal of $\mathcal{O}(C)$, and $\mathfrak{m}_{\mathbf{o}} = \langle \iota(\bar{\mathfrak{m}})\rangle$, which is the unique maximal ideal of $\mathcal{O}_{C,o}$. The composition of the restrictions of π to $\mathfrak{m}$ and ι to $\bar{\mathfrak{m}}$ yields a K-linear map

$$\chi: \; \mathfrak{m} \xrightarrow{\;\pi\;} \bar{\mathfrak{m}} \xrightarrow{\;\iota\;} \mathfrak{m}_o,$$

which induces the map

$$\begin{array}{rccc} \bar{\chi}: & \mathfrak{m}/\big(\langle f\rangle + \mathfrak{m}^2\big) & \longrightarrow & \mathfrak{m}_o/\mathfrak{m}_o^2. \\ & [g] & \longmapsto & [g] \end{array}$$

Note that $\bar{\chi}$ is well-defined since the image $\iota\big(\pi(g)\big)$ of an element $g \in \langle f\rangle + \mathfrak{m}^2$ is in $\iota\big(\pi(\langle f\rangle + \mathfrak{m}^2)\big) = \iota(\bar{\mathfrak{m}}^2) \subset \mathfrak{m}_o^2$. The map $\bar{\chi}$ is injective since $\bar{\chi}([g]) = 0$ implies that $g \in \chi^{-1}(\mathfrak{m}_o^2) = \pi^{-1}(\bar{\mathfrak{m}}^2) = \langle f\rangle + \mathfrak{m}^2$. The map $\bar{\chi}$ is surjective since every $[g] \in \mathfrak{m}_o/\mathfrak{m}_o^2$ can be represented by a linear polynomial $g' = cX + dY$ and thus $[g] = [g'] = \chi\big([cX + dY]\big)$. This shows that $\chi : \mathfrak{m}/(\langle f\rangle + \mathfrak{m}^2) \to T_o^*(C)$ is an isomorphism of K-vector spaces.

The image of the K-linear subspace $(\langle f\rangle + \mathfrak{m}^2)/\mathfrak{m}^2$ of $\mathfrak{m}/\mathfrak{m}^2$ under θ is the K-linear subspace $\langle J_f(\mathbf{o})\rangle$ of K^2, and its dimension is equal to the rank of $J_f(\mathbf{o})$. Thus

$$\begin{aligned}\dim_K\big(T_o^*(C)\big) + \mathrm{rk}\,\big(J_f(\mathbf{o})\big) &= \dim_K\Big(\mathfrak{m}/\big(\langle f\rangle + \mathfrak{m}^2\big)\Big) + \dim_K\Big(\big(\langle f\rangle + \mathfrak{m}^2\big)/\mathfrak{m}^2\Big)\\ &= \dim_K\big(\mathfrak{m}/\mathfrak{m}^2\big) \;=\; \dim_K K^2 \;=\; 2.\end{aligned}$$

If $\mathbf{o}$ is singular, then the rank of $J_f(\mathbf{o})$ is 0 and thus $\dim_K\big(T_o^*(C)\big) = 2$. If $\mathbf{o}$ is nonsingular, then the rank of $J_f(\mathbf{o})$ is 1 and thus $\dim_K\big(T_o^*(C)\big) = 1$. This concludes the proof of the theorem. □

Remark. Let $C = \mathcal{V}(f)$ be a plane curve. The relation between the cotangent space $T_a^*(C)$ and the tangent line $\mathcal{T}_a(f)$ at a nonsingular point $a \in C$ is as follows. We leave the verification of the details in the following explanations as an exercise.

For an arbitrary point $a = (x,y) \in C$, the *tangent space at a* is defined as the dual K-vector space $T_a(C) = \mathrm{Hom}_K\big(T_a^*(C), K\big)$ of the cotangent space. We have a canonical inclusion

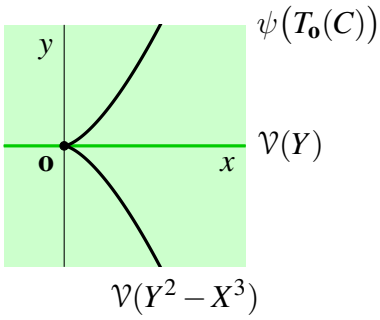
$$\begin{array}{rccc}\psi: & T_a(C) & \longrightarrow & K^2\\ & \alpha : T_a^*(C) \to K & \longmapsto & \big(\alpha([X]), \alpha([Y])\big)\end{array}$$

of K-vector spaces. Since the dimension of a finitely dimensional K-vector space is equal to the dimension of the dual space, Theorem 6.5.6 shows that ψ is surjective if a is singular. If a is nonsingular, then $\dim_K\big(T_a(C)\big) = 1$, and the tangent line $\mathcal{T}_a(f)$ of C at a equals the translation of the image of ψ by the vector $a = (x,y)$, i.e.

$$\mathcal{T}_a(f) \;=\; \big\{a + \psi(\alpha)\,\big|\,\alpha \in T_a(C)\big\}.$$

Example 6.5.7. To conclude, we inspect the tangent space of the curve $C = \mathcal{V}(f)$ with $f = Y^2 - X^3$ at the origin $\mathbf{o} = (0,0)$; cf. Example 6.4.3.(d). The Jacobian is $J_f = (2Y, 3X^2)$, and thus $J_f(\mathbf{o}) = (0,0)$, which shows that $\mathbf{o}$ is a singular point of C. Thus by Theorem 6.5.6, the cotangent space $T_{\mathbf{o}}^*(C)$ and its dual $T_{\mathbf{o}}(C)$ have both dimension 2 and the image of ψ is all of K^2. This shows that the tangent space $T_{\mathbf{o}}(C)$, as defined here, is well-suited to detect singularities, in contrast to the geometric intuition that suggests that the only tangent direction is captured by the line $\mathcal{V}(Y)$, as illustrated

below.



6.6 The stalks of nonsingular points

In this section, we will show that the nonsingular points of a plane curve are characterized by the property that their stalks are discrete valuation rings, which are particularly well-behaved rings whose definition is as follows.

Definition 6.6.1. A **discrete valuation ring** (often just **DVR**) is a principal ideal domain with a unique prime element p, up to associatives. The prime element p is called a **uniformizer**.

Lemma 6.6.2. *Let A be a discrete valuation ring and $p \in A$ a uniformizer. Then the following holds.*

(1) For every nonzero $a \in A$, there are a unique $u \in A^\times$ and $i \in \mathbb{N}$ such that $a = up^i$.

(2) The ring A is local with maximal ideal $\mathfrak{m} = \langle p \rangle$.

(3) Every nonzero ideal I of A is of the form $\mathfrak{m}^i = \langle p^i \rangle$ for some $i \in \mathbb{N}$.

(4) The intersection of all nonzero ideals is $\bigcap_{i \in \mathbb{N}} \mathfrak{m}^i = \{0\}$.

Proof. This is Exercise 1.45. □

For the proof of Theorem 6.6.7, we will apply Nakayama's lemma and Krull's principal ideal theorem, which we state here without a proof.

Theorem 6.6.3 (Nakayama's lemma)**.** *Let A be a ring, I an ideal of A and M a finitely generated A-module. If $IM = M$, then there is an $a \in A$ such that $aM = 0$ and $[a] = [1]$ in A/I.*

Corollary 6.6.4. *Let A be a local ring with maximal ideal $\mathfrak{m}$ and M an A-module. If $\mathfrak{m}M = M$, then $M = 0$.*

Proof. By Nakayama's lemma (Theorem 6.6.3), $\mathfrak{m}M = M$ implies that there is an $a \in A$ such that $aM = 0$ and $[a] = [1]$ in $A/\mathfrak{m}$. Thus $a \notin \mathfrak{m}$, which implies by Lemma 1.8.8 that a is a unit. Thus $M = a^{-1}(aM) = 0$, as claimed. □

Definition 6.6.5. Let A be a ring and I an ideal of A. The height of I is the supremum over the lengths l of properly increasing chains

$$\mathfrak{p}_0 \subsetneq \mathfrak{p}_1 \subsetneq \cdots \subsetneq \mathfrak{p}_l$$

of prime ideals $\mathfrak{p}_0, \ldots, \mathfrak{p}_l$ contained in I.

Theorem 6.6.6 (Krull's principal ideal theorem)**.** *Let A be a Noetherian ring and $a \in A$. Let $\mathfrak{p}$ be a minimal prime ideal of A that contains a, i.e. if $a \in \mathfrak{q} \subset \mathfrak{p}$ for a prime ideal $\mathfrak{q}$, then $\mathfrak{q} = \mathfrak{p}$. Then the height of $\mathfrak{p}$ is at most* 1.

Theorem 6.6.7. *Let C be a plane curve, $a \in C$ and $\mathcal{O}_{C,a}$ its stalk. Then a is nonsingular if and only if $\mathcal{O}_{C,a}$ is a discrete valuation ring.*

Proof. Let us assume that $\mathcal{O}_{C,a}$ is a discrete valuation ring. Let $\mathfrak{m}_a = \langle p \rangle$ be the maximal ideal of $\mathcal{O}_{C,a}$ where p is a uniformizer. Consider the map

$$\begin{array}{rccc} \xi: & K & \longrightarrow & \mathfrak{m}_a/\mathfrak{m}_a^2, \\ & [b] & \longmapsto & [bp] \end{array}$$

which is well-defined since if $[b'] = [b]$ in K, i.e. $b' = b + cp$ for some $c \in A$, then $[b'p] = \left[bp + cp^2\right] = [bp]$ in $\mathfrak{m}_a/\mathfrak{m}_a^2$. That ξ is K-linear follows at once from the fact that it is derived from the A-linear map $A \to \mathfrak{m}_a$ that sends b to bp.

The K-linear map ξ is injective since $\xi\big([b]\big) = 0$ implies that p divides b and thus $[b] = [0]$ in K. It is surjective since by Lemma 6.6.2, every $c \in \mathfrak{m}_a$ is of the form bp for some $b \in A$ and thus $[c] = \xi(b')$ where b' is the inverse image of $[b] \in k(a)$ under the isomorphism $K \to k(a)$ from Lemma 6.5.5. This shows that ξ is an isomorphism of K-vector spaces. Thus $\dim_K \left(\mathfrak{m}_a/\mathfrak{m}_a^2\right) = 1$, which shows that a is nonsingular by Theorem 6.5.6.

Conversely, assume that a is nonsingular. By Theorem 6.5.6, $\mathfrak{m}_a/\mathfrak{m}_a^2$ is a one-dimensional K-vector space and thus generated by a single element $[p]$ where $p \in \mathfrak{m}_a$, i.e. $\mathfrak{m}_a = \langle p \rangle + \mathfrak{m}_a^2$. Then

$$\mathfrak{m}_a \cdot \big(\mathfrak{m}_a/\langle p \rangle\big) \;=\; \big(\mathfrak{m}_a^2 + \langle p \rangle\big)/\langle p \rangle \;=\; \mathfrak{m}_a/\langle p \rangle.$$

Thus Corollary 6.6.4 implies that $\mathfrak{m}_a/\langle p \rangle = 0$, which shows that $\mathfrak{m}_a = \langle p \rangle$ is generated by p.

Our next step is to show that $\mathcal{O}_{C,a}$ is an integral domain. This follows at once from Lemma 6.4.2 if C is irreducible. If not, let $C = \mathcal{V}(f_1 \cdots f_r)$ where $f_1, \ldots, f_r$ are irreducible polynomials in $K[X,Y]$. Let $a = (x,y)$ and $\mathfrak{m} = \langle X - x, Y - y \rangle$ the maximal ideal of $K[X,Y]$ defined by a. Since $a \in C$, we have $\langle f \rangle \subset \mathfrak{m}$ and thus $\mathfrak{p}_i = \langle f_i \rangle \subset \mathfrak{m}$ for some i. Since f_i is irreducible, $\mathfrak{p}_i$ is a prime ideal of $K[X,Y]$. To summarize, we have inclusions $\langle f \rangle \subset \mathfrak{p}_i \subset \mathfrak{m}$.

By Krull's principal ideal theorem (Theorem 6.6.6), the height of the principal ideal $\mathfrak{p}_i = \langle f_i \rangle$ is at most 1. Since the chain of proper inclusions $\langle 0 \rangle \subsetneq \langle X - x \rangle \subsetneq \mathfrak{m}$ of prime ideals shows that the height of $\mathfrak{m}$ is at least 2, we conclude that $\mathfrak{p}_i$ is properly contained in $\mathfrak{m}$. Let $\pi : K[X,Y] \to \mathcal{O}(C)$ be the quotient map. By the third isomorphism theorem for

rings (Theorem 1.4.3), the ideal $\bar{\mathfrak{p}}_i = \pi(\mathfrak{p}_i)$ of $\mathcal{O}(C)$ is properly contained in $\bar{\mathfrak{m}} = \pi(\mathfrak{m})$ and

$$\mathcal{O}(C)/\bar{\mathfrak{p}}_i \;=\; \big(K[X,Y]/\langle f\rangle\big)\big/\big(\mathfrak{p}_i/\langle f\rangle\big) \;\xrightarrow{\sim}\; K[X,Y]/\mathfrak{p}_i$$

is an integral domain, which shows that $\bar{\mathfrak{p}}_i$ is a prime ideal of $\mathcal{O}(C)$. Let $\iota : \mathcal{O}(C) \to \mathcal{O}_{C,a}$ be the canonical map into the localization. By Exercise 1.36, the ideal $\mathfrak{p}_{i,a} = \langle \iota(\bar{\mathfrak{p}}_i)\rangle$ is prime and properly contained in the maximal ideal $\mathfrak{m}_a = \langle \iota(\bar{\mathfrak{m}})\rangle$ of $\mathcal{O}_{C,a}$. Since p generates $\mathfrak{m}_a$, it is not contained in $\mathfrak{p}_{i,a}$. Thus $bp \in \mathfrak{p}_{i,a}$ implies that $b \in \mathfrak{p}_{i,a}$, which means that $\langle p\rangle \cdot \mathfrak{p}_{i,a} = \mathfrak{p}_{i,a}$. Since $\mathfrak{m}_a = \langle p\rangle$ is maximal, Corollary 6.6.4 implies that $\mathfrak{p}_i = 0$. This shows that $\langle 0\rangle$ is a prime ideal of $\mathcal{O}_{C,a}$ and that $\mathcal{O}_{C,a}$ is an integral domain, as claimed.

By Corollary 6.1.3, $\mathcal{O}(C)$ is Noetherian, and by Lemma 6.1.4, $\mathcal{O}_{C,a}$ is Noetherian. Since $\mathfrak{m}_a$ is the minimal prime ideal that contains p, Krull's principal ideal theorem (Theorem 6.6.6) implies that $\mathfrak{m}_a$ has height at most 1. This shows that the only prime ideals of $\mathcal{O}_{C,a}$ are $\langle 0\rangle$ and $\mathfrak{m}_a$.

Our next step is to show that $\mathcal{O}_{C,a}$ is a principal ideal domain. Let I be a nontrivial and proper ideal of $\mathcal{O}_{C,a}$. By Lemma 6.3.4, its radical $\sqrt{I}$ is the intersection of all prime ideals containing I, i.e. $\sqrt{I} = \mathfrak{m}_a$. Thus $p^i \in I$ for some $i \geqslant 1$ and $\mathfrak{m}^i = \langle p^i\rangle \subset I$. If $I = \langle p^i\rangle$, then I is principal.

If $\mathfrak{m}^i$ is properly contained in I, then there is an $j \geqslant 1$ such that $I \subset \mathfrak{m}^j$ and $I \not\subset \mathfrak{m}^{j+1}$. Thus there is a $g = up^j \in I - \mathfrak{m}^{j+1}$ for some $u \in \mathcal{O}_{C,a}^\times$. This implies that $p^j = u^{-1}g \in I \subset \mathfrak{m}^j$, which shows that $I = \langle p^j\rangle$ is principal. This shows that $\mathcal{O}_{C,a}$ is a principal domain, as claimed.

We are left with showing that $\mathcal{O}_{C,a}$ has only one prime element up to associatives. Given two prime elements g and h of $\mathcal{O}_{C,a}$, they both generate a nonzero prime ideal. Since $\mathfrak{m}_a$ is the only nonzero prime ideal of $\mathcal{O}_{C,a}$, we have $\langle g\rangle = \mathfrak{m}_a = \langle h\rangle$, which shows that $g \sim h$. This completes the proof of the theorem. □

Chapter 7

What is a universal property?

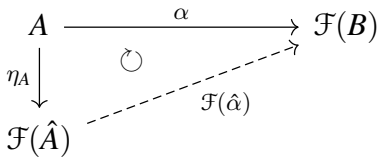
In the previous chapters, we have encountered numerous examples of universal properties. Roughly speaking, a universal property applies in situations where we have a morphism from (or to) an object A to (or from) a second object A', which is related to A by some type of construction. The universal property expresses that the morphisms with certain properties from (or to) A correspond bijectively to the morphisms from (or to) A'.

In this chapter, we will make the concept of a universal property precise in a categorical framework and explain in which sense universal properties are intimately linked to adjoint functors.

7.1 Initial and terminal morphisms to a functor

Throughout the whole chapter, let $\mathcal{C}$ and $\mathcal{D}$ be categories.

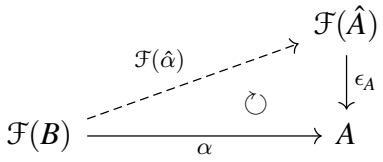
Definition 7.1.1. Let $\mathcal{F} : \mathcal{C} \to \mathcal{D}$ be a covariant functor and A an object in $\mathcal{D}$. An **initial morphism from** A **to** $\mathcal{F}$ is an object $\hat{A}$ in $\mathcal{C}$ together with a morphism $\eta_A : A \to \mathcal{F}(\hat{A})$ that satisfies the following *universal property*: for every object B in $\mathcal{C}$ and every morphism $\alpha : A \to \mathcal{F}(B)$, there is a unique morphism $\hat{\alpha} : \hat{A} \to B$ such that $\alpha = \mathcal{F}(\hat{\alpha}) \circ \eta_A$, i.e. the diagram



commutes.

A **terminal morphism from** $\mathcal{F}$ **to** A is an object $\hat{A}$ in $\mathcal{C}$ together with a morphism $\epsilon_A : \mathcal{F}(\hat{A}) \to A$ that satisfies the following *universal property*: for every object B in $\mathcal{C}$ and every morphism $\alpha : \mathcal{F}(B) \to A$, there is a unique morphism $\hat{\alpha} : B \to \hat{A}$ such that

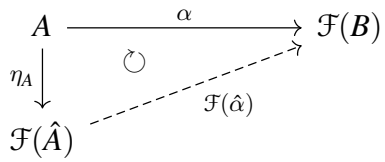
$\alpha = \epsilon_A \circ \mathcal{F}(\hat{\alpha})$, i.e. the diagram



commutes.

Example 7.1.2. For some universal properties from these lecture notes, it is easy to see what the corresponding functor $\mathcal{F} : \mathcal{C} \to \mathcal{D}$ is. In other cases, this is less obvious. We consider some examples in the following, and encourage the reader to think about some other universal properties.

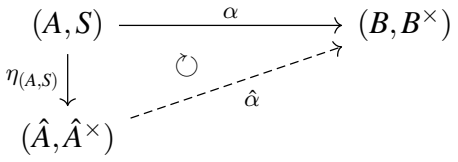
(1) Let $\mathcal{F} : \mathrm{Rings} \to \mathrm{Sets}$ be the forgetful functor, cf. Example 2.5.2, and let A be a set. Then an initial morphism from A to $\mathcal{F}$ is a ring $\hat{A}$ together with a map $\eta_A : A \to \mathcal{F}(\hat{A})$ that satisfies the following universal property: for every ring B and every map $\alpha : A \to \mathcal{F}(B)$, there is a unique ring homomorphism $\hat{\alpha} : \hat{A} \to B$ such that $\alpha = \mathcal{F}(\hat{\alpha}) \circ \eta_A$, i.e. the diagram



commutes. This universal property is satisfied by the polynomial algebra $\hat{A} = \mathbb{Z}[T_i \mid i \in A]$ together with the map $\eta_A : A \to \mathbb{Z}[T_i \mid i \in A]$ with $\eta_A(i) = T_i$.

(2) Let $\mathcal{D}$ be the category whose objects are pairs (A,S) of a ring A together with a multiplicative subset S of A and whose morphisms $\alpha : (A,S) \to (A',S')$ are ring homomorphisms $\alpha : A \to A'$ with $\alpha(S) \subset S'$. Let $\mathcal{F} : \mathrm{Rings} \to \mathcal{D}$ be the functor that sends a ring A to $(A,A^\times)$ and a ring homomorphism $\beta : A \to B$ to itself. Note that $\beta(A^\times) \subset B^\times$, which shows that $\mathcal{F}$ is well-defined.

Let (A,S) be an object in $\mathcal{D}$. An initial morphism from (A,S) to $\mathcal{F}$ is a ring $\hat{A}$ together with a morphism $\eta_{(A,S)} : (A,S) \to (A,A^\times)$ that satisfies the following universal property: for every ring B and every morphism $\alpha : (A,S) \to (B,B^\times)$ there is a unique ring homomorphism $\hat{\alpha} : \hat{A} \to B$ such that the diagram



commutes. This universal property is satisfied by the localization $\hat{A} = S^{-1}A$ together with the canonical morphism $\eta_A = \iota_S : A \to S^{-1}A$, which maps S to $(S^{-1}A)^\times$.

(3) The *product* $\mathcal{C} \times \mathcal{D}$ *of two categories* $\mathcal{C}$ *and* $\mathcal{D}$ is defined as follows: its objects are pairs (A,B) of an object A in $\mathcal{C}$ and an object B in $\mathcal{D}$ and its morphisms $(\alpha,\beta) : (A,B) \to (A',B')$ are pairs of a morphism $\alpha : A \to A'$ in $\mathcal{C}$ and a morphism $\beta : B \to B'$ in $\mathcal{D}$. The composition of two morphisms $(\alpha,\beta) : (A,B) \to (A',B')$ and $(\alpha',\beta') : (A',B') \to (A'',B'')$ is defined as $(\alpha',\beta') \circ (\alpha,\beta) = (\alpha' \circ \alpha, \beta' \circ \beta)$. The *diagonal functor* $\Delta : \mathcal{C} \to \mathcal{C} \times \mathcal{C}$ sends an object A of $\mathcal{C}$ to (A,A) and a morphism $\alpha : A \to A'$ in $\mathcal{C}$ to $(\alpha,\alpha) : (A,A) \to (A',A')$.

Let us consider the diagonal functor $\Delta : \text{Rings} \to \text{Rings} \times \text{Rings}$ and an object (A,B) of $\text{Rings} \times \text{Rings}$. A terminal morphism from Δ to (A,B) is a ring $C = (A,B)\hat{}$ together with a morphism $\epsilon_{(A,B)} : (C,C) \to (A,B)$ in $\text{Rings} \times \text{Rings}$ that satisfies the following universal property: for every ring and every morphism $\alpha : (D,D) \to (A,B)$ in $\text{Rings} \times \text{Rings}$, there is a unique ring homomorphism $\hat{\alpha} : D \to C$ such that the diagram

$$\begin{array}{ccc} & & (C,C) \\ & \overset{(\hat{\alpha},\hat{\alpha})}{\nearrow} & \circlearrowleft \\ (D,D) & \xrightarrow[\alpha]{} & (A,B) \end{array} \qquad \epsilon_{(A,B)} \downarrow$$

commutes. This universal property is satisfied by the product $C = A \times B$ of the rings A and B together with the morphism $\epsilon_{(A,B)} = (\pi_A, \pi_B) : (A \times B, A \times B) \to (A,B)$.

7.2 Natural transformations

Definition 7.2.1. Let $\mathcal{F} : \mathcal{C} \to \mathcal{D}$ and $\mathcal{G} : \mathcal{C} \to \mathcal{D}$ be covariant functors. A **natural transformation** (or **morphism of functors**) $\eta : \mathcal{F} \to \mathcal{G}$ **from** $\mathcal{F}$ **to** $\mathcal{G}$ is a collection η of morphisms $\eta_A : \mathcal{F}(A) \to \mathcal{G}(A)$ in $\mathcal{D}$ where A varies through all objects of $\mathcal{C}$ such that the diagram

$$\begin{array}{ccc} \mathcal{F}(A) & \xrightarrow{\mathcal{F}(\alpha)} & \mathcal{F}(B) \\ \eta_A \downarrow & & \downarrow \eta_B \\ \mathcal{G}(A) & \xrightarrow[\mathcal{G}(\alpha)]{} & \mathcal{G}(B) \end{array}$$

commutes for every morphism $\alpha : A \to B$ in $\mathcal{C}$.

7.3 The unit and the counit of an adjunction

We recall the definition of adjoint functors from section 2.6. Let $\mathcal{F} : \mathcal{C} \to \mathcal{D}$ and $\mathcal{G} : \mathcal{D} \to \mathcal{F}$ be covariant functors. Then $\mathcal{G}$ is left adjoint to $\mathcal{F}$, written $\mathcal{G} \dashv \mathcal{F}$, if for every pair of objects A in $\mathcal{D}$ and B in $\mathcal{C}$, there is a bijection

$$\Phi_{A,B} : \quad \text{Hom}_{\mathcal{C}}\big(\mathcal{G}(A), B\big) \quad \longrightarrow \quad \text{Hom}_{\mathcal{D}}\big(A, \mathcal{F}(B)\big)$$

such that the diagram

$$\begin{array}{ccc} \mathrm{Hom}_{\mathcal{C}}\big(\mathcal{G}(A'),B\big) & \xrightarrow{\Phi_{A',B}} & \mathrm{Hom}_{\mathcal{D}}\big(A',\mathcal{F}(B)\big) \\ {\scriptstyle \beta\circ-\circ\mathcal{G}(\alpha)}\downarrow & & \downarrow{\scriptstyle \mathcal{F}(\beta)\circ-\circ\alpha} \\ \mathrm{Hom}_{\mathcal{C}}\big(\mathcal{G}(A),B'\big) & \xrightarrow{\Phi_{A,B'}} & \mathrm{Hom}_{\mathcal{D}}\big(A,\mathcal{F}(B')\big) \end{array}$$

commutes for every pair of a morphism $\alpha : A \to A'$ in $\mathcal{D}$ and a morphism $\beta : B \to B'$ in $\mathcal{C}$, i.e.

$$\Phi_{A,B'}\big(\beta\circ\gamma\circ\mathcal{G}(\alpha)\big) \;=\; \mathcal{F}(\beta)\circ\Phi_{A',B}(\gamma)\circ\alpha$$

for all morphisms $\gamma : \mathcal{G}(A') \to B$ in $\mathcal{C}$.

Remark. Note that the definition of adjoint functors in section 2.6 was phrased in terms of the inverse bijection $\Psi_{A,B}$ of $\Phi_{A,B}$, with the roles of $\mathcal{F}$ and $\mathcal{G}$ interchanged. For the upcoming explanations, the present convention seems more practical though.

Definition 7.3.1. Let $\mathcal{G} : \mathcal{D} \to \mathcal{C}$ be a left adjoint functor to $\mathcal{F} : \mathcal{C} \to \mathcal{D}$ with adjunction

$$\mathrm{Hom}_{\mathcal{C}}\big(\mathcal{G}(A),B\big) \underset{\Psi_{A,B}}{\overset{\Phi_{A,B}}{\rightleftarrows}} \mathrm{Hom}_{\mathcal{D}}\big(A,\mathcal{F}(B)\big)$$

for $A \in \mathrm{Ob}(\mathcal{D})$ and $B \in \mathrm{Ob}(\mathcal{C})$. The **unit of the adjunction** is the collection η of morphisms

$$\eta_A = \Phi_{A,\mathcal{G}(A)}(\mathrm{id}_{\mathcal{G}(A)}) : \; A \;\longrightarrow\; \mathcal{F}\big(\mathcal{G}(A)\big)$$

where A varies through all objects of $\mathcal{C}$ and where $\mathrm{id}_{\mathcal{G}(A)} \in \mathrm{Hom}_{\mathcal{C}}\big(\mathcal{G}(A),\mathcal{G}(A)\big)$ is the identity of $\mathcal{G}(A)$.

The **counit of the adjunction** is the collection ϵ of morphisms

$$\epsilon_B = \Psi_{\mathcal{F}(B),B}(\mathrm{id}_{\mathcal{F}(B)}) : \; \mathcal{G}\big(\mathcal{F}(B)\big) \;\longrightarrow\; B$$

where B varies through all objects of $\mathcal{D}$ and where $\mathrm{id}_{\mathcal{F}(B)} \in \mathrm{Hom}_{\mathcal{D}}\big(\mathcal{F}(B),\mathcal{F}(B)\big)$ is the identity of $\mathcal{F}(B)$.

In the following, we need to consider compositions of functors, which are defined as follows.

Definition 7.3.2. Let $\mathcal{F} : \mathcal{C} \to \mathcal{D}$ and $\mathcal{G} : \mathcal{D} \to \mathcal{D}'$ be covariant functors. The **composition of $\mathcal{F}$ with** $\mathcal{G}$ is the functor $\mathcal{G}\circ\mathcal{F} : \mathcal{C} \to \mathcal{D}'$ that sends an object A in $\mathcal{C}$ to the object $\mathcal{G}\big(\mathcal{F}(A)\big)$ in $\mathcal{D}'$ and a morphism $\alpha : A \to B$ in $\mathcal{C}$ to the morphism $\mathcal{G}\big(\mathcal{F}(\alpha)\big) : \mathcal{G}\big(\mathcal{F}(A)\big) \to \mathcal{G}\big(\mathcal{F}(B)\big)$ in $\mathcal{D}'$.

Lemma 7.3.3. *Let $\mathcal{G} \dashv \mathcal{F}$ be adjoint functors. Then the unit is a natural transformation $\eta : \mathrm{id}_{\mathcal{D}} \to \mathcal{F}\circ\mathcal{G}$ and the counit is a natural transformation $\epsilon : \mathcal{G}\circ\mathcal{F} \to \mathrm{id}_{\mathcal{C}}$.*

Proof. We want to show that the diagrams

$$\begin{array}{ccc} A & \xrightarrow{\alpha} & A' \\ {\scriptstyle \eta_A}\downarrow & & \downarrow{\scriptstyle \eta_{A'}} \\ \mathcal{F}\circ\mathcal{G}(A) & \xrightarrow{\mathcal{F}\circ\mathcal{G}(\alpha)} & \mathcal{F}\circ\mathcal{G}(A') \end{array} \qquad \text{and} \qquad \begin{array}{ccc} \mathcal{G}\circ\mathcal{F}(B) & \xrightarrow{\mathcal{G}\circ\mathcal{F}(\beta)} & \mathcal{G}\circ\mathcal{F}(B') \\ {\scriptstyle \epsilon_B}\downarrow & & \downarrow{\scriptstyle \epsilon_{B'}} \\ B & \xrightarrow{\beta} & B' \end{array}$$

commute for all morphisms $\alpha : A \to A'$ in $\mathcal{D}$ and $\beta : B \to B'$ in $\mathcal{C}$. Let

$$\mathrm{Hom}_{\mathcal{C}}\big(\mathcal{G}(A), B\big) \underset{\Psi_{A,B}}{\overset{\Phi_{A,B}}{\rightleftarrows}} \mathrm{Hom}_{\mathcal{D}}\big(A, \mathcal{F}(B)\big)$$

be the adjunction between $\mathcal{F}$ and $\mathcal{G}$. Using the definition of the adjunction and of the unit yields

$$\begin{aligned} \big(\mathcal{F}\circ\mathcal{G}(\alpha)\big)\circ\eta_A &= \mathcal{F}\big(\mathcal{G}(\alpha)\big)\circ\Phi_{A,\mathcal{G}(A)}(\mathrm{id}_{\mathcal{G}(A)})\circ\mathrm{id}_A \\ &= \Phi_{A,\mathcal{G}(A')}\big(\mathcal{G}(\alpha)\circ\mathrm{id}_{\mathcal{G}(A)}\circ\mathrm{id}_{\mathcal{G}(A)}\big) \\ &= \Phi_{A,\mathcal{G}(A')}\big(\mathrm{id}_{\mathcal{G}(A')}\circ\mathrm{id}_{\mathcal{G}(A')}\circ\mathcal{G}(\alpha)\big) \\ &= \mathrm{id}_{\mathcal{F}\circ\mathcal{G}(A')}\circ\Phi_{A',\mathcal{G}(A')}(\mathrm{id}_{\mathcal{G}(A')})\circ\alpha \\ &= \eta_{A'}\circ\alpha, \end{aligned}$$

which shows that the first diagram commutes. The proof of the commutativity of the second diagram is similar. □

Proposition 7.3.4. *Let $\mathcal{G}\dashv\mathcal{F}$ be an adjunction with unit $\eta : \mathrm{id}_{\mathcal{D}} \to \mathcal{F}\circ\mathcal{G}$ and counit $\epsilon : \mathcal{G}\circ\mathcal{F} \to \mathrm{id}_{\mathcal{C}}$. Then the diagrams*

$$\begin{array}{ccc} \mathcal{G}(A) & \xrightarrow{\mathcal{G}(\eta_A)} & \mathcal{G}\circ\mathcal{F}\circ\mathcal{G}(A) \\ & {\scriptstyle \mathrm{id}_{\mathcal{G}(A)}}\searrow & \downarrow{\scriptstyle \epsilon_{\mathcal{G}(A)}} \\ & & \mathcal{G}(A) \end{array} \qquad \textit{and} \qquad \begin{array}{ccc} \mathcal{F}(B) & \xrightarrow{\eta_{\mathcal{F}(B)}} & \mathcal{F}\circ\mathcal{G}\circ\mathcal{F}(B) \\ & {\scriptstyle \mathrm{id}_{\mathcal{F}(B)}}\searrow & \downarrow{\scriptstyle \mathcal{F}(\epsilon_B)} \\ & & \mathcal{F}(B) \end{array}$$

commute for every object A in $\mathcal{D}$ and every object B in $\mathcal{C}$.

Proof. Let

$$\mathrm{Hom}_{\mathcal{C}}\big(\mathcal{G}(A), B\big) \underset{\Psi_{A,B}}{\overset{\Phi_{A,B}}{\rightleftarrows}} \mathrm{Hom}_{\mathcal{D}}\big(A, \mathcal{F}(B)\big)$$

be the adjunction between $\mathcal{F}$ and $\mathcal{G}$. By the definitions of ϵ and η, we have $\Phi_{\mathcal{F}\circ\mathcal{G}(A),\mathcal{G}(A)}(\epsilon_{\mathcal{G}(A)}) = \mathrm{id}_{\mathcal{F}\circ\mathcal{G}(A)}$ and $\eta_A = \Phi_{A,\mathcal{G}(A)}(\mathrm{id}_{\mathcal{G}(A)})$, and thus

$$\begin{aligned} \Phi_{A,\mathcal{G}(A)}\big(\mathcal{G}(\eta_A)\circ\epsilon_{\mathcal{G}(A)}\circ\mathrm{id}_{\mathcal{G}(A)}\big) &= \eta_A\circ\Phi_{\mathcal{F}\circ\mathcal{G}(A),\mathcal{G}(A)}(\epsilon_{\mathcal{G}(A)})\circ\mathrm{id}_{\mathcal{F}\circ\mathcal{G}(A)} \\ &= \Phi_{A,\mathcal{G}(A)}(\mathrm{id}_{\mathcal{G}(A)}). \end{aligned}$$

Since $\Phi_{A,\mathcal{G}(A)}$ is injective, $\mathcal{G}(\eta_A)\circ\epsilon_{\mathcal{G}(A)} = \mathrm{id}_{\mathcal{G}(A)}$, which shows that the first diagram commutes.

By the definitions of η and ϵ, we have $\Psi_{\mathcal{F}(B),\mathcal{G}\circ\mathcal{F}(B)}(\eta_{\mathcal{F}(B)}) = \mathrm{id}_{\mathcal{G}\circ\mathcal{F}(B)}$ and $\epsilon_B = \Psi_{B,\mathcal{F}(B)}(\mathrm{id}_{\mathcal{F}(B)})$, and thus

$$\begin{aligned}\Psi_{\mathcal{F}(B),B}\big(\mathrm{id}_{\mathcal{F}(B)}\circ\eta_{\mathcal{F}(B)}\circ\mathcal{F}(\epsilon_B)\big) &= \mathrm{id}_{\mathcal{G}\circ\mathcal{F}(B)}\circ\Psi_{\mathcal{F}(B),\mathcal{G}\circ\mathcal{F}(B)}(\eta_{\mathcal{F}(B)})\circ\epsilon_B\\ &= \Psi_{\mathcal{F}(B),B}\big(\mathrm{id}_{\mathcal{F}(B)}\big).\end{aligned}$$

Since $\Psi_{\mathcal{F}(B),B}$ is injective, $\eta_{\mathcal{F}(B)}\circ\mathcal{F}(\epsilon_B) = \mathrm{id}_{\mathcal{F}(B)}$, which shows that the second diagram commutes. □

7.4 The relation between universal properties and adjoint functors

Theorem 7.4.1. *Let $\mathcal{F}:\mathcal{C}\to\mathcal{D}$ be a functor. Then the following claims are equivalent.*

(1) Every object A in $\mathcal{D}$ has an initial morphism to $\mathcal{F}$.

(2) The functor $\mathcal{F}$ has a left adjoint $\mathcal{G}:\mathcal{D}\to\mathcal{C}$.

Proof. Assume (1). We construct the functor $\mathcal{G}:\mathcal{D}\to\mathcal{C}$ as follows. For every object A in $\mathcal{D}$, we define $\mathcal{G}(A)=\hat{A}$ where $\hat{A}\in\mathrm{Ob}(\mathcal{C})$ together with $\eta_A : A\to\mathcal{F}(\hat{A})$ is the initial morphism from A to $\mathcal{F}$. By the universal property of the initial morphism η_A applied to $\eta_{A'}\circ\alpha : A\to\mathcal{F}(\hat{A}')$, there is a unique morphism $\hat{\alpha}:\hat{A}\to\hat{A}'$ such that $\eta_{A'}\circ\alpha = \mathcal{F}(\hat{\alpha})\circ\eta_A$, i.e. the diagram

$$\begin{array}{ccc} A & \xrightarrow{\ \alpha\ } & A' \\ {\scriptstyle\eta_A}\downarrow & \searrow^{\eta_{A'}\circ\alpha} & \downarrow{\scriptstyle\eta_{A'}} \\ \mathcal{F}(\hat{A}) & \xrightarrow{\ \mathcal{F}(\hat{\alpha})\ } & \mathcal{F}(\hat{A}') \end{array}$$

commutes. We define $\mathcal{G}(\alpha)=\hat{\alpha}$. We leave it as an exercise to verify that this defines indeed a functor $\mathcal{G}:\mathcal{D}\to\mathcal{C}$. The commutaivity of the above diagrams shows that the morphisms $\eta_A : A\to\mathcal{F}\circ\mathcal{G}(A)$ define a natural transformation $\eta:\mathrm{id}_{\mathcal{D}}\to\mathcal{F}\circ\mathcal{G}$.

For $A\in\mathrm{Ob}(\mathcal{D})$ and $B\in\mathrm{Ob}(\mathcal{C})$, we define the map

$$\begin{array}{rccc} \Phi_{A,B}: & \mathrm{Hom}_{\mathcal{C}}\big(\mathcal{G}(A),B\big) & \longrightarrow & \mathrm{Hom}_{\mathcal{D}}\big(A,\mathcal{F}(B)\big). \\ & \beta & \longmapsto & \mathcal{F}(\beta)\circ\eta_A \end{array}$$

Given a morphism $\alpha : A\to\mathcal{F}(B)$ in the image of $\Phi_{A,B}$, there is a unique morphism $\hat{\alpha}:\hat{A}\to B$ such that $\alpha = \mathcal{F}(\hat{\alpha})\circ\eta_A = \Phi_{A,B}(\hat{\alpha})$ by the universal property of the initial morphism $\eta_A : A\to\mathcal{F}(\hat{A})$ from A to $\mathcal{F}$. This shows that $\Phi_{A,B}$ is a bijection.

In order to show that $\mathcal{G}$ is a left adjoint to $\mathcal{F}$, we have to show that the diagram

$$\begin{array}{ccc} \mathrm{Hom}_{\mathcal{C}}\big(\mathcal{G}(A'),B\big) & \xrightarrow{\ \Phi_{A',B}\ } & \mathrm{Hom}_{\mathcal{D}}\big(A',\mathcal{F}(B)\big) \\ {\scriptstyle\beta\circ-\circ\mathcal{G}(\alpha)}\downarrow & & \downarrow{\scriptstyle\mathcal{F}(\beta)\circ-\circ\alpha} \\ \mathrm{Hom}_{\mathcal{C}}\big(\mathcal{G}(A),B'\big) & \xrightarrow{\ \Phi_{A,B'}\ } & \mathrm{Hom}_{\mathcal{D}}\big(A,\mathcal{F}(B')\big) \end{array}$$

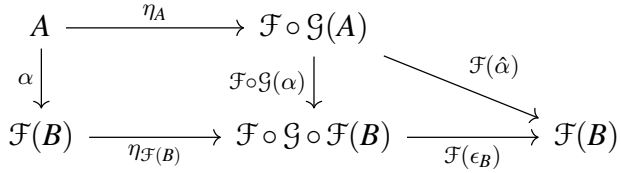
commutes for every pair of morphisms $\alpha : A \to A'$ in $\mathcal{D}$ and $\beta : B \to B'$ in $\mathcal{C}$. This follows from the identity

$$\begin{aligned}\Phi_{A,B'}\big(\beta \circ \gamma \circ \mathcal{G}(\alpha)\big) &= \mathcal{F}\big(\beta \circ \gamma \circ \mathcal{G}(\alpha)\big) \circ \eta_A \\ &= \mathcal{F}(\beta) \circ \mathcal{F}(\gamma) \circ \big(\mathcal{F} \circ \mathcal{G}(\alpha)\big) \circ \eta_A \\ &= \mathcal{F}(\beta) \circ \mathcal{F}(\gamma) \circ \eta_{A'} \circ \alpha \\ &= \mathcal{F}(\beta) \circ \Phi_{A',B}(\gamma) \circ \alpha\end{aligned}$$

for every morphism $\gamma : \mathcal{G}(A') \to B$ in $\mathcal{C}$ where we use that $\eta : \mathrm{id}_{\mathcal{D}} \to \mathcal{F} \circ \mathcal{G}$ is a natural transformation to conclude that $\big(\mathcal{F} \circ \mathcal{G}(\alpha)\big) \circ \eta_A = \eta_{A'} \circ \alpha$. This completes the proof the $\mathcal{G}$ is left adjoint to $\mathcal{F}$ and thus (2).

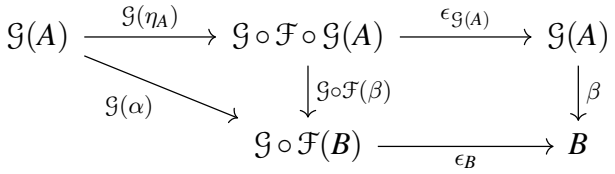
Assume (2). Let $\mathcal{G} : \mathcal{D} \to \mathcal{C}$ be a left adjoint to $\mathcal{F}$. Let $\eta : \mathrm{id}_{\mathcal{D}} \to \mathcal{F} \circ \mathcal{G}$ be the unit and $\epsilon : \mathcal{G} \circ \mathcal{F} \to \mathrm{id}_{\mathcal{C}}$ be the counit of the adjunction. Consider an object A in $\mathcal{D}$. We will verify the universal property of an initial morphism from A to $\mathcal{F}$ for $\hat{A} = \mathcal{G}(A)$ and the morphism $\eta_A : A \to \mathcal{F}(\hat{A})$.

Let B be an object of $\mathcal{C}$ and $\alpha : A \to \mathcal{F}(B)$ a morphism in $\mathcal{D}$. We define $\hat{\alpha} = \epsilon_B \circ \mathcal{G}(\alpha)$, and consider the diagram



whose square on the left hand side commutes by Lemma 7.3.3 and whose triangle on the right hand side commutes by the definition of $\hat{\alpha}$. By Proposition 7.3.4, we have $\mathcal{F}(\epsilon_B) \circ \eta_{\mathcal{F}(B)} = \mathrm{id}_{\mathcal{F}(B)}$ and thus $\alpha = \mathcal{F}(\epsilon_B) \circ \eta_{\mathcal{F}(B)} \circ \alpha = \mathcal{F}(\hat{\alpha}) \circ \eta_A$, which shows that $\hat{\alpha}$ satisfies the condition of the universal property of an initial morphism.

To show the uniqueness of $\hat{\alpha}$, let $\beta : \mathcal{G}(A) \to B$ be a morphism such that $\alpha = \mathcal{F}(\beta) \circ \eta_A$. Consider the diagram



whose triangle on the left hand side commutes by our assumption on β and whose square on the right hand side commutes by Lemma 7.3.3. By Proposition 7.3.4, we have $\epsilon_{\mathcal{G}(A)} \circ \mathcal{G}(\eta_A) = \mathrm{id}_{\mathcal{G}(A)}$, which implies that $\beta = \beta \circ \epsilon_{\mathcal{G}(A)} \circ \mathcal{G}(\eta_A) = \epsilon_B \circ \mathcal{G}(\alpha) = \hat{\alpha}$. This shows that $\hat{\alpha}$ is unique and that $\eta_A : A \to \mathcal{F}(\hat{A})$ is indeed an initial morphism from A to $\mathcal{F}$. Thus (1), which completes the proof of the theorem. □

Similarly, one can establish the following variant of Theorem 7.4.1, whose proof we omit.

Theorem 7.4.2. *Let $\mathcal{F} : \mathcal{C} \to \mathcal{D}$ be a functor. Then the following claims are equivalent.*

(1) Every object A in $\mathcal{D}$ has a terminal morphism from $\mathcal{F}$.

(2) The functor $\mathcal{F}$ has a right adjoint $\mathcal{G} : \mathcal{D} \to \mathcal{C}$.

Example 7.4.3. We illustrate Theorems 7.4.1 and 7.4.2 in the cases of the functors $\mathcal{F} : \mathcal{C} \to \mathcal{D}$ considered in Example 7.1.2.

(1) Let $\mathcal{F} : \mathrm{Rings} \to \mathrm{Sets}$ be the forgetful functor. As shown in Example 7.1.2.(1), every set A has an initial morphism to $\mathcal{F}$, namely the polynomial ring $\mathbb{Z}[T_i \mid i \in A]$ together with the map $\eta_A : A \to \mathbb{Z}[T_i \mid i \in A]$. Thus Theorem 7.4.1 shows that the functor $\mathcal{F}$ has a left adjoint $\mathcal{G} : \mathrm{Sets} \to \mathrm{Rings}$, which sends a set A to the ring $\mathcal{G}(A) = \mathbb{Z}[T_i \mid i \in A]$ and a map $\alpha : A \to B$ to the ring homomorphism $\mathcal{G}(\alpha) : \mathbb{Z}[T_i \mid i \in A] \to \mathbb{Z}[T_i \mid i \in B]$ that maps T_i to $T_{\alpha(i)}$.

(2) Let $\mathcal{D}$ be the category whose objects are pairs (A,S) of a ring A with a multiplicative subset S of A and whose morphisms $\alpha : (A,S) \to (A'S')$ are ring homomorphisms $\alpha : A \to A'$ with $\alpha(S) \subset S'$; cf. Example 7.1.2.(2). Let $\mathcal{F} : \mathrm{Rings} \to \mathcal{D}$ be the functor that sends a ring A to $(A, A^\times)$ and a ring homomorphism to itself.

By Theorem 7.4.1, $\mathcal{F}$ has a left adjoint $\mathcal{G}$, which sends an element (A,S) of $\mathcal{D}$ to the ring $S^{-1}A$ and that sends a morphism $\alpha : (A,S) \to (A',S')$ in $\mathcal{D}$ to the ring homomorphism $\alpha_S : S^{-1}A \to (S')^{-1}A'$ that maps $\frac{a}{s}$ to $\frac{\alpha(a)}{\alpha(s)}$.

(3) Let $\Delta : \mathrm{Rings} \to \mathrm{Rings} \times \mathrm{Rings}$ be the diagonal functor; cf. Example 7.1.2.(3). Since there is a terminal morphism from Δ to every object (A,B) of $\mathrm{Rings} \times \mathrm{Rings}$, Theorem 7.4.2 shows that Δ has a right adjoint $\Pi : \mathrm{Rings} \times \mathrm{Rings} \to \mathrm{Rings}$, which sends an object (A,B) in $\mathrm{Rings} \times \mathrm{Rings}$ to the ring $\Pi(A,B) = A \times B$ and a morphism $(\alpha,\beta) : (A,B) \to (A',B')$ to the ring homomorphism $(\alpha,\beta) : A \times B \to A' \times B'$.

Appendix A

Background and complementary topics

A.1 Zorn's Lemma

Zorn's Lemma is reformulation of the axiom of choice. In this lecture, we assume the validity of the axiom of choice, and we will use it in the form of Zorn's Lemma at a few instances. In the following, we introduce the necessary notions and formulate Zorn's Lemma.

Definition A.1.1. Let S be a set. A **partial order on** S is a relation $\leqslant$, which a subset R of $S \times S$ where we write $a \leqslant b$ if $(a,b) \in R$, that satisfies that

(1) $a \leqslant a$; (*reflexive*)

(2) $a \leqslant b$ and $b \leqslant a$ implies $a = b$; (*anti-symmetric*)

(3) $a \leqslant b$ and $b \leqslant c$ implies $a \leqslant c$ (*transitive*)

for all $a,b,c \in S$. A **total order on** S is a partial order such that

(4) $a \leqslant b$ or $b \leqslant a$ (*total*)

for all $a,b \in A$. A **partially ordered set** is a set S together with a partial order R. A **chain in** S is a totally ordered subset C of S, with respect to the restriction $R \cap (C \times C)$ of R to C. An **upper bound of** C **in** S is an element $b \in S$ such that $a \leqslant b$ for all $a \in C$. A **maximal element of** S is an element $a \in S$ such that $a \leqslant b$ implies $a = b$ for all $b \in S$.

Theorem A.1.2 (Zorn's Lemma)**.** *Let S be a partially ordered set. If every chain in S has an upper bound, then S has a maximal element.*

The typical situation where we apply Zorn's Lemma is that of a collection S of subsets of a set X, together with the partial order that is defined by inclusion, i.e. $A \leqslant B$ for $A,B \in S$ if and only if $A \subset B$ as subsets of X.

A.2 Topological spaces

Some exercises use the notion of topological spaces. We provide the necessary definitions in this section.

Definition A.2.1. Let X be a set. A **topology for** X is a collection $\mathcal{T}$ subsets of X that satisfies the following axioms.

(1) Both $\varnothing$ and X are in $\mathcal{T}$.

(2) Finite intersections of subsets in $\mathcal{T}$ are in $\mathcal{T}$.

(3) Arbitrary unions of subsets in $\mathcal{T}$ are in $\mathcal{T}$.

A **topological space** is a set X together with a topology $\mathcal{T}$, which we usual suppress from the notation. An **open subset of** X is an element of $\mathcal{T}$, and a **closed subset of** X is the complement of an open subset. A **basis for the** X is a subset $\mathcal{B}$ of $\mathcal{T}$ such that every open subset is the union of open subsets in $\mathcal{B}$.

Let X and Y be topological spaces. A **continuous map from** X **to** Y is a map $f : X \to Y$ such that $f^{-1}(U)$ is an open subset of X for every open subset U of Y.

Example A.2.2. A typical example is $\mathbb{R}$ together with the usual notion of open subsets, for which the collection of bounded open intervals $(a,b) = \{c \in | \ a < c < b\}$ (with $a,b \in \mathbb{R}$) forms a basis. A map $f : \mathbb{R} \to \mathbb{R}$ is continuous in the sense of Definition A.2.1 if and only if for every $\epsilon > 0$ there is a $\delta > 0$ such that $|f(a) - f(b)| < \delta$ for all $a,b \in \mathbb{R}$ with $|a-b| < \epsilon$. The latter characterization is the approach to continuous functions that one typically sees as a first definition in a course in Analysis. We leave the proof of these claims as an exercise.